МАТЕМАТИЧЕСКИЕ И КОМПЬЮТЕРНЫЕ

ОСНОВЫ КРИПТОЛОГИИ

Допущено Министерством образования Республики Беларусь в качестве учебного пособия для студентов математических и инженерно-технических специальностей высших учебных заведений

МИНСК ООО «НОВОЕ ЗНАНИЕ» 2003

УДК003.26:51:004(075.8)
ББК22.18я73+32.811.4я73
М34

Авторы:
Ю.С. Харин, В.И. Берник, Г.В. Матвеев, С.В. Агиевич

Рецензенты:
кафедра информатики Белорусского государственного университета информатики и радиоэлектроники (заведующий кафедрой — доктор физико-математических наук, профессор *Л.И. Минченко);*

доктор технических наук, профессор, директор РУП «Научно-исследовательский институт технической защиты информации» *В.Ф. Голиков*

М34 **Математические** и компьютерные основы криптологии: Учеб. пособие / Ю.С. Харин, В.И. Берник, Г.В. Матвеев, С.В. Агиевич. — Мн.: Новое знание, 2003. — 382 с.
ISBN 985-475-016-7.

Изложены математические основы криптографической защиты информации в компьютерных сетях и системах связи. Приведены описания многих известных компьютерных криптографических алгоритмов и анализ их свойств. Представлено более 300 упражнений и заданий для выполнения лабораторного практикума. Важным достоинством пособия является демонстрация органичной взаимосвязи математических основ шифрования и новейших компьютерных технологий защиты информации.

Для студентов и аспирантов, обучающихся по математическим и инженерно-техническим специальностям, слушателей факультетов переподготовки и повышения квалификации, а также для всех желающих познакомиться с основами криптологии.

УДК 003.26:51:004(075.8)
ББК 22.18я73+32.811.4я73

ISBN 985-475-016-7

© Коллектив авторов, 2003
© Оформление: ООО «Новое знание», 2003

SXJKN8EYHR3I4HR987YDWFUIHW9E8
YR384H5O34596EUF9O8WE76R83Y43
9846UHFE98YDSFJH8374H4BR8276R
UBNWDF87Y6234BN9WE872H3WRTJOJ
RTGM905JRO9G70954TOIRF7ERFH23

ОГЛАВЛЕНИЕ

SXJKN8EYHR3I4HR987YDWFUIHW9E8
YR384H5O34596EUF9O8WE76R83Y43
9846UHFE98YDSFJH8374H4BR8276R
UBNWDF87Y6234BN9WE872H3WRTJOJ
RTGM905JRO9G70954TOIRF7ERFH23

Предисловие 9

Основные обозначения 13

Часть 1. Математические основы криптологии

Глава 1. Докомпьютерная криптология 17

Задания 22

Глава 2. Арифметические основы 25

2.1. Алгоритм деления с остатком 25
2.2. Наибольший общий делитель 25
2.3. Взаимно простые числа 27
2.4. Наименьшее общее кратное 27
2.5. Простые числа 28
2.6. Сравнения 29
2.7. Классы вычетов 30
2.8. Функция Эйлера 30
2.9. Сравнения первой степени 32
2.10. Система сравнений первой степени 32
2.11. Первообразные корни 33
2.12. Существование первообразных корней 35
2.13. Индексы по модулям p^k и $2p^k$ 35
2.14. Символ Лежандра 36
2.15. Квадратичный закон взаимности 37
2.16. Символ Якоби 39
2.17. Цепные дроби 41
2.18. Подходящие дроби 42
2.19. Подходящие дроби в качестве наилучших приближений 44
2.20. Задания 46

Глава 3. Алгебраические основы 49

3.1. Понятие группы 49
3.2. Подгруппы групп 50
3.3. Циклические группы 52
3.4. Гомоморфизмы групп 54

3.5. Группы подстановок ... 55
3.6. Действие группы на множестве ... 57
3.7. Кольца и поля ... 58
3.8. Подкольца ... 59
3.9. Гомоморфизмы колец ... 61
3.10. Евклидовы кольца ... 62
3.11. Простые и максимальные идеалы ... 63
3.12. Конечные расширения полей ... 65
3.13. Поле разложения ... 67
3.14. Конечные поля ... 68
3.15. Порядки неприводимых многочленов ... 69
3.16. Линейные рекуррентные последовательности ... 70
3.17. Последовательности максимального периода ... 71
3.18. Задания ... 71

Глава 4. Введение в эллиптические кривые ... 79
4.1. Уравнение Вейерштрасса, дискриминант и j-инвариант ... 79
4.2. Сложение точек эллиптической кривой ... 82
4.3. Эллиптические кривые над конечными полями ... 84
4.4. Задания ... 88

Глава 5. Вероятностно-статистические модели сообщений и их энтропийные свойства ... 89
5.1. Источники дискретных сообщений и их вероятностные модели ... 89
5.2. Функционал энтропии и его свойства ... 91
5.3. Условная энтропия и ее свойства ... 93
5.4. Удельная энтропия стационарной символьной последовательности ... 99
5.5. Энтропийные характеристики марковских символьных последовательностей ... 103
5.6. Источники непрерывных сообщений и их энтропийные свойства ... 108
5.7. Оптимизация функционала энтропии на классе вероятностных распределений ... 117
5.8. Задания ... 121

Глава 6. Методы теории информации в криптологии ... 125
6.1. Асимптотические свойства стационарного источника дискретных сообщений ... 125
6.2. Энтропийная устойчивость случайных символьных последовательностей ... 130
6.3. Количество информации по Шеннону и его свойства ... 135
6.4. Шенноновские модели криптосистем ... 140
6.5. Теоретико-информационные оценки стойкости симметричных криптосистем ... 145
6.6. Задания ... 150

Глава 7. Статистическое тестирование случайных и псевдослучайных последовательностей ... 153
7.1. Равномерно распределенная случайная последовательность и ее свойства ... 153
7.2. Универсальный алгоритм статистического тестирования случайных и псевдослучайных последовательностей ... 156
7.3. Тест *n*-серий ... 158
7.4. Тест интервалов ... 159
7.5. Обобщенный покер-тест ... 159
7.6. Тест «собирателя купонов» ... 160
7.7. Тест перестановок ... 161
7.8. Тест пересекающихся *n*-грамм ... 162
7.9. Тест, основанный на рангах двоичных матриц ... 163
7.10. Спектральные тесты ... 164
7.11. Тесты случайного блуждания ... 167
7.12. Универсальный статистический тест Маурера ... 169
7.13. Тесты на основе приращений энтропии ... 170
7.14. Тест, основанный на алгоритме сжатия Лемпеля — Зива ... 172
7.15. Тест, основанный на линейной сложности ... 173
7.16. Тест на основе экстремальной статистики скалярного произведения ... 175
7.17. Тест на основе экстремальной статистики дельта-произведения ... 178
7.18. Алгоритмическое определение случайности ... 180
7.19. Задания ... 183

Часть 2. Математические и компьютерные методы криптологии

Глава 8. Алгоритмы генерации псевдослучайных последовательностей ... 193
8.1. Классификация алгоритмов генерации ... 193
8.2. Линейные и мультипликативные конгруэнтные генераторы ... 193
8.3. Нелинейные конгруэнтные генераторы ... 196
8.4. Рекурренты в конечном поле ... 197
8.5. Последовательности, порождаемые линейными регистрами сдвига с обратной связью (LFSR) ... 198
8.6. Генераторы Фибоначчи ... 200
8.7. Криптостойкие генераторы на основе односторонних функций ... 200
8.8. Криптостойкие генераторы, основанные на проблемах теории чисел ... 204
8.9. Методы «улучшения» элементарных псевдослучайных последовательностей ... 205
8.10. Комбинирование алгоритмов генерации методом Макларена — Марсальи ... 207
8.11. Комбинирование LFSR-генераторов ... 207
8.12. Конгруэнтный генератор со случайными параметрами ... 210
8.13. Задания ... 210

Глава 9. Поточные криптосистемы 215
9.1. Основные понятия 215
9.2. Рекуррентные последовательности 216
9.3. Линейные рекуррентные последовательности 216
9.4. Оценка параметров и распознавание ЛРП 218
9.5. Линейная сложность 219
9.6. Определение начального состояния ЛРП 220
9.7. Комбинирование последовательностей 221
9.8. Корреляционный криптоанализ 223
9.9. Задания 225

Глава 10. Математические модели стандартных блочных криптосистем 229
10.1. Криптосистема DES и ее свойства 229
10.2. Криптосистема IDEA 235
10.3. Криптосистема ГОСТ 28147—89 239
10.4. Общая структура алгоритма Rijndael 240
10.5. Использование алгебры полиномов в алгоритме Rijndael 241
10.6. Задания 243

Глава 11. Математические методы криптоанализа симметричных систем 245
11.1. Задачи и принципы криптоанализа 245
11.2. Метод «опробования» и его вычислительная сложность 247
11.3. Методы криптоанализа с использованием теории статистических решений 250
11.4. Разностный криптоанализ 261
11.5. Линейный криптоанализ 265
11.6. Задания 272

Глава 12. Криптосистемы с открытым ключом 275
12.1. Описание RSA-криптосистемы 275
12.2. Возможные атаки на криптосистему RSA 277
12.3. Стойкость RSA против метода повторного шифрования 278
12.4. Поиск секретного ключа d и факторизация модуля N 279
12.5. Биты в RSA-криптосистеме 281
12.6. Система Рабина 282
12.7. Рюкзачный метод шифрования 283
12.8. Стойкость рюкзачного шифра 284
12.9. Теорема М. Винера о малом секретном ключе 286
12.10. Арифметика больших чисел 287
12.11. Модулярная арифметика 288
12.12. Признаки простоты 290
12.13. Алгоритмы генерации простых чисел 291
12.14. Задача факторизации 297
12.15. Задания 300

Глава 13. Функции хэширования ... 305

13.1. Определение и свойства ... 305
13.2. Блочно-итерационные функции хэширования ... 306
13.3. Использование блочных криптосистем ... 306
13.4. Атака «дней рождения» ... 307
13.5. Криптосистемы аутентификации ... 309
13.6. Функция хэширования СТБ 1176.1—99 ... 309
13.7. Задания ... 313

Глава 14. Электронная цифровая подпись ... 315

14.1. Обобщенная модель ЭЦП ... 315
14.2. Схема ЭЦП Рабина ... 317
14.3. Схема Диффи — Лампорта ... 318
14.4. Вероятностная схема подписи Рабина ... 319
14.5. Стандарт ЭЦП DSS ... 320
14.6. Схема ЭЦП Эль-Гамаля ... 322
14.7. Арифметические свойства российского стандарта цифровой подписи ... 323
14.8. Эквивалентность задач фальсификации подписи в DSS и схеме Эль-Гамаля ... 329
14.9. Электронная цифровая подпись СТБ 1176.2—99 ... 330
14.10. Задача дискретного логарифмирования ... 332
14.11. Задания ... 333

Глава 15. Эллиптические кривые в криптографии ... 335

15.1. Цифровая подпись на эллиптических кривых ... 335
15.2. Особенности скалярного умножения на эллиптических кривых ... 338
15.3. Вычисление порядка эллиптической кривой ... 338
15.4. Задания ... 341

Глава 16. Протоколы управления криптографическими ключами ... 343

16.1. Протоколы генерации ключей ... 343
16.2. Протоколы взаимной аутентификации ... 344
16.3. Протоколы прямого обмена ключами ... 346
16.4. Протоколы распределения сеансовых ключей с использованием центра распределения ключей ... 348

Глава 17. Новые направления в криптологии ... 351

17.1. Возможности квантовой криптографии ... 351
17.2. Математическое разделение секрета ... 354
17.3. Стеганография и ее применение ... 356
17.4. Активный криптоанализ ... 358

Приложение А. Таблицы 363

Приложение Б. Архив дискретных последовательностей 365

Приложение В. Компьютерный практикум «КриптоЛаборатория» 367

Общие сведения 367
Рабочий план «КриптоЛаборатории» 367
Инструменты «КриптоЛаборатории» 369
Задания «КриптоЛаборатории» 370

Литература 371

Предметный указатель 379

SXJKN8EYHR3I4HR987YDWFUIHW9E8
YR384H5O34596EUF9O8WE76R83Y43
9846UHFE98YDSFJH8374H4BR8276R
UBNWDF87Y6234BN9WE872H3WRTJOJ
RTGM905JRO9G70954TOIRF7ERFH23

ПРЕДИСЛОВИЕ

SXJKN8EYHR3I4HR987YDWFUIHW9E8
YR384H5O34596EUF9O8WE76R83Y43
9846UHFE98YDSFJH8374H4BR8276R
UBNWDF87Y6234BN9WE872H3WRTJOJ
RTGM905JRO9G70954TOIRF7ERFH23

Среди способов защиты информации наиболее важным считается криптографический. Он предусматривает такое преобразование информации, при котором она становится доступной для прочтения лишь обладателю некоторого секретного параметра (ключа). В последние годы область применения криптографии значительно расширилась. Ее стали повседневно использовать многие организации, коммерческие фирмы, частные лица. При этом законного пользователя того или иного криптографического средства прежде всего беспокоит его надежность. Одним из способов оценки надежности является попытка «взлома», т.е. получение доступа к информации без знания ключа. Подобные задачи призвано решать смежное научное направление, называемое криптоанализом. Криптоанализ и криптография объединены общим названием — криптология.

Длительное время криптографическая защита информации применялась лишь в государственных интересах, и поэтому криптология была чисто ведомственной, закрытой наукой. В последние годы положение коренным образом изменилось. Одновременно с расширением сферы применения произошел качественный сдвиг в развитии. В криптологии все чаще стали применяться результаты из теории информации, теории чисел, алгебры и математической статистики. Одновременно она становится все более университетской наукой. *Криптология* — это раздел прикладной математики, в котором изучаются модели и методы, алгоритмы, программные и аппаратные средства криптографической защиты информации, а также оценки ее эффективности.

За рубежом чтение курсов по криптологии уже более двадцати лет ведется студентам, специализирующимся в области математики, прикладной математики, информатики и электроники. В последние годы эти курсы в обязательном порядке включаются в учебные планы подготовки всех специалистов, деятельность которых связана с информационными процессами. В России с 1995 г. ведется подготовка специалистов по четырем специальностям: 075100 — Криптография (с квалификацией «Математик»), 075200 — Компьютерная безопасность («Математик»), 075500 — Комплексное обеспечение информационной безопасности автоматизированных систем («Специалист по защите информации»), 075600 — Информационная безопасность телекоммуникационных систем («Специалист по защите информации»). В Республике Беларусь в 1997 г. начата подготовка специалистов по прикладной математике с новой специализацией Н.08.01.14 — Математическое и программное обеспечение криптографии и анализа данных (квалификация «Математик-программист»).

В пособии изложены математические и компьютерные основы криптографической защиты информации в компьютерных сетях и системах связи. Учебное пособие состоит из семнадцати глав, объединенных в две части, трех приложений и списка литературы.

Часть 1 включает семь глав, представляющих математические основы криптологии. Глава 1 является вводной и содержит основные понятия криптологии, ее историю и описание элементарных криптосистем. В главах 2 и 3 излагаются арифметические и алгебраические основы криптологии, а в главе 4 — введение в эллиптические кривые, которые активно применяются в криптологии. Вероятностно-статистические модели сообщений и их энтропийные свойства изучаются в главе 5. Глава 6 посвящена изложению классического теоретико-информационного подхода К. Шеннона к проблемам криптологии. Важные для теории и практики проблемы статистического тестирования случайных и псевдослучайных последовательностей рассмотрены в главе 7.

Часть 2 посвящена математическим и компьютерным методам криптологии и состоит из десяти глав. В главе 8 излагаются алгоритмы генерации псевдослучайных последовательностей, а в главе 9 эти алгоритмы и их комбинации используются для построения поточных криптосистем. В главе 10 дается описание современных стандартных блочных криптосистем и анализируются их свойства. Глава 11 посвящена изложению широко применяемых за рубежом методов криптоанализа симметричных криптосистем. Криптосистемы с открытым ключом и их надежность изучаются в главе 12. В главе 13 содержится описание и анализ свойств функций хэширования. Глава 14 посвящена системам электронной цифровой подписи, а в главе 15 рассматривается подход к построению электронной цифровой подписи, основанный на эллиптических кривых. В главе 16 изучаются важные вопросы политики безопасности использования криптосистем. Глава 17 посвящена новым направлениям в криптологии.

В приложении А приводятся вспомогательные таблицы, в приложении Б — архив дискретных последовательностей для выполнения заданий компьютерного практикума, а в приложении В описывается специально разработанный ППП «КриптоЛаборатория» и методика его использования для проведения компьютерного практикума по криптологии.

Учебное пособие снабжено обширным списком литературы, необходимой не только для учебной, но и для научно-исследовательской работы студентов и аспирантов.

Главы 1 (совместно с Г.В. Матвеевым), 5—8, 11, 17 (совместно с Г.В. Матвеевым) написал Ю.С. Харин; главы 1 (совместно с Ю.С. Хариным), 2, 3, 10, 12 (совместно с В.И. Берником), 17 (совместно с Ю.С. Хариным) — Г.В. Матвеев; главы 4, 12 (совместно с Г.В. Матвеевым), 14, 15 — В.И. Берник; главы 9, 13 — С.В. Агиевич; главу 16 — В.А. Галинский.

Авторы благодарят зам. директора ННИЦ ППМИ В.А. Галинского за подготовку главы 16 и сотрудников ННИЦ ППМИ, ГЦБИ, НИИ ТЗИ за поддержку и советы.

Авторы будут признательны всем, кто поделится своими соображениями по совершенствованию данного пособия.

Предложения и замечания направляйте по адресу:

Кафедра математического моделирования и анализа данных
Факультет прикладной математики и информатики
Белорусский государственный университет
пр. Ф. Скорины, 4
220050, Минск
Республика Беларусь

Телефоны: (375-17) 226-57-04, 209-50-71
Факс: (375-17) 226-55-48, 209-50-54
E-mail: kharin@bsu.by

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3

$|A|$ — модуль числа A, или мощность множества A, или определитель матрицы A

$S(A)$ — симметрическая группа, действующая на множестве A

S_n — симметрическая группа степени n

id — тождественная подстановка

$\mathbb{N}$ — множество натуральных чисел

$\mathbb{Z}$ — кольцо целых чисел

$\mathbb{R}$ — поле вещественных чисел

$\mathbb{Z}_n = \mathbb{Z}/(n)$ — кольцо классов вычетов целых чисел по модулю n

$\mathbb{F}_q$ — конечное поле из q элементов

V_n — n-мерное векторное пространство над полем $\mathbb{F}_2$

$\mathcal{F}_n$ — множество всех булевых функций от n переменных, т.е. отображений $V_n \to \mathbb{F}_2$

$\oplus$ — операция сложения в поле $\mathbb{F}_2$

$\bar{v}$ — число $2^{n-1}v_1 + 2^{n-2}v_2 + \ldots + v_n$ для вектора $v = (v_1, v_2, \ldots, v_n) \in V_n$

$\boxplus$ — операция сложения в V_n по правилу: $c = a \boxplus b$, если $\bar{c} = (\bar{a} + \bar{b}) \bmod 2^n$

$v \lll d$ — циклический сдвиг компонентов вектора $v \in V_n$ на d позиций влево

$\|$ — конкатенация

$[x], \lfloor x \rfloor$ — целая часть числа $x \in \mathbb{R}$, наибольшее целое $n \le x$

$\lceil x \rceil$ — наименьшее целое $n \ge x$

$\varphi(n)$ — функция Эйлера

$\mu(n)$ — функция Мёбиуса

$(a,b) = \text{НОД}(a,b)$ — наибольший общий делитель чисел a и b

$[a,b] = \text{НОК}(a,b)$ — наименьшее общее кратное чисел a и b

$\text{ind}_a b$ — индекс (дискретный логарифм) b по основанию a

$\mathbf{1}(x)$ — единичная функция Хэвисайда

$\mathbf{I}\{A\}$ — индикатор наступления события A

$\mathbf{1}_A\{x\} = \mathbf{I}\{x \in A\}$ — индикаторная функция множества A

$\delta_{ij} = \mathbf{I}\{i = j\}$ — символ Кронекера

$\mathbf{P}\{A\}$ — вероятность наступления случайного события A

$\mathbf{E}\{\xi\}$ — математическое ожидание случайной величины ξ

$\mathbf{D}\{\xi\}$ — дисперсия случайной величины ξ

$\mathbf{Cov}\{\xi,\eta\}$ — ковариация случайных величин ξ,η

$\mathfrak{L}\{\xi\}$ — закон распределения вероятностей случайной величины ξ

$\xrightarrow{\text{п н}}$ — сходимость «почти наверное»

$\xrightarrow{\text{P}}$ — сходимость по вероятности

$O(\varepsilon)$, $o(\varepsilon)$ — символы Ландау

$\binom{n}{m}$ — $\dfrac{n!}{m!(n-m)!}$, если $0 \le m \le n$, и 0 — в противном случае

$u \leftarrow a$ — присвоить значение a переменной u

$u \leftrightarrow v$ — поменять местами значения переменных u и v

$\square$ — конец доказательства

Теоремы, леммы, следствия, замечания, формулы и задания имеют автономную тройную нумерацию Г.П.О, где Г — номер главы, П — номер параграфа, О — номер объекта. Заданиям текущего параграфа дается одинарный номер О. Используется двойная автономная нумерация Г.О таблиц и рисунков.

kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654ughet
r873yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr98374h34yr
98e76d9w8yedyt8354753griewgfr87we68r73yt4irg384687efiuweghr834rh
iufh8e7r63gh4fiye86r8re7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4r
hw9e76r98e6r98y3i4u5h3 9kdfjhgjdhrgjkldfkvjsier
tuesrlojeoiruyt9866593 74uhreifhei654ughetr873y
f5u3eiry9rewyejhrf9879 4358yoihoir7f84uy5ihjlie
hjr98374h34yr98e76d9w8 yedyt8354753griewgfr87we
68r73yt4irg384687efiuweghr834rhiufh8e7r63gh4fiye86r8re7yrfiuehri
t763549y43iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h394kdfjhgjdh
rgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654ughetr873yf5u3
eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr98374h34yr98e76d9w8
yedyt835 4753griewgf
r87we68r 73yt4irg384
687efiuw eghr834rhiu
fh8e7r63 gh4fiye86r8
re7yrfiu ehrit763549
y43iurhi euyf98e4i5u
h3i4rhw9 e76r98e6r98
y3i4u5h394kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifh
ei654ughetr873yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr
98374h34yr98e76d9w8yedyt8354753griewgfr87we68r73yt4irg384687efiu
weghr834rhiufh8e7r63gh4fiye8r8re7yrfiuehrit763549y43iurhieuyf98e
4i5uh3i4rhw9e76r98e6r98y3i4u5h394kdfjhgjdhrgjkldfkvjsiertuesrloj
eoiruyt986659374uhreifhei654ughetr873yf5u3eiry9rewyejhrf98794358
yoihoir7f84uy5ihjliehjr98374h34yr98e76d9w8yedyt8354753griewgfr87
we68r73yt4irg384687efiuweghr834rhiufh8e7r63gh4fiye86r8re7yrfiueh
rit763549y43iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h394kdfjhgj
dhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654ughetr873yf5
u3eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr98374h34yr98e76d9
w8yedyt8354753griewgfr87we68r73yt4irg384687efiuweghr834rhiufh8e7
r63gh4fiye86r8re7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4rhw9e76r
98e6r98y3i4u5h394kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986659374
uhreifhei654ughetr873yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5ih
jliehjr98374h34yr98e76d9w8yedyt8354753griewgfr87we68r73yt4irg384
687efiuweghr834rhiufh8e7r63gh4fiye86r8re7yrfiuehrit763549y43iurh
ieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h394kdfjhgjdhrgjkldfkvjsier
tuesrlojeoiruyt986659374uhreifhei654ughetr873yf5u3eiry9rewyejhrf
98794358yoihoir7f84uy5ihjliehjr98374h34yr98e76d9w8yedyt8354753gr
iewgfr87we68r73yt4irg384687efiuweghr834rhiufh8e7r63gh4fiye86r8re
7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h39
4kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654ughe
tr873yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5iwgfr87we68r73yt4i
rg384687efiuweghr834rhiufh8e7r63gh4fiye86r8re7yrfiuehrit763549y4
3iurhieuyf98e4i5uh3i4rhw9eflkgnoernerfpewcokwe75tkngfery65i54ht9
87rfyh6r98e6r98y3i4u5h394kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt9
86659374uhreiierugeriugeuggiuwqefhei654ughetr873yf5u3eiry9rewyej
hrf98794358yoihoir7f84uy5ihjliehjr98374h34yr98e76d9w8yedyt835475
3griewgfr87we68r73yt4irg384687efiuweghr834rrhw9e76r98e6r98y3i4u5
h394kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654u
ghetr873yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr98374h
34yr98e76d9w8yedyt8354753griewgfr87we68r73yt4irg384687efiuwe63gh

Часть 1

МАТЕМАТИЧЕСКИЕ ОСНОВЫ КРИПТОЛОГИИ

SXJKN8EYHR3I4HR987YDWFUIH
W9E8YR384H5O34596EUF9O8WE
76R83Y439846UHFE98YDSFJH8
374H4BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM905JR
O9G70954TOIRF7ERFH2347FQ7
WERF9UJ1349YUR7DFKJHGJDFB
VDGFJVBDSKFJHKJFVKBDFKJVJ
KGFKJSBDFJBVKJDSGFJKVBKSD
JBFKVBDGFVAKDLKWEF90UQ3I4

Глава 1

ДОКОМПЬЮТЕРНАЯ КРИПТОЛОГИЯ

SXJKN8EYHR3I4HR987YDWFUIH
W9E8YR384H5O34596EUF9O8WE
76R83Y439846UHFE98YDSFJH8
374H4BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM905JR
O9G70954TOIRF7ERFH2347FQ7
WERF9UJ1349YUR7DFKJHGJDFB
VDGFJVBDSKFJHKJFVKBDFKJVJ
KGFKJSBDFJBVKJDSGFJKVBKSD
JBFKVBDGFVAKDLKWEF90UQ3I4

Защита информации — это всевозможные средства и функции, обеспечивающие доступность, конфиденциальность или целостность информации или связи, исключая средства и функции, предохраняющие от неисправностей. Защита информации включает в себя криптографию, криптоанализ, защиту от собственного излучения и защиту (компьютера) от несанкционированного доступа.

Криптография — это раздел прикладной математики, изучающий модели, методы, алгоритмы, программные и аппаратные средства преобразования информации (шифрования) в целях сокрытия ее содержания, предотвращения видоизменения или несанкционированного использования.

Криптосистема — это система, реализованная программно, аппаратно или программно-аппаратно и осуществляющая криптографическое преобразование информации.

Криптоанализ — это раздел прикладной математики, изучающий модели, методы, алгоритмы, программные и аппаратные средства анализа криптосистемы или ее входных и выходных сигналов с целью извлечения конфиденциальных параметров, включая открытый текст.

Из данных определений видно, что криптоанализ занимается задачами, которые в математическом смысле обратны задачам криптографии. Система криптографии и криптоанализа образует новую науку *криптологию*.

В развитии криптологии выделяют три этапа. *Первый этап* (с древних времен до 1949 г.) характеризуется частными, узкоспециальными и вычислительно простыми алгоритмами криптографии и криптоанализа без использования компьютеров. Его часто называют этапом докомпьютерной криптографии. *Второй этап* (1949—1976 гг.) принято отсчитывать с момента публикации работы американского математика-прикладника К. Шеннона «Теория связи в секретных системах». В этот период активно проводились систематические исследования по криптологии с использованием компьютера. Криптология становится математической наукой. Однако потребителями результатов криптологии являлись лишь службы связи и информации в дипломатических и военных организациях, поэтому криптология являлась «закрытой» наукой. *Третий этап* (1976 г. — настоящее время), который можно назвать этапом открытой криптологии, принято отсчитывать с момента публикации работы американских математиков У. Диффи, М. Хеллмана «Новые направления в криптографии».

В этой работе показано, что «секретная» передача информации возможна (в отличие от результатов К. Шеннона) без предварительной передачи «секретного ключа». Главной особенностью этого этапа становится массовое применение криптографии в банковском деле, электронной торговле, компьютерных сетях (например, в Интернете) и других приложениях.

Во многих странах в развитие криптологии вкладываются значительные государственные средства. Например, в США ежегодные расходы на криптологию составляют около 15 млрд долларов. Характерно также, что криптология стимулирует развитие информатики.

Современная криптология широко использует теорию вероятностей, математическую статистику, алгебру, теорию чисел, теорию алгоритмов и сложности вычислений, а сам процесс шифрования осуществляется на мощных специализированных компьютерах, называемых *устройствами шифрования*. На Западе известны устройства *B-CRYPT*, *IBM-4755*, *Datacryptor* и др.

Приведем ряд интересных фактов из истории развития криптологии. При раскопках в Месопотамии был найден относящийся к XX в. до н.э. один из самых древних шифртекстов. Он был написан клинописью на глиняной дощечке и содержал коммерческую тайну: рецепт глазури для покрытия гончарных изделий. В XVII в. кардинал Ришелье создал первую шифрослужбу. Отметим, что задачами, имеющими непосредственное отношение к криптологии, занимались известные математики Ньютон, Лейбниц, Эйлер, Гаусс, Кардано и др.

В криптологии общеприняты следующие понятия.

Пространство сообщений PT — множество всевозможных сообщений pt (plaintext, message). Для сообщения используется также обозначение m.

Пространство ключей K. Каждый ключ $k \in K$ определяет некоторую подстановку E_k (encryption) на пространстве PT и обратное преобразование D_k (decryption).

Пространство зашифрованных сообщений CT, состоящее из зашифрованных сообщений ct (ciphertext), $ct = E_k(pt)$. Используется также обозначение c.

Обычно к криптосистеме предъявляются требования: 1) $E_k(pt)$, $D_k(ct)$ — легко вычислимы; 2) не зная k, невозможно вычислить pt при известном ct.

В классических криптосистемах «секретный ключ» k определяет преобразования E_k и D_k, причем справедливо тождество $D_k(E_k(pt)) = D_k(ct) = pt$.

Основной задачей специалиста по криптоанализу (криптоаналитика) является поиск ключа k. При этом ему могут предоставиться, в частности, следующие возможности для атаки: 1) он знает лишь зашифрованный текст (Ciphertext only attack); 2) известен незашифрованный и зашифрованный текст (Known plaintext attack); 3) имеется возможность получить пару $(pt, E_k(pt))$, где pt выбрано криптоаналитиком (Chosen plaintext attack).

Криптография и криптоанализ развиваются параллельно. Криптографы всегда пытаются создать такую криптосистему, которая была бы стойкой ко всем известным методам криптоанализа.

Остается уточнить понятие текста. При этом обычно фиксируют некоторую систему Σ символов, называемую *алфавитом*. Это может быть английский, русский или какой-либо другой алфавит. Часто в качестве алфавита используются

натуральные числа или символы 0 и 1. *Словом* называется упорядоченный набор букв данного алфавита. Множество слов обозначают через Σ^*. *Текст* — набор слов. Например, все, что написано далее, является текстом.

Одним из самых старых шифров является шифр Юлия Цезаря. При шифровании с его помощью каждая буква латинского алфавита сдвигается циклически вправо на $k = 3$ позиций. Таким образом, имеем подстановку (замену):

A	B	C	D	E	F	G	H	...	T	U	V	W	X	Y	Z
D	E	F	G	H	I	J	K	...	W	X	Y	Z	A	B	C

Шифрование осуществляется в соответствии с этой подстановкой. Например, $E_3(\text{CUT}) = \text{FXW}$. Понятно, что выбор $k = 3$ не является единственно возможным. При других ключах k имеем

$$E_{25}(\text{IBM}) = \text{HAL}, \quad E_6(\text{IDMUP}) = \text{OJSAV}.$$

Нетрудно показать, что $D_k = E_{26-k}$, $D_k E_k = E_0 = D_0$, а ключ k определен по модулю 26.

Шифр Цезаря является примером *шифра подстановки,* или *замены.* Его еще называют *шифром простой подстановки.* Криптоанализ такого шифра очень прост. Дело в том, что для любого современного языка вычислены частотные характеристики букв, т.е. относительные частоты их появления в «нормальных» текстах. Приведем эти частоты в процентах (с упорядочением) для английского языка:

Высокий		Средний		Низкий	
E	12.31	L	4.03	B	1.62
T	9.59	D	3.65	G	1.61
A	8.05	C	3.20	V	0.93
O	7.94	U	3.10	K	0.52
N	7.19	P	2.29	Q	0.20
I	7.18	F	2.28	X	0.20
S	6.59	H	2.25	J	0.10
R	6.03	W	2.03	Z	0.09
H	5.14	Y	1.88		

Можно сказать, что при шифровании простой заменой буквы текста заменяются буквами этого или другого алфавита в соответствии с некоторой подстановкой.

Еще одним примером шифра простой замены является *модулярный шифр.* Выберем число a, взаимно простое с модулем $m = 26$. Пусть p — буква английского алфавита, отождествленная со своим порядковым номером (0, 1, ..., 25). Тогда $E_a(p) = ap + k(\text{mod}\, m)$, где k — фиксировано. В этом случае ключом явля-

ется пара чисел (a, k). Условие взаимной простоты необходимо для обратимости шифра. Конечно, буквы можно заменять и какими-то другими символами. К семейству шифров замены относятся гомофонические, полиграммные и многоалфавитные шифры.

Гомофоническое шифрование — один из способов защиты от частотной криптоатаки. Каждая буква текста шифруется несколькими символами этого или другого алфавита. Число этих символов пропорционально частотной характеристике шифруемой буквы. Например:

Буква	Гомофония								
A	17	19	34	41	56	60	67	83	
I	08	22	53	65	88	90			
L	03	44	76						
N	02	09	15	27	32	40	59		
O	01	11	23	42	54	70	80		
P	33	91							
T	05	10	20	29	45	58	64	78	99

Одна из возможностей зашифровать текст PLAIN PILOT: 91 44 56 65 59 33 08 76 28 78.

При *полиграммном шифровании* заменяются не буквы текста, а их комбинации. Если заменяются пары букв, то имеем *биграммное шифрование*. Примером биграммного шифрования является шифр Плейфера. Образуем из английского алфавита какой-нибудь квадрат 5×5 и будем хранить его, как всякий ключ, в секрете. Например:

H	A	R	P	S
I	C	O	D	B
E	F	G	K	L
M	N	Q	T	U
V	W	X	Y	Z

Здесь буква J не употребляется или отождествляется с буквой I.

Замена биграмм проводится по правилам:

1) если m_1 и m_2 находятся в одной строке, то биграмма m_1m_2 шифруется биграммой c_1c_2, где буквы c_1 и c_2 являются правыми соседями букв m_1 и m_2 соответственно; если правого соседа нет, то берется первая буква строки;

2) если m_1 и m_2 — в одном столбце, то берутся нижние соседи с аналогичной оговоркой;

3) если $m_1 = m_2$, то в незашифрованном тексте между ними вставляется незначащая буква (например, X);

4) при нечетном количестве букв в незашифрованном тексте к нему дописывается незначащая буква;

5) в наиболее вероятном случае, когда m_1 и m_2 расположены в разных столбцах и строках, c_1 и c_2 выбираются, как показано на схеме:

$$\begin{array}{ccccccccccc} m_1 & \cdots & c_1 & c_2 & \cdots & m_2 & c_1 & \cdots & m_1 & m_2 & \cdots & c_2 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ c_2 & \cdots & m_2 & m_1 & \cdots & c_1 & m_2 & \cdots & c_2 & c_1 & \cdots & m_1 \end{array}$$

Покажем это на примере:

$$\begin{array}{rcccccc} m & = & \text{RE} & \text{NA} & \text{IS} & \text{SA} & \text{NC} & \text{EX} \\ E(m) & = & \text{HG} & \text{WC} & \text{BH} & \text{HR} & \text{WF} & \text{GV} \end{array}$$

Еще одна биграммная криптосхема, принадлежащая Хиллу, основана на линейной алгебре. Осуществим цифровую кодировку букв английского алфавита: A = 0, B = 1, C = 2, ..., Z = 25. Выберем какую-нибудь обратимую по модулю 26 квадратную матрицу M порядка 2. Это — ключ. Пусть, например,

$$M = \begin{pmatrix} 2 & 5 \\ 3 & 3 \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} 17 & 15 \\ 9 & 20 \end{pmatrix}.$$

Биграммы будем записывать в виде матриц-столбцов. Например:

$$P_1 = \begin{pmatrix} \text{H} \\ \text{E} \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \end{pmatrix}, \quad P_2 = \begin{pmatrix} \text{L} \\ \text{P} \end{pmatrix} = \begin{pmatrix} 11 \\ 15 \end{pmatrix}.$$

Шифрование биграмм определим формулой $C = MP$. Зашифруем, для примера, слово $m = P_1P_2$: $m = \text{HELP}$, $c = \text{IHUX}$.

К биграммному шифрованию, так же как и к шифру простой подстановки, применима частотная криптоатака. Приведем в связи с этим наиболее распространенные английские биграммы с указанием процентов их встречаемости:

TH	6.3	AK	2.0	HA	1.7
IN	3.1	EN	2.0	OU	1.4
ER	2.7	TI	2.0	IT	1.4
RE	2.5	TE	1.9	ES	1.4
AN	2.2	AT	1.8	ST	1.4
HE	2.2	ON	1.7	OR	1.4

При *многоалфавитном подстановочном шифровании* задается d шифров простой подстановки, определяемых функциями $f_1, f_2, \ldots, f_d$, а сообщение $m = m_1, m_2, \ldots, m_d, m_{d+1}, \ldots, m_{2d}, \ldots$ шифруется по правилу

$$E_k(m) = f_1(m_1), f_2(m_2), \ldots, f_d(m_d), f_1(m_{d+1}), \ldots, f_d(m_{2d}), \ldots$$

К таким шифрам относится шифр Виженера. Ключ образуется последовательностью букв $k_1, k_2, \ldots, k_d$, при этом для буквы a i-го алфавита функция $f_i(a) = (a + k_i)(\mathrm{mod}\, m)$. Например:

$$\begin{aligned} m &= \text{RENA} \quad \text{ISSA} \quad \text{NCE} \\ k &= \text{BAND} \quad \text{BAND} \quad \text{BAN} \\ E_k(m) &= \text{SEAD} \quad \text{JSFD} \quad \text{OCR} \end{aligned}$$

Наряду с подстановочными шифрами известны так называемые *перестановочные* (*транспозиционные*) шифры. При этом буквы сообщения остаются прежними, но меняют свое расположение в тексте. Приведем два примера.

Сообщение можно разбить на группы букв, скажем, по три буквы, а затем в каждой группе сделать одну и ту же перестановку. Например:

LETUSGOTOLONDON → ETLSGUTOOONLOND.

То же сообщение можно записать в прямоугольнике 3 × 5:

L	E	T	U	S
G	O	T	O	L
O	N	D	O	N

Затем можно переписать его по столбцам: LGOEONTTDUOOSLN, что и будет криптограммой (шифртекстом).

В заключение отметим еще один шифр, предложенный Вернамом. Сообщение m обычно записывают в виде последовательности нулей и единиц. Длина ключа k равна длине сообщения. Шифрование состоит в применении к m и k *операции XOR*, $E_k(m) = m \oplus k$. Очевидно, $D_k = E_k$, так как $(m \oplus k) \oplus k = m \oplus (k \oplus k) = m$. Шифр Вернама считается практически нераскрываемым, так как данное сообщение с помощью подбора ключа, к сожалению, слишком большого, можно преобразовать в любое другое. Основная проблема состоит в хранении и передаче ключа.

В современной компьютерной криптографии используются многие перечисленные шифры как составные части сложных криптосистем, рассматриваемых в главах 9, 10, 12—16.

Задания

1. Сколько модулярных шифров $p \to ap + k \pmod{26}$ существует в английском алфавите?

2. Показать, что для обратимости модулярного шифра необходимо и достаточно условие взаимной простоты $(a, m) = 1$. Как найти обратное преобразование? Будет ли оно модулярным шифром?

3. Описать алгоритм D для шифра Плейфера.

4. Описать алгоритм D для шифра Виженера и найти соответствующий ключ.

5. Какие докомпьютерные шифры являются групповыми (замкнуты относительно операции повторного шифрования), т.е. для любых двух ключей k_1, k_2 существует ключ k_3 такой, что $E_{k_2} E_{k_1} = E_{k_3}$.

6. Пусть словами являются булевы n-ки, т.е. $m \in V_n$. Сформулировать условие перестановочности шифров перестановки и Вернама.

7. Показать, что шифр перестановки является линейным преобразованием в пространстве V_n.

8. Пусть P — криптопреобразование простой перестановки. При каком условии верно равенство $P^2 = id$?

9. Что для криптографии означает условие $E^2 = id$?

10. Пусть P — перестановка, а V — преобразование Вернама. Какая система шифрования предпочтительнее: PV или VP?

11. При каком условии перестановочное шифрование с помощью прямоугольника будет тождественным?

12. Сколько всего существует нелинейных криптопреобразований E: $V_3 \to V_3$?

SXJKN8EYHR3I4HR987YD
WFUIHW9E8YR384H5O345
96EUF9O8WE76R83Y4398
46UHFE98YDSFJH8374H4
BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7E
RFH2347FQ7WERF9UJ134

Глава 2

АРИФМЕТИЧЕСКИЕ ОСНОВЫ

SXJKN8EYHR3I4HR987YD
WFUIHW9E8YR384H5O345
96EUF9O8WE76R83Y4398
46UHFE98YDSFJH8374H4
BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7E
RFH2347FQ7WERF9UJ134

2.1. Алгоритм деления с остатком

Множество всех натуральных чисел 1, 2, 3, ... будем обозначать через $\mathbb{N}$, а через $\mathbb{Z}$ — множество целых чисел 0, ±1, ±2, ...

Пусть a, b — элементы из $\mathbb{Z}$, $a \neq 0$. Говорят, что *a делит b* (записывают $a \mid b$), если существует такое целое c, что $b = ac$. Это отношение рефлексивно, транзитивно, но не симметрично. Легко доказываются следующие свойства.

Свойство 2.1.1. $a \mid b,\ a \mid c \Rightarrow a \mid b \pm c$.

Свойство 2.1.2. $a \mid b \Rightarrow a \mid bc,\ c \in \mathbb{Z}$.

Свойство 2.1.3. $a \mid b,\ b \mid a \Leftrightarrow a = \pm b$.

Теорема 2.1.1. Пусть a — целое, b — натуральное. Тогда существуют такие однозначно определенные $q, r \in \mathbb{Z}$, $0 \le r < b$, что $a = bq + r$.

Доказательство. Возьмем наибольшее bq, не превосходящее a. Положим $r = a - bq$. Условие $b(q+1) > a$ влечет $r < b$. Допустив также, что $a = bq_1 + r_1$, $0 \le r_1 < b$, получим $0 = b(q - q_1) + r - r_1$. Отсюда следует, что $r - r_1$ кратно b. Но так как $|r - r_1| < b$, то последнее возможно лишь при $r - r_1 = 0$, т.е. при $r = r_1$, откуда вытекает также $q = q_1$. □

Теорема справедлива для любого целого $b \neq 0$ при условии, что ограничение $r < b$ заменяется на $r < |b|$.

2.2. Наибольший общий делитель

Всякое целое, делящее числа a и b, называется их *общим делителем*. Наибольший из общих делителей для чисел a и b называется *наибольшим общим делителем* (НОД) и обозначается (a, b). Ввиду конечности числа делителей данного числа существование и единственность наибольшего общего делителя очевидны. Если $(a,b) = 1$, то числа a и b называются *взаимно простыми*. Найдем НОД двух чисел.

Лемма 2.2.1. $b \mid a \Rightarrow (a,b) = b$.

Доказательство. Легко видеть, что совокупность общих делителей a и b совпадает с совокупностью делителей b. □

Лемма 2.2.2. $a = bq + c \Rightarrow (a,b) = (b,c)$.

Доказательство. По свойствам делимости из п. 2.1 у пары a, b те же общие делители, что и у пары b, c. □

Эти две леммы лежат в основе алгоритма нахождения наибольшего общего делителя (алгоритм Евклида). Выполним следующие деления с остатком:

$$\begin{aligned} a &= bq_1 + r_1, & 0 \le r_1 &< b, \\ b &= r_1 q_2 + r_2, & 0 \le r_2 &< r_1, \\ r_1 &= r_2 q_3 + r_3, & 0 \le r_3 &< r_2, \\ &\vdots & &\vdots \\ r_{n-2} &= r_{n-1} q_n + r_n, & 0 \le r_n &< b_{n-1}, \\ r_{n-1} &= r_n q_{n+1}. \end{aligned} \tag{2.2.1}$$

Тогда $(a,b) = (b,r_1) = (r_1,r_2) = \ldots = (r_{n-1},r_n) = r_n$. Последнее равенство выполняется в силу леммы 2.2.1, а все предыдущие — в силу леммы 2.2.2.

Таким образом, НОД двух чисел равен последнему отличному от нуля остатку в алгоритме Евклида.

Пример. Применим алгоритм Евклида к нахождению (175, 77).

$$175 = 77 \cdot 2 + 21,\ \ 77 = 21 \cdot 3 + 14,\ \ 21 = 14 \cdot 1 + 7,\ \ 14 = 7 \cdot 2.$$

Последний положительный остаток — $r_3 = 7$. Значит, $(175, 77) = 7$.

Умножая все равенства (2.2.1) на натуральное m, убеждаемся в справедливости свойства $(am, bm) = (a,b)m$. Пусть δ — делитель a и b. Тогда

$$(a,b) = \left(\frac{a}{\delta}\delta, \frac{b}{\delta}\delta\right) = \left(\frac{a}{\delta}, \frac{b}{\delta}\right)\delta.$$

Отсюда получаем

$$\left(\frac{a}{\delta}, \frac{b}{\delta}\right) = \frac{(a,b)}{\delta},\ \ \left(\frac{a}{(a,b)}, \frac{b}{(a,b)}\right) = 1. \tag{2.2.2}$$

Понятие наибольшего общего делителя можно ввести и для нескольких чисел $a_1, a_2, \ldots, a_n$. Его обозначают $(a_1, a_2, \ldots, a_n)$. Наибольший общий делитель нескольких чисел можно вычислять последовательно. Например, $(a_1, a_2, a_3) = ((a_1, a_2), a_3)$, $(a_1, a_2, a_3, a_4) = ((a_1, a_2, a_3), a_4)$ и т.д.

2.3. Взаимно простые числа

Теорема 2.3.1 (критерий взаимной простоты чисел). Два целых a и b будут взаимно простыми тогда и только тогда, когда найдутся целые u и v такие, что $au + bv = 1$.

Доказательство. Необходимость вытекает из алгоритма Евклида. Из равенства $r_{n-2} = r_{n-1}q_n + r_n$ находим линейное выражение наибольшего общего делителя $r_n = r_{n-1}q_n + r_{n-2}$ через предыдущие остатки. Затем вставляем сюда аналогичное выражение для r_{n-1} и т.д. В итоге получается так называемое *линейное разложение НОДа*: $d = au + bv$. Достаточность вытекает из свойств делимости:

$$(a,b)\,|\,a,\ (a,b)\,|\,b \Rightarrow (a,b)\,|\,au + bv = 1 \Rightarrow (a,b) = 1. \qquad \square$$

Из доказанного критерия взаимной простоты вытекают следующие свойства.

Свойство 2.3.1. $(a,b) = 1,\ (a,c) = 1 \Rightarrow (a,bc) = 1.$

Свойство 2.3.2. $a\,|\,bc,\ (a,c) = 1 \Rightarrow a\,|\,c.$

Свойство 2.3.3. $a\,|\,c,\ b\,|\,c,\ (a,b) = 1 \Rightarrow ab\,|\,c.$

Например, первое из них выводится так. Умножим почленно равенство $au + bv = 1$ на c. Получим $acu + bcv = c$. Если теперь натуральное d делит a и bc, то оно должно делить и c. Поэтому $d = 1$.

2.4. Наименьшее общее кратное

Если $a\,|\,M$, $b\,|\,M$, то число $M \in \mathbb{N}$ называют *общим кратным* целых чисел $a,b \in \mathbb{Z}$. *Наименьшее общее кратное* (НОК) чисел a и b принято обозначать $[a,b]$.

Теорема 2.4.1. Если M — общее кратное целых a и b, то $[a,b]\,|\,M$.

Доказательство. Разделим с остатком M на $[a,b]$: $M = [a,b]q + r$. Ввиду того, что $r = M - [a,b]q$ и $a\,|\,M$, $a\,|\,[a,b]$, имеем $a\,|\,r$. Аналогично, $b\,|\,r$. Поскольку $0 \le r < [a,b]$, то это возможно лишь при $r = 0$. $\square$

Теорема 2.4.2. Справедливо соотношение $[a,b] = ab/(a,b)$.

Доказательство. Пусть сначала $(a,b) = 1$. Тогда по свойству 2.3.3 имеем $a\,|\,[a,b]$, $b\,|\,[a,b] \Rightarrow ab\,|\,[a,b] \Rightarrow [a,b] = ab$. Пусть теперь $(a,b) = d \ne 1$. В этом случае $(a/d,\ b/d) = 1$. Следовательно, $[a/d, b/d] = ab/d^2$. Если два числа умножить на d, то их НОК и НОД также приобретут этот множитель. Поэтому $[a,b] = dab/d^2 = ab/(a,b)$. $\square$

Наименьшее общее кратное нескольких чисел можно вычислять последовательно: $[a_1, a_2, a_3] = [[a_1, a_2], a_3]$, $[a_1, a_2, a_3, a_4] = [[a_1, a_2, a_3], a_4]$ и т.д.

2.5. Простые числа

Натуральное число $p > 1$ называется *простым*, если оно не имеет других натуральных делителей, кроме 1 и p. Простым числом будет наименьший, отличный от единицы делитель целого a, $a > 1$.

Теорема 2.5.1 (Евклид). Существует бесконечно много простых чисел.

Доказательство. Пусть $p_1, p_2, \ldots, p_n$ — различные простые числа. Простой делитель числа $p_1 p_2 \cdots p_n + 1$ не может совпадать ни с одним из чисел $p_1, p_2, \ldots, p_n$. □

Отметим еще несколько свойств простых чисел (p — простое).

Свойство 2.5.1. $(p, a) \neq 1 \Rightarrow p \mid a$.

Действительно, (p, a), будучи делителем p, может быть равен или 1, или p. Следовательно, $(p, a) = p \Rightarrow p \mid a$.

Свойство 2.5.2. $p \mid ab \Rightarrow p \mid a$ либо $p \mid b$.

Если бы имело место $(p, a) = (p, b) = 1$, то по свойству 2.3.1 $(p, ab) = 1$. Поэтому хотя бы один из множителей делится на p. Свойство легко обобщается на случай нескольких чисел $a, b, c, \ldots$

Теорема 2.5.2. Всякое целое, больше единицы, разложимо в произведение простых множителей. Это разложение единственно с точностью до порядка следования множителей.

Доказательство. Если целое a не простое, то $a = p_1 a_1$, где p_1 — наименьший простой делитель. В случае, если a_1 не простое, то $a_1 = p_2 a_2$. В итоге придем к случаю $a_n = 1$. Следовательно,

$$a = p_1 p_2 \cdots p_n. \tag{2.5.1}$$

Пусть есть еще одно разложение $a = q_1 q_2 \cdots q_s$. Тогда $p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_s$. По свойству 2.5.2 p_1 делит хотя бы одно из чисел в правой части. Пусть, например, $p_1 \mid q_1$. Это значит $p_1 = q_1$. Сокращая обе части на p_1, имеем $p_2 p_3 \cdots p_n = q_2 q_3 \cdots q_s$. Аналогичным образом доказывают совпадение p_2 с каким-либо множителем в правой части. В итоге получим тождественность двух разложений. □

В разложении (2.5.1) некоторые множители могут повторяться. Если объединить повторения, то получается так называемое *каноническое разложение* числа a на простые множители:

$$a = p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s}. \tag{2.5.2}$$

2.6. Сравнения

Будем рассматривать целые числа в связи с остатками от их деления на натуральное m, называемое *модулем*. Если два целых a и b имеют одинаковые остатки от деления на m, то они называются *сравнимыми по модулю m*. Сравнимость чисел a и b записывают в виде

$$a \equiv b (\mathrm{mod}\, m).$$

Отметим следующие легко доказываемые либо очевидные свойства.

Свойство 2.6.1. $a \equiv b(\mathrm{mod}\, m) \Leftrightarrow m \mid a - b$.

Свойство 2.6.2. $a \equiv b(\mathrm{mod}\, m),\ b \equiv c(\mathrm{mod}\, m) \Rightarrow a \equiv c(\mathrm{mod}\, m)$.
Действительно, $m \mid a - b, m \mid b - c \Rightarrow m \mid ((a - b) + (b - c)) \Rightarrow m \mid a - c$.

Свойство 2.6.3. Сравнения можно почленно складывать.
Пусть $a_1 \equiv b_1(\mathrm{mod}\, m)$, $a_2 \equiv b_2(\mathrm{mod}\, m)$. Тогда $m \mid a_1 - b_1$, $m \mid a_2 - b_2$. В силу 2.1.1 $m \mid a_1 - b_1 + a_2 - b_2$, т.е. $m \mid (a_1 + a_2) - (b_1 + b_2)$.

Свойство 2.6.4. Сравнения можно почленно перемножать.
Пусть $a_1 = mq_1 + r_1$, $b_1 = mq_2 + r_1$, $a_2 = mq_3 + r_2$, $b_2 = mq_4 + r_2$. Тогда $a_1 a_2 = m(mq_2 + q_1 r_1 + q_2 r_1) + r_1 r_2$. Следовательно, $a_1 a_2 \equiv r_1 r_2 (\mathrm{mod}\, m)$. Аналогично, $b_1 b_2 \equiv r_1 r_2 (\mathrm{mod}\, m)$.

Свойство 2.6.5. К обеим частям сравнения можно прибавить одно и то же число.

Свойство 2.6.6. Обе части сравнения можно умножить на одно и то же число.

Свойство 2.6.7. Обе части сравнения можно разделить на их общий делитель, если он взаимно прост с модулем.
Пусть $a_1 d \equiv b_1 d(\mathrm{mod}\, m)$. Тогда $m \mid d(a_1 - b_1)$. В силу свойства 2.3.2 $m \mid a_1 - b_1$, поскольку $(m, d) = 1$.

Свойство 2.6.8. Обе части сравнения и модуль можно сокращать на их общий делитель.

Свойство 2.6.9. $a \equiv b(\mathrm{mod}\, m_1),\ a \equiv b(\mathrm{mod}\, m_2) \Rightarrow a \equiv b(\mathrm{mod}[m_1, m_2])$.
Вытекает из теоремы 2.4.1.

Свойство 2.6.10. $a \equiv b(\mathrm{mod}\, m),\ d \mid b,\ d \mid m \Rightarrow d \mid a$.
Действительно, $m \mid a - b \Rightarrow d \mid a - b \Rightarrow d \mid a - b + b \Rightarrow d \mid a$.

Свойство 2.6.11. $a \equiv b(\mathrm{mod}\, m) \Rightarrow (a, m) = (b, m)$.
Ввиду $a = b + m$, непосредственно следует из леммы 2.2.2.

2.7. Классы вычетов

Числа, сравнимые по модулю m, образуют *класс вычетов по модулю m*. Все числа из одного класса имеют один и тот же остаток r от деления на m. Любое число a из класса вычетов называется *вычетом по модулю m*. Соответствующий класс обозначается через $\overline{a}$. Поскольку отношение $a \equiv b(\mathrm{mod}\, m)$ является бинарным отношением эквивалентности, то имеем разбиение целых чисел на *классы эквивалентности* (*классы вычетов*). Всего имеется m классов вычетов по модулю m: $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{m-1}$.

Свойство 2.7.1. $a \equiv b(\mathrm{mod}\, m) \Leftrightarrow \overline{a} = \overline{b}$.

Свойство 2.7.2. $a \not\equiv b(\mathrm{mod}\, m) \Leftrightarrow \overline{a} \cap \overline{b} = \varnothing$.

Взяв из каждого класса по одному вычету, получим *полную систему вычетов*. Например, наряду с $0, 1, 2, \ldots, m-1$ полной системой вычетов будет $1, 2, \ldots, m$.

Свойство 2.7.3. Любые m чисел, попарно несравнимые по модулю m, образуют полную систему вычетов.

Свойство 2.7.4. Если $(a,m) = 1$ и x пробегает полную систему вычетов по модулю m, то $ax + b$, где b — любое целое, также пробегает полную систему вычетов по модулю m.

Доказательство. В силу свойства 2.7.3 достаточно проверить, что $ax_1 + b \not\equiv ax_2 + b(\mathrm{mod}\, m)$ при $x_1 \not\equiv x_2(\mathrm{mod}\, m)$. Допустив $ax_1 + b \equiv ax_2 + b(\mathrm{mod}\, m)$, получим $ax_1 \equiv ax_2(\mathrm{mod}\, m) \Rightarrow x_1 \equiv x_2(\mathrm{mod}\, m)$. □

Согласно 2.6.11, числа одного класса вычетов имеют с модулем m один и тот же общий делитель. Рассмотрим те классы, для которых этот делитель равен единице. Взяв от каждого такого класса по одному вычету, получим *приведенную систему вычетов*. Например, приведенная система по модулю 42 будет 1, 5, 11, 13, 17, 19, 23, 25, 29, 31, 37, 41.

Свойство 2.7.5. Если $(a,m) = 1$ и x пробегает приведенную систему вычетов по модулю m, то ax также будет пробегать приведенную систему вычетов по модулю m.

Достаточно показать, что числа ax попарно несравнимы (2.6.7) и взаимно просты с модулем m. Второе следует из $(m, a) = 1$, $(m, x) = 1$.

2.8. Функция Эйлера

Количество классов вычетов в приведенной системе вычетов обозначают через $\varphi(m)$ и называют *функцией Эйлера*. Она определена для всех натуральных чисел и представляет собой количество чисел ряда $0, 1, \ldots, a-1$ $(1, 2, \ldots, a)$, взаимно простых с a.

Примеры. $\varphi(1)=1$, $\varphi(2)=1$, $\varphi(3)=2$, $\varphi(4)=2$, $\varphi(5)=4$, $\varphi(6)=2$.

Очевидны следующие свойства (p — простое).

Свойство 2.8.1. $\varphi(p)=p-1$.

Свойство 2.8.2. $\varphi(p^k)=p^k-p^{k-1}$, $k\in\mathbb{N}$.

При $k=1$ все ясно. Пусть $k>1$. Тогда в ряду $i=1, 2, 3, \ldots, p^k$ условие $(i,p^k)=1$ нарушается лишь для каждого p-го члена. Всего их $p^k/p=p^{k-1}$.

Лемма 2.8.1 (мультипликативность функции Эйлера).

$$(a,b)=1 \Rightarrow \varphi(ab)=\varphi(a)\varphi(b).$$

Доказательство. Поместим числа 1, 2, ..., ab в таблицу:

1	2	3	$\ldots$	4
$b+1$	$b+2$	$b+3$	$\ldots$	$2b$
$2b+1$	$2b+2$	$2b+3$	$\ldots$	$3b$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$(a-1)b+1$	$(a-1)b+2$	$(a-1)b+3$	$\ldots$	ab

Числа, взаимно простые с b, могут быть лишь в столбцах, номера которых взаимно простые с b. Все числа такого столбца взаимно простые с b. Таких столбцов всего $\varphi(b)$. По свойству 2.7.4 любой такой столбец представляет полную систему вычетов по модулю a. Поэтому он содержит $\varphi(a)$ чисел, взаимно простых с a. Воспользуемся тем, что $(i,ab)=1 \Leftrightarrow (i,a)=(i,b)=1$ (2.3.1). Поэтому $\varphi(ab)=\varphi(a)\varphi(b)$. □

Используя лемму и каноническое разложение числа на простые множители $a=p_1^{k_1}p_2^{k_2}\cdots p_s^{k_s}$, имеем

$$\varphi(a)=(p_1^{k_1}-p_1^{k_1-1})(p_2^{k_2}-p_2^{k_2-1})\cdots(p_s^{k_s}-p_s^{k_s-1})$$

или

$$\varphi(a)=p^{k_1}p^{k_2}\cdots p^{k_s}\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_s}\right),$$

$$\varphi(a)=a\prod_{p|a}\left(1-\frac{1}{p}\right).$$

Функция Эйлера используется в теории сравнений.

Теорема 2.8.1 (Ферма). $a^{p-1}\equiv 1 \pmod p$, если $(a,p)=1$.

Теорема 2.8.2 (Эйлер). $a^{\varphi(m)} \equiv 1 (\mathrm{mod}\, m)$, если $(a,m) = 1$.

Первая теорема вытекает из второй, которая будет доказана в следующей главе.

2.9. Сравнения первой степени

Рассмотрим сравнение

$$ax \equiv b (\mathrm{mod}\, m) \tag{2.9.1}$$

при условии $(a,m) = 1$. Под *решением* любого сравнения понимают класс вычетов по модулю m, один элемент которого (а значит, и все) удовлетворяет сравнению. В нашем случае найдутся целые u, v такие, что $au + mv = 1$. Следовательно, $au \equiv 1 (\mathrm{mod}\, m)$. Будем называть *$u$ обратным* к *a по модулю m*. Умножим обе части сравнения 2.9.1 на u. Получим

$$x \equiv bu (\mathrm{mod}\, m). \tag{2.9.2}$$

Следовательно, сравнение имеет единственное решение по модулю m.

Пусть $(a,m) = d > 1$. По свойству 2.6.10 условие $d \mid b$ является необходимым условием разрешимости сравнения 2.9.1. Будем считать его выполненным. Пусть $a = a_1 d$, $b = b_1 d$, $m = m_1 d$. Тогда наше сравнение равносильно $a_1 x \equiv$ $\equiv b_1 (\mathrm{mod}\, m_1)$. Имеем одно решение $x \equiv x_1 (\mathrm{mod}\, m_1)$. По модулю m имеем d решений:

$$x_1, x_1 + m_1, x_1 + 2m_1, \ldots, x_1 + (d-1)m_1.$$

Теорема 2.9.1. Пусть $(a,m) = d$. Сравнение $ax \equiv b (\mathrm{mod}\, m)$ разрешимо тогда и только тогда, когда $d \mid b$. В этом случае оно имеет d решений.

При небольшом m сравнение $ax \equiv b (\mathrm{mod}\, m)$ решается подбором. Для этого достаточно найти число u такое, что $au \equiv 1 (\mathrm{mod}\, m)$; это можно сделать с помощью алгоритма Евклида. В качестве u можно также взять $u = a^{\varphi(m)-1}$ (способ Эйлера).

2.10. Система сравнений первой степени

Система сравнений

$$\begin{cases} a_1 x \equiv b_1 (\mathrm{mod}\, m_1); \\ a_2 x \equiv b_2 (\mathrm{mod}\, m_2); \\ \quad \vdots \\ a_n x \equiv b_n (\mathrm{mod}\, m_n) \end{cases} \tag{2.10.1}$$

сводится к системе вида

$$\begin{cases} x \equiv b_1 (\bmod m_1); \\ x \equiv b_2 (\bmod m_2); \\ \quad \vdots \\ x \equiv b_n (\bmod m_n). \end{cases} \tag{2.10.2}$$

Для решения последней достаточно уметь решать систему

$$\begin{cases} x \equiv b_1 (\bmod m_1); \\ x \equiv b_2 (\bmod m_2). \end{cases} \tag{2.10.3}$$

Из первого сравнения $x = b_1 + m_1 t$. Подставляя это выражение во второе сравнение, получаем $m_1 t \equiv b_2 - b_1 (\bmod m)$. Поэтому критерием разрешимости (2.10.3) является условие $(m_1, m_2) \mid b_2 - b_1$. В этом случае имеем одно решение по модулю

$$\frac{m_2}{(m_1, m_2)} : t \equiv t_0 \left(\bmod \frac{m_2}{(m_1, m_2)} \right).$$

Поэтому

$$x = b_1 + m_1 \left(t_0 + \frac{m_2}{(m_1, m_2)} l \right) = b_0 + \frac{m_1 m_2}{(m_1, m_2)} t = b_0 + [m_1, m_2] t.$$

Таким образом, система (2.10.3) в случае ее разрешимости имеет единственное решение по модулю $[m_1, m_2]$. Исходная система (2.10.1) в случае ее разрешимости имеет единственное решение по модулю $[m_1, m_2, \ldots, m_n]$.

В случае, когда все модули $m_1, m_2, \ldots, m_n$ попарно взаимно простые, к системе (2.10.2) применим так называемый *китайский способ*. Определим числа $x_1, x_2, \ldots, x_n$ из условий $m_2 m_3 \cdots m_n x_1 \equiv 1 \ (\bmod m_1)$, $m_1 m_3 \cdots m_n x_2 \equiv 1 \ (\bmod m_2)$, ..., $m_1 m_2 \cdots m_{n-1} x_n \equiv 1 \ (\bmod m_n)$. Тогда решением системы (2.10.2) будет число

$$x = m_2 m_3 \cdots m_n x_1 b_1 + m_1 m_3 \cdots m_n x_2 b_2 + \ldots + m_1 m_2 \cdots m_{n-1} x_n b_n.$$

Теорема 2.10.1 (китайская теорема об остатках). Система сравнений (2.10.2) при попарно взаимно простых модулях имеет единственное решение по модулю произведения.

2.11. Первообразные корни

Говорят, что число a, взаимно простое с модулем m, *принадлежит показателю* δ, если δ — такое наименьшее натуральное число, что выполняется сравнение $a^\delta \equiv 1 (\bmod m)$. Справедливы следующие свойства.

Свойство 2.11.1. Числа $a^0, a^1, \ldots, a^{\delta-1}$ попарно несравнимы по модулю m.

Действительно, $a^l \equiv a^k (\bmod m)$, $l > k \Rightarrow a^{l-k} \equiv 1 (\bmod m)$, где $l - k \in \mathbb{N}, l - k < \delta$.

Свойство 2.11.2. $a^{\gamma} \equiv a^{\gamma'} (\mathrm{mod}\, m) \Leftrightarrow \gamma \equiv \gamma' (\mathrm{mod}\, \delta)$.

Разделим γ, γ' на δ с остатками $\gamma = \delta q + r$, $\gamma' = \delta q' + r'$. Тогда $a^{\gamma} \equiv a^{\gamma'} \Leftrightarrow a^{\delta q + r} \equiv a^{\delta q' + r'} \Leftrightarrow d^{r} \equiv d^{r'} \Leftrightarrow r' = r$. Отсюда вытекает следующее свойство.

Свойство 2.11.3. $\delta \mid \varphi(m)$.

Число, принадлежащее показателю $\varphi(m)$, называется *первообразным корнем* по модулю m.

Свойство 2.11.4. По любому простому модулю p существует первообразный корень.

Это свойство будет доказано в следующей главе. Гауссом доказано существование первообразных корней по модулям p^k и $2p^k$ при любом нечетном простом p. Легко убедиться, что при $m = 4$ первообразный корень также существует. Таким образом, первообразные корни существуют по модулям 2, 4, p^{α}, $2p^{\alpha}$, где p — нечетное простое, $\alpha \in \mathbb{N}$.

Докажем отсутствие первообразных корней по всем остальным модулям. Если m не является степенью 2, $m = m_1 m_2$, $m_1 > 2$, $m_2 > 2$, $(m_1, m_2) = 1$, то $\varphi(m_1) \equiv \varphi(m_2) \equiv 0 (\mathrm{mod}\, 2)$. Поэтому

$$a^{\varphi(m)/2} = (a^{\varphi(m_1)})^{\varphi(m_2)/2} \equiv 1 (\mathrm{mod}\, m_1),$$

аналогично $a^{\varphi(m)/2} \equiv 1 (\mathrm{mod}\, m_2)$. Следовательно, $a^{\varphi(m)/2} \equiv 1 (\mathrm{mod}\, m)$ ввиду (2.6.9). Если $m = 2^i$, то при нечетном a: $a^2 \equiv 1 (\mathrm{mod}\, 8)$. Следовательно, $a^{2^{i-2}} \equiv 1 (\mathrm{mod}\, 2^i)$.

Нахождение первообразных корней упрощает следующее свойство.

Свойство 2.11.5. Пусть $c = \varphi(m)$ и $q_1, q_2, \ldots, q_k$ — различные простые делители числа c. Число a, взаимно простое с модулем m, будет первообразным корнем тогда и только тогда, когда не выполнено ни одно из следующих сравнений:

$$a^{c/q_1} \equiv 1 (\mathrm{mod}\, m),\ a^{c/q_2} \equiv 1 (\mathrm{mod}\, m), \ldots,\ a^{c/q_k} \equiv 1 (\mathrm{mod}\, m). \qquad (2.11.1)$$

Необходимость следует из того, что $a^{\varphi(m)} \equiv 1 (\mathrm{mod}\, m)$ и сравнение не имеет места при меньших показателях степени. Обратно, допустим, что a не удовлетворяет ни одному из сравнений, и пусть a принадлежит показателю $\delta < c$. Тогда $\delta \mid c \Rightarrow c = \delta n$. Обозначим через q простой делитель u. Тогда легко получить противоречие:

$$a^{c/q} = a^{\delta u/q} = (a^{\delta})^{u/q} \equiv 1 (\mathrm{mod}\, m).$$

Пример. Пусть $m = 41$. Имеем $c = \varphi(41) = 40 = 2^3 \cdot 5$. Итак, первообразный корень не должен удовлетворять двум сравнениям

$$a^8 \equiv 1 (\mathrm{mod}\, 41),\ a^{20} \equiv 1 (\mathrm{mod}\, 41).$$

Испытываем числа 2, 3, 4, ...: $2^8 \equiv 10$, $2^{20} \equiv \mathbf{1}$, $3^8 \equiv \mathbf{1}$, $4^8 \equiv \mathbf{18}$, $4^{20} \equiv \mathbf{1}$, $5^8 \equiv \mathbf{18}$, $5^{20} \equiv \mathbf{1}$, $6^8 \equiv 10$, $6^{20} \equiv 40$. Отсюда видим, что 6 является наименьшим первообразным корнем по модулю 41.

2.12. Существование первообразных корней

Теорема 2.12.1. Пусть p — нечетное простое. Тогда по модулям вида p^k и $2p^k$, $k \geq 1$ существуют первообразные корни.

Доказательство. Пусть a — первообразный корень по модулю p. Покажем, что существует такое целое x, что $a' = a + px$ будет первообразным корнем сразу по всем модулям вида p^k, $k > 1$. Поскольку a — первообразный корень по модулю p, то $a^{p-1} = 1 + py$ при некотором целом y. Применим формулу бинома Ньютона:

$$(a')^{p-1} = (a + px)^{p-1} = a^{p-1} + (p-1)a^{p-2}px + \ldots =$$

$$= 1 + py + (p-1)a^{p-2}px + \ldots = 1 + p(y + (p-1)a^{p-2}x + \ldots).$$

В скобках коэффициент при x не делится на p. Поэтому x можно выбрать так, что содержимое скобок не будет делиться на p.

В этом случае $(a')^{p-1} = 1 + pz$, где $(z, p) = 1$. Пусть a' принадлежит показателю d по модулю p^k. Тогда d делит $p^{k-1}(p-1)$.

С другой стороны, $p-1$ делит d. Поэтому $d = p^l(p-1)$ при некотором $l < k$. Далее,

$$(a')^d = (1 + pz)^{p^l} = 1 + p^{l+1}u, \quad (u, p) = 1.$$

Это означает, что $l = k - 1$, т.е. $d = p^{k-1}(p-1) = \varphi(p^k)$.

Рассмотрим модуль $2p^k$. Возьмем первообразный корень a по модулю p^k. Одно из чисел a или $a + p^k$ будет нечетным. Оно также является и первообразным корнем по модулю $2p^k$, так как $\varphi(2p^k) = \varphi(p^k)$. □

Таким образом, первообразные корни существуют лишь по модулям 2, 4, p^k, $2p^k$, где p — простое нечетное.

2.13. Индексы по модулям p^k и $2p^k$

Обозначим через m модуль вида p^k или $2p^k$, а через g — первообразный корень по этому модулю. Положим $c = \varphi(m)$. Из свойства 2.12.1 вытекает следующее.

Свойство 2.13.1. Если число γ принимает последовательно значения 0, 1, ..., $c-1$, то g^{γ} пробегает приведенную систему вычетов по модулю m.

Для чисел a, взаимно простых с m, введем понятие индекса, называемого иногда *дискретным логарифмом.*

Пусть $a \equiv g^{\gamma} (\mathrm{mod}\, m)$. Число $\gamma (\gamma \geq 0)$ называется *индексом числа a по модулю m при основании g.* Используются обозначения $\gamma = \mathrm{ind}_g a$ или $\gamma = \mathrm{ind}\, a$. В силу теоремы Эйлера индекс определен по модулю c. Тем самым было бы правильнее говорить о классе вычетов по модулю c.

Свойство 2.13.2. $\mathrm{ind}\, ab \equiv \mathrm{ind}\, a + \mathrm{ind}\, b (\mathrm{mod}\, c)$.

Перемножая сравнения $a \equiv g^{\mathrm{ind}\, a} (\mathrm{mod}\, m)$ и $b \equiv g^{\mathrm{ind}\, b} (\mathrm{mod}\, m)$, получаем требуемое.

Свойство 2.13.3. $\mathrm{ind}\, a^n \equiv n\, \mathrm{ind}\, a (\mathrm{mod}\, c)$.

Если воспользоваться таблицами индексов, то можно решать показательные и степенные сравнения путем их индексирования (дискретного логарифмирования). В самом деле, степенное сравнение $x^n \equiv a (\mathrm{mod}\, m)$ равносильно сравнению $n\, \mathrm{ind}\, x \equiv \mathrm{ind}\, a (\mathrm{mod}\, c)$, решение которого при наличии таблиц не составляет труда. Положим $d = (n, c)$.

Свойство 2.13.4. Сравнение $x^n \equiv a (\mathrm{mod}\, m)$ разрешимо тогда и только тогда, когда d делит $\mathrm{ind}\, a$. В случае разрешимости имеется d решений.

2.14. Символ Лежандра

В п. 2.9 мы изучали сравнение $ax \equiv b\, (\mathrm{mod}\, m)$. Рассмотрим сравнение $ax^2 + bx + c \equiv 0\, (\mathrm{mod}\, m)$. Путем выделения квадрата приведем его к виду $(2ax + b)^2 \equiv b^2 - 4ac\, (\mathrm{mod}\, 4am)$. Полагая $y = 2ax + b$, $d = b^2 - 4ac$, имеем сравнение $y^2 \equiv d\, (\mathrm{mod}\, 4am)$. Фактически исходное сравнение сведено к сравнению вида $x^2 \equiv a\, (\mathrm{mod}\, m)$. Рассмотрим случай, когда m — простое.

Пусть p — нечетное простое число, $(a, p) = 1$. *Символ Лежандра* $\left(\frac{a}{p}\right)$ определяется равенством $\left(\frac{a}{p}\right) = 1$, если сравнение $x^2 \equiv a\, (\mathrm{mod}\, p)$ разрешимо, и $\left(\frac{a}{p}\right) = -1$ в противном случае. Говорят также, что в первом случае a является *квадратичным вычетом* по модулю p и *квадратичным невычетом* во втором. Таким образом,

$$\left(\frac{a}{p}\right) = \pm 1.$$

Пример. Квадратичные вычеты по mod 7 — это 1, 2, 4; невычеты — 3, 5, 6.

Если g — первообразный корень по mod p, то каждое целое g^{2k} — квадратичный вычет, а каждое g^{2k+1} — квадратичный невычет. В самом деле, сравнение $(g^t)^2 \equiv g^{2k+1} (\bmod p)$ влечет $g^{|2k+1-2t|} \equiv 1 \ (\bmod p)$, а значит $2k+1-2t$ должно делиться на $p-1$. Противоречие, так как первое число нечетно, а второе четно. Следовательно, существует $(p-1)/2$ квадратичных вычетов и столько же невычетов по mod p. Представление вычетов в форме g^{2k}, а невычетов в форме g^{2k+1} влечет важное свойство.

Свойство 2.14.1. $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$, где $(a, p) = (b, p) = 1$.

Теорема 2.14.1 (критерий Эйлера). Если $(a, p) = 1$, то

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} (\bmod p).$$

Доказательство.

$$a^{p-1} - 1 = \left(a^{\frac{p-1}{2}} - 1\right)\left(a^{\frac{p-1}{2}} + 1\right) \equiv 0 (\bmod p).$$

Множители $a^{\frac{p-1}{2}} - 1$ и $a^{\frac{p-1}{2}} + 1$ различаются на 2. Поэтому только один из них делится на p. Для $a = g^{2k}$ это будет первый множитель.

Свойство 2.14.2. $\left(\frac{-1}{p}\right) = \begin{cases} 1, & p \equiv 1 \ (\bmod 4); \\ -1, & p \equiv 3 \ (\bmod 4). \end{cases}$

2.15. Квадратичный закон взаимности

Гаусс доказал следующую лемму, упрощающую вычисление символа Лежандра.

Лемма 2.15.1. $\left(\frac{a}{p}\right) = (-1)^{\sum_{i=1}^{p_1}\left[\frac{2ai}{p}\right]}$, где $p_1 = \frac{p-1}{2}$.

Доказательство. Рассмотрим числа a, $2a$, ..., $p_1 a$ и их абсолютно наименьшие вычеты $\varepsilon_1 r_1, \varepsilon_2 r_2, \ldots, \varepsilon_{p_1} r_{p_1}$, где $\varepsilon_j = \pm 1$, $|\varepsilon_j r_j| = r_j$. Числа $\pm a, \pm 2a, \ldots, \pm p_1 a$ образуют приведенную систему вычетов по mod p, среди которых содержатся r_1,

$r_2, \ldots, r_{p_1}$. Последние лишь порядком отличаются от чисел $1, 2, \ldots, p_1$. Перемножая почленно сравнения $ia \equiv \varepsilon_i r_i \pmod p$, $i = 1, 2, \ldots, p_1$ и сокращая произведение на $1, 2, \ldots, p_1 = r_1, r_2, \ldots, p_1$, получаем $a^{\frac{p-1}{2}} \equiv \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{p_1} \pmod p$. В силу

$$\left[\frac{2ia}{p}\right] = \left[2\left[\frac{ia}{p}\right] + 2\left\{\frac{ia}{p}\right\}\right] = 2\left[\frac{ia}{p}\right] + \left[2\left\{\frac{ia}{p}\right\}\right]$$

видно, что $\varepsilon_i = 1$ при четном $\left[\frac{2ia}{p}\right]$ и $\varepsilon_i = -1$ в противном случае. Отсюда легко следует утверждение леммы. □

При нечетном a лемму можно сформулировать в другом виде:

$$\left(\frac{2a}{p}\right) = \left(\frac{4[(a+p)/2p]}{p}\right) = \left(\frac{(a+p)/2}{p}\right) = (-1)^{\sum_{i=1}^{p_1}\left[\frac{(a+p)i}{p}\right]} = (-1)^{\sum_{i=1}^{p_1}\left[\frac{ai}{p}\right] + \sum_{i=1}^{p_1} i}.$$

Поэтому $\left(\frac{2}{p}\right)\left(\frac{a}{p}\right) = (-1)^{\sum_{i=1}^{p_1}\left[\frac{ai}{p}\right] + \frac{p^2-1}{8}}$.

При $a = 1$ имеем $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$, а при нечетном a

$$\left(\frac{a}{p}\right) = (-1)^{\sum_{i=1}^{p}\left[\frac{ai}{p}\right]}.$$

Эта формула необходима при доказательстве следующей весьма важной теоремы.

Теорема 2.15.1 (квадратичный закон взаимности). Для любых нечетных простых p, q выполнено равенство

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right).$$

Доказательство. Для вычисления $\left(\frac{q}{p}\right)$ используем лемму Гаусса. Пусть $q_1 = \frac{q-1}{2}$. Рассмотрим теперь $p_1 q_1$ пар чисел pi и qj, где $i = 1, 2, \ldots, p_1$; $j = 1, 2, \ldots, q_1$. Все они попарно различны. Образуем из них S_1 пар с условием

$qi < pj$ и S_2 пар с условием $q_i > p_j$. Первое условие означает также, что $i \le pj/q$. Поэтому

$$S_1 = \sum_{j=1}^{q_1} \left[\frac{p}{q} j \right].$$

Аналогично

$$S_2 = \sum_{i=1}^{p_1} \left[\frac{q}{p} i \right].$$

В силу леммы Гаусса

$$\left(\frac{p}{q}\right) = (-1)^{S_1}, \quad \left(\frac{q}{p}\right) = (-1)^{S_2}.$$

Перемножая эти равенства, получаем

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{p_1 q_1}. \qquad \square$$

2.16. Символ Якоби

Символ Якоби является обобщением символа Лежандра и служит для упрощения вычисления последнего. Пусть P — нечетное натуральное число, $P = p_1 p_2 \cdots p_s$ — его разложение на простые множители. Для всякого целого a, $(a, P) = 1$, символ Якоби определяется по формуле

$$\left(\frac{a}{P}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right) \cdots \left(\frac{a}{p_s}\right).$$

Отметим следующие свойства.

Свойство 2.16.1. $a \equiv b (\mathrm{mod}\, P) \Rightarrow \left(\frac{a}{P}\right) = \left(\frac{b}{P}\right)$.

Доказательство. Поскольку $a \equiv b (\mathrm{mod}\, P) \Rightarrow a \equiv b (\mathrm{mod}\, p_i)$, то $\left(\frac{a}{p_i}\right) = \left(\frac{b}{p_i}\right)$, $i = 1, 2 \ldots, s$. $\square$

Свойство 2.16.2. $\left(\frac{a_1 a_2 \cdots a_k}{P}\right) = \left(\frac{a_1}{P}\right)\left(\frac{a_2}{P}\right) \cdots \left(\frac{a_k}{P}\right)$.

Доказательство. По определению левая часть равна

$$\left(\frac{a_1 a_2 \cdots a_k}{p_1}\right)\left(\frac{a_1 a_2 \cdots a_k}{p_2}\right) \cdots \left(\frac{a_1 a_2 \cdots a_k}{p_s}\right).$$

Первый множитель по следствию 2.14.1 равен

$$\left(\frac{a_1}{p_1}\right)\left(\frac{a_2}{p_1}\right)\ldots\left(\frac{a_k}{p_1}\right).$$

Применяя это свойство ко всем множителям и группируя их, получаем правую часть. □

Свойство 2.16.3. $\left(\frac{-1}{P}\right)=(-1)^{\frac{p-1}{2}}$.

Доказательство. Заметим, что

$$\left(\frac{-1}{P}\right)=\left(\frac{-1}{p_1}\right)\left(\frac{-1}{p_2}\right)\ldots\left(\frac{-1}{p_s}\right)=(-1)^{\frac{p_1-1}{2}+\frac{p_2-1}{2}+\ldots+\frac{p_s-1}{2}}.$$

С другой стороны, можно многократно воспользоваться легко проверяемым утверждением. Пусть a и b нечетные. Тогда

$$\frac{1}{2}(ab-1)\equiv\frac{1}{2}(a-1)+\frac{1}{2}(b-1)\pmod 2.$$ □

Свойство 2.16.4. $\left(\frac{1}{P}\right)=1$.

Оно вытекает из предыдущих свойств.

Заметим, что из равенства

$$\left(\frac{a}{P}\right)=1$$

не следует, что a является квадратичным вычетом по модулю P. В действительности, a является квадратичным вычетом по модулю P тогда и только тогда, когда a — квадратичный вычет по модулю каждого простого p_i, $i = 1, 2, \ldots, s$. В то же время из равенства $\left(\frac{a}{P}\right)=-1$ следует, что a — квадратичный невычет по mod P.

Теорема 2.16.1. Пусть P и Q — положительные взаимно простые нечетные модули. Тогда

$$\left(\frac{Q}{P}\right)=(-1)^{\frac{P-1}{2}\frac{Q-1}{2}}\left(\frac{P}{Q}\right).$$

Доказательство. Во-первых,

$$\left(\frac{Q}{P}\right)=\prod_{i=1}^{r}\prod_{j=1}^{s}\left(\frac{q_i}{p_j}\right)=(-1)^{\sum_i\sum_j\frac{p_j-1}{2}\frac{q_i-1}{2}}\prod_i\prod_j\left(\frac{p_j}{q_i}\right).$$

Во-вторых, правую часть легко преобразовать к виду

$$(-1)^{\left(\sum_{j=1}^{s}\frac{p_j-1}{2}\right)\left(\sum_{i=1}^{r}\frac{q_i-1}{2}\right)\left(\frac{P}{Q}\right)}.$$

Осталось заметить, что

$$\sum_{j=1}^{s}\frac{p_j-1}{2}\equiv\frac{P-1}{2}\pmod 2,\quad \sum_{i=1}^{r}\frac{q_i-1}{2}\equiv\frac{Q-1}{2}\pmod 2. \qquad \square$$

2.17. Цепные дроби

Пусть α — положительное вещественное число. Положим $q_1=[\alpha]$. Тогда, если α нецелое, то

$$\alpha=q_1+\frac{1}{\alpha_2},$$

где $\alpha_2>1$. Продолжая этот процесс, получаем

$$\alpha_2=q_2+\frac{1}{\alpha_3},\quad \alpha_3=q_3+\frac{1}{\alpha_4},\quad \ldots\quad \alpha_s=q_s+\frac{1}{\alpha_{s+1}}.$$

Следовательно,

$$\alpha=q_1+\cfrac{1}{q_2+\cfrac{1}{q_3+\ldots+\cfrac{1}{q_s+\cfrac{1}{\alpha_{s+1}}}}}.$$

Представление α в указанном виде называется разложением α в цепную (непрерывную) дробь. При иррациональном α цепная дробь, очевидно, оказывается бесконечной. При рациональном α дробь будет конечной. Покажем это.

Пусть $\alpha=\frac{a}{b}$. Применим алгоритм Евклида:

$$\begin{aligned}
&a=bq_1+r_2, && \frac{a}{b}=q+\frac{1}{b/r_2},\\
&b=r_2q_2+r_3, && \frac{b}{r_2}=q_2+\frac{1}{r_2/r_2},\\
&\ldots && \ldots\\
&r_{n-2}=r_{n-1}q_{n-1}+r_n, && \frac{r_{n-2}}{r_{n-1}}=q_{n-1}+\frac{1}{(r_n-1)/r_r},\\
&r_{n-1}=r_nq_n, && \frac{r_n}{r_n}=q_n.
\end{aligned}$$

Следовательно,

$$\frac{a}{b} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \ldots + \cfrac{1}{q_n}}}.$$

Теорема 2.17.1. Всякое иррациональное α разлагается в бесконечную цепную дробь. Всякое рациональное α разлагается в конечную цепную дробь.

2.18. Подходящие дроби

Далее будем использовать запись

$$\alpha = [q_1, q_2, \ldots, q_s, \alpha_{s+1}].$$

Числа $q_1, q_2, \ldots, q_s$ называются неполными частными, а рациональное число $\delta_s = [q_1, q_2, \ldots, q_s]$ — s-й подходящей дробью числа α. Через P_s обозначается числитель, а через Q_s — знаменатель подходящей дроби δ_s.

Например, $P_1 = q_1$, $Q_1 = 1$. Удобно считать, что $P_0 = 1$, $Q_0 = 0$. Имеет место следующий закон образования подходящих дробей.

Теорема 2.18.1. $P_s = q_s P_{s-1} + P_{s-2}$, $Q_s = q_s Q_{s-1} + Q_{s-2}$.

Доказательство. Заметим, что δ_s получается из δ_{s-1} заменой в формуле $\delta_{s-1} = [q_1, q_2, \ldots, q_{s-1}]\, q_{s-1}$ на $q_{s-1} + \frac{1}{q_s}$. Поэтому

$$\delta_1 = \frac{P_1}{Q} = \frac{q_1}{1},$$

$$\delta_2 = \frac{P_2}{Q_2} = q_1 + \frac{1}{q_2} = \frac{q_2 q_1 + 1}{q_2 1 + 0} = \frac{q_2 P_1 + P_0}{q_2 Q_1 + 0}.$$

Применим индукцию:

$$\delta_s = \frac{P_s}{Q_s} = \frac{\left(q_{s-1} + \frac{1}{q_s}\right) P_{s-2} + P_{s-3}}{\left(q_{s-1} + \frac{1}{q_s}\right) Q_{s-2} + Q_{s-3}} = \frac{\frac{1}{q_s} P_{s-2} + P_{s-1}}{\frac{1}{q_s} Q_{s-2} + Q_{s-1}} = \frac{q_s P_{s-1} + P_{s-2}}{q_s Q_{s-1} + Q_{s-2}}. \quad \square$$

Теорема 2.18.2. $\delta_s - \delta_{s-1} = \frac{(-1)^s}{Q_s Q_{s-1}}$, $P_s Q_{s-1} - Q_s P_{s-1} = (-1)^s$.

Доказательство.

$$\delta_s - \delta_{s-1} = \frac{P_s}{Q_s} - \frac{P_{s-1}}{Q_{s-1}} = \frac{P_s Q_{s-1} - Q_s P_{s-1}}{Q_s Q_{s-1}}.$$

$$P_s Q_{s-1} - Q_s P_{s-1} = (q_s P_{s-1} + P_{s-2}) Q_{s-1} - (q_s Q_{s-1} + Q_{s-2}) Q_{s-1} =$$

$$= -(\delta_{s-1} - \delta_{s-2}) = (-1)^{s-1}(P_1 Q_0 - Q_1 P_0) = (-1)^s. \qquad \square$$

Следствие 2.18.1. Все подходящие дроби несократимы.

Теорема 2.18.3. При $\delta_s \neq \alpha$ подходящие дроби $\delta_1, \delta_2, \ldots, \delta_s$ с четными номерами являются последовательными приближениями α, с одной стороны, а с нечетными номерами — с другой, т.е.

$$\alpha_1 < \alpha_3 < \ldots < \alpha < \ldots < \alpha_4 < \alpha_2.$$

Доказательство. Рассмотрим последовательность

$$\alpha = q_1 + \frac{1}{\alpha_2},\ \alpha_2 = q_2 + \frac{1}{\alpha_3},\ \ldots,\ \alpha_{s-2} = q_{s-2} + \frac{1}{\alpha_{s-1}},\ \alpha_{s-1} = q_{s-1} + \frac{1}{\alpha_s}.$$

Если $\frac{1}{\alpha_s}$ заменить на ноль, то последнее число α_{s-1} уменьшится, α_{s-2} увеличится и т.д. Если α_s заменить на q_s, то характер увеличений и уменьшений будет обратным. Это означает, что все подходящие дроби с четными номерами больше α, а с нечетными меньше.

Докажем убывание четных номеров:

$$\frac{P_{2m}}{Q_{2m}} - \frac{P_{2m-2}}{Q_{2m-2}} = \frac{P_{2m} Q_{2m-2} - P_{2m-2} Q_{2m}}{Q_{2m} Q_{2m-2}}.$$

В правой части сделаем замену:

$$P_{2m} = q_{2m} P_{2m-1} + P_{2m-2},$$

$$Q_{2m} = q_{2m} P_{2m-1} + P_{2m-1}.$$

Тогда

$$\delta_{2m} - \delta_{2m-2} = \frac{q_{2m} P_{2m-1} + P_{2m-2} Q_{2m-2} - P_{2m-2}(q_{2m} Q_{2m-1} + Q_{2m-2})}{Q_{2m} Q_{2m-2}} =$$

$$= \frac{q_{2m}(P_{2m-1} Q_{2m-2} - P_{2m-2} Q_{2m-1})}{Q_{2m} Q_{2m-2}} = q_{2m} \frac{(-1)^{2m-1}}{Q_{2m} Q_{2m-2}} < 0.$$

Аналогично доказывается возрастание нечетных номеров. $\square$

2.19. Подходящие дроби в качестве наилучших приближений

Несократимую дробь $\frac{a}{b}(b > 0)$ назовем *наилучшим приближением первого рода* числа $\alpha \in R$, если

$$\frac{a}{b} \neq \frac{c}{d}, \quad 0 < d \leq b \Rightarrow \left|\alpha - \frac{c}{d}\right| > \left|\alpha - \frac{a}{b}\right|.$$

В случае, когда выполняется условие

$$\frac{a}{b} \neq \frac{c}{d}, \quad 0 < d < b \Rightarrow |\alpha d - c| > |\alpha b - a|,$$

говорят о *наилучшем приближении второго рода*, которое является и наилучшим приближением первого рода. В самом деле, перемножая неравенства

$$\left|\alpha - \frac{c}{d}\right| \leq \left|\alpha - \frac{a}{b}\right|, \quad d \leq b,$$

получаем $|\alpha d - c| \leq |\alpha b - c|$. Обратное утверждение неверно. Легко проверить, что 1/3 является наилучшим приближением лишь первого рода числа 1/5.

Теорема 2.19.1. Всякая подходящая дробь δ_s, $s > 1$ есть наилучшее приближение второго рода.

Доказательство. При $\alpha = \delta_s$ утверждение очевидно. Пусть $\alpha \neq \delta_s$. Требуется доказать, что

$$0 < Q \leq Q_s, \quad \frac{P}{Q} \neq \frac{P_s}{Q_s} \Rightarrow |\alpha Q - P| > |\alpha Q_s - P_s|.$$

Представим P и Q в виде

$$P = uP_s + vP_{s-1}, \quad Q = uQ_s + vQ_{s-1},$$

где u, v — целые. Это возможно, так как по теореме 2.18.2

$$\begin{vmatrix} P_s & P_{s-1} \\ Q_s & Q_{s-1} \end{vmatrix} = P_s Q_{s-1} - Q_s P_{s-1} = -1,$$

т.е. применимо правило Крамера. Поэтому

$$|\alpha Q - P| = |u(\alpha Q_s - P_s) + v(\alpha Q_{s-1} - P_{s-1})| > |\alpha Q_s P_s|.$$

Последнее неравенство имеет место, так как $\alpha Q_s - P_s \neq 0$ либо $\delta_s \neq \alpha$. По той же причине $\alpha Q_{s-1} - P_{s-1} \neq 0$. Более того, эти величины разных знаков. Числа u, v также ненулевые. Например, $v = 0$ влечет

$$\frac{P}{Q} = \frac{P_s}{Q_s}.$$

Более того, числа u, v разных знаков, так как $Q \leq Q_s$. □

Теорема 2.19.2. $|\alpha Q_{s-1}P_{s-1}| > |\alpha Q_s - P_s|$, а значит, $\left|\alpha - \frac{P_{s-1}}{Q_{s-1}}\right| > \left|\alpha - \frac{P_s}{Q_s}\right|$.

Доказательство. Представим число α в виде

$$\alpha = [q_1, q_2, \ldots, q_s, \alpha_{s-4}],$$

либо α_{s+1} нет, либо $\alpha_{s+1} > 1$. В первом случае $|\alpha Q_s - P_s| = 0$, а $|\alpha Q_{s-1} - P_{s-1}| > 0$. Во втором случае

$$\alpha = \frac{P_{s+1}}{Q_{s+1}} = \frac{\alpha_{s+1}P_s + P_{s-1}}{\alpha_{s+1}Q_s + Q_{s-1}} \Rightarrow \alpha\alpha_{s+1}Q_s + \alpha Q_{s-1} = \alpha_{s+1}P_s + P_{s-1}.$$

Поэтому $\alpha_{s+1}(\alpha Q_s - P_s) = -\alpha Q_{s-1} + P_{s-1}$. Поскольку $\alpha_{s+1} > 0$, то $|\alpha Q_s - P_s| < |\alpha Q_{s-1} - P_{s-1}|$. □

Теорема 2.19.3. Если $\frac{P}{Q}$ — несократимая дробь, $Q > 0$, такая, что

$$\left|\alpha - \frac{P}{Q}\right| < \frac{1}{2Q^2},$$

то $\frac{P}{Q}$ — подходящая дробь числа α.

Доказательство. Выберем s такое, что $Q_s \le Q < Q_{s+1}$. Тогда

$$\left|\frac{P}{Q} - \frac{P_s}{Q_s}\right| \le \left|\alpha - \frac{P_s}{Q_s}\right| = \frac{1}{Q}|\alpha Q - P| + \frac{1}{Q_s}|\alpha Q_s - P_s| \le$$

$$\le \left(\frac{1}{Q} + \frac{1}{Q_s}\right)|\alpha Q - P| < \left(\frac{1}{Q} + \frac{1}{Q_s}\right)\frac{1}{2Q} \le \frac{1}{Q_s Q}.$$

Значит, число в левой части равно нулю. □

В заключение отметим факт, который потребуется в дальнейшем:

$$Q_k \ge 2^{\frac{1}{2}(k-1)}.$$

В самом деле,

$$Q_k = q_k Q_{k-1} + Q_{k-2} \ge q_{k-1} + q_{k-2} \ge 2q_{k-2}.$$

Это означает, в частности, что при разложении рационального числа P/Q в цепную дробь длина последней ограничена сверху величиной $2\log_2 Q + 1$.

2.20. Задания

1. Показать, что $30 \mid m^5 - m$, $6 \mid m(m^2 + 5)$, $42 \mid m^7 - m$, $30 \mid mn(m^4 - n^4)$ при любых натуральных m, n.

2. Показать, что при натуральном m произведение $(m+1)(m+2)\cdots(m+m)$ делится на 2^m.

3. Будут ли целыми числа $1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$, $1 + \frac{1}{3} + \frac{1}{5} + \ldots + \frac{1}{2n+1}$, $n > 1$?

4. Пусть $(a, b) = 1$. Показать, что $(a + b, a - b) \le 2$.

5. Вычислить НОДы: (549, 387), (589, 343), (12606, 6494), (6188, 4709) и найти их линейные разложения.

6. Доказать, что НОД можно определить как такой общий делитель, который делится на любой другой общий делитель.

7. Пусть $(a,b) = 1$, $ab = c^2$. Показать, что числа a и b будут квадратами.

8. Показать, что $p^2 - q^2$ кратно 24, где p, q — простые числа, большие 3.

9. Доказать бесконечность множества простых чисел вида $4m + 3$.

10. Доказать бесконечность множества простых чисел вида $6m + 5$.

11. Пусть k — натуральное. Доказать, что в натуральном ряду имеется бесконечно много отрезков $m, m+1, \ldots, m+k$, не содержащих простых чисел.

12. Найти канонические разложения чисел 82798848, 81057226635.

13. Разложить на простые множители числа 10!, 15!, 20!, 30!.

14. Сколькими нулями оканчиваются числа 50!, 100!?

15. Найти функцию Эйлера для чисел 375, 720, 957, 988, 1200, 4320.

16. Сколько чисел в интервале от 1 до 120 не взаимно простых с 30?

17. Дано $\varphi(a) = 120$, $a = pq$, где p, q — простые. Найти a, если $p - q = 2$.

18. Доказать, что уравнение $15x^2 - 7y^2 = 9$ не имеет решений в целых числах.

19. Решить в целых числах уравнение $x^2 + y^2 = z^2$.

20. Доказать, что сумма квадратов пяти последовательных целых чисел не может быть точным квадратом.

21. Решить в целых числах уравнения $53x + 47y = 1$, $22x + 32y = 18$.

22. Путем перебора решить сравнения:

1) $5x^2 - 15x + 22 \equiv 0 \pmod 3$,
2) $x^3 - 12 \equiv 0 \pmod 5$,
3) $3x \equiv 1 \pmod 5$,
4) $7x \equiv 31 \pmod 6$,
5) $12x \equiv 1 \pmod 7$,
6) $6x + 5 \equiv 6 \pmod 7$.

23. Способом Эйлера решить сравнения:

1) $5x \equiv 0 \pmod 7$,
2) $25x \equiv 15 \pmod 7$,
3) $5x \equiv 7 \pmod 7$,
4) $5x \equiv 26 \pmod{12}$.

24. Решить системы сравнений:

$$\begin{cases} x \equiv 3 \pmod 8, \\ x \equiv 11 \pmod{20}, \\ x \equiv 1 \pmod{15}, \end{cases} \quad \begin{cases} x \equiv 2 \pmod{17}, \\ 5x \equiv 3 \pmod 9, \\ 8x \equiv 4 \pmod{14}. \end{cases}$$

25. Решить системы сравнений китайским способом:

$$\begin{cases} x \equiv 2 \pmod 7, \\ x \equiv 9 \pmod{11}, \\ x \equiv 3 \pmod{13}, \end{cases} \quad \begin{cases} x \equiv 3 \pmod 8, \\ x \equiv 2 \pmod 9, \\ x \equiv 5 \pmod 7. \end{cases}$$

26. Составить таблицы индексов:
1) по mod 29 с основанием 2,
2) по mod 23 с основанием 5.
Путем индексирования решить сравнения:

1) $2^x \equiv \pmod{67}$,
2) $52^x \equiv 38 \pmod{29}$,
3) $13^x \equiv \pmod{47}$,
4) $12^x \equiv 17 \pmod{31}$,
5) $37x^{16} \equiv 62 \pmod{73}$,
6) $2x^3 \equiv 17 \pmod{41}$,
7) $5x^4 \equiv 3 \pmod{11}$,
8) $27x^5 \equiv 2 \pmod{31}$.

27. Пусть a принадлежит показателю δ, b — показателю γ, $(\delta, \gamma) = 1$. Показать, что ab принадлежит показателю $\delta\gamma$.

28. Пусть a принадлежит показателю δ, b — показателю γ. Как построить элемент, принадлежащий показателю $[\delta, \gamma]$?

29. Пусть a принадлежит показателю δ. Какому показателю принадлежит a^{γ}?

30. Пусть g — первообразный корень по модулю m. Сколько всего первообразных корней по этому модулю?

31. Вычислить символы Лежандра и Якоби:

$$\left(\frac{47}{125}\right), \left(\frac{131}{283}\right), \left(\frac{123}{917}\right).$$

32. Найти все квадратичные вычеты по модулю числа p:

$$p=11, \quad p=13, \quad p=17.$$

33. Доказать, что при $p=4k+1$ числа a и $p-a$ — одновременно квадратичные вычеты или невычеты, а при $p=4k+3$ — наоборот.

34. Решить сравнения: $x^2 \equiv 19 \pmod{31}$, $x^2 \equiv 15 \pmod{53}$, $x^2 \equiv 11 \pmod{59}$.

35. Решить сравнения: $x^2+8x-20 \equiv 0 \pmod{45}$, $5x^2+x+4 \equiv 0 \pmod{10}$.

36. Пользуясь леммой Гаусса и представлением

$$\left(\frac{2}{p}\right)=\left(\frac{4[(p+1)/2]}{p}\right)=\left(\frac{(1+p)/2}{p}\right),$$

доказать еще раз, что $\left(\frac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}$.

37. Показать, что для любого целого d и любого нечетного простого p количество решений сравнения $x^2 \equiv d \pmod{p}$ равно $1+\left(\frac{d}{p}\right)$.

38. Найти способ решения сравнений вида $x^2 \equiv 4 \pmod{m}$.

39. Доказать, что количество решений уравнения $x^2+y^2=p$, $(x,y)=1$, $x>0$, $y>0$ равно количеству решений сравнения $z^2+1 \equiv 0 \pmod{p}$.

40. Разложить в цепные дроби числа $\frac{170}{109}, \frac{99}{170}, \frac{125}{92}, \sqrt{2}$.

SXJKN8EYHR3I4HR987YD
WFUIHW9E8YR384H5O345
96EUF9O8WE76R83Y4398
46UHFE98YDSFJH8374H4
BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7E

Глава 3

АЛГЕБРАИЧЕСКИЕ ОСНОВЫ

SXJKN8EYHR3I4HR987YD
WFUIHW9E8YR384H5O345
96EUF9O8WE76R83Y4398
46UHFE98YDSFJH8374H4
BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7E

3.1. Понятие группы

Группой называется непустое множество G с алгебраической операцией $*$ на нем, для которой выполняются первые три из четырех следующих аксиом.

1. Операция $*$ ассоциативна, т.е. для любых $a, b, c \in G$

$$a*(b*c) = (a*b)*c.$$

2. В G имеется *единичный элемент* (или *единица*) e такой, что для любого $a \in G$

$$a*e = e*a = a.$$

3. Для каждого $a \in G$ существует *обратный элемент* $a^{-1} \in G$ такой, что

$$a*a^{-1} = a^{-1}*a = e.$$

4. Для любых $a, b \in G$

$$a*b = b*a.$$

Если дополнительно группа удовлетворяет четвертой аксиоме, то группа называется *абелевой* (или *коммутативной*).

Для групповой операции будем использовать мультипликативное обозначение и вместо $a*b$ писать ab, называя этот элемент *произведением* элементов a и b. Иногда для групповой операции используют аддитивную запись: $a + b$. В этом случае вместо единицы пишут ноль, а вместо a^{-1} — $-a$. Такие обозначения обычно резервируют для абелевых групп.

В группе имеется лишь один единичный элемент. Действительно, если e' — еще одна единица, то $e' = e'e = e$. Для любого элемента имеется лишь один обратный. Пусть x и y — обратные элементы для $a \in G$. Тогда по ассоциативности

$$x = xe = x(ay) = (xa)y = ey = y.$$

Примеры. Множество $\mathbb{Z}$ целых чисел образует группу относительно операции сложения. То же можно сказать относительно рациональных чисел $\mathbb{Q}$, вещественных чисел $\mathbb{R}$ и комплексных чисел $\mathbb{C}$.

Обозначим через $\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}$ классы вычетов по модулю 5. Определим их сложение по модулю 5. Например, $\overline{1} + \overline{2} = \overline{3}$, $\overline{2} + \overline{3} = \overline{0}$, $\overline{2} + \overline{4} = \overline{1}$, т.е. осуществляется обычное сложение и при необходимости берется остаток от деления

на 5. Эта группа обозначается через $\mathbb{Z}_5$ и называется (аддитивной) *группой классов вычетов по модулю* 5. Аналогично строится группа классов вычетов $\mathbb{Z}_m$ по любому модулю m. Если взять все классы вычетов, взаимно простые с модулем m, и определить их умножение по модулю m, то получается группа, обозначаемая через $\mathbb{Z}_m^*$. Отметим, что существование обратного элемента для $\bar{a} \in \mathbb{Z}_m^*$ вытекает из разрешимости сравнения $ax \equiv 1(\mathrm{mod}\, m)$ при $a, m = 1$.

Число элементов конечной группы G называется *порядком группы* и обозначается через $|G|$. Например, $|\mathbb{Z}_m| = m$, $|\mathbb{Z}_m^*| = \varphi(m)$.

Мультипликативная группа G называется *циклической*, если она порождена одним элементом, т.е. в ней имеется такой элемент a (образующий), что любой другой элемент b представим в виде $b = a^n$, $n \in \mathbb{Z}$. Если n — отрицательное, то под a^n понимается произведение $(a^{-1})^{|n|}$. Циклическими являются группы $\mathbb{Z}$ и $\mathbb{Z}_m$. Группа $\mathbb{Z}_m^*$ — циклическая лишь в случае, когда по модулю m существует первообразный корень. В циклической группе, конечной или нет, может быть несколько образующих элементов. В аддитивной группе $\mathbb{Z}$ образующими будут элементы 1 и –1. Циклическая группа всегда коммутативна.

Существует удобный способ задания конечной группы — в виде таблицы. Обычно она называется *таблицей Кэли*. Ее строки и столбцы помечаются элементами группы, и на пересечении строки, помеченной элементом a, и столбца, помеченного элементом b, ставится элемент ab. Например, таблица Кэли группы $\mathbb{Z}_5$ имеет вид

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

3.2. Подгруппы групп

Подмножество H группы G называется *подгруппой* этой группы, если H образует группу относительно операции группы G.

Подгруппы группы G, отличные от тривиальных подгрупп $\{e\}G$, называются *собственными* подгруппами.

Теорема 3.2.1. Подмножество H группы G будет ее подгруппой тогда и только тогда, когда выполняются условия:

1) $a, b \in H \Rightarrow ab \in H$;
2) $a \in H \Rightarrow a^{-1} \in H$.

Доказательство. Необходимость следует из определения группы. Докажем достаточность. Аксиома 1 в определении группы (ассоциативность) автоматически выполняется и на любом подмножестве. Осталось показать, что $e \in H$. Возьмем произвольный элемент $a \in H$. Тогда $a^{-1} \in H$, а значит, $e = aa^{-1} \in H$. □

Условия 1 и 2 в теореме 3.2.1 эквивалентны условию

$$a,b \in H \Rightarrow ab^{-1} \in H.$$

Теорема 3.2.2. Если H — подгруппа группы G, то отношение R_H на G, определяемое условием

$$(a,b) \in R_H \Leftrightarrow a = bh \text{ для некоторого } h \in H,$$

является бинарным отношением эквивалентности.

Классы эквивалентности по отношению R_H называются *левыми смежными классами* группы G по подгруппе H и обозначаются

$$aH = \{ah \mid h \in H\}.$$

Смежные классы группы по подгруппе либо совпадают, либо не пересекаются. Аналогично определяются *правые смежные классы* по подгруппе H, которые имеют вид $H\,a = \{ha \mid h \in H\}$. Для абелевой группы эти два понятия идентичны.

Пусть, например, $G = \mathbb{Z}$, $H = 2\mathbb{Z}$ — подгруппа четных чисел. Тогда имеем два смежных класса: $2\mathbb{Z}$ — четные числа, $1 + 2\mathbb{Z}$ — нечетные.

Теорема 3.2.3. Если H — конечная группа, то каждый (левый или правый) смежный класс по ней содержит $|H|$ элементов.

Доказательство. $a_1h = a_2h \Rightarrow a_1 = a_2$, так как обе части первого равенства можно умножить на h^{-1}. □

Если количество смежных классов группы G по подгруппе H конечно, то оно называется *индексом* подгруппы H в группе G и обозначается через $[G:H]$. В этом случае имеет место так называемое *разложение группы по подгруппе*:

$$G = H \cup g_1H \cup g_2H \cup \ldots \cup g_\rho H,$$

где все смежные классы попарно не пересекаются, $\rho = (G:H) - 1$. Аналогично можно ввести разложение на правые смежные классы по подгруппе.

Теорема 3.2.4. Пусть G — конечная группа. Тогда

$$|G| = [G:H]\,|H|.$$

Следствие 3.2.1 (Лагранж). Порядок конечной группы делится на порядок любой ее подгруппы.

Пусть $a \in G$. Положим $a^n = aa \cdots a$, если n — натуральное, $a^n = (a^{-1})^{|n|}$, если n — целое отрицательное, и, наконец, $a^0 = e$. Таким образом можно рассмотреть подмножество

$$\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}.$$

Оно, как легко показать, является подгруппой группы G. Эта подгруппа называется *циклической подгруппой*, порожденной элементом a. Ее порядок называется *порядком элемента* a. Иными словами, элемент $a \in G$ называется элементом порядка $m \in \mathbb{N}$, если $a^m = e$, где m — наименьшее натуральное с этим условием. Легко показать, что $m \mid l$, если $a^l = e$. Если такого m нет, то элемент a называется *элементом бесконечного порядка*. Из теоремы Лагранжа вытекает, что порядок конечной группы G делится на порядок любого ее элемента a. Поэтому $a^{|G|} = e$. Поскольку $\left|\mathbb{Z}_m^*\right| = \varphi(m)$, то в качестве еще одного следствия получается теорема Эйлера о том, что если $(a,m) = 1$, то $a^{\varphi(m)} \equiv 1 (\mathrm{mod}\, m)$.

В дальнейшем нам потребуется утверждение о порядках элементов в абелевой группе G. Пусть $a, b \in G$ и их порядки m и n соответственно, причем $(m,n) = 1$. Покажем, что порядок произведения ab равен mn. Обозначим порядок ab через δ. Тогда $(ab)^\delta$, а значит, $(ab)^{\delta m} = 1 \Rightarrow a^{\delta m} b^{\delta m} = 1 \Rightarrow b^{\delta m} = 1 \Rightarrow n \mid \delta m \Rightarrow n \mid \delta$. Аналогично показывается, что $m \mid \delta$. Поэтому $mn \mid \delta$. Последнее, в силу минимальности δ, означает $\delta = mn$. Доказанное является частным случаем следующего утверждения.

Теорема 3.2.5. Пусть даны два элемента $a, b \in G$ абелевой группы. Тогда в группе найдется элемент порядка $[m,n]$.

Доказательство. Случай $(m,n) = 1$ уже рассмотрен. Пусть $(m,n) = d$. Легко построить разбиение $d = d_1 d_2$ такое, что $(m/d_1, n/d_2) = 1$. Элементы a^{d_1} и a^{d_2} имеют порядки m/d_1 и n/d_2 соответственно. Их произведение имеет порядок $(m/d_1)(n/d_2) = mn/d = [m,n]$. □

3.3. Циклические группы

Пусть $\langle a \rangle$ — циклическая группа, порожденная элементом a.

Теорема 3.3.1. Каждая подгруппа циклической группы также является циклической.

Доказательство. Пусть H — подгруппа циклической группы $\langle a \rangle$ такая, что $H \neq \{e\}$. Если $a^n \in H$, то $a^{-n} \in H$. Поэтому H содержит степень a с натуральным показателем. Пусть d — наименьший натуральный показатель, для которого

$a^d \in H$. Пусть $a^s \in H$. Положим $s = dq + r$, $0 \le r < d$, $q, r \in \mathbb{Z}$. Тогда $a^r = a^s(a^{-d})^q \in H$, что противоречит минимальности d, если $r \ne 0$. Поэтому $H = \langle a^d \rangle$. □

Теорема 3.3.2. В конечной циклической группе $\langle a \rangle$ порядка m элемент a^k порождает подгруппу порядка $m(k,m)^{-1}$.

Доказательство. Пусть $d = (k,m)$. Порядок группы $\langle a^k \rangle$ — наименьшее натуральное n такое, что $a^{kn} = e$. Условие $a^{kn} = e$ справедливо лишь тогда, когда $m \mid kn$, т.е. тогда и только тогда, когда $(m/d) \mid n$. Наименьшее натуральное n с таким свойством есть m/d. □

Теорема 3.3.3. Если d — положительный делитель m, то G содержит единственную подгруппу индекса d.

Доказательство. Если d задано, то $\langle a^d \rangle$ является подгруппой порядка m/d, и потому она имеет индекс d. Если $\langle a^k \rangle$ — другая подгруппа индекса d, то ее порядок равен m/d. Поэтому в силу теоремы 3.3.2 $\frac{m}{d} = \frac{m}{(k,m)} \Rightarrow d = (k,m) \Rightarrow d \mid k$. Следовательно, $a^k \in \langle a^d \rangle$. Но так как обе группы одного порядка, то они совпадают. □

Теорема 3.3.4. Для любого положительного делителя l числа m группа $\langle a \rangle$ содержит в точности одну подгруппу порядка l.

Доказательство. В силу предыдущей теоремы подгруппами порядка l являются только подгруппы, индексы которых равны m/l. □

Теорема 3.3.5. Пусть l — положительный делитель порядка конечной циклической группы $\langle a \rangle$. Тогда $\langle a \rangle$ содержит $\varphi_{t(l)}$ элементов порядка l.

Доказательство. Пусть $|\langle a \rangle| = m = dl$. В силу теоремы 3.3.2 элемент a^k имеет порядок l лишь в случае, если $(k,m) = d$. Поэтому количество элементов порядка l равно количеству целых k, $1 \le k \le m$, для которых $(k,m) = d$. Поэтому $k = dn$, где $1 \le n \le l$. Теперь условие $(k,m) = d$ равносильно условию $(k,l) = 1$. Количество таких чисел равно $\varphi(l)$. □

Попутно нами доказана следующая теорема.

Теорема 3.3.6. Конечная циклическая группа $\langle a \rangle$ порядка m содержит $\varphi(m)$ образующих. Элемент a^l является образующим лишь при условии $(l,m) = 1$.

3.4. Гомоморфизмы групп

Отображение $f: G \to G'$ группы G в группу G' называется *гомоморфизмом*, если оно согласовано с операциями на группах G и G', т.е. $f(ab) = f(a)f(b)$ для любых двух элементов $a, b \in G$. Если это отображение сюръективное, то оно называется *эпиморфизмом*. В этом случае группа G' называется *гомоморфным образом* группы G. Приставка «моно» употребляется в случае, когда гомоморфизм инъективен. Биективный гомоморфизм называется *изоморфизмом*. Для изоморфных групп употребляется обозначение $G \cong H$. Изоморфизм группы G на себя называется *автоморфизмом*.

Примеры. Обозначим через $GL(n, R)$ группу по умножению всех невырожденных матриц n-го порядка с вещественными элементами. Тогда отображение $A \to \det A, A \in GL(n, R)$ будет эпиморфизмом на мультипликативную группу поля вещественных чисел $\mathbb{R}^*$.

Еще один пример эпиморфизма дает отображение $\varphi: \mathbb{Z} \to \mathbb{Z}_m$, при котором $\varphi(a) = \overline{a}$, т.е. элемент $a \in \mathbb{Z}$ отображается в соответствующий класс вычетов по модулю m.

Ядром гомоморфизма $f: G \to H$ называется множество

$$\ker \varphi = \{a \in G \mid f(a) = e'\},$$

где e' — единичный элемент группы H.

В случае гомоморфизма $GL(n, \mathbb{R}) \to \mathbb{R}^*$ ядром будет подгруппа матриц с единичным определителем. Ядром во втором примере является группа чисел $m\mathbb{Z}$, кратных модулю m.

Сохранив для аддитивной группы поля вещественных чисел обозначение $\mathbb{R}$ и обозначив через $\mathbb{R}^+$ мультипликативную группу положительных вещественных чисел, имеем изоморфизм $\mathbb{R}^+ \cong R$, заданный функцией $y = \ln x$.

Легко показать, что ядро любого гомоморфизма является подгруппой H группы G с важным дополнительным условием: $g^{-1}Hg = H$ для любого элемента $g \in G$. Такие подгруппы называются *нормальными подгруппами* (*нормальными делителями*). Используется обозначение $H \lhd G$. Условие нормальности, как нетрудно видеть, можно переписать в виде $gH = Hg$, или $gHg^{-1} = H$. В абелевой группе все подгруппы являются нормальными.

Если H — нормальная подгруппа группы G, то множество смежных классов группы G по подгруппе H можно наделить групповой структурой. Соответствующая группа называется *фактор-группой* группы G по подгруппе H и обозначается G/H. Определим композицию смежных классов по формуле $(g_1H)(g_2H) = g_1g_2H$. Докажем корректность. Пусть g_1h_1 и g_2h_2 — другие представители смежных классов. Тогда $g_1h_1g_2h_2$ можно представить в виде $g_1g_2h_1'h_2$, так как в силу $gH = Hg$ произведение h_1g_2 можно представить в виде g_2h_1'. Поэтому

$$(g_1h_1g_2h_2)H = (g_1g_2)((h_1'h_2)H) = g_1g_2H.$$

Теорема 3.4.1 (о гомоморфизме). Пусть $f: G \to G_1$ — эпиморфизм. Тогда $\ker f \lhd G$, причем группа G_1 изоморфна фактор-группе $G/\ker f$. Если H — нормальная подгруппа группы G, то $f: G \to G/H$, определяемое условием $f(a) = aH$, является эпиморфизмом, причем $\ker f = H$.

Доказательство. Первое утверждение уже доказано. Покажем, что $G_1 \cong G/\ker f$. Легко убедиться, что гомоморфизм $f: G \to G_1$ постоянен на смежных классах. В самом деле, если $f(g) = g'$, то $f(gh) = f(g)f(h) = f(g) = g'$, где $h \in \ker f$. Поэтому существует сюръекция $\bar{f}$ группы $G/\ker f$ на группу G_1. Докажем ее инъективность. Если $f(g_1) = f(g_2)$, то $f(g_1)f(g_2)^{-1} = f(g_1 g_2^{-1}) = e$, а значит, $g_1 \in g_2 \ker f$. Таким образом, нами построена биекция с условием $\bar{f}(g \ker f) = f(g)$. Условие $\bar{f}(g_1 \ker f\, g_2 \ker f) = \bar{f}(g_1 \ker f)\bar{f}(g_2 \ker f)$ легко получается из условия $f(g_1 g_2) = f(g_1)f(g_2)$. Аналогично проверяется третье утверждение. □

3.5. Группы подстановок

Обозначим через X конечное множество, а его элементы — через $1, 2, \ldots, n$. Рассмотрим все биекции (подстановки) $\sigma: X \to X$. Легко видеть, что они образуют группу относительно операции композиции отображений. Эта группа называется *симметрической группой n-й степени* и обозначается через S_n или через $S(X)$. Нетрудно показать, что $|S_n| = n!$. Так, например, группа S_3 состоит из шести подстановок:

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}, \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}, \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},$$

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}, \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}, \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}.$$

В нижней строке указаны образы элементов 1, 2, 3, расположенных в верхней строке. Условимся при вычислении произведения подстановок $\sigma_1\sigma_2$ выполнять отображения справа налево, т.е. сначала отображение σ_2, а затем σ_1. Например:

$$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix} = \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}.$$

Подгруппы симметрической группы называются *группами подстановок.*

Подстановку вида $1 \to 2 \to 3 \to \ldots \to k \to 1$ назовем *циклом длиной k* и обозначим $(1, 2, \ldots, k)$. Два цикла называются *независимыми*, если перемещаемые ими элементы попарно различны. Независимые циклы коммутируют, т.е. для них выполнено условие $\sigma_1\sigma_2 = \sigma_2\sigma_1$. Цикл длиной 2 называется *транспозицией.*

Теорема 3.5.1. Каждая подстановка единственным образом разложима в произведение независимых циклов.

Доказательство. Пусть π — произвольная подстановка, i — любой элемент из X. Рассмотрим последовательность $i, \pi(i), \pi^2(i), \ldots$ При некотором k, в силу конечности множества X, $\pi^k(i) = i$. Таким образом, имеем цикл длиной k: $i \to \pi(i) \to \pi^2(i) \to \pi^3(i) \to \ldots \to \pi^k(i) \to i$. Если $k < n$, то берем один из оставшихся элементов и поступаем точно так же. В итоге получаем разложение в произведение независимых циклов $\pi = \sigma_1\sigma_2\cdots\sigma_s$.

Пусть $\pi = \alpha_1\alpha_2\cdots\alpha_r$ — еще одно разложение, i — элемент, не остающийся на месте под действием π. Тогда найдется по одному циклу, например, σ_1 и α_1, обладающему тем же свойством, что и π. Для них имеем $\pi^k(i) = \sigma_1^k(i) = \alpha_1^k(i)$. Но цикл определяется действием своих степеней на любой элемент, который не остается на месте. Поэтому $\sigma_1 = \alpha_1$. Далее применима индукция по r или по s. □

Теорема 3.5.2. Каждая подстановка $\tau \in S_n$ является произведением транспозиций.

Доказательство. В силу предыдущей теоремы достаточно рассмотреть случай цикла. Но это можно сделать, например, так:

$$(1, 2, 3, \ldots, k) = (1, k)(1, k-1)\cdots(1, 3)(1, 2). \qquad \square$$

Ни о какой единственности не может быть и речи хотя бы потому, что для любой транспозиции τ и подстановки σ имеем $\sigma\tau^2 = \sigma$. Тем не менее, характер четности числа k в разложении подстановки в произведении транспозиций $\pi = \tau_1\tau_2\cdots\tau_k$ определяется подстановкой π однозначно. В самом деле, рассмотрим очевидное равенство

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ a_1 & a_2 & \ldots & a_n \end{pmatrix}(1, 2) = \begin{pmatrix} 1 & 2 & \ldots & n \\ a_2 & a_1 & \ldots & a_n \end{pmatrix},$$

которое показывает, что умножение подстановки на транспозицию меняет характер четности перестановки $a_1, \ldots, a_n$ на противоположный. Поэтому, если транспозиции $\tau_1, \ldots, \tau_k$ приводят перестановку $a_1, \ldots, a_n$ к виду $1, \ldots, n$, то $\pi = \tau_k\cdots\tau_1$, и наоборот, поэтому характер четности подстановки π совпадает с характером четности перестановки $a_1, \ldots, a_n$ Подстановка называется *четной* или *нечетной* в зависимости от четности числа k.

Теорема 3.5.3. При $n > 1$ количество четных подстановок равно количеству нечетных подстановок и равно $n!/2$.

Доказательство. Утверждение вытекает из аналогичного свойства перестановок. □

Нетрудно показать, что все четные перестановки образуют подгруппу группы S_n. Эта подгруппа называется *знакопеременной группой* и обозначается через A_n. При $n > 1$ имеем разложение $S_n = A_n \cup (1, 2)A_n$. Поэтому $[S_n : A_n] = 2$.

Для любой подстановки $\pi \in S_n$ смежные классы πA_n и $A_n\pi$ состоят из всех четных или всех нечетных подстановок в зависимости от четности подстановки π. Поэтому $A_n \lhd S_n$.

Теорема 3.5.4 (Кэли). Всякая конечная группа G изоморфна подгруппе симметрической группы S_n, где $n = |G|$.

Доказательство. Обозначим через $e, g_1, \ldots, g_{n-1}$ элементы группы G. Рассмотрим отображение, где a — фиксированный элемент, $a \in G$. Имеем подстановку

$$\tau_a = \begin{pmatrix} e & g_1 & \cdots & g_{n-1} \\ a & ag_1 & \ldots & ag_{n-1} \end{pmatrix}.$$

Вместе с тем получаем отображение $G \to S_n$, при котором $a \to \tau_a$. Убедимся, что это — искомый мономорфизм. Это отображение инъективно, поскольку различным элементам a и b отвечают различные подстановки τ_a и τ_b. В первом случае $e \to a$, а во втором $e \to b$. Осталось проверить условие $\tau_{ab} = \tau_a\tau_b$. Убедиться в равенстве двух подстановок можно, проверив их действие на произвольном элементе группы:

$$\tau_{ab}(x) = abx, \quad (\tau_a\tau_b)x = \tau_a(\tau_b(x)) = abx. \qquad \square$$

3.6. Действие группы на множестве

Пусть G — произвольная конечная группа, X — конечное множество из n элементов. Будем говорить, что *G действует на X*, если задан любой гомоморфизм $G \to S(X)$. Тем самым задано отображение декартова произведения $G \times X$ в множество X. Если $g \to \pi \in S(X)$, то $(g, x) \to \pi(x)$. Вместо (g, x) будем писать $gx = \pi(x)$. При этом выполняются очевидные свойства

$$e(x) = x,\ x \in X;\ (gh)x = g(h(x)),\ g,\ h, \in G.$$

Два элемента x, $x' \in X$ называются *эквивалентными относительно группы G*, действующей на X, если $x' = gx$. Легко проверяются свойства рефлексивности, транзитивности и симметричности. Соответствующие классы эквивалентности называются *орбитами*. Орбиту, содержащую элемент x_0, удобно обозначать символом $G(x_0)$, т.е. $G(x_0) = \{gx_0 \mid g \in G\}$. Например, $S_n(1) = \{1, 2, \ldots, n\}$.

Пусть x_0 — элемент из X. Рассмотрим множество

$$St(x_0) = \{g \in G \mid gx_0 = x_0\}.$$

Легко убедиться, что $St(x_0)$ — подгруппа в G. Она называется *стабилизатором* элемента x_0. Для рассмотренного примера $St(1)$ — множество всех подстановок, оставляющих элемент 1 на месте. Очевидно, $St(1) \cong S_{n-1}$, т.е. это фактически симметрическая группа на множестве $2, 3, \ldots, n$.

Теорема 3.6.1. Card $G(x_0) = [G : St(x_0)]$.

Доказательство. $gx_0 = g_1x_0 \Leftrightarrow x_0 = g^{-1}g_1x_0 \Leftrightarrow g^{-1}g_1 \in St(x_0) \Leftrightarrow g_1 \in gSt(x_0)$. Поэтому левые смежные классы находятся во взаимно однозначном соответствии с элементами орбиты $G(x_0)$. □

Из этой теоремы и теоремы Лагранжа следует, что длина любой орбиты конечной группы является делителем порядка группы.

Группа перестановок $G \subset S_n$, действующая на множестве $X = \{1, 2, \ldots, n\}$, называется *транзитивной*, если орбита некоторой (а значит, и любой) точки совпадает со всем множеством X. Транзитивной будет вся группа S_n и, как нетрудно убедиться, знакопеременная группа A_n. Определим действие группы G на левых смежных классах по подгруппе H по правилу $g(g_1H) = (g_1g_2)H$. В этом случае также имеем дело с транзитивностью. В самом деле, если g_1H и g_2H — два смежных класса, то $g_2g_1^{-1}g_1H = g_2H$.

3.7. Кольца и поля

Кольцом называется множество R с двумя бинарными операциями, обозначаемыми символами «+» и «·», такими, что:

1) R — абелева группа относительно операции «+»;
2) операция умножения ассоциативна, т.е. для всех $a, b, c \in R$ $(ab)c = a(bc)$;
3) выполняются законы дистрибутивности, т.е. для всех $a, b, c \in R$

$$a(b + c) = ab + ac \quad \text{и} \quad (b + c)a = ba + ca.$$

Условимся называть нейтральный элемент аддитивной группы кольца *нулем* и обозначать его символом 0. Противоположный к a элемент обозначают через $-a$. Вместо $a + (-b)$ обычно пишут $a - b$. Легко доказываются свойства $a0 = 0a = 0$ для всех $a \in R$. Из этого следует, что $(-a)b = a(-b) = -ab$ для всех $a, b \in R$. Простейшими примерами колец являются кольца целых чисел $\mathbb{Z}$ и многочленов $\mathbb{R}[x]$ с вещественными коэффициентами.

Кольцо называется *кольцом с единицей*, если оно имеет мультипликативную единицу, т.е. такой элемент e, что $ae = ea = a$ для любого $a \in R$.

Кольцо называется *коммутативным*, если операция умножения коммутативна.

Два элемента кольца $a \neq 0$, $b \neq 0$ называются *делителями нуля*, если $ab = 0$. Приведем пример делителей нуля. Рассмотрим кольцо классов вычетов $\mathbb{Z}_m$ по модулю m. Оно состоит из элементов $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{m-1}$. Операция сложения над этими элементами была определена ранее. Аналогично определяется умножение. Выполняем обычное умножение чисел и при необходимости берем остаток от деления на m. Если m — составное, $m = ab$, то делителями нуля будут $\overline{a}, \overline{b}$.

Кольцо называется *областью целостности*, если оно является коммутативным кольцом с единицей и без делителей нуля.

Коммутативное кольцо называется *полем*, если его ненулевые элементы образуют группу относительно операции умножения. Очевидно, всякое поле содержит не менее двух элементов. Отметим простейшие свойства полей.

Свойство 3.7.1. В поле нет делителей нуля.
Равенство $ab = 0$ при $a \neq 0$ влечет $a^{-1}ab = a^{-1}0$, а значит, $b = 0$.

Свойство 3.7.2. В поле второй закон дистрибутивности вытекает из первого.

Теорема 3.7.1. Конечная область целостности является полем.

Доказательство. Обозначим через $a_1, a_2, .., a_n$ элементы кольца R. Фиксируем ненулевой элемент $a \in R$ и рассмотрим последовательность элементов $aa_1, aa_2, ..., aa_n$. Все они попарно различны, так как $aa_1 = aa_2$ влечет $a_1 = a_2$. Поэтому в нашей последовательности имеется единичный элемент $e = aa_i$. Тем самым доказана обратимость любого элемента $a \neq 0$, так как $a(a_1 - a_2) = 0$ и нет делителей нуля. □

Теорема 3.7.2. Кольцо классов вычетов $\mathbb{Z}_m$ будет областью целостности, а значит, и полем лишь при простом m.

Доказательство. Если m — составное, то в кольце $\mathbb{Z}_m$ есть делители нуля, а значит, оно не может являться полем. Обратно, если $m = p$ — простое, то для любого ненулевого элемента $a \in \mathbb{Z}_m$ можно указать обратный, что легко следует из разрешимости сравнения $ax \equiv 1 (\mathrm{mod}\ p)$, так как $a \in \{1, 2, .., p-1\}$, т.е. $(a, p) = 1$. □

Поле $\mathbb{Z}_p$ называется *полем Галуа* порядка p и обозначается через $\mathbb{F}_p$.

3.8. Подкольца

Подмножество S кольца R называется *подкольцом* этого кольца, если оно замкнуто относительно имеющихся операций сложения и умножения и само образует кольцо относительно этих операций.

Подкольцо H кольца R называется *идеалом* (*двусторонним идеалом*) этого кольца, если для всех $a \in H$, $r \in R$ имеет место $ar \in H$, $ra \in H$.

Примеры. Множество целых чисел $\mathbb{Z}$ является подкольцом поля рациональных чисел $\mathbb{Q}$, но не идеалом. Легко подобрать одно целое и одно рациональное число, произведение которых не будет целым.

Пусть R — коммутативное кольцо, $a \in R$. Положим $H = \{ar \mid r \in R\}$. Тогда H — идеал кольца R. В частности, числа, кратные данному модулю m в кольце $\mathbb{Z}$, образуют идеал $m\mathbb{Z}$.

Пусть R — коммутативное кольцо. Идеал H кольца R называется *главным идеалом* кольца R, если существует элемент $a \in R$ такой, что $H = \{ar \mid r \in R\}$. В этом случае H называют также *главным идеалом, порожденным элементом a*.

Пусть кольцо R содержит единичный элемент. Рассмотрим циклическую подгруппу аддитивной группы кольца, порожденную единицей. Она автоматически будет подкольцом, так как

$$\underbrace{(1+\ldots+1)}_{m}\underbrace{(1+\ldots+1)}_{n}=\underbrace{(1+\ldots+1)}_{mn}.$$

Она автоматически будет изоморфна либо аддитивной группе кольца $\mathbb{Z}$, либо аддитивной группе одного из колец вычетов $\mathbb{Z}_m$. В первом случае говорят, что *характеристика* кольца R равна нулю, char $R = 0$. Во втором случае полагают char $R = m$.

Теорема 3.8.1. Характеристика области целостности либо равна нулю, либо является простым числом.

Доказательство. Пусть циклическая группа, порождаемая единицей кольца R, есть $\mathbb{Z}_m$. Если бы m было составным, $m = m_1m_2$, то в силу очевидного равенства

$$\underbrace{(1+\ldots+1)}_{m_1}\underbrace{(1+\ldots+1)}_{m_2}=\underbrace{(1+\ldots+1)}_{m}=0$$

это кольцо имело бы делители нуля. □

Центром кольца R называется множество всех его элементов $a \in R$, для которых

$$ax = xa \quad \text{при всех} \quad x \in R.$$

Центр коммутативного кольца R совпадает с R.

Теорема 3.8.2. Центр любого кольца является его подкольцом.

Доказательство. Действительно, если $ax = xa$, $bx = xb$ при всех x, то

$$(a \pm b)x = ax \pm bx = xa \pm xb = x(a \pm b),$$

$$(ab)x = a(bx) = a(xb) = (ax)b = (xa)b = x(ab).$$

При наличии единицы e верно $ex = xe$. □

Если H — идеал кольца R, то факторгруппа R/H также наделяется структурой кольца. Определим умножение смежных классов по формуле

$$(a + H)(b + H) = ab + H.$$

Убедимся в корректности определения. Пусть $a + h_1$ и $b + h_2$ — другие представители смежных классов. Тогда $(a + h_1)(b + h_2) = ab + ah_2 + bh_1 + h_1h_2$. Три последних слагаемых принадлежат идеалу, поэтому имеем тот же смежный класс. Смежные классы по идеалу называют также *классами вычетов по модулю идеала*. При таком определении классы вычетов наследуют аксиомы кольца, и поэтому они образуют кольцо, называемое *фактор-кольцом* кольца R по идеалу H. Оно обозначается R/H.

Фактически мы уже имели дело с фактор-кольцом $\mathbb{Z}/m\mathbb{Z}$. Эти фактор-кольца можно задавать с помощью таблиц Кэли. В случае $p = 3$ имеем таблицы умножения и сложения поля Галуа $\mathbb{F}_3$:

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Следует отметить, что фактор-кольцо наследует не все свойства кольца. Например, в кольце $\mathbb{Z}/6\mathbb{Z}$ есть делители нуля $\overline{2} \cdot \overline{3} = 0$, хотя в самом кольце $\mathbb{Z}$ их нет.

3.9. Гомоморфизмы колец

В дальнейшем мы чаще всего будем рассматривать кольца с единицами.

Пусть R и S — кольца. *Гомоморфизмом* $\varphi: R \to S$ называется отображение, для которого

$$\varphi(a + b) = \varphi(a) + \varphi(b), \quad \varphi(ab) = \varphi(a)\varphi(b), \quad \varphi(e) = e'$$

при всех $a, b \in R$.

Термины «ядро», «образ», «эпиморфизм», «мономорфизм», «изоморфизм», «автоморфизм» и «эндоморфизм» имеют тот же смысл, что и для групп.

Теорема 3.9.1. Если φ — гомоморфизм кольца R на кольцо S, то $\ker\varphi$ — идеал кольца R, причем кольцо S изоморфно фактор-кольцу $R/\ker\varphi$. Обратно, если H — идеал кольца R, то отображение $\overline{\varphi}: R \to R/H$, определяемое условием $\varphi(a) = a + H$, является эпиморфизмом R на R/H с ядром H.

Доказательство этой теоремы аналогично доказательству соответствующей теоремы для групп.

Теорема 3.9.2. В любой области целостности R положительной харатеристики p отображение $x \to x^p$ является мономорфизмом $\varphi : R \to R$.

Доказательство. Воспользуемся формулой бинома:

$$\varphi(a + b) = (a + b)^p = a^p + C_p^1 a^{p-1} b + \ldots + C_p^p b^b.$$

Нетрудно показать, что все биномиальные коэффициенты, кроме крайних, делятся на p. Поэтому соответствующие им члены в разложении вообще отсутствуют. Итак, $\varphi(a + b) = a^p + b^p = \varphi(a) + \varphi(b)$. Далее, $\varphi(ab) = (ab)^p = a^p b^p = \varphi(a)\varphi(b)$, $\varphi(e) = e$ в любом коммутативном кольце. □

В заключение отметим, что в кольце $\mathbb{Z}$ любой идеал главный, так как является аддитивной подгруппой бесконечной циклической группы.

3.10. Евклидовы кольца

Теория делимости целых чисел может быть значительно обобщена. В частности, ее можно развить для колец многочленов над произвольными полями.

Евклидовым кольцом называется область целостности R вместе с нормой $\nu: R^* \to \mathbb{N} \cup \{0\}$ (R^* — множество ненулевых элементов кольца, $\mathbb{N} \cup \{0\}$ — неотрицательные целые числа), которая удовлетворяет следующим условиям:

$$\nu(ab) \geq \nu(a) \text{ для любых } x, y \in R^*;$$

для любых $a \in R$, $b \in R^*$ существуют элементы q и r такие, что

$$a = bq + r, \text{ где либо } r = 0, \text{ либо } \nu(r) < \nu(b).$$

Евклидово кольцо — это кольцо Z вместе с нормой, являющейся обычным модулем. Кольцо многочленов $F[x]$ над любым полем F также является евклидовым. В качестве нормы следует взять степень многочлена. Менее очевидно, что евклидово кольцо образуют целые гауссовские числа, т.е. числа вида $a + bi$, $a, b \in \mathbb{Z}$. Здесь в качестве нормы следует взять квадрат модуля $\nu(a + bi) = a^2 + b^2$. Эти числа обозначают $\mathbb{Z}[i]$.

Теорема 3.10.1. В евклидовом кольце R для любых $a, b \in R^*$

$$\nu(ab) = \nu(a),$$

если b — обратим, и $\nu(ab) > \nu(a)$ в противном случае.

Доказательство. Если b обратим, то

$$\nu(a) = \nu(abb^{-1}) \geq \nu(ab) \Rightarrow \nu(ab) = \nu(a),$$

так как $\nu(ab) \geq \nu(q)$. Пусть теперь b не обратим. В этом случае ab не делит a. В самом деле, $ab \mid a \Rightarrow a = (ab)q$. После сокращения на a имеем $bq = e$. Поэтому

$$a = (ab)q + r, \text{ где } r \neq 0, \ \nu(r) < \nu(ab).$$

Кроме того, $r = a - (ab)q = a(e - qb)$. Поэтому $\nu(r) \geq \nu(a)$. Окончательно получаем $\nu(ab) > \nu(r) \geq \nu(a)$. □

Следствие 3.10.1. В евклидовом кольце элемент a обратим тогда и только тогда, когда $\nu(a) = \nu(e)$.

Например, обратимые элементы кольца $\mathbb{Z}$ — это ± 1. В кольце целых гауссовских чисел обратимыми будут четыре элемента: ± 1, $\pm i$.

Теорема 3.10.2. В евклидовом кольце все идеалы главные.

Доказательство. Оно аналогично доказательству теоремы о том, что каждая подгруппа циклической группы циклическая. Если H — идеал, то он будет порожден элементом a с наименьшим значением нормы $\nu(a)$, $a \in H$. □

В силу наличия в евклидовом кольце алгоритма деления с остатком для него можно построить теорию делимости, аналогичную теории делимости для кольца целых чисел. В частности, можно ввести понятия НОК и НОД двух элементов. Для нахождения НОД можно использовать алгоритм Евклида. В евклидовом кольце любой элемент можно разложить в произведение простых элементов. Однако это разложение менее определенное, чем каноническое разложение целого числа, из-за наличия в произвольном евклидовом кольце обратимых элементов. Отметим в связи с этим, что все обратимые элементы евклидова кольца образуют группу по умножению. Она называется *группой единиц кольца* и обозначается через $U(R)$. Например, $U(\mathbb{Z}) = \pm 1$, $U(\mathbb{Z}[i]) = [\pm 1, \pm i]$, $U(R[x]) = R^*$. Назовем элемент $p \in R^*$ *простым*, если он необратим и не раскладывается в произведение двух необратимых множителей. Два простых элемента называются *ассоциированными*, если они отличаются обратимым множителем, т.е. $p = uq$, где $u \in U(R)$. Например, 5 и –5 — простые ассоциированные элементы в кольце Z. Отношение ассоциированности является бинарным отношением эквивалентности. Поэтому имеем разбиение простых элементов на непересекающиеся классы ассоциированных.

Теорема 3.10.3. В евклидовом кольце всякий ненулевой необратимый элемент можно представить в виде произведения степеней попарно неассоциированных простых элементов. В этом разложении классы простых элементов и их степени определены однозначно.

Доказательство этой теоремы аналогично доказательству теоремы о каноническом разложении натурального числа.

3.11. Простые и максимальные идеалы

Идеал H кольца R называется *простым*, если $ab \in H \Rightarrow a \in H$ либо $b \in H$. Идеал H кольца R называется *максимальным*, если он не содержится ни в каком большем идеале, кроме самого кольца R.

Теорема 3.11.1. Пусть R — коммутативное кольцо с единицей, H — идеал кольца R.

1. Идеал H прост тогда и только тогда, когда фактор-кольцо R/H является областью целостности.
2. Идеал H максимален тогда и только тогда, когда фактор-кольцо R/H является полем.

Доказательство. Пусть идеал H — простой. Необходимо доказать отсутствие делителей нуля в фактор-кольце. Их наличие означало бы существование элементов $a, b \in H$ таких, что $(a + h)(b + H) = H$, где $a \notin H$, $b \notin H$. Поскольку $(a + H)(b + H) = ab + H \Rightarrow ab \in H$, то получено противоречие с простотой идеала H. Аналогично доказывается обратное утверждение.

Пусть H максимален в кольце R. Докажем обратимость любого ненулевого элемента фактор-кольца $a + H$, $a \notin H$. С этой целью рассмотрим множество элементов $S = \{h + ax \mid h \in H, x \in R\}$, которое, как легко убедиться, является идеалом, большим идеала H, поскольку $a \in S$, $a \notin H$. Поэтому $S = R$. Следовательно, $e = h + ax$ для некоторых $h \in H$, $x \in R$. Поэтому

$$e + H = h + ax + H = ax + H = (a + H)(x + H),$$

т.е. в фактор-кольце $(a + H)^{-1} = (x + H)$. Обратно, пусть R/H — поле. Обозначим через S строго больший идеал, чем H. Пусть $a \in S$, $a \notin H$. В силу того, что R/H — поле, уравнение

$$(a + H)(x + H) = b + H$$

разрешимо для любого $b \in R$, поскольку $H \neq a + H$. Поэтому $ax + H = b + H$. Но $ax + H \subset S$, так как $H \subset S$. Следовательно, $ax \in H$, т.е. $S = b + S$ для любого b. Но это возможно лишь при $S = R$. Поэтому H является максимальным идеалом. □

Следствие 3.11.1. Всякий максимальный идеал прост.

Доказанная теорема позволяет строить новые конечные поля, отличные от полей типа $\mathbb{F}_p$.

Теорема 3.11.2. В евклидовом кольце идеал $H = aR$, являющийся автоматически главным, максимален тогда и только тогда, когда a — простой элемент.

Доказательство. Пусть идеал $M = aR$ максимален. Если бы при этом элемент a допускал разложение $a = a_1 a_2$, то имело бы место строгое включение $aR \subset a_1 R$. Обратно, пусть A — простой элемент. Если при этом $H = aR$ содержится в строго большем собственном идеале bR, то имеем делимость $b \mid a$. □

Пример. Многочлен $x^2 + x + 1$ является простым элементом кольца многочленов над полем $\mathbb{F}_2$, поскольку его приводимость означает наличие корней в этом поле. Здесь, конечно, мы используем то обстоятельство, что он может разлагаться лишь на два множителя первой степени. Представителями классов вычетов являются элементы 0, 1, x, $x + 1$. Над ними можно производить сложение и умножение по модулю $x^2 + x + 1$. Тем самым построено поле $\mathbb{F}_4$. Приведем таблицы умножения и сложения в этом поле.

$\cdot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

3.12. Конечные расширения полей

Пусть F и P — два поля, причем $F \subset P$. Тогда F называется *подполем* поля P, которое в свою очередь называется *надполем* поля F, или его *расширением*. Каждое поле содержит так называемое *простое подполе*, т.е. поле, порожденное единицей. Если характеристика поля положительна (равна p), то простое подполе изоморфно $\mathbb{F}_p$. Если характеристика равна нулю, то простое подполе изоморфно полю рациональных чисел. Таким образом, каждое поле является расширением своего простого подполя.

Поле P можно рассматривать как векторное пространство над полем F. Размерность этого пространства называется *степенью расширения* $F \subset P$ и обозначается $[P:F]$. Например: $[\mathbb{C}:\mathbb{R}] = 2$, $[\mathbb{R}:\mathbb{Q}] = \infty$, $[\mathbb{F}_4:\mathbb{F}_2] = 2$.

Теорема 3.12.1. Пусть даны два расширения $F \subset P$, $P \subset K$ конечной степени. Тогда

$$[K:F] = [P:F][K:P].$$

Доказательство. Пусть элементы $p_1, p_2, \ldots, p_s$ образуют базис P над F, а элементы $k_1, k_2, \ldots, k_r$ — базис K над P. Покажем, что rs элементов вида $p_i k_j$ образуют базис K над F. Любой элемент $k \in K$ можно записать в виде

$$k = \sum_i \alpha_i k_i, \ \alpha_i \in P.$$

В свою очередь,

$$\alpha_i = \sum_j \beta_{ij} p_j.$$

Значит,

$$k = \sum_{i,j} \beta_{ij} p_j k_i.$$

Поэтому элементы $p_i k_j$ являются системой образующих пространства K над полем F. Осталось доказать их линейную независимость. Пусть

$$\sum_{i,j} \alpha_{ij} p_i k_j = 0$$

при некоторых $\alpha_{ij} \in F$. Тогда

$$0 = \sum_{i,j} \alpha_{ij} p_i k_j = \sum_j \left(\sum_i \alpha_{ij} p_i \right) k_j \Rightarrow \sum_i \alpha_{ij} f_i = 0 \Rightarrow \alpha_{ij} = 0. \qquad \square$$

Элемент $\alpha \in P$ называется *алгебраическим* над полем F, если он является корнем многочлена $f(x) \in F(x)$, $f(\alpha) = 0$. Многочлен $f(x)$ называется *аннулирующим многочленом* элемента α. Среди всех аннулирующих многочленов можно выбрать многочлен наименьшей степени со старшим коэффициентом, равным единице. Такой многочлен называется *минимальным многочленом* элемента α. Минимальный многочлен очевидно неприводим, т.е. не разлагается в произ-

ведение многочленов меньшей степени. Покажем, что минимальный многочлен определен однозначно. Пусть $m_1(x)$ и $m_2(x)$ — различные минимальные многочлены. Поскольку их старшие коэффициенты одинаковые (равны единице), а их степени равны, то их разность $\varphi(x) = m_1(x) - m_2(x)$ будет ненулевым многочленом меньшей степени. Вместе с тем, $\varphi(\alpha) = m_1(\alpha) - m_2(\alpha) = 0$, что противоречит условию.

Теорема 3.12.2. Всякое расширение $F \subset P$ конечной степени алгебраично, т.е. все элементы поля P алгебраичны над полем F.

Доказательство. Для произвольного элемента $\alpha \in P$ рассмотрим его степени $1, \alpha, \alpha^2, \ldots, \alpha^n$, где $n = [P:F]$. Поскольку их количество превышает размерность пространства, они линейно зависимы, т.е. найдутся скаляры $f_0, f_1, \ldots, f_n \in F$ такие, что $f_0 + f_1\alpha + f_2\alpha^2 + \ldots + f_n\alpha^n = 0$, тем самым найден аннулирующий многочлен. □

Теорема 3.12.3. Все аннулирующие многочлены для данного элемента $\alpha \in P$ образуют главный идеал в кольце $F[x]$. Он порожден минимальным многочленом $m(x)$, т.е. имеет вид $m(x)F[x]$.

Доказательство. Достаточно показать, что любой аннулирующий многочлен делится на минимальный. Пусть $f(\alpha) = 0$. Разделим $f(x)$ с остатком на $m(x)$. Пусть $f(x) = m(x)g(x) + r(x)$. Тогда $f(\alpha) = m(\alpha)g(\alpha) + r(\alpha) = 0$. Поэтому $r(\alpha) = 0$, что возможно лишь при $r(x) = 0$. □

Пусть $F \subset P$ — произвольное расширение, $\alpha \in P$. Обозначим через $F[\alpha]$ наименьшее подкольцо поля P, содержащее элемент α и поле F, а через $F(\alpha)$ — наименьшее подполе с аналогичным условием. Очевидно, что

$$F[\alpha] = \{f(\alpha) \mid f(x) \in F[x]\},$$

$$F(\alpha) = \left\{\frac{f(\alpha)}{g(\alpha)} \mid f(x), g(x) \in F[x], g(\alpha) \neq 0\right\}.$$

Теорема 3.12.4. Если элемент α алгебраичен над F, то $F[\alpha] = F(\alpha) \cong$ $\cong F[x]/m(x)F[x]$, где $m(x)$ — минимальный многочлен для α.

Доказательство. Рассмотрим эпиморфизм колец $F[x] \to F[\alpha]$, определенный подстановкой $x \to \alpha$. Его ядро, очевидно, является идеалом $m(x)F[x]$. Осталось применить теорему о гомоморфизме кольца на фактор-кольцо, а поскольку идеал $m(x)F[x]$ максимален в силу неприводимости многочлена $m(x)$, то кольцо $F[\alpha]$ является полем и поэтому должно совпасть с $F(\alpha)$. □

$F \subset F[\alpha]$ называется *простым алгебраическим расширением*. Элемент α называется *примитивным элементом* расширения. Степень этого расширения равна

степени минимального многочлена. Пусть $m(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n$. Тогда n элементов $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ образуют базис расширения. В самом деле, они линейно независимы, так как их меньше, чем степень минимального многочлена. Пусть $f(\alpha)$ — произвольный элемент поля $F[\alpha]$. Разделим многочлен $f(x)$ с остатком на $m(x)$. Имеем $f(x) = m(x)g(x) + r(x)$. Поэтому $f(\alpha) = r(\alpha)$, что линейно выражается через $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$.

3.13. Поле разложения

Пусть $f(x)$ — неприводимый многочлен над полем F. Покажем, что для него существует конечное расширение поля F, содержащее все корни многочлена $f(x)$. Такое расширение принято называть *полем разложения*. Более точно расширение P поля F называют полем разложения многочлена $f(x)$, если:

1) $f(x)$ разлагается на линейные множители в $P[x]$;
2) P порождено как поле корнями многочлена $f(x)$ и полем F.

Поле разложения — это наименьшее поле, в котором $f(x)$ разлагается на линейные множители.

Теорема 3.13.1. Пусть многочлен $m(x)$ неприводим над полем F. Тогда над расширением $P = F[x]/m(x)F[x]$ многочлен $m(x)$ имеет корень $\bar{x} = \varphi(x)$, где φ — канонический эпиморфизм на фактор-кольцо $\varphi: F[x] \to F[x]/m(x)F[x]$.

Доказательство. Достаточно применить канонический эпиморфизм к уравнению $m(x) = 0$. □

Теорема 3.13.2. Пусть F — поле, $f(x)$ — произвольный ненулевой многочлен $f(x) \in F[x]$. Тогда его поле разложения существует и имеет степень не больше чем $n!$ над F.

Доказательство. Разложим $f(x) = p_1(x)p_2(x)\cdots p_r(x)$ на неприводимые множители. Построим поле $P_1 = F[x]/p_1(x)F[x]$ для первого множителя. Получим поле, которое содержит корень α_1 многочлена $p_1(x)$, а значит, и $f(x)$. Пусть над этим полем $f(x) = (x - \alpha_1)f_1(x)$. Применим тот же прием к многочлену $f_1(x)$ над полем P_1 и т.д. Степени построенных полей не будут превышать соответственно $n, n-1, \ldots, 2, 1$. □

Теорема 3.13.3. Поле разложения данного многочлена над полем F определяется с точностью до изоморфизма.

Доказательство. Поскольку многочлен $f(x)$ разлагается на неприводимые множители, достаточно рассмотреть случай, когда $f(x)$ — неприводим. Пусть u и v — корни неприводимого многочлена в полях G и H.

Тогда отображение

$$a_0 + a_1 u + \ldots + a_{n-1} u^{n-1} \to a_0 + a_1 v + \quad + a_{n-1} v^{n-1}$$

определяет изоморфизм $F(u) \to F(v)$:

$$F(u) \cong F[x]/f(x) \cong F(v),$$

где $F(u)$ и $F(v)$ — подполя G и H. Применяя это рассуждение несколько раз, устанавливаем изоморфизм $G \cong H$. □

3.14. Конечные поля

Всякое конечное поле P содержит простое подполе $F_p = \mathbb{F}_p$. Поскольку поле P конечное, оно имеет конечную степень n над простым подполем. Обозначим через $p_1, p_2, \ldots, p_n$ базис P над F_p. Тогда любой элемент $x \in P$ можно однозначно записать в виде $x = f_1 p_1 + f_2 p_2 + \ldots + f_n p_n$, $f_i \in F_p$. Для каждого коэффициента f_i имеется p возможностей выбора. Поэтому порядок поля p необходимо равен p^n. Однако мы не знаем, для каких n поле $\mathbb{F}_{p^n}$ действительно существует.

Наша ближайшая цель — доказать существование *поля Галуа* $\mathbb{F}_{p^n}$ для любого простого p и натурального n. Положим $q = p^n$.

Лемма 3.14.1. Многочлен $f(x) = x^q - x$ не имеет кратных корней в любом поле характеристики p, в котором он разлагается на линейные множители.

Доказательство. Вычислим производную этого многочлена и покажем, что она взаимно проста с ним: $(x^q - x)' = qx^{q-1} - 1 = p^n x^{q-1} - 1 = -1$. Поэтому $(f(x), f'(x)) = 1$. □

Заметим, что для любого ненулевого корня многочлена $x^q - x$ верно $\alpha^{q-1} = 1$.

Теорема 3.14.1. Поле разложения многочлена $x^q - x$ содержит в точности p^n элементов.

Доказательство. Достаточно показать, что p^n корней многочлена $x^q - x$ сами образуют поле. Пусть α, β — ненулевые корни этого многочлена, а значит, и многочлена $x^{q-1} - 1$. Тогда $(\alpha\beta)^{q-1} = \alpha^{q-1}\beta^{q-1} = 1 \cdot 1 = 1$. Вместе с тем в силу доказанной формулы бинома в характеристике p имеем $(\alpha + \beta)^p = \alpha^p + \beta^p$. Следовательно, $(\alpha + \beta)^q = \alpha^q + \beta^q = \alpha + \beta$, т.е. $\alpha + \beta$ также является корнем $x^q - x$. □

Следствие 3.14.1. Существует лишь одно с точностью до изоморфизма поле $\mathbb{F}_{p^n}$.

Теорема 3.14.2. Мультипликативная группа любого конечного поля P циклична.

Доказательство. Эта группа — абелева порядка $q-1$. Допустим, что она не циклична, тогда по теореме 3.2.5 НОК порядков ее элементов r будет собственным делителем $q-1$. Поэтому $x_i^r = 1$ для всех $x_i \in P^*$. Следовательно,

$$\prod_i (x - x_i) \mid (x^r - 1),$$

но это невозможно, так как произведение имеет большую степень, чем r. □

Образующий элемент циклической группы $\mathbb{F}_{p^n}^*$ называется *примитивным элементом* этого поля. Его можно взять в качестве примитивного элемента расширения $\mathbb{F}_p \subset \mathbb{F}_{p^n}$. Поэтому верна следующая теорема.

Теорема 3.14.3. Всякое конечное поле характеристики p является простым алгебраическим расширением поля $\mathbb{F}_p$.

Минимальный многочлен примитивного элемента поля $\mathbb{F}_{p^n}$ имеет степень n. Поэтому можно сказать, что это поле изоморфно фактор-кольцу $\mathbb{F}_p[x]/m(x)\mathbb{F}_p[x]$ для некоторого неприводимого многочлена $m(x)$ степени n. При этом $m(x) \mid (x^{q-1} - 1)$.

3.15. Порядки неприводимых многочленов

Теорема 3.15.1. Поле $\mathbb{F}_{p^n}$ является полем разложения всякого неприводимого многочлена $f(x)$ степени n над полем $\mathbb{F}_{p^n}$.

Доказательство. Пусть α — корень $f(x)$. Тогда $[\mathbb{F}_p(\alpha):\mathbb{F}_p] = n$. Поэтому $\mathbb{F}_p(\alpha) = \mathbb{F}_{p^n}$. Применяя формулу бинома в характеристике p, получаем

$$f(\alpha^p) = (a_0\alpha^n + a_1\alpha^{n-1} + \ldots + a_0)^p = (f(\alpha))^p = 0.$$

Следовательно, элементы $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}$ являются корнями многочлена $f(x)$. Остается показать, что все они различны. Пусть $\alpha^{p^j} = \alpha^{p^k}$, где $0 < j < k \le n-1$. Возводя это равенство в степень p^{n-k}, получаем

$$\alpha^{p^{n-k+j}} = \alpha^{p^n} = \alpha.$$

Следовательно, неприводимый многочлен $f(x)$ делит $x^{n-k+j} - x$. Это означает, что поле $\mathbb{F}_{p^n}$ является подполем поля $\mathbb{F}_{p^l}$, где $l = n - k + j < n$. Противоречие. □

Следствие 3.15.1. Порядки всех корней неприводимого многочлена равны.

Доказательство. Корни многочлена $\alpha, \alpha^p, \ldots, \alpha^{p^{n-1}}$ имеют одинаковый порядок в мультипликативной группе поля $\mathbb{F}_{p^n}$ в силу теоремы 3.3.2. □

Назовем порядком многочлена $f(x) \in F_p[x]$, $f(0) \neq 0$ наименьшее натуральное e, при котором многочлен $f(x)$ делит $x^e - 1$. Порядок обозначается $ord(f(x))$.

Следствие 3.15.2. Пусть $f(x)$ — неприводимый многочлен степени n над полем F_p. Порядок этого многочлена совпадает с порядком любого его корня в мультипликативной группе поля разложения.

3.16. Линейные рекуррентные последовательности

Последовательность элементов $s_0, s_1, \ldots$ поля $\mathbb{F}_q$, удовлетворяющих условию

$$s_{n+k} = a_{k-1}s_{n+k-1} + a_{k-2}s_{n+k-2} + \ldots + a_0 s_k, \tag{3.16.1}$$

где $a_{k-1}, a_{k-2}, \ldots, a_0$ — фиксированные элементы поля, называется *линейной рекуррентой* (ЛРП) k-го порядка над полем $\mathbb{F}_q$. Эта последовательность полностью определяется вектором начального состояния $S_0 = (s_0, s_1, \ldots, s_{k-1})$ и коэффициентами $a_{k-1}, a_{k-2}, \ldots, a_0$.

С линейной рекуррентой можно связать матрицу

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & a_0 \\ 1 & 0 & \ldots & 0 & a_1 \\ 0 & 1 & \ldots & 0 & a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & a_{k-1} \end{pmatrix}.$$

Рассмотрим теперь последующие состояния ЛРП $S_1 = (s_1, s_2, \ldots, s_k)$, $S_2 = (s_2, s_3, \ldots, s_{k+1})$, ... Определение (3.16.1) можно переписать в виде

$$S_i = S_{i-1}A, \quad i = 1, 2, \ldots \tag{3.16.2}$$

Далее рассматриваются лишь ЛРП с условием $a_0 \neq 0$. В этом случае матрица A является элементом группы $GL(k, F_q)$ всех невырожденных матриц k-го порядка с элементами из поля $\mathbb{F}_q$. Поскольку эта группа конечна, то матрица A имеет конечный порядок как элемент группы.

Теорема 3.16.1. Любая линейная рекуррента при $a_0 \neq 0$ является чисто периодической последовательностью.

Доказательство. Существует натуральное l такое, что $S^l = I_k$, где I_k — единичная матрица. Следовательно,

$$S_{i+l} = S_i A^l = S_i I_k = S_i.$$

Это означает, что период ЛРП не превосходит l. □

Более подробно линейные рекурренты изучаются в п. 8.4, 8.5, 9.3.

3.17. Последовательности максимального периода

Под периодом ЛРП будем понимать ее минимальный период.

Рассмотрим характеристический многочлен $f(x) = |xE - S|$ матрицы S. Легко подсчитать, что $f(x) = x^k - s_{k-1}x^{k-1} - \ldots - s_1 x - s_0$. Можно показать, что он будет также и минимальным многочленом матрицы S. Этот многочлен называется характеристическим многочленом ЛРП (3.16.1).

Теорема 3.17.1. Пусть многочлен $f(x) = |xE - A|$ неприводим. Тогда его порядок совпадает с порядком матрицы S как элемента группы $GL(k, Fq)$.

Доказательство. Корни многочлена $f(x)$ являются характеристическими числами матрицы S. Если l — порядок матрицы, то $S^l = E$ и, значит, для любого корня α многочлена выполняется $\alpha^l = 1$. Поэтому $\operatorname{ord} f(x) \le l$. С другой стороны, у матрицы $S^{\operatorname{ord} f(x)}$ все собственные значения равны 1, а ее минимальный многочлен неприводим. Поэтому $S^{\operatorname{ord} f(x)} = E \Rightarrow l \le \operatorname{ord} f(x)$. □

Теорема 3.17.2. Минимальный период линейной рекурренты с неприводимым характеристическим многочленом $f(x)$, $f(0) \ne 0$ при ненулевом начальном состоянии равен порядку многочлена $f(x)$.

Доказательство. Пусть $l = \operatorname{ord} f(x) = \operatorname{ord} S$. По теореме 3.16.1 l — период. Необходимо лишь установить его минимальность. Пусть, например, $A_i = A_j$, $i < j$, $j - i < l$. Это означает, что $A_0 S^i = A_0 S^j$, т.е. $A_0 S^{j-i} = A_0$. Противоречие: у матрицы S^{j-i} не может быть единичного собственного значения. □

Данная теорема показывает, как для генерации последовательностей максимального периода можно использовать примитивные многочлены.

3.18. Задания

1. Выяснить, обладают ли свойствами ассоциативности и коммутативности операции $*$ на множестве A, если

$A = N$, $x * y = x + 2y$; $A = N$, $x * y = x^y$;

$A = N$, $x * y = 3xy$; $A = N$, $x * y = \text{НОД}(x, y)$;

$A = Z$, $x * y = x - y$; $A = Z$, $x * y = x^2 + y^2$;

$A = R$, $x * y = \sin x \cos y$; $A = R$, $x * y = x^{|y|}$.

2. Какие из указанных числовых множеств являются группами относительно заданных операций:
   1) множество степеней данного вещественного числа с целыми показателями относительно умножения;
   2) множество комплексных чисел с фиксированным модулем относительно операции умножения;
   3) множество положительных действительных чисел относительно операции умножения;
   4) отрезок [0,1] относительно операции умножения;
   5) отрезок [0,1] относительно операции $\alpha * \beta = \{\alpha + \beta\}$;
   6) корни всех степеней из единицы относительно умножения?

3. Какие из указанных множеств квадратных матриц фиксированного порядка образуют группу относительно операции умножения:
   1) множество симметрических матриц с вещественными элементами;
   2) множество невырожденных матриц с вещественными элементами;
   3) множество целочисленных матриц с определителем, равным ±1;
   4) множество верхних треугольных матриц с вещественными элементами;
   5) множество ортогональных матриц;
   6) множество диагональных матриц?

4. Доказать, что множество функций вида $y = (ax + b)(cx + d)^{-1}$, где $a, b, c, d \in \mathbb{R}$, $ad - bc \neq 0$, является группой относительно операции композиции функций.

5. Доказать, что если в группе G выполнено условие $x^2 = e, x \in G$, то группа G коммутативна.

6. Доказать, что в любой группе $(ab)^{-1} = b^{-1}a^{-1}$ и вообще $(a_1 a_2 \cdots a_n)^{-1} = a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}$.

7. Доказать, что множество всех квадратных матриц данного порядка, в каждой строке и каждом столбце которых один элемент равен 1, а остальные — 0, образует группу.

8. *Составить таблицу Кэли для циклической группы пятого порядка.*

9. Найти с точностью до изоморфизма все группы порядков 3, 4, 6. Составить их таблицы Кэли.

10. Доказать, что в конечной группе любое подмножество H, замкнутое по умножению ($h_1, h_2 \in H \Rightarrow h_1 h_2 \in H$), будет подгруппой.

11. Доказать, что для нетривиального смежного класса $gH (e \notin gH)$ выполнено условие $h_1, h_2 \in gH \Rightarrow h_1 h_2 \notin H$.

12. Доказать, что пересечение двух подгрупп будет подгруппой.

13. Доказать, что в любой группе подстановок, содержащей хотя бы одну нечетную подстановку, количество четных подстановок равно количеству четных.

14. Найти все подгруппы групп S_3, Z_6, Z_{24}.

15. Разложить:
    1) аддитивную группу вещественных чисел по подгруппе целых чисел;
    2) аддитивную группу комплексных чисел по подгруппе целых гауссовских чисел;
    3) симметрическую группу S_n по подгруппе подстановок, оставляющей элемент 1 на месте;
    4) группу Z_6 по своим подгруппам;
    5) группу S_3 по своим подгруппам.

16. Пусть $a \in G$. Доказать, что множество элементов $\{x; xa = ax, x \in G\}$, называемое централизатором элемента a, является подгруппой.

17. Доказать, что множество элементов $\{x \mid xa = ax, \forall a \notin G\}$, называемое центром группы, является нормальной подгруппой.

18. Найти централизаторы элементов

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 5 & 4 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

    в группе $GL(2, R)$.

19. Найти центр группы $GL(2, R)$.

20. Показать, что порядки элементов a и a^{-1} равны.

21. Показать, что порядки элементов ab и ba равны.

22. Доказать, что аддитивные группы вещественных и рациональных чисел не являются циклическими.

23. Пусть A и B нормальны в G, $A \cap B = e$, тогда каждый элемент $a \in A$ перестановочен с каждым элементом $b \in B$.

24. Подгруппа, порожденная элементами вида $aba^{-1}b^{-1}$ (коммутаторами), называется коммутантом. Доказать, что
    1) $aba^{-1}b^{-1} = e \Leftrightarrow ab = ba$;
    2) $K \triangleleft G$;

3) фактор-группа G/K — абелева;
4) если G/H — абелева, то $K \subset H$.

25. Определить с точностью до изоморфизма все абелевы группы порядка 8.

26. Пусть H — подмножество группы G с условием $h_1, h_2 \in H \Rightarrow h_1 h_2 \notin H$. Доказать, что $|H| \le 0.5|G|$.

27. Вычислить следующие произведения подстановок:
1) (1,2) (1,3) (1,4) (2,3) (3,5);
2) (125) (124) (129);
3) f^{100}, где $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 5 & 4 & 6 & 9 & 7 & 1 & 10 & 8 & 2 \end{pmatrix}$.

28. Найти подстановку X, если $AXB^2 = C$,

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 3 & 2 & 1 & 6 & 5 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 1 & 2 & 7 & 4 & 5 & 6 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 1 & 3 & 6 & 4 & 7 & 2 \end{pmatrix}.$$

29. Доказать, что две подстановки сопряжены тогда и только тогда, когда имеют одинаковое количество циклов каждой длины.

30. Доказать, что в булевом кольце ($x^2 = x$): 1) умножение коммутативно; 2) $x + x = 0$; 3) кольцо не является областью целостности.

31. Доказать, что в определении кольца с единицей не обязательно требовать, чтобы операция «+» была коммутативной.

32. Доказать, что следующие подмножества являются подкольцами в кольце $M_n(R)$:
1) диагональные матрицы;
2) верхнетреугольные матрицы.

33. Найти все подкольца колец вычетов $\mathbb{Z}_7$, $\mathbb{Z}_{10}$, $\mathbb{Z}_{12}$.

34. Может ли в кольце, не являющемся полем, содержаться некоторое поле?

35. Показать, что эндоморфизмы абелевой группы образуют кольцо.

36. Показать, что биекция $a + b\sqrt{2} \leftrightarrow a + b\sqrt{3}$ не является изоморфизмом полей $Q(\sqrt{2})$, $Q(\sqrt{3})$ и что эти поля вообще неизоморфны.

37. Докажите, что $Q(\sqrt{2}+\sqrt{3}) = Q(\sqrt{2}-\sqrt{3}) = Q(\sqrt{2},\sqrt{3})$.

38. Установить, что I — идеал кольца K. Является ли фактор-кольцо полем?
1) $K = \{a + bi \mid a,b \in \mathbb{Z}\}$, $I = \{a + bi \mid a,b \in 3\mathbb{Z}\}$.
2) $K = \{a + b\sqrt{2} \mid a,b \in \mathbb{Z}\}$, $I = \{a + b\sqrt{2} \mid a,b \in 3\mathbb{Z}\}$.
3) $K = \mathbb{Z}_3[x]$, $I = (x^2 + 1)$.
4) $K = \mathbb{Z}_2[x]$, $I = (x^2 + x)$.
5) $K = \mathbb{R}[x]$, $I = (x^2 + 1)$.

39. Найдите идеал, порожденный множеством M, если
1) $M = \{3,5\}$ в кольце $\mathbb{Z}$;
2) $M = \{4,10\}$ в кольце $\mathbb{Z}$;
3) $M = \{x^6 - 1, x^4 - 1\}$ в кольце $\mathbb{R}[x]$;
4) $M = \{x, x+1\}$ в кольце $\mathbb{R}[x]$.

40. Доказать, что фактор-кольцо $R[x]/(x^4 + x^3 + x + 1)$ не может быть полем ни для какого коммутативного кольца R.

41. Вычислить образ $(2x+1)^{-1}$ в фактор-кольце $F[x]/(x^3 - 2)$, где
1) $F = Q$;
2) $F = \mathbb{Z}_5$;
3) $F = F_7$.

42. Доказать, что $(x^m - 1) \mid (x^n - 1) \Leftrightarrow m \mid n$ над любым полем коэффициентов.

43. Является ли $C[0,1]$ областью целостности? Показать, что отображение $f \to f(a)$ является эпиморфизмом, а ядро — максимальным идеалом.

44. Показать, что если $p(x)$ приводим, то идеал $(p(x))$ немаксимален.

45. Показать, что $x^2 + x + 1$, $x^3 + x + 1$, $x^4 + x + 1$ неприводимы над F_2 и что нет других неприводимых многочленов 2-й и 3-й степени.

46. Составить таблицы умножения и сложения для колец $\mathbb{Z}_2[x]/(x^2 + x + 1)$, $\mathbb{Z}_2[x]/(x^3 + x + 1)$, $\mathbb{Z}_2[x]/(x^3 + x^2 + 1)$.

47. Применив алгоритм Евклида, найти наибольшие общие делители многочленов с коэффициентами из поля F:
1) $F = F_2$, $x^7 + 1$, $x^5 + x^3 + 1$;
2) $F = F_2$, $x^5 + x + 1$, $x^6 + x^5 + x^4 + 1$;
3) $F = F_3$, $x^8 + 2x^5 + x^3 + x^2 + 1$, $2x^6 + x^5 + 2x^3 + 2x^2 + 2$.

48. Вычислить $f(3)$, если $f(x) = x^{214} + 3x^{152} + 2x^{47} + 2 \in F_5[x]$.

49. Решить, если возможно, сравнения:
1) $(x^2+1)f(x)=1(\mathrm{mod}(x^3+1))$ в $F^3[x]$;
2) $(x^4+x^3+x^2+1)f(x)=x^2+1(\mathrm{mod}(x^3+1))$ в $F_2[x]$.

50. Пусть $f(x)\in F_p[x]$, тогда $(f(x))^p=f(x^p)$. Доказать.

51. Доказать, что в коммутативном кольце характеристики p

$$(a+b)^{p^n}=a^{p^n}+b^{p^n}.$$

Доказать неприводимость многочленов x^2+1, x^2+x+4 над полем F_{11} и построить изоморфизм фактор-колец

$$F_{11}[x]/(x^2+1)\cong F_{11}[x]/(x^2+x+4).$$

52. Найти порядки многочленов: $x^{10}+x^9+x^3+x^2\in F_2[x]$, $x^8+x^7+x^3+x+$ $+1\in F_2[x]$, $x^7+x^6+x^4-x^2+x\in F_3[x]$, $(x^2+x+1)^5(x^3+x+1)\in F_2[x]$. Какие из них неприводимы?

53. Найти примитивный многочлен степени 6 над полем F_2.

54. Найти примитивный элемент α в $\mathbb{Z}[x]/(x^2-2)$, представить степени α^2, ..., α^8 в виде $a+b\sqrt{2}$, $a,b\in F_3$. Однозначно ли такое представление?

55. Можно ли вложить F_4 в F_8?

56. При каком условии поле F_{p^n} можно изоморфно вложить в поле F_{p^m}?

57. Найти все примитивные элементы полей F_7, F_9, F_{17}.

58. Найти базис поля F_{25} над простым подполем. Разложить все элементы по этому базису. Найти примитивный элемент α этого поля и для любого элемента $\beta\in F_{25}$ найти n такое, что $a=\beta^n$.

59. Показать, что каждый элемент конечного поля F_{p^n} имеет в нем только один корень p-й степени.

60. Доказать, что если I — идеал кольца K, то $I[x]$ — идеал кольца $K[x]$.

61. Показать, что элемент $\sqrt{2}+i$ имеет степень 4 над Q и степень 2 над R. Найти его минимальные многочлены.

62. Доказать, что если многочлен $F(x)$ неприводим в $F_q[x]$, то $F(ax+b)$ также неприводим, где $a,b\in F_q$, $a\neq 0$.

63. Разложить многочлены на неприводимые множители: $x^9 + x + 1$ над F_2, $x^7 + x^6 + x^5 - x^3 + x^2 - x - 1$ над F_3.

64. С помощью матриц дать представление для элементов поля F_8, используя многочлен $x^3 + x + 1$. Дать аналогичное представление для F_{16}, F_9.

65. Доказать неприводимость многочлена $x^4 + x + 1$ над F_2 и построить таблицы операций для его поля корня.

66. Показать, что поле корня $x^3 + x + 1$ над F_2 есть поле разложения.

67. Найти поля разложения $(x^2 - 3)(x^3 + 1)$ над Q, $(x^2 - 3)(x^2 - 2x - 2)$ над Q, $x^3 + x + 1$ над F_2.

68. Найти изоморфизм полей разложения $x^3 + 2x + 1$ и $x^3 + 2x + 2$ над F_3.

69. Найти степени неприводимых множителей многочлена $x^{17} - 1$ над F_2 и его поле разложения.

70. Установить примитивность многочленов $x^6 + x^5 + x^2 + x + 1$ над F_2, $x^5 - x + 1$ над F_3.

71. Найти хотя бы один примитивный многочлен степени 3 над полем F_4.

72. Установить неприводимость, примитивность и найти порядок $x^4 + x^3 + x^2 - x - 1$ над полем F_3.

73. Пусть A — любое множество автоморфизмов поля F. Показать, что элементы, инвариантные относительно всех $\alpha \in A$, образуют подполе $S(A) \subset F$.

74. Над полем F_2 рассматривается рекуррентное уравнение $s_i - s_{i-2} + s_{i-3} = 0$. Вычислить периоды всех 8 последовательностей и их внутрипериодные значения.

75. Построить регистр сдвига с обратной связью, реализующий соотношение из предыдущего задания.

76. Вычислить порядки матриц

$$A_1 = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

над полями F_2 и F_3.

77. Вычислить характеристический и минимальный многочлены матрицы S из п. 3.17.

78. Построить последовательность максимального периода над полем F_3 с периодом не меньше 50.

79. Найти примитивный многочлен четвертой степени над F_2. Построить соответствующий регистр сдвига и последовательность периода 15.

SXJKN8EYHR3I4HR987Y
DWFUIHW9E8YR384H5O3
4596EUF9O8WE76R83Y4
39846UHFE98YDSFJH83
74H4BR8276RUBNWDF87
Y6234BN9WE872H3WRTJ
OJRTGM905JRO9G70954
TOIRF7ERFH2347FQ7WE
RF9UJ1349YUR7DFKJHG
JDFBVDGFJVBDSKFJHKJ

Глава 4

ВВЕДЕНИЕ В ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ

SXJKN8EYHR3I4HR987Y
DWFUIHW9E8YR384H5O3
4596EUF9O8WE76R83Y4
39846UHFE98YDSFJH83
74H4BR8276RUBNWDF87
Y6234BN9WE872H3WRTJ
OJRTGM905JRO9G70954
TOIRF7ERFH2347FQ7WE
RF9UJ1349YUR7DFKJHG
JDFBVDGFJVBDSKFJHKJ

4.1. Уравнение Вейерштрасса, дискриминант и j-инвариант

В последнее время в криптографии все чаще используются эллиптические кривые (см. главу 15). Цель данной главы — ознакомить читателя с основными фактами этой увлекательной теории. К сожалению, из-за ограниченности объема книги дать полное изложение вопроса невозможно. Пропущенные доказательства можно найти в книге [184]. Для понимания содержания главы необходимо знать понятие алгебраического замыкания поля K, в котором каждое алгебраическое уравнение имеет корень. Об этом можно прочитать в любом серьезном учебнике по алгебре, например, Б.Л. Ван дер Вардена.

Определение 4.1.1. *Эллиптической кривой над полем* K называется алгебраическая кривая, которая задается *уравнением Вейерштрасса*:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{4.1.1}$$

где $a_1, a_2, a_3, a_4, a_6 \in K$.

Для обозначения эллиптической кривой используют запись E или E/K, чтобы подчеркнуть, что кривая определена над K. Точки, лежащие на эллиптической кривой, являются решениями (x, y) уравнения (4.1.1) (это так называемые *аффинные точки*); единственная, лежащая на E (в проективной плоскости) *точка на бесконечности*, в дальнейшем обозначается $\mathcal{O}$.

Определение 4.1.2. Точка эллиптической кривой, координаты которой принадлежат F, где F — некоторое расширение поля K, называется *F-рациональной*. Множество всех F-рациональных точек эллиптической кривой E будем обозначать $E(F)$. Под E будем также понимать множество $\bar{K}$-рациональных точек, где $\bar{K}$ — алгебраическое замыкание K.

Определение 4.1.3. Точка $P = (x_0, y_0) \in E$ называется *невырожденной* (*гладкой*), если для многочлена $f(x,y) = y^2 + a_1xy + a_3y - (x^3 + a_2x^2 + a_4x + a_6)$ верно по крайней мере одно из условий

$$\frac{\partial f}{\partial x}(x_0, y_0) \neq 0 \quad \text{или} \quad \frac{\partial f}{\partial y}(x_0, y_0) \neq 0. \tag{4.1.2}$$

Эллиптическая кривая E/K называется *невырожденной* (*гладкой*), если каждая аффинная точка $E(\overline{K})$ является невырожденной.

Определение 4.1.4. *Дискриминантом эллиптической кривой* E называется величина

$$\Delta = -d_2^2 d_8 - 8d_4^3 - 27d_6^2 + 9d_2 d_4 d_6.$$

При $\Delta \neq 0$ *j-инвариантом эллиптической кривой* E называется величина

$$j(E) = c_4^3/\Delta,$$

где $d_2 = a_1^2 + 4a_2$, $d_4 = 2a_4 + a_1 a_3$, $d_6 = a_3^2 + 4a_6$, $d_8 = a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2$, $c_4 = d_2^2 - 24d_4$.

Теорема 4.1.1. Эллиптическая кривая E невырождена тогда и только тогда, когда $\Delta(E) \neq 0$.

Будем рассматривать только невырожденные эллиптические кривые. Уравнение (4.1.1) является наиболее общим, однако в зависимости от характеристики поля, над которым задана кривая, оно может быть приведено к более простому виду путем линейной замены переменных. При изучении эллиптических кривых допускаются не всякие линейные замены, а лишь те, которые осуществляют изоморфизм эллиптических кривых как проективных многообразий. Определим изоморфизм эллиптических кривых следующим образом.

Определение 4.1.5. Две эллиптические кривые E и $\tilde{E}$ над полем K, заданные уравнениями

$$E: y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6,$$

$$\tilde{E}: y^2 + \tilde{a}_1 xy + \tilde{a}_3 y = x^3 + \tilde{a}_2 x^2 + \tilde{a}_4 x + \tilde{a}_6,$$

будем называть изоморфными над K, если существуют $u, r, s, t \in K$, $u \neq 0$ такие, что замена переменных

$$(x, y) \mapsto (u^2 x + r,\ u^3 y + u^2 s x + t) \tag{4.1.3}$$

переводит кривую E в $\tilde{E}$.

Эллиптические кривые, заданные над полем K, могут быть не изоморфны над этим полем, но могут стать изоморфными над расширением. Изоморфность эллиптических кривых также можно описать в терминах j-инварианта.

Теорема 4.1.2. Две эллиптические кривые E_1 и E_2 изоморфны над алгебраическим замыканием поля K тогда и только тогда, когда $j(E_1) = j(E_2)$.

Замена переменных (4.1.3) в уравнении (4.1.1) называется *допустимой*. Легко проверить, что это преобразование обратимо и обратное преобразование также допустимо. Кроме того, допустима тождественная замена переменных, поэтому всякая эллиптическая кривая изоморфна самой себе. Также несложно проверить, что композиция допустимых преобразований является допустимой.

Таким образом, изоморфность эллиптических кривых является отношением эквивалентности, и множество эллиптических кривых разбивается на классы эквивалентных, причем каждый класс эллиптических кривых, изоморфных над алгебраическим замыканием данного поля, однозначно определяется величиной j-инварианта.

Теорема 4.1.3. Всякая эллиптическая кривая над полем характеристики 2 изоморфна некоторой эллиптической кривой одного из следующих двух видов:

$$y^2 + xy = x^3 + b_2x^2 + b_6 \tag{4.1.4}$$

или

$$y^2 + b_3y = x^3 + b_4x + b_6. \tag{4.1.5}$$

Доказательство. Пусть K — поле характеристики 2, E — эллиптическая кривая, заданная уравнением (4.1.1). Если $a_1 \neq 0$, то замена переменных $(x,y) \mapsto (a_1^2x + a_1^{-1}a_3, a_1^3y)$ переводит E в кривую

$$y^2 + xy = x^3 + \tilde{a}_2x^2 + \tilde{a}_4x + \tilde{a}_6, \tag{4.1.6}$$

которая переводится в (4.1.4) заменой $(x,y) \mapsto (x, y + \tilde{a}_4)$.

Если $a_1 = 0$, то замена переменных $(x,y) \mapsto (x + a_2, y)$ переводит E в кривую вида (4.1.5). □

Теорема 4.1.4. Всякая эллиптическая кривая над полем характеристики $\neq 2$ изоморфна некоторой эллиптической кривой вида

$$y^2 = x^3 + b_2x^2 + b_4x + b_6. \tag{4.1.7}$$

Доказательство. Пусть K — поле характеристики $\neq 2$, E — эллиптическая кривая, заданная уравнением (4.1.1). Тогда замена переменных

$$(x,y) \mapsto \left(x, y - \frac{a_1}{2}x - \frac{a_3}{2}\right)$$

переводит кривую E в кривую (4.1.7). □

Теорема 4.1.5. Всякая эллиптическая кривая над полем характеристики > 3 изоморфна некоторой эллиптической кривой вида

$$y^2 = x^3 + ax + b. \tag{4.1.8}$$

Доказательство. Пусть K — поле характеристики > 3, E — эллиптическая кривая, заданная уравнением (4.1.1). Согласно теореме 4.1.4, данная эллиптическая кривая изоморфна кривой (4.1.7). Тогда замена переменных

$$(x,y) \mapsto (x - b_2/3, y)$$

переводит кривую (4.1.7) в (4.1.8). □

Для эллиптической кривой, заданной уравнением (4.1.8), формулы для дискриминанта и j-инварианта принимают вид

$$\Delta = -16(4a^3 + 27b^2),$$

$$j(E) = -1728\frac{64a^3}{\Delta} = 1728\frac{4a^3}{4a^3 + 27a^2}. \tag{4.1.9}$$

4.2. Сложение точек эллиптической кривой

Существует ряд эквивалентных способов для описания группового закона сложения точек на эллиптической кривой. Геометрически этот закон можно сформулировать следующим образом: *три коллинеарные (лежащие на одной прямой) точки на эллиптической кривой дают нулевую сумму.* Известно, что операция сложения точек на эллиптической кривой, введенная по данному правилу, превращает эллиптическую кривую в коммутативную группу. Далее будут рассмотрены формулы сложения точек эллиптической кривой, которые вытекают из данного правила.

В качестве нулевого элемента выбирается точка $\mathcal{O}$ на бесконечности. Таким образом, для любой точки $P \in E$ имеют место равенства $P + \mathcal{O} = \mathcal{O} + P = \mathcal{O}$.

Вертикальные прямые в плоскости xy проходят через точку $\mathcal{O}$[1]. Пусть $P \neq \mathcal{O}$ — точка эллиптической кривой E и l — вертикальная прямая, проходящая через точку $P = (x_1, y_1)$. Эта прямая пересекает E еще и в точке Q (с учетом кратности). Таким образом, $-P = Q$. Пусть координаты точки Q равны (x_2, y_2). Так как она лежит на l, то $x_2 = x_1$. Тогда y_2 является решением уравнения

$$y^2 + f_1(x_1)y - f_3(x_1) = 0, \tag{4.2.1}$$

где $f_1(x) = a_1x + a_3$, $f_3(x) = x^3 + a_2x^2 + a_4x + a_6$.

Уравнение (4.2.1) имеет два корня: y_1 и y_2. По теореме Виета $y_1 + y_2 = -f_1(x_1)$, откуда получаем $y_2 = -y_1 - f_1(x_1)$. Итак,

$$-(x_1, y_1) = (x_1, -y_1 - a_1x_1 - a_3). \tag{4.2.2}$$

---

[1] Точка $\mathcal{O}$ задается в проективной плоскости координатами (0:1:0); всякая прямая в проективной плоскости имеет вид $ax + by + cz = 0$; таким образом, так как $\mathcal{O}$ лежит на прямой, то $b = 0$, т.е. прямая вертикальная.

Может оказаться, что $y_2 = y_1$ и, соответственно, $-P = P$. Это возможно, если вертикальная прямая, проходящая через P, касается кривой E в точке P.

Рассмотрим, как вычисляется сумма двух различных точек $P_1 = (x_1, y_1)$ и $P_2 = (x_2, y_2)$ на эллиптической кривой. Предположим, что $P_1, P_2 \neq O$ и $P_1 \neq \pm P_2$. Тогда $x_1 \neq x_2$. Пусть l — прямая, проходящая через точки P_1 и P_2. Уравнение l имеет вид

$$y = \alpha x + \beta. \tag{4.2.3}$$

Коэффициенты α и β — решения системы

$$\begin{cases} y_1 = \alpha x_1 + \beta, \\ y_2 = \alpha x_2 + \beta. \end{cases}$$

Тогда $\alpha = \dfrac{y_1 - y_2}{x_1 - x_2}$, $\beta = y_1 - \alpha x_1 = y_2 - \alpha x_2$.

Прямая l пересекает эллиптическую кривую E в некоторой третьей точке $Q = -(P_1 + P_2) = (\tilde{x}_3, \tilde{y}_3)$. Подставляя (4.2.3) в уравнение Вейерштрасса, получаем

$$(\alpha x + \beta)^2 + (\alpha x + \beta) f_1(x) - f_3(x) = 0.$$

Это кубическое уравнение имеет три корня: $x_1, x_2, \tilde{x}_3$. По теореме Виета сумма $x_1 + x_2 + \tilde{x}_3$ равна $\alpha^2 + a_1\alpha - a_2$. Таким образом,

$$\begin{cases} x_3 = -x_1 - x_2 + \alpha^2 + a_1\alpha - a_2, \\ \tilde{y}_3 = \alpha\tilde{x}_3 + \beta = y_1 - \alpha(x_1 - \tilde{x}_3). \end{cases}$$

Используя (4.2.2), находим

$$P_1 + P_2 = (x_3, y_3), \text{ где } \begin{cases} x_3 = -x_1 - x_2 + \alpha^2 + a_1\alpha - a_2, \\ y_3 = -y_1 + \alpha(x_1 - x_3) - a_1 x_3 - a_3. \end{cases} \tag{4.2.4}$$

Найдем правило вычисления удвоенной точки $P + P$. Будем предполагать, что $P \neq -P$. Чтобы найти $P + P$, проведем касательную через точку P. Эта прямая задается уравнением (4.2.3), только α равно отрицательному отношению частных производных многочлена, задающего в неявном виде эллиптическую кривую:

$$\alpha = \frac{3x_1^2 + 2a_2 x_1 + a_4 - a_1 y_1}{2y_1 + a_1 x_1 + a_3}. \tag{4.2.5}$$

Таким образом, удвоенная точка также вычисляется по формулам (4.2.4), но с α, найденным по (4.2.5).

Теорема 4.2.1. Пусть K — поле, E/K — эллиптическая кривая. Тогда для любого расширения F поля K относительно введенной выше операции «+» множество $E(F)$ образует абелеву группу. Операция сложения задается по следующим правилам.

1. Для любой точки $P \in E$ верно $P + O = O + P = P$.

2. Для любой точки $P = (x, y) \neq O$ точка $-P$ находится по формуле $-P = (x, -y - a_1 x - a_3)$.

3. Для любых $P_1, P_2 \neq O$ с координатами $P_i = (x_i, y_i)$, $i = 1, 2$ таких, что $P_1 \neq -P_2$, сумма этих точек равна точке $P_1 + P_2 = (x_3, y_3)$ такой, что

$$\begin{cases} \alpha = \begin{cases} \dfrac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3}, & \text{если } x_1 = x_2, \\ \dfrac{y_1 - y_2}{x_1 - x_2}, & \text{если } x_1 \neq x_2, \end{cases} \\ x_3 = -x_1 - x_2 + \alpha^2 + a_1\alpha - a_2, \\ y_3 = -y_1 + \alpha(x_1 - x_3) + a_1x_3 - a_3. \end{cases} \tag{4.2.6}$$

Следствие 4.2.1. Для кривой, заданной уравнением (4.1.8), формулы (4.2.6) принимают упрощенный вид:

$$\begin{cases} \alpha = \begin{cases} \dfrac{3x_1^2 + a}{2y_1}, & \text{если } x_1 = x_2, \\ \dfrac{y_1 - y_2}{x_1 - x_2}, & \text{если } x_1 \neq x_2, \end{cases} \\ x_3 = -x_1 - x_2 + \alpha^2, \\ y_3 = -y_1 + \alpha(x_1 - x_3). \end{cases} \tag{4.2.7}$$

Кроме того, противоположная к $P = (x, y)$ точка равна $(x, -y)$.

4.3. Эллиптические кривые над конечными полями

Здесь изложены основные факты, связанные с эллиптическими кривыми, заданными над конечными полями. Преимущественно будем рассматривать эллиптические кривые над полями характеристики больше 3. Прежде всего разберем вопрос о количестве точек на эллиптической кривой.

Пусть задана эллиптическая кривая E:

$$y^2 = x^3 + ax + b \tag{4.3.1}$$

над полем $\mathbb{F}_p$, где p — простое число больше l. Величины a и b можно считать целыми числами, а аффинные точки кривой E рассматривать как решения сравнения

$$y^2 \equiv x^3 + ax + b \pmod{p}. \tag{4.3.2}$$

Количество точек на данной эллиптической кривой равно

$$1 + \sum_{x=0}^{p-1} \{\text{количество решений (4.3.2) при данном } x\}. \tag{4.3.3}$$

Суммируемая величина равна 0 (если $x^3 + ax + b$ — квадратичный невычет), 1 (если $x^3 + ax + b \equiv 0 \pmod p$) и 2 (если $x^3 + ax + b$ — квадратичный вычет). Используя символ Лежандра, эту величину можно записать в виде

$$1 + \left(\frac{x^3 + ax + b}{p}\right).$$

Подставляя это выражение в (4.3.3), получаем следующую теорему.

Теорема 4.3.1. Пусть $a, b \in \mathbb{Z}$, p — простое число больше l и над полем $\mathbb{F}_p$ задана эллиптическая кривая (4.2.3). Тогда количество $\mathbb{F}_p$-рациональных точек на этой кривой равно

$$\#E\,\mathbb{F}_p = p + 1 + \sum_{x=0}^{p-1}\left(\frac{x^3 + ax + b}{p}\right). \tag{4.3.4}$$

Определение 4.3.1. Пусть $\mathbb{F}_q$ — конечное поле и задана эллиптическая кривая $E/\mathbb{F}_q$. Пусть $N_r = \#E(\mathbb{F}_{q^r})$. Ряд

$$Z(E/\mathbb{F}_q; T) = e^{\sum_{r=1}^{\infty} N_r T^r / r} \tag{4.3.5}$$

называется *дзета-функцией кривой* $E/\mathbb{F}_q$.

Замечание 4.3.1. Поскольку при каждом $x \in F_{q^r}$ уравнение (4.3.1) имеет не более двух решений y, то $N_r \le 1 + 2q^r$. Таким образом, ряд (4.3.5) сходится абсолютно и равномерно на отрезке $|T| \le T_0$ при любом $T_0 < 1/q$, и, следовательно, этот ряд определяет аналитическую функцию.

Следующая теорема показывает, что дзета-функция эллиптической кривой имеет довольно простой вид.

Теорема 4.3.2 (Хассе). Дзета-функция эллиптической кривой (4.3.1) — рациональная функция от T вида

$$Z(E/\mathbb{F}_q; T) = \frac{1 - tT + qT^2}{(1 - T)(1 - qT)}, \tag{4.3.6}$$

где число t определяется равенством $N_1 = q + 1 - t$, а дискриминант числителя — отрицательный либо 0.

Теорема Хассе имеет два важных следствия. Первое позволяет точно вычислять количество $\mathbb{F}_{q^r}$-рациональных точек на данной эллиптической кривой,

если известно количество $\mathbb{F}_q$-рациональных точек. Второе показывает, что порядок эллиптической кривой лежит в сравнительно небольшом интервале около q.

Следствие 4.3.1. Квадратный трехчлен $T^2 - tT + q$ имеет комплексно сопряженные корни α и $\overline{\alpha}$. При этом

$$N_r = \left|\alpha^r - 1\right|^2 = q^r + 1 - \alpha^r - \overline{\alpha}^r.$$

Доказательство. Согласно теореме 4.3.2, дзета-функция кривой имеет вид

$$Z(E/\mathbb{F}_q; T) = \frac{(1-\alpha T)(1-\overline{\alpha} T)}{(1-T)(1-qT)}.$$

Логарифмируя это равенство, получаем

$$\sum_{r=1}^{\infty} N_r T^r / r = \ln Z(E/\mathbb{F}_q; T) =$$

$$= \ln(1-\alpha T) + \ln(1-\overline{\alpha} T) - \ln(1-T) - \ln(1-qT). \qquad (4.3.7)$$

Далее, используя тождество $\ln(1-x) = -\sum_{r=1}^{\infty} x^r / r$, из (4.3.7) получаем

$$\sum_{r=1}^{\infty} N_r T^r / r = -\sum_{r=1}^{\infty} \alpha^r T^r / r - \sum_{r=1}^{\infty} \overline{\alpha}^r T^r / r + \sum_{r=1}^{\infty} T^r / r + \sum_{r=1}^{\infty} q^r T^r / r. \qquad (4.3.8)$$

Приравнивая коэффициенты при одинаковых степенях T, получаем $N_r = q^r + 1 - \alpha^r - \overline{\alpha}^r$. Поскольку α и $\overline{\alpha}$ — комплексно сопряженные корни $T^2 - tT + q$, то по теореме Виета $\alpha \cdot \overline{\alpha} = q$. Тогда

$$\left|\alpha^r - 1\right|^2 = (\alpha^r - 1)(\overline{\alpha}^r - 1) = (\alpha\overline{\alpha})^r + 1 - \alpha^r - \overline{\alpha}^r = q^r + 1 - \alpha^r - \overline{\alpha}^r. \qquad \square$$

Следствие 4.3.2. Количество N q-рациональных точек кривой $E/\mathbb{F}_q$ удовлетворяет неравенству $|N - (q+1)| \le 2\sqrt{q}$.

Доказательство. Поскольку α и $\overline{\alpha}$ — комплексно сопряженные корни $T^2 - tT + q$, то по теореме Виета $\alpha \cdot \overline{\alpha} = q$ и $|\alpha| = |\overline{\alpha}| = \sqrt{q}$. Таким образом,

$$|N - (q+1)| = |N_1 - (q+1)| = |\alpha + \overline{\alpha}| \le |\alpha| + |\overline{\alpha}| = 2\sqrt{q}. \qquad \square$$

Следующая теорема Ленстры [122] говорит о том, что порядки эллиптических кривых над данным полем $\mathbb{F}_p$ имеют распределение на интервале $[p+1-2\sqrt{p},\ p+1+2\sqrt{p}]$, близкое к равномерному.

Теорема 4.3.3. Существуют эффективно вычисляемые постоянные $c_1, c_2 > 0$ такие, что для каждого простого $p > 3$ для любого подмножества S целых чисел интервала $[p+1-2\sqrt{p},\ p+1+2\sqrt{p}]$ вероятность ρ_S того, что случайно выбранная пара $(a,b) \in \mathbb{F}_p \times \mathbb{F}_p$ определяет эллиптическую кривую $E: y^2 = x^3 + ax + b$, удовлетворяющую условию $\# E(\mathbb{F}_p) \in S$, ограничена следующим образом:

$$\frac{\#S-2}{2[\sqrt{p}]+1} c_1 \log^{-1} p \le \rho_S \le \frac{\#S}{2[\sqrt{p}]+1} c_2 \log p (\log\log p)^2.$$

Определение 4.3.2. Пусть $E: y^2 = x^3 + ax + b$ — эллиптическая кривая, заданная над конечным полем $\mathbb{F}_q$, $\nu \in \mathbb{F}_q$ — некоторый квадратичный невычет (т.е. уравнение $x^2 = \nu$ не имеет решений в $\mathbb{F}_q$). Эллиптическая кривая E':

$$y^2 = x^3 + a'x + b',$$

где $a' = \nu^2 a$, $b' = \nu^3 b$, называется *скручиванием* E над полем $\mathbb{F}_q$.

Теорема 4.3.4. Пусть E — эллиптическая кривая над $\mathbb{F}_q$, E' — скручивание эллиптической кривой E над $\mathbb{F}_q$. Тогда

$$\# E(\mathbb{F}_q) + \# E'(\mathbb{F}_q) = 2q + 2.$$

Доказательство. Пусть $g(x) = x^3 + ax + b$. Тогда кривая E задается уравнением $y^2 = g(x)$, а кривая E' — уравнением $y^2 = \nu^3 g(x/\nu)$. Когда x пробегает все значения из $\mathbb{F}_q$, x/ν также пробегает все значения из $\mathbb{F}_q$. Таким образом, каждому корню многочлена g соответствует одна точка на каждой из эллиптических кривых E и E'. Каждому значению g, которое является квадратичным вычетом, соответствуют две точки на E и ни одной на E', так как $\nu^3 g$ будет квадратичным невычетом. Обратно, каждому значению g, которое является квадратичным невычетом, соответствуют две точки на E' и ни одной на E, так как $\nu^3 g$ будет квадратичным вычетом. Поэтому каждое значение g в сумме дает две точки на кривых E и E'. Поскольку всего имеется q значений g (с учетом кратности), то получаем аффинные $2q$ точки на кривых E и E'. Так как каждая из этих кривых содержит $\mathcal{O}$, то общее количество элементов на E и E' равно $2q + 2$. □

Известно также, как устроена группа точек на эллиптической кривой над конечным полем.

Теорема 4.3.5. Пусть $E/\mathbb{F}_q$ — эллиптическая кривая, заданная над конечным полем $\mathbb{F}_q$. Тогда $E(\mathbb{F}_q)$ либо является циклической группой, либо изоморфна произведению двух циклических групп:

$$E(\mathbb{F}_q) \cong (\mathbb{Z}/d_1\mathbb{Z})(\mathbb{Z}/d_2\mathbb{Z}).$$

При этом $d_1 \mid d_2$ и $d_1 \mid (q-1)$.

Дополнительно об эллиптических кривых можно прочитать в [56, 70, 115, 116, 122, 142, 183].

4.4. Задания

1. Доказать, что всякая эллиптическая кривая над полем характеристики 3 изоморфна некоторой эллиптической кривой вида $y^2 = x^3 + ax^2 + b$.

2. Найти дискриминант и j-инвариант для кривых (4.1.4) и (4.1.5) над полем характеристики 2.

3. Найти дискриминант и j-инвариант для кривых (4.1.7) и вида, определенного в задании 1, над полем характеристики 3.

4. Найти формулы сложения и обратной точки для кривых, заданных уравнениями (4.1.4) и (4.1.5), над полями характеристики 2.

5. Найти формулы сложения и обратной точки для кривых, вида (4.1.7) и вида, определенного в задании 1, над полями характеристики 3.

6. Найти количество $\mathbb{F}_{2^r}$-рациональных точек на кривой $y^2 + y = x^3$, заданной над $\mathbb{F}_2$, $r = 1, 2, 3, \ldots$

7. Доказать, что две эллиптические кривые $y^2 = x^3 + ax + b$ и $y^2 = x^3 + a'x + b'$ изоморфны над полем K тогда и только тогда, когда $a' = u^4 a$, $b' = u^6 b$ для некоторого $u \in K^*$.

8. Пусть E — эллиптическая кривая над $\mathbb{F}_q$, E' — скручивание эллиптической кривой E над $\mathbb{F}_q$. Доказать, что E и E' не изоморфны над $\mathbb{F}_q$, но изоморфны над $\mathbb{F}_{q^2}$.

9. Пусть E — эллиптическая кривая над $\mathbb{F}_q$, а E' и E — скручивания эллиптической кривой E над $\mathbb{F}_q$. Доказать, что E' и E изоморфны над $\mathbb{F}_q$.

SXJKN8EYHR3I4HR98
7YDWFUIHW9E8YR384
H5O34596EUF9O8WE7
6R83Y439846UHFE98
YDSFJH8374H4BR827
6RUBNWDF87Y6234BN
9WE872H3WRTJOJRTG
M905JRO9G70954TOI
RF7ERFH2347FQ7WER
F9UJ1349YUR7DFKJH
GJDFBVDGFJVBDSKFJ
HKJFVKBDFKJVJKGFK
JSBDFJBVKJDSGFJKV

Глава 5

ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКИЕ МОДЕЛИ СООБЩЕНИЙ И ИХ ЭНТРОПИЙНЫЕ СВОЙСТВА

SXJKN8EYHR3I4HR98
7YDWFUIHW9E8YR384
H5O34596EUF9O8WE7
6R83Y439846UHFE98
YDSFJH8374H4BR827
6RUBNWDF87Y6234BN
9WE872H3WRTJOJRTG
M905JRO9G70954TOI
RF7ERFH2347FQ7WER
F9UJ1349YUR7DFKJH
GJDFBVDGFJVBDSKFJ
HKJFVKBDFKJVJKGFK
JSBDFJBVKJDSGFJKV

5.1. Источники дискретных сообщений и их вероятностные модели

Пусть рассматривается произвольный источник сообщений. Каждое сообщение представляет собой некоторую последовательность символов (например, букв русского алфавита, точек и тире в телеграфии, нулей и единиц в компьютерной логике и т.д.). Отдельный *символ сообщения* будем обозначать ξ и предполагать числовой величиной, принимающей всевозможные значения из некоторого множества $A \subset \mathbb{R}$, $\xi \in A$.

Множество A значений символа сообщения ξ принято называть *алфавитом сообщений*. Если алфавит A является конечным множеством мощности $2 \le N < \infty$:

$$A = \{a^{(1)}, a^{(2)}, \ldots, a^{(N)}\},$$

то принято говорить, что имеет место *источник дискретных сообщений* (ИДС). В противном случае говорят об *источнике непрерывных сообщений* (ИНС).

В этом параграфе будем рассматривать модели ИДС, наиболее часто используемые в современных криптосистемах, а модели ИНС будут изучены в п. 5.6.

Величины $a^{(1)}, \ldots, a^{(N)}$ принято называть *символами алфавита*, а число N — *мощностью алфавита*. Появление в сообщении любого символа алфавита характеризуется высокой степенью неопределенности. Для математического описания этой неопределенности будем использовать дискретную вероятностную модель. Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — основное вероятностное пространство, описывающее случайный «эксперимент» по появлению (регистрации) символа ξ. Тогда каждому элементарному исходу $\omega \in \Omega$ этого эксперимента ставится в соответствие значение символа $\xi = \xi(\omega) \in A$. Таким образом, символ $\xi = \xi(\omega)$ — дискретная случайная величина, полностью определяемая дискретным распределением вероятностей:

$$\mathsf{P}\{\xi = a\} = p(a),\ a \in A, \tag{5.1.1}$$

причем

$$0 < p(a) < 1,\ \sum_{i=1}^{N} p(a^{(i)}) = 1. \tag{5.1.2}$$

В качестве примера в таблице представлено распределение вероятностей символов русского алфавита (упорядоченных в порядке убывания $p(a^{(i)})$):

Символ	Пробел	о	е, ё	а	и	т	н	с
Вероятность	0.175	0.090	0.072	0.062	0.062	0.053	0.053	0.045

Символ	р	в	л	к	м	д	п	у
Вероятность	0.040	0.038	0.035	0.028	0.026	0.025	0.023	0.021

Символ	я	ы	з	ь, ъ	б	г	ч	й
Вероятность	0.018	0.016	0.016	0.014	0.014	0.013	0.012	0.010

Символ	х	ж	ю	ш	ц	щ	э	ф
Вероятность	0.009	0.007	0.006	0.006	0.004	0.003	0.003	0.002

Таким образом, ИДС в случае односимвольного сообщения описывается дискретной вероятностной моделью $\langle A, p(a)\rangle$. Эта модель описывает лишь одиночный случайный символ сообщения. Сообщение, порождаемое ИДС, — это в общем случае последовательность $n \geq 1$ случайных символов:

$$\Xi_n = (\xi_1, \ldots, \xi_n) \in A^n.$$

При этом полное вероятностное описание ИДС задается вероятностной моделью случайного временного ряда (случайного процесса) Ξ_n с дискретным временем $t \in \mathbb{N}$ и дискретным пространством состояний A ($n = 1, 2, \ldots$):

$$\left\langle A^n, p_n(a_1, \ldots, a_n)\right\rangle,$$
$$p_n(a_1, \ldots, a_n) = \mathbf{P}\{\xi_1 = a_1, \ldots, \xi_n = a_n\}, a_1, \ldots, a_n \in A, \tag{5.1.3}$$

где $p_n(a_1, \ldots, a_n)$ — n-мерное дискретное распределение вероятностей n-символьного сообщения. Отметим, что n-мерные распределения вероятностей (5.1.3) удовлетворяют условию самосогласованности ($1 \leq k_1 < \ldots < k_m \leq n, 1 \leq m < n$):

$$p_m(a_{k_1}, \ldots, a_{k_m}) = \sum_{a_i \in A,\, i \in \{1, \ldots, n\} \setminus \{k_1, \ldots, k_m\}} p_n(a_1, \ldots, a_n). \tag{5.1.4}$$

ИДС называется *стационарным*, если случайный процесс Ξ_n является стационарным (в узком смысле), т.е. если конечномерные распределения (5.1.3) инвариантны относительно сдвига начала отсчета времени.

Стационарный ИДС называется *источником без памяти*, если для любых $a_1, \ldots, a_n \in A$ справедлива факторизация n-мерного распределения вероятностей:

$$p_n(a_1, \ldots, a_n) = \prod_{i=1}^{n} p(a_i), \tag{5.1.5}$$

т.е. порождаемые ИДС случайные символы $\xi_1, \ldots, \xi_n$ независимы в совокупности и одинаково распределены.

5.2. Функционал энтропии и его свойства

Пусть ИДС описывается некоторой дискретной вероятностной моделью $(A, p(a))$. Тогда *энтропией* ИДС (или энтропией случайного символа ξ) называется величина, определяемая функционалом

$$\mathsf{H}\{\xi\} = h(p(a^{(1)}), \ldots, p(a^{(N)})) :: = \mathsf{E}\{-\log_b p(\xi)\} = -\sum_{i=1}^{N} p(a^{(i)}) \log_b p(a^{(i)}), \quad (5.2.1)$$

где $\mathsf{E}\{\cdot\}$ — символ математического ожидания. Если в (5.2.1) логарифм берется по основанию $b = 2$, то энтропия измеряется в *битах* (bit = binary digit), а если используется натуральный логарифм по основанию $b = e$, то энтропия измеряется в *натах* (nat = natural digit).

Заметим, что встречающаяся в (5.2.1) неопределенность $0 \log 0$ разрешается следующим образом: $0 \log 0 :: = 0$. Мы будем пользоваться логарифмом по основанию $b = 2$ и, следовательно, измерять энтропию в битах.

Сформулируем и докажем основные свойства функционала энтропии.

Свойство 5.2.1. Функционал энтропии принимает неотрицательные значения: $\mathsf{H}\{\xi\} \geq 0$; он обращается в 0 только для вырожденного распределения:

$$\exists a' \in A,\ p(a') = 1,\ p(a) = 0,\ a \neq a'. \quad (5.2.2)$$

Доказательство. Так как $0 \leq p(\xi) \leq 1$, то $\eta = -\log p(\xi) \geq 0$. Согласно свойству математического ожидания из (5.2.1) имеем

$$\mathsf{H}\{\xi\} = \mathsf{E}\{\eta\} \geq 0,$$

причем $\mathsf{H}\{\xi\} = 0$ тогда и только тогда, когда $\eta \overset{\text{п.н.}}{=} 0$. Последнее соотношение, очевидно, выполняется лишь в случае (5.2.2). □

Заметим, что вырожденное распределение (5.2.2) соответствует случаю, когда символ ξ не является случайным: $\xi = a'$ с вероятностью 1 ($\xi \overset{\text{п.н.}}{=} \text{const}$).

Свойство 5.2.2. Энтропия ИДС с алфавитом мощности $N < \infty$ имеет максимальное значение

$$\max_{p(\ )} H\{\xi\} = \log N, \quad (5.2.3)$$

которое достигается, если дискретное распределение вероятностей $p(\cdot)$ — равномерное, т.е. все N значений символов равновероятны:

$$p(a) = \frac{1}{N},\ a \in A. \quad (5.2.4)$$

Доказательство. Воспользуемся неравенством Иенсена:

$$\mathsf{E}\{f(\zeta)\} \leq f(\mathsf{E}\{\zeta\}), \quad (5.2.5)$$

которое справедливо для любой случайной величины ζ и произвольной выпуклой функции $y = f(x)$. Положим в (5.2.5):

$$\zeta = \frac{1}{p(\xi)}, \; f(x) = \log_2 x, \tag{5.2.6}$$

причем $f(x) = -\dfrac{1}{(\ln 2)x^2} < 0$, так что $f(\cdot)$ — выпуклая. Используя (5.2.1) и (5.2.6), имеем

$$\mathbf{E}\{\zeta\} = \sum_{i=1}^{N} p(a^{(i)}) \frac{1}{p(a^{(i)})} = N, \; f(\mathbf{E}\{\zeta\}) = \log_2 N;$$

$$\mathbf{E}\{f(\zeta)\} = \sum_{i=1}^{N} p(a^{(i)}) \log_2 \frac{1}{p(a^{(i)})} = \mathbf{H}\{\xi\}.$$

Подставляя эти выражения в (5.2.5), получаем (5.2.3). Подставляя (5.2.4) в (5.2.1), находим

$$h\left(\frac{1}{N}, \ldots, \frac{1}{N}\right) = -\sum_{i=1}^{N} \frac{1}{N} \log \frac{1}{N} = \log N. \qquad \square$$

Отметим, что впервые энтропия была введена Хартли в 1928 г. в виде

$$\mathbf{H}\{\xi\} = \log N \tag{5.2.7}$$

для случайного символа ξ с N равновероятными значениями, и поэтому (5.2.7) иногда и теперь называют *энтропией Хартли*. Энтропия в общем виде (5.2.1) называется *энтропией Шеннона*.

Следствие 5.2.1. Чем больше мощность алфавита N, тем больше энтропия Хартли (максимально возможная энтропия).

Свойство 5.2.3 (аддитивности). Если случайные символы $\xi_1 \in A_1$, $\xi_2 \in A_2$ сообщения независимы, то совместная энтропия равна сумме энтропий:

$$\mathbf{H}\{\xi_1, \xi_2\} = \mathbf{H}\{\xi_1\} + \mathbf{H}\{\xi_2\}. \tag{5.2.8}$$

Доказательство. Построим на $(\Omega, \mathcal{F}, P)$ случайный вектор $\Xi_2 = (\xi_1, \xi_2) \in A_1 \cdot A_2$ с дискретным распределением вероятностей:

$$p_2(a_1, a_2) = \mathbf{P}\{\xi_1 = a_1, \xi_2 = a_2\}, \quad \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} p_2(a_1, a_2) = 1.$$

В силу независимости ξ_1, ξ_2 (ИДС без памяти) имеем

$$p_2(a_1, a_2) = p(a_1)p(a_2), \; a_1 \in A_1, \; a_2 \in A_2. \tag{5.2.9}$$

Тогда из (5.2.1) и (5.2.9) следует (с учетом условия нормировки):

$$\mathbf{H}\{\xi_1, \xi_2\} = -\sum_{a_1 \in A_1} \sum_{a_2 \in A_2} p_2(a_1, a_2) \log p_2(a_1, a_2) =$$

$$= -\sum_{a_1 \in A_1} p(a_1) \sum_{a_2 \in A_2} p(a_2) (\log p(a_1) + \log p(a_2)) =$$

$$= -\sum_{a_1 \in A_1} p(a_1) \log p(a_1) - \sum_{a_2 \in A_2} p(a_2) \log p(a_2) = \mathbf{H}\{\xi_1\} + \mathbf{H}\{\xi_2\},$$

что совпадает с (5.2.8). □

Следствие 5.2.2. Если независимы в совокупности $n \geq 2$ случайных символов $\xi_1 \in A_1, \ldots, \xi_n \in A_n$, то их совместная энтропия аддитивна:

$$\mathbf{H}\{\xi_1, \ldots, \xi_n\} = \mathbf{H}\{\xi_1\} + \ldots + \mathbf{H}\{\xi_n\}. \tag{5.2.10}$$

Доказательство. Оно состоит в $(n-1)$-кратном применении свойства 5.2.3. □

Свойство 5.2.4. Добавление к алфавиту символов одного символа с нулевой вероятностью, а следовательно, и любого количества таких символов не изменяет энтропии ИДС.

Доказательство. Пусть алфавит A расширен, как указано в условии:

$$A = \{a^{(1)}, \ldots, a^{(N)}\},\ A' = A \cup \{a^{(N+1)}\},\ \mathbf{P}\{\xi = a^{(N+1)}\} = 0.$$

Тогда согласно (5.2.1)

$$\mathbf{H}\{\xi'\} = -\sum_{i=1}^{N+1} p(a^{(i)}) \log p(a^{(i)}) = -\sum_{i=1}^{N} p(a^{(i)}) \log p(a^{(i)}) = \mathbf{H}\{\xi\}.$$ □

5.3. Условная энтропия и ее свойства

Чтобы изучить новые важные свойства энтропии, используемые в криптосистемах, нам понадобится понятие условной энтропии.

Пусть определен случайный n-вектор символов $\xi = (\xi_j) \in A_1 \cdots A_n$ с некоторым n-мерным дискретным распределением вероятностей

$$p_n(a_1, \ldots, a_n) = \mathbf{P}\{\xi_1 = a_1, \ldots, \xi_n = a_n\}, \tag{5.3.1}$$

где $a_1 \in A_1, \ldots, a_n \in A_n$. Пусть задано натуральное число $2 \leq k \leq n$ и определено $(n-k+1)$-мерное условное распределение вероятностей подвектора $\xi' = (\xi_k, \ldots, \xi_n) \in A_k \cdots A_n$ при условии, что фиксирован подвектор $\xi = (\xi_1, \ldots, \xi_{k-1}) \in A_1 \cdots A_{k-1}$:

$$p_{n-k+1}(a_k, \ldots, a_n \mid a_1, \ldots, a_{k-1}) = \frac{p_n(a_1, \ldots, a_n)}{p_{k-1}(a_1, \ldots, a_{k-1})}. \tag{5.3.2}$$

Здесь использована формула умножения вероятностей.

Условной энтропией подвектора ξ' *при условии*, что фиксирован подвектор ξ, называется функционал

$$\mathbf{H}\{\xi_k, \ldots, \xi_n \mid \xi_1 = a_1, \ldots, \xi_{k-1} = a_{k-1}\} =$$
$$= -\sum_{a_k \in A_k} \ldots \sum_{a_n \in A_n} p_{n-k+1}(a_k, \ldots, a_n \mid a_1, \ldots, a_{k-1}) \times$$
$$\times \log p_{n-k+1}(a_k, \ldots, a_n \mid a_1, \ldots, a_{k-1}). \qquad (5.3.3)$$

Условной энтропией случайного подвектора символов $\xi' = (\xi_k, \ldots, \xi_n)$ *относительно* случайного подвектора символов $\xi = (\xi_1, \ldots, \xi_{k-1})$ называется функционал, получающийся усреднением (5.3.3):

$$\mathbf{H}\{\xi' \mid \xi\} = \sum_{a_1 \in A_1} \ldots \sum_{a_{k-1} \in A_{k-1}} p_{k-1}(a_1, \ldots, a_{k-1}) \times$$
$$\times \mathbf{H}\{\xi_k, \ldots, \xi_n \mid \xi_1 = a_1, \ldots, \xi_{k-1} = a_{k-1}\} = -\sum_{a_1 \in A_1} \ldots \sum_{a_n \in A_n} p_n(a_1, \ldots, a_n) \times$$
$$\times \log p_{n-k+1}(a_k, \ldots, a_n \mid a_1, \ldots, a_{k-1}) \geq 0. \qquad (5.3.4)$$

Продолжим исследование свойств энтропии и условной энтропии с учетом введенных понятий.

Теорема 5.3.1. Если подвекторы случайных символов ξ', ξ независимы, то условная энтропия совпадает с безусловной:

$$\mathbf{H}\{\xi' \mid \xi\} = \mathbf{H}\{\xi'\}. \qquad (5.3.5)$$

Доказательство. В силу независимости ξ', ξ условное распределение вероятностей совпадает с безусловным:

$$p_{n-k+1}(a_k, \ldots, a_n \mid a_1, \ldots, a_{k-1}) = p_{n-k+1}(a_k, \ldots, a_n).$$

Подставляя это выражение в (5.3.4) и используя свойство самосогласованности распределений

$$\sum_{a_1 \in A_1} \ldots \sum_{a_{k-1} \in A_{k-1}} p_n(a_1, \ldots, a_n) = p_{n-k+1}(a_k, \ldots, a_n),$$

получаем (5.3.5). □

Теорема 5.3.2. Для любой последовательности случайных символов сообщения $\xi_1, \ldots, \xi_n$ энтропия обладает *свойством иерархической аддитивности*:

$$\mathbf{H}\{\xi_1, \ldots, \xi_n\} = \mathbf{H}\{\xi_1\} + \mathbf{H}\{\xi_2 \mid \xi_1\} +$$
$$+ \mathbf{H}\{\xi_3 \mid (\xi_1, \xi_2)\} + \ldots + \mathbf{H}\{\xi_n \mid (\xi_1, \ldots, \xi_{n-1})\}. \qquad (5.3.6)$$

Доказательство. Воспользуемся обобщенной формулой умножения вероятностей (свойством иерархической мультипликативности вероятностей):

$$p_n(a_1, \ldots, a_n) = p_1(a_1) p_1(a_2 \mid a_1) \ldots p_1(a_n \mid (a_1, \ldots, a_{n-1})).$$

Тогда по определению энтропии с учетом свойств вероятности имеем

$$\mathbf{H}\{\xi_1, \ldots, \xi_n\} = -\sum_{a_1 \in A_1} p_1(a_1)\log_2 p_1(a_1) - \sum_{a_1 \in A_1}\sum_{a_2 \in A_2} p_2(a_1, a_2)\log_2 p_1(a_2|a_1) - \ldots$$

$$\ldots - \sum_{a_1 \in A_1}\ldots\sum_{a_n \in A_n} p_n(a_1, \ldots, a_n)\log_2 p_1(a_n|(a_1, \ldots, a_{n-1})).$$

Учитывая (5.3.4), приходим к (5.3.6). □

Следствие 5.3.1. Если случайные символы $\xi_1, \ldots, \xi_n$ сообщения независимы в совокупности, то выполняется доказанное в предыдущем параграфе свойство аддитивности (5.2.10):

$$\mathbf{H}\{\xi_1, \ldots, \xi_n\} = \sum_{i=1}^{n}\mathbf{H}\{\xi_i\}.$$

Доказательство. Согласно теореме (5.3.1), условные энтропии совпадают с безусловными:

$$\mathbf{H}\{\xi_i \mid (\xi_1, \ldots, \xi_{i-1})\} = \mathbf{H}\{\xi_i\}, (i = \overline{1, n}).$$

Используя с учетом этого факта теорему 5.3.1, приходим к доказываемому. □

Следует заметить, что свойство иерархической аддитивности (5.3.6) и его следствие порождаются наличием логарифмической функции в определении энтропии; это свойство является ключевым для функционала энтропии.

Теорема 5.3.3. Пусть $\xi = (\xi_i) \in A$ — произвольный случайный вектор символов с дискретным распределением вероятностей $p_n(a)$, $a = (a_i) \in A$, а $q_n(a)$, $a \in A$ — некоторое дискретное распределение вероятностей. Тогда справедливо неравенство

$$J(p_n : q_n) = \sum_{a \in A} p_n(a)\log_2 \frac{p_n(a)}{q_n(a)} \geq 0. \tag{5.3.7}$$

Равенство нулю имеет место тогда и только тогда, когда распределение $q_n(\cdot)$ совпадает с $p_n(\cdot)$:

$$q_n(a) = p_n(a), \; a \in A. \tag{5.3.8}$$

Доказательство. Воспользуемся известным неравенством Иенсена, справедливым для произвольной случайной величины ζ и выпуклой функции $f(x)$:

$$\mathbf{E}\{f(\zeta)\} \leq f(\mathbf{E}\{\zeta\}). \tag{5.3.9}$$

Положим в (5.3.9)

$$\zeta ::= \frac{q_n(\xi)}{p_n(\xi)} \geq 0, \; f(x) ::= \log_2 x.$$

Тогда, используя условие нормировки, имеем

$$\mathbf{E}\{\zeta\} = \sum_{a\in A} p_n(a)\frac{q_n(a)}{p_n(a)} = \sum_{a\in A} q_n(a) = 1,$$

$$\mathbf{E}\{f(\zeta)\} = \sum_{a\in A} p_n(a)\log_2\frac{q_n(a)}{p_n(a)} = -J(p_n : q_n).$$

Подставляя эти выражения в (5.3.9), получаем неравенство

$$J(p_n : q_n) \geq 0.$$

Как известно, равенство в неравенстве Иенсена имеет место тогда и только тогда, когда $\zeta \equiv \text{const} = c$. В силу условия нормировки константа c может быть равна только единице:

$$c = \frac{q_n(a)}{p_n(a)} \equiv 1,$$

что и означает $q_n(\cdot) \equiv p_n(\cdot)$. □

Теорема 5.3.4. Условная энтропия не может превосходить безусловную:

$$\mathbf{H}\{\xi' \mid \xi\} \leq \mathbf{H}\{\xi'\}. \tag{5.3.10}$$

Доказательство. Воспользуемся теоремой 5.3.3 и положим в (5.3.7)

$$p(a) ::= \mathbf{P}\{\xi' = a \mid \xi = b\},\ \ q(a) ::= \mathbf{P}\{\xi' = a\},\ a \in A, b \in B.$$

Тогда получим

$$\sum_{a\in A}\mathbf{P}\{\xi' = a \mid \xi = b\}\log_2\frac{\mathbf{P}\{\xi' = a \mid \xi = b\}}{\mathbf{P}\{\xi' = a\}} \geq 0,$$

или (что эквивалентно)

$$-\sum_{a\in A}\mathbf{P}\{\xi' = a \mid \xi = b\}\log_2\mathbf{P}\{\xi' = a \mid \xi = b\} \leq -\sum_{a\in A}\mathbf{P}\{\xi' = a \mid \xi = b\}\log_2\mathbf{P}\{\xi' = a\}.$$

Умножим обе части этого неравенства на $\mathbf{P}\{\xi = b\} \geq 0$ и просуммируем по всевозможным $b \in B$ (с учетом формулы умножения вероятностей):

$$-\sum_{a\in A}\sum_{b\in B}\mathbf{P}\{\xi' = a \mid \xi = b\}\log_2\mathbf{P}\{\xi' = a \mid \xi = b\} \leq$$

$$\leq -\sum_{a\in A}\sum_{b\in B}\mathbf{P}\{\xi' = a \mid \xi = b\}\log_2\mathbf{P}\{\xi' = a\} = -\sum_{a\in A}\mathbf{P}\{\xi' = a\}\log_2\mathbf{P}\{\xi' = a\},$$

что совпадает с (5.3.10). □

Следствие 5.3.2. При добавлении условий условная энтропия не увеличивается:

$$\mathbf{H}\{\xi \mid (\eta, \zeta)\} \leq \mathbf{H}\{\xi \mid \eta\}. \tag{5.3.11}$$

Доказательство. Неравенство (5.3.11) доказывается аналогично теореме (5.3.4). □

Следствие 5.3.3. Энтропия n-символьной последовательности $\xi_1, \ldots, \xi_n$ случайных символов сообщения не превосходит суммы энтропий всех этих символов, рассматриваемых по отдельности:

$$\mathbf{H}\{\xi_1, \ldots, \xi_n\} \le \sum_{i=1}^{n} \mathbf{H}\{\xi_i\}. \tag{5.3.12}$$

Доказательство. Согласно теореме 5.3.4, справедливы неравенства

$$\mathbf{H}\{\xi_i \mid (\xi_1, \ldots, \xi_{i-1})\} \le \sum_{i=1}^{n} \mathbf{H}\{\xi_i\} \quad (i = \overline{2, n}).$$

Подставляя их в (5.3.6), получаем (5.3.12). □

В криптологии дискретные сообщения $\xi \in A$ часто подвергаются некоторому *дискретному функциональному преобразованию* (ДФП):

$$\eta = f(\xi),\ \xi \in A, \eta \in B, \tag{5.3.13}$$

где A, B — некоторые конечные множества. Исследуем, как изменяется энтропия сообщения при таком функциональном преобразовании.

Теорема 5.3.5. При дискретном функциональном преобразовании вида (5.3.13) энтропия не возрастает:

$$\mathbf{H}\{\eta\} \le \mathbf{H}\{\xi\}, \tag{5.3.14}$$

причем равенство в (5.3.14) достигается тогда и только тогда, когда ДФП (5.3.13) — биекция.

Доказательство. Рассмотрим «составное» сообщение $\begin{pmatrix}\xi\\ \eta\end{pmatrix} \in A \times B$. По теореме 5.3.2

$$\mathbf{H}\{\xi, \eta\} \le \mathbf{H}\{\xi\} + \mathbf{H}\{\eta \mid \xi\} = \mathbf{H}\{\eta\} + \mathbf{H}\{\xi \mid \eta\}. \tag{5.3.15}$$

В силу функциональной зависимости (5.3.13) условное распределение η при условии, что $\xi = a$ фиксировано, является вырожденным:

$$\mathbf{P}\{\eta = b \mid \xi = a\} = \mathbf{P}\{f(a) = b \mid \xi = a\} = \delta_{f(a),b},\ b \in B,$$

где $\delta_{i,j}$ — символ Кронеккера. Согласно свойству 5.2.1, энтропия для вырожденного распределения обращается в ноль:

$$\mathbf{H}\{\eta \mid \xi\} = 0.$$

Тогда из последнего равенства (5.3.15) имеем

$$\mathbf{H}\{\eta\} = \mathbf{H}\{\xi\} - \mathbf{H}\{\xi \mid \eta\}.$$

Поскольку $\mathbf{H}\{\xi|\eta\} \geq 0$ по свойству энтропии, то отсюда следует (5.3.13). Равенство в (5.3.14) будет тогда и только тогда, когда $\mathbf{H}\{\xi|\eta\} = 0$. Это возможно лишь в случае функциональной зависимости ξ от η: $\xi = f^{-1}(\eta)$, т.е. когда (5.3.13) — биекция. □

Введенный и исследованный в п. 5.2 и 5.3 функционал энтропии Шеннона

$$\mathbf{H}\{\xi\} = h_N(p_1, \ldots, p_N) = -\sum_{i=1}^{N} p_i \log p_i \tag{5.3.16}$$

для ИДС $\langle A, \{p_i\}\rangle$, $p_i = \mathbf{P}\{\xi = a^{(i)}\}$, $i = \overline{1, N}$ обладает всеми свойствами, которые необходимо требовать от количественной меры неопределенности. Поэтому энтропию Шеннона $\mathbf{H}\{\xi\}$ и используют в криптологии как количественную меру неопределенности сообщения $\xi \in A$. Однако является ли (5.3.16) единственным функционалом, обладающим изученными свойствами? Положительный ответ на этот вопрос дает теорема единственности, приведенная без доказательства [30].

Теорема 5.3.6. Пусть для любого натурального числа N функция N переменных

$$h = h_N(p_1, \ldots, p_N), \quad p_1, \ldots, p_N \geq 0, \quad \sum_{i=1}^{N} p_i = 1$$

непрерывна по совокупности аргументов и обладает следующими тремя свойствами:

а) максимальное значение функции достигается при равномерном распределении:

$$h_N(p_1, \ldots, p_N) \leq h_N\left(\frac{1}{N}, \ldots, \frac{1}{N}\right);$$

б) иерархическая аддитивность:

$$\mathbf{H}\{\xi|\eta\} = \mathbf{H}\{\xi\} + \mathbf{H}\{\eta|\xi\};$$

в) добавление к алфавиту, состоящему из N символов, одного символа с нулевой вероятностью ($p_{N+1} = 0$) не изменяет значения энтропии:

$$h_{N+1}(p_1, \ldots, p_N, 0) = h_N(p_1, \ldots, p_N).$$

Тогда эта функция $h_N(\cdot)$ необходимо имеет шенноновский вид:

$$h_N(p_1, \ldots, p_N) = -\lambda \sum_{i=1}^{N} p_i \log p_i,$$

где λ — произвольная положительная константа.

5.4. Удельная энтропия стационарной символьной последовательности

Пусть рассматривается стационарный ИДС с некоторым алфавитом $A = \{a^{(1)}, \ldots, a^{(N)}\}$, порождающий (генерирующий) n-символьные случайные сообщения $\Xi_n = (\xi_1, \ldots, \xi_n) \in A^n$. Каково асимптотическое поведение энтропии $\mathbf{H}\{\Xi_n\}$ при $n \to \infty$? Это основной вопрос, исследуемый в данном параграфе.

В силу свойства стационарности (в узком смысле) n-мерное дискретное распределение вероятностей

$$p_n(a_1, \ldots, a_N) = \mathbf{P}\{\xi_1 = a_1, \ldots, \xi_N = a_N\}, \; a_1, \ldots, a_N \in A \tag{5.4.1}$$

не меняется при сдвиге $\tau \geq 1$ начала отсчета времени:

$$\mathbf{P}\{\xi_{\tau+1} = a_1, \ldots, \xi_{\tau+n} = a_n\} = p_n(a_1, \ldots, a_n) = \underset{\tau}{\mathrm{invar}}. \tag{5.4.2}$$

В силу (5.4.1) и (5.4.2) энтропия n-символьного «сдвинутого» сообщения $\tilde{\Xi}_{n\tau} = (\xi_{\tau+1}, \ldots, \xi_{\tau+n}) \in A^n$ не зависит от τ:

$$\mathbf{H}\{\tilde{\Xi}_{n\tau}\} = \mathbf{H}\{\Xi_n\} = \underset{\tau}{\mathrm{invar}}. \tag{5.4.3}$$

Поэтому в дальнейшем при исследовании энтропии мы можем не делать различий между случайными сообщениями Ξ_n и $\tilde{\Xi}_{n\tau}$.

Удельной энтропией (*плотностью энтропии*) стационарного ИДС называется предел

$$h = \lim_{n \to \infty} \frac{\mathbf{H}\{\Xi_n\}}{n}, \tag{5.4.4}$$

если он существует.

Согласно определению (5.4.4), h представляет собой энтропию, приходящуюся на один символ и вычисленную по бесконечно длинному случайному сообщению. Если $h > 0$, то с увеличением длины сообщения $n \to \infty$ энтропия растет линейно:

$$\mathbf{H}\{\Xi_n\} \sim hn. \tag{5.4.5}$$

Знак «~» означает, что $\mathbf{H}\{\Xi_n\}/(hn) \to 1$. Выясним условия существования предела в (5.4.4), т.е. асимптотического поведения (5.4.5) энтропии.

Пусть $\Xi_{n-1} \in A^{n-1}$ и $\Xi_n = (\Xi_{n-1} \,||\, \xi_n) \in A^n$ — случайные сообщения длиной $n-1$ и n соответственно (здесь $||$ — символ конкатенации (присоединения)). Обозначим условную энтропию символа ξ_n относительно случайного сообщения Ξ_{n-1}, состоящего из $n-1$ предыдущих символов ($n = 1, 2, \ldots$):

$$\mathbf{H}\{\xi_n \mid \Xi_{n-1}\} = -\sum_{a_1, \ldots, a_n \in A} p_n(a_1, \ldots, a_n) \log p_1(a_n \mid a_1, \ldots, a_{n-1}). \tag{5.4.6}$$

При этом в случае $n = 1$ полагается $\mathbf{H}\{\xi_1 \mid \Xi_0\} ::= \mathbf{H}\{\xi_1\}$.

Теорема 5.4.1. Для произвольного стационарного ИДС числовая последовательность условных энтропий $H\{\xi_n|\Xi_{n-1}\}$, $n=1, 2, \ldots$, определяемых (5.4.6), имеет конечный предел:

$$\mathbf{H}\{\xi\,|\,\Xi_\infty\} ::= \lim_{n\to\infty} \mathbf{H}\{\xi_n\,|\,\Xi_{n-1}\}. \tag{5.4.7}$$

Доказательство. Докажем, что исследуемая последовательность не возрастает и ограничена снизу. Действительно, в силу свойства стационарности (5.4.3) $\forall i, j \in \mathbb{N}$:

$$\mathbf{H}\{\xi_j\,|\,\xi_{j-i}, \ldots, \xi_{j-1}\} = \mathbf{H}\{\xi_n\,|\,\xi_{n-i}, \ldots, \xi_{n-1}\}.$$

Поэтому исследуемая последовательность совпадает с последовательностью

$$\mathbf{H}\{\xi_n\}, \mathbf{H}\{\xi_n\,|\,\xi_{n-1}\}, \ldots, \mathbf{H}\{\xi_n\,|\,\xi_1, \ldots, \xi_{n-1}\}.$$

В силу свойств условной энтропии:

$$\mathbf{H}\{\xi_n\} \ge \mathbf{H}\{\xi_n\,|\,\xi_{n-1}\} \ge \ldots \ge \mathbf{H}\{\xi_n\,|\,\xi_1, \ldots, \xi_{n-1}\} \ge 0.$$

Как известно, любая невозрастающая ограниченная снизу числовая последовательность имеет предел, который обозначается согласно (5.4.7). □

Теорема 5.4.2. Для произвольного стационарного ИДС удельная энтропия (5.4.4) существует и совпадает с предельным значением (5.4.7):

$$h = \mathbf{H}\{\xi\,|\,\Xi_\infty\}. \tag{5.4.8}$$

Доказательство. Во-первых, докажем, что числовая последовательность

$$h_n ::= \frac{\mathbf{H}\{\Xi_n\}}{n},\ n=1, 2, \ldots, \tag{5.4.9}$$

из определения удельной энтропии (5.4.4) $\left(h ::= \lim\limits_{n\to\infty} h_n\right)$, является невозрастающей и ограниченной снизу. Для $(n+1)$-символьного сообщения $\Xi_{n+1} = (\Xi_n \,||\, \xi_{n+1}) \in A^{n+1}$ в силу свойства иерархической аддитивности энтропии имеем

$$\mathbf{H}\{\Xi_{n+1}\} = \mathbf{H}\{\Xi_n\} + \mathbf{H}\{\xi_{n+1}\,|\,\Xi_n\}. \tag{5.4.10}$$

В силу свойства условной энтропии (следствие 5.3.2) и свойства стационарности (5.4.3) имеем оценку для второго слагаемого в (5.4.10):

$$\mathbf{H}\{\xi_{n+1}\,|\,\Xi_n\} = \mathbf{H}\{\xi_{n+1}\,|\,\xi_1, \ldots, \xi_n\} \le \mathbf{H}\{\xi_{n+1}\,|\,\xi_2, \ldots, \xi_n\} =$$
$$= \mathbf{H}\{\xi_n\,|\,\xi_1, \ldots, \xi_{n-1}\} = \mathbf{H}\{\xi_n\,|\,\Xi_{n-1}\}.$$

Подставляя это выражение в (5.4.10), получаем неравенство

$$\mathbf{H}\{\Xi_{n+1}\} \le \mathbf{H}\{\Xi_n\} + \mathbf{H}\{\xi_n\,|\,\Xi_{n-1}\}. \tag{5.4.11}$$

В силу свойства иерархической аддитивности энтропии (теорема 5.4.2) и свойств условной энтропии имеем

$$\mathbf{H}\{\Xi_n\} = \sum_{i=1}^{n} \mathbf{H}\{\xi_i \mid \Xi_{i-1}\} \geq n\mathbf{H}\{\xi_n \mid \Xi_{n-1}\},$$

так что

$$\mathbf{H}\{\xi_n \mid \Xi_{n-1}\} \leq \frac{1}{n}\mathbf{H}\{\Xi_n\} = h_n. \tag{5.4.12}$$

Подставляя (5.4.12) в (5.4.11), получаем

$$0 \leq \mathbf{H}\{\Xi_{n+1}\} \leq \left(1 + \frac{1}{n}\right)\mathbf{H}\{\Xi_n\} = \frac{n+1}{n}\mathbf{H}\{\Xi_n\}.$$

Разделив обе части этого неравенства на $n+1$ и использовав обозначение (5.4.9), находим:

$$0 \leq h_{n+1} \leq h_n,$$

т.е. $\{h_n\}$ — невозрастающая числовая последовательность, ограниченная снизу. Следовательно, ее предел (5.4.4) — удельная энтропия h — существует.

Во-вторых, покажем справедливость равенства (5.4.8) для этого предела. Для произвольного $1 \leq k < n$ по свойствам энтропии с учетом (5.4.9) имеем

$$h_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{H}\{\xi_i \mid \Xi_{i-1}\} \equiv \frac{1}{n}\sum_{i=1}^{k} \mathbf{H}\{\xi_i \mid \Xi_{i-1}\} +$$

$$+\frac{1}{n}\sum_{i=k+1}^{n} \mathbf{H}\{\xi_i \mid \Xi_{i-1}\} \leq \frac{k}{n}\mathbf{H}\{\xi_1\} + \frac{n-k}{n}\mathbf{H}\{\xi_{k+1} \mid \Xi_k\}. \tag{5.4.13}$$

Воспользуемся произволом k и выберем $k = k(\varepsilon)$ таким, чтобы для любого наперед заданного $\varepsilon > 0$ выполнялось неравенство

$$\mathbf{H}\{\xi_{k+1} \mid \Xi_k\} - \mathbf{H}\{\xi \mid \Xi_\infty\} \leq \frac{\varepsilon}{2}. \tag{5.4.14}$$

Это всегда можно сделать, так как по теореме 5.4.1

$$\mathbf{H}\{\xi_{k+1} \mid \Xi_k\} \to \mathbf{H}\{\xi \mid \Xi_\infty\} + 0.$$

По выбранному таким образом k определим $\bar{n} = \bar{n}(k, \varepsilon)$ так, чтобы при любом $n > \bar{n}$ выполнялось неравенство

$$\frac{k}{n}\mathbf{H}\{\xi_1\} \leq \frac{\varepsilon}{2}. \tag{5.4.15}$$

Тогда из (5.4.13), (5.4.14), (5.4.15) следует, что $\forall\varepsilon > 0\ \exists\bar{n}_1 = \bar{n}_1(\varepsilon)$ такое, что для всех $n > \bar{n}_1$ справедлива оценка сверху:

$$h_n \leq \mathbf{H}\{\xi \mid \Xi_\infty\} + \varepsilon. \tag{5.4.16}$$

С другой стороны, имеем оценку снизу:

$$h_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{H}\{\xi_i \mid \Xi_{i-1}\} \geq \mathbf{H}\{\xi_n \mid \Xi_{n-1}\} \geq \mathbf{H}\{\xi \mid \Xi_\infty\}. \tag{5.4.17}$$

Объединяя (5.4.16) и (5.4.17), имеем

$$\mathbf{H}\{\xi \mid \Xi_\infty\} \le h_n \le \mathbf{H}\{\xi \mid \Xi_\infty\} + \varepsilon.$$

Так как ε — произвольное положительное число, то из этих неравенств следует, что

$$h = \lim_{n\to\infty} h_n = \mathbf{H}\{\xi \mid \Xi_\infty\},$$

т.е. выполняется (5.4.8). □

Из теорем 5.4.1 и 5.4.2 следует *асимптотическое разложение энтропии* n-символьного сообщения:

$$\mathbf{H}\{\Xi_n\} = nh + o(n). \tag{5.4.18}$$

Уточним остаточный член в (5.4.17).

Теорема 5.4.3. Для энтропии произвольной дискретной стационарной случайной последовательности $\Xi_n \in A^n$ при увеличении количества символов $n \to \infty$ справедлива асимптотика

$$\mathbf{H}\{\Xi_n\} = nh + 2b + o(1), \tag{5.4.19}$$

где

$$b = \frac{1}{2}\lim_{m,n\to\infty}(\mathbf{H}\{\Xi_m\} + \mathbf{H}\{\Xi_n\} - \mathbf{H}\{\Xi_{m+n}\}) \ge 0. \tag{5.4.20}$$

Доказательство. Обозначим

$$B_{mn} = \mathbf{H}\{\Xi_m\} + \mathbf{H}\{\Xi_n\} - \mathbf{H}\{\Xi_{m+n}\},\ \ m, n \in \mathbf{N}. \tag{5.4.21}$$

Во-первых, заметим, что B_{mn} — симметричная функция относительно m, n.

Во-вторых, по свойству иерархической аддитивности и свойству стационарности из (5.4.7) имеем

$$B_{mn} = \mathbf{H}\{\Xi_m\} + \mathbf{H}\{\xi_{m+1}, \ldots, \xi_{m+n}\} - (\mathbf{H}\{\Xi_m\} + \mathbf{H}\{\xi_{m+1}, \ldots, \xi_{m+n}\} \mid \Xi_m) =$$

$$= \mathbf{H}\{\xi_{m+1}, \ldots, \xi_{m+n}\} - \mathbf{H}\{\xi_{m+1}, \ldots, \xi_{m+n}\} \mid \Xi_m\}. \tag{5.4.22}$$

В силу известного свойства энтропии для $\eta = (\xi_{m+1}, \ldots, \xi_{m+n})$

$$\mathbf{H}\{\eta\} \ge \mathbf{H}\{\eta \mid \xi_1\} \ge \ldots \ge \mathbf{H}\{\eta \mid \xi_1, \ldots, \xi_m\},$$

при фиксированном n и растущем m энтропия

$$\mathbf{H}\{\xi_{m+1}, \ldots, \xi_{m+n} \mid \Xi_m\}$$

не возрастает. Следовательно, согласно (5.4.22), при фиксированном n последовательность B_{mn} — неубывающая по $m = 1, 2, \ldots$ Аналогичное свойство монотонности выполняется для B_{mn} по n, так как B_{mn} симметрична относительно m, n.

В силу установленного свойства монотонности существует предел (5.4.20) — конечный или бесконечный.

Чтобы доказать (5.4.19), воспользуемся (5.4.20), (5.4.22) и свойством стационарности:

$$2b = \lim_{m,n\to\infty} (\mathsf{H}\{\Xi_n\} - \mathsf{H}\{\xi_{m+1}, \ldots, \xi_{m+n} \mid \xi_1, \ldots, \xi_m\}). \qquad (5.4.23)$$

Применим еще раз свойство иерархической аддитивности энтропии:

$$\mathsf{H}\{\xi_{m+1}, \ldots, \xi_{m+n} \mid \xi_1, \ldots, \xi_m\} = \sum_{i=1}^{n} \mathsf{H}\{\xi_{m+i} \mid \xi_1, \ldots, \xi_{m+i-1}\}. \qquad (5.4.24)$$

Устремим $m \to \infty$ в выражении (5.4.24) и воспользуемся теоремой 5.4.2 ($i = \overline{1,n}$):

$$\mathsf{H}\{\xi_{m+1} \mid \xi_1, \ldots, \xi_{m+i-1}\} \to h, \; m \to \infty.$$

В результате из (5.4.24) получаем

$$\lim_{m\to\infty} \mathsf{H}\{\xi_{m+1}, \ldots, \xi_{m+n} \mid \xi_1, \ldots, \xi_m\} = nh. \qquad (5.4.25)$$

Переходя в (5.4.23) к пределу при $m \to \infty$ и используя (5.4.25), имеем

$$\lim_{n\to\infty} (\mathsf{H}\{\Xi_n\} - nh) = 2b,$$

что эквивалентно (5.4.19). □

Асимптотическое разложение энтропии (5.4.19) имеет следующую содержательную интерпретацию. Главный член разложения nh — это n-кратная удельная энтропия. Второй член разложения $2b$ — энтропия, обусловленная краевыми (граничными) эффектами. Третий член $o(1)$ — остаточный член разложения.

Следствие 5.4.1. Для стационарного ИДС без памяти соотношение (5.4.19) обращается в точное равенство:

$$\mathsf{H}\{\Xi_n\} = nh, \qquad (5.4.26)$$

где

$$h = \mathsf{H}\{\xi_1\} = -\sum_{a_1 \in A} p_1(a_1) \log p_1(a_1)$$

есть энтропия единичного случайного символа.

Доказательство. Соотношение (5.4.26) следует из свойств условной энтропии. Кроме того, $\mathsf{H}\{\Xi_{m+n}\} = \mathsf{H}\{\Xi_m\} + \mathsf{H}\{\Xi_n\}$, поэтому согласно (5.4.20) граничные эффекты отсутствуют: $b = 0$. □

5.5. Энтропийные характеристики марковских символьных последовательностей

В системах обработки и защиты информации символы, образующие сообщение $\Xi_n = (\xi_1, \ldots, \xi_n) \in A^n$, обычно стохастически зависимы. Одной из распространенных моделей стохастической зависимости символов печатных текстов,

речевых сообщений и изображений является *однородная цепь Маркова* (ОЦМ) с дискретным временем и пространством состояний $A = \{a^{(1)}, \ldots, a^{(N)}\}$, $N < \infty$.

Пусть случайная символьная последовательность, порожденная источником дискретных сообщений, $\xi_1, \ldots, \xi_n, \ldots \in A$ является однородной цепью Маркова. Это означает, что распределение вероятностей будущих значений (состояний) при фиксированных настоящих и прошлых значениях не зависит от прошлых значений:

$$\mathbf{P}\{\xi_{t+1} = a_{t+1} \mid \xi_t = a_t, \ldots, \xi_1 = a_1\} = p_{a_t, a_{t+1}}, \tag{5.5.1}$$

$a_1, \ldots, a_{t+1} \in A$, $t = 1, 2, \ldots$ Известно также, что все конечномерные распределения и все вероятностные характеристики ОЦМ полностью выражаются лишь через вектор-столбец начальных вероятностей:

$$\pi = \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_N \end{pmatrix}, \pi_i = \mathbf{P}\{\xi_1 = a^{(i)}\}, i = \overline{1, N}, \tag{5.5.2}$$

и через $(N \times N)$-матрицу вероятностей одношаговых переходов:

$$P = (p_{i,j}), p_{i,j} = \mathbf{P}\{\xi_{t+1} = a^{(j)} \mid \xi_t = a^{(i)}\}, i, j = \overline{1, N} \ (t = 1, 2, \ldots). \tag{5.5.3}$$

Вероятности (5.5.2) и (5.5.3) удовлетворяют условиям нормировки:

$$\sum_{i=1}^{N} \pi_i = 1, \sum_{j=1}^{N} p_{i,j} = 1 \ (i = \overline{1, N}). \tag{5.5.4}$$

Для получения дальнейших результатов введем N-вектор-столбец *стационарных вероятностей* $\pi^* = (\pi_i^*)$, $i = \overline{1, N}$, являющийся решением системы линейных алгебраических уравнений:

$$\begin{cases} \sum_{i=1}^{n} \pi_i p_{ij} = \pi_j, (i = \overline{1, N}); \\ \pi_1 + \ldots + \pi_N = 1. \end{cases} \tag{5.5.5}$$

Обозначим

$$\mathbf{H}^*\{\xi_1\} = -\sum_{i=1}^{N} \pi_i^* \log \pi_i^*$$

энтропию стационарного распределения вероятностей;

$$h = -\sum_{i=1}^{N} \pi_i^* \sum_{j=1}^{N} p_{ij} \log p_{ij}. \tag{5.5.6}$$

Теорема 5.5.1. Если случайная символьная последовательность является ОЦМ со стационарным начальным распределением π^* и матрицей вероятностей одношаговых переходов P, то энтропия n-символьного сообщения Ξ_n равна:

$$\mathbf{H}\{\Xi_n\} = \mathbf{H}^*\{\xi_1\} + (n-1)h; \tag{5.5.7}$$

при этом величина h, определяемая (5.5.6), является удельной энтропией.

Доказательство. В силу марковского свойства (5.5.1) имеем следующее представление для n-мерного дискретного распределения вероятностей:

$$p_n(a^{(i_1)}, \ldots, a^{(i_n)}) = \mathbf{P}\{\xi_1 = a^{(i_1)}\} \prod_{t=1}^{n-1} \mathbf{P}\{\xi_{t+1} = a^{(i_t+1)} \mid \xi_t = a^{(i_t)}, \ldots, \xi_1 = a^{(i_1)}\} =$$

$$= \pi_{i_1} \prod_{t=1}^{n-1} p_{i_t, i_{t+1}} (i_1, \ldots, i_n \in \{1, \ldots, N\}). \tag{5.5.8}$$

Тогда с учетом (5.5.8) энтропия n-символьного сообщения $\Xi_n = (\xi_1, \ldots, \xi_n)$ будет иметь вид

$$\mathbf{H}\{\Xi_1\} = -\sum_{i_1, \ldots, i_n = 1}^{N} p_n(a^{(i_1)}, \ldots, a^{(i_n)}) \left(\log \pi_{i_1} + \sum_{t=1}^{n-1} \log p_{i_t, i_{t+1}} \right).$$

В силу свойства согласованности многомерных вероятностных распределений имеем

$$\mathbf{H}\{\Xi_n\} = -\sum_{i_1=1}^{N} \pi_{i_1} \log \pi_{i_1} - \sum_{t=1}^{n-1} \sum_{i_t, i_{t+1}=1}^{N} \mathbf{P}\{\xi_t = a^{(i_t)}, \xi_{t+1} = a^{(i_{t+1})}\} \log p_{i_t, i_{t+1}} =$$

$$= -\sum_{i=1}^{N} \pi_i \log \pi_i - \sum_{t=1}^{n-1} \sum_{i_t, i_{t+1}=1}^{N} \mathbf{P}\{\xi_t = a^{(i_t)}\} p_{i_t, i_{t+1}} \log p_{i_t, i_{t+1}}. \tag{5.5.9}$$

По условию теоремы начальное распределение ОЦМ совпадает с ее стационарным распределением: $\pi_i = \pi_i^*$ $(i = \overline{1, N})$. Тогда из (5.5.5) легко показать, что одномерное распределение ОЦМ не изменяется с течением времени:

$$\mathbf{P}\{\xi_t = a^{(i)}\} = \pi_i^*, t = 1, 2, \ldots (i = \overline{1, N}).$$

Учитывая это в (5.5.9) и используя обозначения (5.5.6), получаем

$$\mathbf{H}\{\Xi_n\} = \mathbf{H}^*\{\xi_1\} + \sum_{t=1}^{n-1} \sum_{i=1}^{N} \left(-\pi_i^* \sum_{j=1}^{N} p_{ij} \log p_{ij} \right) =$$

$$= \mathbf{H}^*\{\xi_1\} + (n-1)h,$$

что совпадает с (5.5.7).

Найдем удельную энтропию, используя определение и (5.5.7):

$$\lim_{n\to\infty}\frac{\mathsf{H}\{\Xi_n\}}{n}=\lim_{n\to\infty}\frac{\mathsf{H}^*\{\xi_t\}}{n}+h\lim_{n\to\infty}\left(1-\frac{1}{n}\right)=h,$$

что и требовалось доказать. □

Отметим, что если нарушено условие теоремы о стационарности ОЦМ: $\pi_i=\pi_i^*$ $(i=\overline{1,N})$, то формула (5.5.9) примет вид

$$\mathsf{H}\{\Xi_n\}=\mathsf{H}\{\xi_1\}+\sum_{t=1}^{n-1}\sum_{i,j=1}^{N}(-\pi_i^{(t)}p_{ij}\log p_{ij}), \tag{5.5.10}$$

где

$$\pi_i^{(t)}=\mathsf{P}\{\xi_t=a^{(i)}\}.$$

Из теории цепей Маркова известно, что при $t\to\infty$ и выполнении некоторых ограничений на P согласно эргодической теореме имеет место экспоненциальная сходимость распределения вероятностей $\pi^{(t)}=(\pi_i^{(t)})$ к стационарному распределению $\pi^*=(\pi_i^*)$:

$$\pi_i^{(t)}=\pi_i^*+O(\rho^t),\ i=\overline{1,N},$$

где $0<\rho<1$. Поэтому из (5.5.10) имеем

$$\frac{\mathsf{H}\{\Xi_n\}}{n}=\frac{\mathsf{H}\{\Xi_1\}}{n}+\frac{1}{n}\left((n-1)h+\sum_{t=1}^{n-1}O(\rho^t)\right)\to h,\ n\to\infty.$$

Таким образом, удельная энтропия ОЦМ даже при нарушении условия стационарности определяется (5.5.6), хотя соотношение (5.5.7) при этом не выполняется.

Следствие 5.5.1. В условиях теоремы 5.5.1 соотношение (5.4.19) обращается в точное равенство:

$$\mathsf{H}\{\Xi_n\}=nh+2b, \tag{5.5.11}$$

где

$$2b=\sum_{i,j=1}^{N}\pi_i^*p_{ij}\log\frac{p_{ij}}{\pi_j^*}. \tag{5.5.12}$$

Доказательство. Представляя (5.5.7) в виде (5.5.11), имеем

$$2b=\mathsf{H}^*\{\xi_1\}-h.$$

Тогда из (5.5.6) следует:

$$2b=-\sum_{i=1}^{N}\pi_i^*\log\pi_i^*+\sum_{i=1}^{N}\pi_i^*\sum_{j=1}^{N}p_{ij}\log p_{ij}=\sum_{i=1}^{N}\pi_i^*\sum_{j=1}^{N}p_{ij}\log\frac{p_{ij}}{\pi_j^*}.$$

Используя (5.5.5), можно показать, что это выражение совпадает с правой частью (5.5.12). □

Пример. Рассмотрим двоичную (бинарную) ОЦМ $\xi_1, \xi_2, \ldots \in A = \{0, 1\}$, т.е. $N = 2$, $a^{(1)} = 0$, $a^{(2)} = 1$. Вероятностные характеристики ОЦМ:

$$\pi = \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix}, \quad P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix},$$

где $\alpha, \beta \in [0, 1]$. Заметим, что при $\alpha + \beta = 1$, $\pi_1 = 1 - \alpha$, $\pi_2 = \alpha$ имеем схему независимых испытаний. Для нахождения стационарного распределения π^*, согласно (5.5.5), имеем систему

$$\begin{cases} \pi_1^*(1-\alpha) + \pi_2^*\beta = \pi_1^*, \\ \pi_1^* + \pi_2^* = 1. \end{cases}$$

Решение ее единственно:

$$\pi^* = \begin{pmatrix} \dfrac{\beta}{\alpha+\beta} \\ \dfrac{\alpha}{\alpha+\beta} \end{pmatrix}.$$

По формулам (5.5.6) находим

$$\mathsf{H}^*\{\xi_1\} = -\frac{\beta}{\alpha+\beta}\log\frac{\beta}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}\log\frac{\alpha}{\alpha+\beta},$$

$$h = -\frac{\beta}{\alpha+\beta}(\alpha\log\alpha + (1-\alpha)\log(1-\alpha)) - \frac{\alpha}{\alpha+\beta}(\beta\log\beta + (1-\beta)\log(1-\beta)).$$

По формуле (5.5.12) получаем

$$2b = 2\frac{\alpha\beta}{\alpha+\beta}\log\beta + \log\left(1 + \frac{\alpha}{\beta}\right) + \frac{\beta}{\alpha+\beta}(1-\alpha)\log(1-\alpha) + \frac{\alpha}{\alpha+\beta}(1-\beta)\log(1-\beta).$$

При $\beta = \alpha$ (симметричная ОЦМ) находим $\pi_i^* = 1/2$,

$$\mathsf{H}^*\{\xi_1\} = \log 2 = 1,$$

$$h = -((1-\alpha)\log(1-\alpha) + \alpha\log\alpha),$$

$$2b = \log 2 + (1-\alpha)\log(1-\alpha) + \alpha\log\alpha = 1 - h.$$

При $\alpha = 1/2$ имеем схему независимых испытаний (удельная энтропия максимальна). При $\alpha \to 0$ или $\alpha \to 1$ удельная энтропия $h \to 0$, что согласуется с фактом уменьшения неопределенности двоичных сообщений.

5.6. Источники непрерывных сообщений и их энтропийные свойства

До сих пор мы предполагали, что источник сообщений порождает дискретные символьные последовательности со значениями в дискретном алфавите. В приложениях, однако, часто встречаются источники непрерывных сообщений $\xi \in \mathcal{A}$, где $\mathcal{A} \subseteq \mathbb{R}$ — некоторое подмножество мощности континуум. Например, речевой сигнал в каждый момент времени можно рассматривать как величину звукового давления ξ в заданной точке пространства (в данном случае $\mathcal{A} = [0, b]$, где b — некоторая максимально допустимая величина давления). Другой пример — оптические сигналы. Для того чтобы рассмотреть эти более сложные случаи, построим обобщенную математическую модель и обобщим понятие энтропии.

Пусть (для фиксированного момента времени) сообщение принимает значение, описываемое случайной величиной $\xi \in \mathcal{A}$, заданной на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ и имеющей распределение вероятностей (индуцированную вероятностную меру)

$$\mathbf{P}_\xi(B) :: = \mathbf{P}(\xi(\omega) \in B),\ B \in \mathcal{B}, \tag{5.6.1}$$

где $\mathcal{B}$ — борелевская σ-алгебра на числовой прямой; при этом $\mathbf{P}_\xi(\mathcal{A}) = 1$. Для обобщения понятия энтропии потребуем, чтобы на измеримом пространстве $(\mathbb{R}, \mathcal{B})$, кроме $\mathbf{P}_\xi(\cdot)$, была определена еще одна вспомогательная мера $\nu = \nu(B)$, $B \in \mathcal{B}$, такая, что мера $\mathbf{P}_\xi(\cdot)$ абсолютно непрерывна относительно $\nu(\cdot)$. Напомним, что мера $\mathbf{P}_\xi(\cdot)$ называется абсолютно непрерывной относительно $\nu(\cdot)$, если $\forall B \in \mathcal{B}\ \nu(B) = 0 \Rightarrow \mathbf{P}_\xi(B) = 0$.

Согласно известной теореме Радона — Никодима из условия абсолютной непрерывности меры $\mathbf{P}_\xi(\cdot)$ относительно меры $\nu(\cdot)$ вытекает существование борелевской функции $f(x)$, $x \in \mathcal{A}$, обозначаемой

$$\frac{d\mathbf{P}_\xi}{d\nu}(x)$$

и называемой *производной Радона — Никодима*. Она определена везде в $\mathcal{A}$, за исключением подмножества $\mathcal{A}_0$, где $\nu(\mathcal{A}_0) = 0$, а значит, и $\mathbf{P}_\xi(\mathcal{A}_0) = 0$. При этом справедливо интегральное представление:

$$\mathbf{P}_\xi(B) = \int_B f(x)\nu(dx),\ B \in \mathcal{B}.$$

Энтропией случайного сообщения ξ с распределением $\mathbf{P}_\xi(\cdot)$ называется величина интеграла Лебега:

$$\mathbf{H}(\xi) = -\int_{\mathcal{A}} \log\frac{d\mathbf{P}_\xi}{d\nu}(x)\,\mathbf{H}_\xi(dx). \tag{5.6.2}$$

В частности, если $\mathcal{A}$ — дискретное множество, а $\nu(\cdot)$ — «считающая мера» (например, при $|\mathcal{A}| < \infty$, $\nu(B) = |B|$ — мощность множества B), то (5.6.2) превращается в определение Шеннона:

$$\mathbf{H}\{\xi\} = -\sum_{a \in \mathcal{A}} (\log p_\xi(a)) p_\xi(a), \tag{5.6.3}$$

где $p_\xi(a) ::= \mathbf{P}\{\xi = a\}$, $a \in \mathcal{A}$ — дискретное распределение вероятностей случайного символа.

Рассмотрим другой частный случай (5.6.2), когда случайное сообщение ξ имеет абсолютно непрерывное распределение вероятностей (относительно меры Лебега):

$$\mathbf{P}_\xi(B) = \int_B p_\xi(x)\, dx, \quad B \in \mathcal{B} \tag{5.6.4}$$

с плотностью распределения $p_\xi(x) \geq 0$, удовлетворяющей условию нормировки:

$$\int_{\mathcal{A}} p_\xi(x)\, dx = 1.$$

Учитывая (5.6.4), в качестве меры $\nu(\cdot)$ в (5.6.2) примем меру Лебега:

$$\nu(B) = \frac{\mathrm{mes}(B)}{\mathrm{mes}(\mathcal{A})}, \; B \in \mathcal{A}; \tag{5.6.5}$$

предполагается, что $\mathrm{mes}(\mathcal{A}) < \infty$.

Тогда из (5.6.4) и (5.6.5) следует:

$$\frac{d\mathbf{P}_\xi}{d\nu}(x) = p_\xi(x)\, \mathrm{mes}(\mathcal{A}),$$

а интеграл Лебега (5.6.2) выражается через интеграл Римана:

$$\mathbf{H}\{\xi\} = -\log \mathrm{mes}(\mathcal{A}) - \int_{\mathcal{A}} p_\xi(x) \log p_\xi(x)\, dx. \tag{5.6.6}$$

Дифференциальной (относительной) энтропией случайного сообщения $\xi \in \mathcal{A}$ с плотностью распределения $p_\xi(x)$ называется значение функционала

$$\mathbf{H}_d\{\xi\} = -\int_{\mathcal{A}} p_\xi(x) \log p_\xi(x)\, dx. \tag{5.6.7}$$

Замечание 5.6.1. Термин «относительная» показывает, что она вычислена *относительно* меры Лебега (5.6.5).

Замечание 5.6.2. Если «доопределить» функцию плотности на всей числовой прямой:

$$p_\xi(x) = 0, \quad x \in \mathbb{R} \setminus \mathcal{A},$$

то в определении (5.6.7) всегда будем полагать интегрирование на $\mathbb{R} = (-\infty, \infty)$.

Замечание 5.6.2. Аналогично (5.6.7) определяется энтропия и в случае, когда случайный символ $\xi \in \mathbb{R}^N$ — N-мерный и описывается N-мерной плотностью распределения $p_\xi(x)$, $x \in \mathbb{R}^N$.

Отметим, что функционалы (5.6.6) и (5.6.7) отличаются на константу:

$$\mathbf{H}\{\xi\} = \mathbf{H}_d\{\xi\} - \log \operatorname{mes}(\mathcal{A}).$$

Пользоваться функционалом (5.6.6) менее удобно, чем (5.6.7), так как при $\mathcal{A} = \mathbb{R}$ получаем $\operatorname{mes}(\mathcal{A}) = +\infty$ и $\mathbf{H}(\xi) = -\infty$ для любой плотности $p_\xi(\cdot)$, в то время как дифференциальная энтропия (5.6.7) конечна. Поэтому свойства обобщенной энтропии будем выражать через свойства дифференциальной.

Свойство 5.6.1. Дифференциальная энтропия не изменяется при сдвиге распределения вероятностей, «зеркальных отражениях» и «перестановках фрагментов».

Доказательство. Проиллюстрируем схему доказательства для преобразований типа «сдвиг»:

$$\tilde{p}_\xi(x) = p_\xi(x - c),$$

где $c \in \mathbb{R}$ — некоторая константа, определяющая величину сдвига. Тогда из (5.6.7), делая замену переменных $y = x - c$, имеем

$$\tilde{\mathbf{H}}_d\{\xi\} = -\int_{-\infty}^{\infty} p_\xi(x - c)\log p_\xi(x - c)\,dx = -\int_{-\infty}^{\infty} p_\xi(y)\log p_\xi(y)\,dy = \mathbf{H}_d\{\xi\}.$$

Для других преобразований доказательство предлагается провести самостоятельно. □

Условной дифференциальной энтропией случайного символа $\xi \in \mathbb{R}$ относительно случайного символа $\eta \in \mathbb{R}$ называется величина двойного интеграла:

$$\mathbf{H}_d\{\xi|\eta\} = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{\xi,\eta}(x,y)\log p_{\xi|\eta}(x|y)\,dx\,dy, \tag{5.6.8}$$

где $p_{\xi,\eta}(x,y)$ — совместная плотность распределения вероятностей случайных символов (ξ,η); $p_{\xi|\eta}(x|y) = p_{\xi,\eta}(x,y)/p_\eta(y)$ — условная плотность распределения ξ при условии $\eta = y$. Это определение похоже на (5.3.4), справедливое для ИДС.

Свойство 5.6.2. Справедливо свойство *иерархической аддитивности* дифференциальной энтропии для произвольной системы случайных символов $(\xi,\eta) \in \mathbb{R}^2$:

$$\mathbf{H}_d\{\xi,\eta\} = \mathbf{H}_d\{\eta\} + \mathbf{H}_d\{\xi|\eta\} = \mathbf{H}_d\{\xi\} + \mathbf{H}_d\{\eta|\xi\}. \tag{5.6.9}$$

Доказательство. Оно проводится аналогично доказательству подобного свойства в дискретном случае с использованием формулы умножения плотностей и условия нормировки. □

Лемма 5.6.1. Для любых плотностей распределения вероятностей $p(x)$, $q(x)$, $x \in \mathbb{R}^N$ выполняется неравенство

$$J(p(\cdot): q(\cdot)) ::= \int_{\mathbb{R}^N} p(x) \log \frac{p(x)}{q(x)} dx \geq 0. \tag{5.6.10}$$

Доказательство. Оно проводится с использованием неравенства Иенсена по аналогии с дискретным случаем (см. теорему 5.3.3). □

Пусть $p_X(x)$, $x \in \mathbb{R}$ — произвольная плотность распределения вероятностей, а $a(x,y)$, $x, y \in \mathbb{R}$ — произвольная весовая функция, удовлетворяющая следующим условиям:

$$a(x,y) \geq 0, \quad \int_{-\infty}^{\infty} a(x,y)dx = \int_{-\infty}^{\infty} a(x,y)dy = 1. \tag{5.6.11}$$

Тогда интегральное преобразование $p_X(\cdot) \to q(\cdot)$:

$$q(y) = \int_{-\infty}^{\infty} a(x,y) p_X(x) dx, \quad y \in \mathbb{R} \tag{5.6.12}$$

называется *преобразованием линейного сглаживания* (*усреднения*), или *линейной фильтрации*.

Замечание 5.6.4. Если $a(x,y) = \delta(y - x)$ — обобщенная δ-функция Дирака, то преобразование (5.6.12) становится тождественным: $q(\cdot) \equiv p_X(\cdot)$.

Замечание 5.6.5. Легко убедиться, что функция $q(y)$, определяемая (5.6.12), с учетом (5.6.11) удовлетворяет свойствам плотности распределения вероятностей.

Свойство 5.6.3. При линейном сглаживании плотности распределения вероятностей дифференциальная энтропия не убывает.

Доказательство. Пусть случайный символ $\xi \in \mathbb{R}$ имеет исходную плотность $p_X(\cdot)$, а $\eta \in \mathbb{R}$ — сглаженную плотность $q(\cdot)$. Вычислим разность их дифференциальных энтропий с учетом (5.6.12) и (5.6.11):

$$\Delta H ::= \mathsf{H}_d\{\eta\} - \mathsf{H}_d\{\xi\} = -\int_{-\infty}^{\infty} q(y) \log q(y) dy + \int_{-\infty}^{\infty} p_X(x) \log p_X(x) dx =$$

$$= \int_{-\infty}^{\infty} p_X(x) \log p_X(x) dx \int_{-\infty}^{\infty} a(x,y) dy - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(x,y) p_X(x) \log q(y) dx dy =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q(x,y) p_X(x) \log \frac{a(x,y) p_X(x)}{a(x,y) q(y)} dx dy.$$

В силу (5.6.11) функции $p_1(x,y) = a(x,y)p_X(x) \geq 0$ и $q_1(x,y) = a(x,y)q(y) \geq 0$ удовлетворяют условиям нормировки:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_1(x,y)dxdy = \int_{-\infty}^{\infty} p_X(x)\int_{-\infty}^{\infty} a(x,y)dydx = 1,$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q_1(x,y)dxdy = \int_{-\infty}^{\infty} q(y)\int_{-\infty}^{\infty} a(x,y)dxdy = 1$$

и, следовательно, являются двумерными плотностями некоторых распределений вероятностей. Тогда в силу предыдущей леммы имеем доказываемое неравенство:

$$\Delta H = \int_{\mathbb{R}^2} p_1(x,y)\log\frac{p_1(x,y)}{q_1(x,y)}dxdy \geq 0. \qquad \square$$

Свойство 5.6.4. Пусть $\xi \in \mathbb{R}^N$ — случайный символ с плотностью распределения $p_\xi(x)$, $x \in \mathbb{R}^N$, имеющий дифференциальную энтропию $\mathbf{H}_d\{\xi\}$, а $y = f(x) : \mathbb{R}^N \to \mathbb{R}^N$ — взаимно-однозначное непрерывно дифференцируемое функциональное преобразование. Тогда случайный символ $\eta = f(\xi)$ имеет дифференциальную энтропию

$$\mathbf{H}_d\{\eta\} = \mathbf{H}_d\{\xi\} + \mathbf{E}\{\log|J_f(\xi)|\}, \tag{5.6.13}$$

где $J_f(x) = \left|\frac{Df(x)}{Dx}\right|$ — якобиан преобразования $y = f(x)$.

Доказательство. Обозначим $x = f^{-1}(y)$ — обратное функциональное преобразование, а $J_{f^{-1}}(y) = \left|\frac{Df^{-1}(y)}{Dy}\right|$ — его якобиан. По правилам функционального преобразования многомерных случайных величин имеем

$$p_\eta(y) = p_\xi(f^{-1}(y))\left|J_{f^{-1}}(y)\right|. \tag{5.6.14}$$

Тогда из (5.6.14) по формуле (5.6.7)

$$\mathbf{H}_d\{\eta\} = \mathbf{E}\{-\log p_\eta(\xi)\} = \mathbf{E}\{-\log p_\xi(f^{-1}(\eta))\} - \mathbf{E}\{\log\left|J_{f^{-1}}(\eta)\right|\}.$$

Воспользовавшись свойством математического ожидания функции от случайных величин и свойством якобиана прямого $y = f(x)$ и обратного $x = f^{-1}(y)$ преобразований, имеем

$$\left|J_{f^{-1}}(y)\right|_{y=f(x)} = \left|J_f(x)\right|^{-1},$$

$$\mathsf{H}_d\{\eta\} = \mathsf{E}\{-\log p_\xi(\xi)\} - \mathsf{E}\{\log|J_f(\xi)|^{-1}\} = \mathsf{H}_d\{\xi\} + \mathsf{E}\{\log|J_f(\xi)|\},$$

что совпадает с (5.6.13). □

Следствие 5.6.1. При взаимно-однозначном функциональном преобразовании $y = f(x) : \mathbb{R}^N \to \mathbb{R}^N$ дифференциальная энтропия может возрастать, убывать и оставаться неизменной. Она неизменна тогда и только тогда, когда плотность распределения $p_\xi(\cdot)$ и якобиан преобразования $J_f(x)$ обладают специальным свойством:

$$\mathsf{E}\{\log|J_f(\xi)|\} = \int_{\mathbb{R}^N} p_\xi(x)\log|J_f(x)|dx = 0. \tag{5.6.15}$$

Следствие 5.6.2. Если функциональное преобразование $f(\cdot)$ линейное:

$$y = f(x) = Ax + b,$$

где $b \in \mathbb{R}^N$ — произвольный вектор, $A = (a_{ij})$ — произвольная невырожденная $N \times N$-матрица, то

$$\mathsf{H}_d\{\eta\} = \mathsf{H}_d\{\xi\} + \log|A|. \tag{5.6.16}$$

Доказательство. Согласно (5.6.13), имеем

$$J_f(x) = \left|\frac{Dy}{Dx}\right| = A,$$

$$\mathsf{E}\{\log|J_f(\xi)|\} = \mathsf{E}\{\log|A|\} = \log|A|. \qquad \square$$

Замечание 5.6.6. Свойство 5.6.4 определяет существенное различие между энтропией ИДС и дифференциальной энтропией. Как известно, при взаимно-однозначных функциональных преобразованиях энтропия ИДС неизменна. Это различие — результат определения дифференциальной энтропии *относительно* меры Лебега.

Следствие 5.6.3. В условиях следствия 5.6.2 дифференциальная энтропия неизменна, если преобразование имеет единичный якобиан: $|A| = 1$.

Следствие 5.6.4. При ортогональном преобразовании случайного сообщения $\xi \in \mathbb{R}^N$:

$$\eta = A\xi, \quad AA^T = I_N,$$

дифференциальная энтропия не изменяется:

$$\mathsf{H}_d\{\eta\} = \mathsf{H}_d\{\xi\}. \tag{5.6.17}$$

Доказательство. Для ортогонального преобразования $|A| = 1$. Поэтому, согласно следствию 5.6.3, из (5.6.16) получаем (5.6.17). □

Пример 1. Если случайный символ ξ на $[a,b]$ имеет равномерное распределение $\mathcal{L}\{\xi\} = R[a,b]$, то дифференциальная энтропия равна

$$\mathsf{H}_d\{\xi\} = \log(b-a).$$

Пример 2. Если случайный символ ξ имеет гауссовское распределение $\mathcal{L}\{\} = \mathcal{N}_1(\mu,\sigma^2)$ со средним μ и дисперсией $\sigma^2 > 0$, то

$$\mathsf{H}_d\{\xi\} = \log\sqrt{2\pi e\sigma^2}.$$

Пример 3. Если случайное N-символьное ($N \geq 1$) сообщение имеет N-мерное гауссовское распределение $\mathcal{L}\{\xi\} = \mathcal{N}_N(\mu,\Sigma)$ с вектором математического ожидания $\mu = (\mu_i)$ и ковариационной $(N \times N)$-матрицей $\Sigma = (\sigma_{ij})$, то

$$\mathsf{H}_d\{\xi\} = \log\sqrt{(2\pi e)^N|\Sigma|}.$$

В заключение рассмотрим ситуацию, когда имеется стационарный источник непрерывных сообщений, порождающий случайный процесс с дискретным временем:

$$\xi_1, \xi_2, \ldots \in \mathbb{R}.$$

Исследуем дифференциальную энтропию отрезка этого процесса длительностью n ($n = 1, 2, \ldots$):

$$\Xi_n = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n.$$

Будем предполагать, что исследуемый случайный процесс является *стационарным гауссовским процессом*. Условие гауссовости означает, что для любого n распределение вероятностей случайного n-вектора Ξ_n является n-мерным гауссовским с плотностью

$$p_n(x) = (2\pi)^{-n/2}|\Sigma_n|^{-1/2}\exp\left(-\frac{1}{2}(x-1)^T\Sigma_n^{-1}(x-1)\right), \quad x = (x_i) \in \mathbb{R}^n, \quad (5.6.18)$$

где $a = (a_i) \in \mathbb{R}^n$ — n-вектор-столбец математических ожиданий $\mathsf{E}\{\xi_i\} = a_i$; $\Sigma_n = (\sigma_{ij})$ — это ковариационная $(n \times n)$-матрица; $\sigma_{ij} = \mathsf{Cov}\{\xi_i, \xi_j\} = \mathsf{E}\{(\xi_i - a_i)(\xi_j - a_j)\}$ — ковариация случайных величин $\xi_i, \xi_j (i, j = \overline{1,n})$.

Согласно примеру 3, в случае (5.6.18) имеем

$$\mathsf{H}_d\{\Xi_n\} = \log\sqrt{(2\pi e)^n|\Sigma_n|},$$

поэтому

$$\frac{\mathsf{H}_d\{\Xi_n\}}{n} = \frac{1}{2n}(n\log(2\pi e) + \log|\Sigma_n|) = \log\sqrt{2\pi e} + \frac{1}{2n}\log|\Xi_n|. \quad (5.6.19)$$

Свойство стационарности гауссовского случайного процесса проявляется в специальных свойствах ковариационной матрицы Σ_n:

а) дисперсия случайного процесса ξ_t не зависит от времени t, т.е.

$$\mathbf{D}\{\xi_t\} = \mathbf{Cov}\{\xi_t, \xi_t\} = \sigma_{tt} = \sigma_0 = \operatorname*{invar}_t;$$

б) ковариация значений случайного процесса ξ_t, $\xi_{t'}$ зависит лишь от разности моментов времени:

$$\sigma_{t,t'} = \mathbf{Cov}\{\xi_t, \xi_{t'}\} = \sigma_{|t-t'|}.$$

В силу указанных следствий стационарности исследуемого случайного процесса имеем

$$D_n = |\Sigma_n| = \begin{vmatrix} \sigma_0 & \sigma_1 & \sigma_2 & \dots & \sigma_{n-1} \\ \sigma_1 & \sigma_0 & \sigma_1 & \dots & \sigma_{n-2} \\ \sigma_2 & \sigma_1 & \sigma_0 & \dots & \sigma_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n-1} & \sigma_{n-2} & \sigma_{n-3} & \dots & \sigma_0 \end{vmatrix}, \tag{5.6.20}$$

где Σ_n — так называемая *симметричная теплицева матрица*. Математиками Пойа (в 1915 г.) и Сеге (в 1921 г.) было установлено следующее асимптотическое поведение обобщенной дисперсии (5.6.20) при $n \to \infty$:

$$\lim_{n\to\infty} \log(D_n)^{1/n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log S_\xi(\lambda)d\lambda, \tag{5.6.21}$$

где

$$S_\xi(\lambda) = \sum_{\tau=-\infty}^{\infty} \sigma_\tau \cos(\lambda\tau), \quad \lambda \in [-\pi, \pi] \tag{5.6.22}$$

есть *спектральная плотность* случайного процесса ξ_t, определяемая как косинус-преобразование Фурье *ковариационной функции* $\sigma_\tau = \sigma_{-\tau} = \mathbf{Cov}\{\xi_t, \xi_{t+\tau}\}$, $\tau \in \mathbf{Z}$.

Введем *нормированную корреляционную функцию*

$$\rho_\tau = \frac{\sigma_\tau}{\sigma_0} = \mathbf{Corr}\{\xi_t, \xi_{t+\tau}\}, \quad \tau \in \mathbf{Z} \tag{5.6.23}$$

и *нормированную спектральную плотность*

$$s_\xi(\lambda) = \frac{S_\xi(\lambda)}{\sigma_0} = \sum_{\tau=-\infty}^{\infty} \rho_\tau \cos(\lambda\tau), \quad \lambda \in [-\pi, \pi]. \tag{5.6.24}$$

Используя (5.6.21)—(5.6.24), найдем удельную энтропию гауссовского стационарного случайного процесса. Из (5.6.19) с помощью эквивалентных преобразований имеем

$$\frac{\mathbf{H}_d\{\Xi_n\}}{n} = \log\sqrt{2\pi e\sigma_0} + \frac{1}{2n}\log\left|\frac{1}{\sigma_0}\Sigma_n\right|.$$

Согласно (5.6.20) и (5.6.23), элементами матрицы $\frac{1}{\sigma_0}\Sigma_n$ являются значения нормированной корреляционной функции:

$$d_n = \left|\frac{1}{\sigma_0}\Sigma_n\right| = \begin{vmatrix} \rho_0 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & \rho_0 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & \rho_0 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & \rho_{n-0} \end{vmatrix}, \quad \rho_0 = 1.$$

Тогда в силу (5.6.21) и (5.6.24)

$$\lim_{n\to\infty} \log(d_n)^{1/n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log s_\xi(\lambda)d\lambda.$$

Получаем выражение удельной энтропии:

$$h = \lim_{n\to\infty} \frac{\mathbf{H}_d\{\Xi_n\}}{n} = \log\sqrt{2\pi e\sigma_0} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \log s_\xi(\lambda)d\lambda =$$

$$= \mathbf{H}\{\xi_1\} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \log s_\xi(\lambda)d\lambda, \qquad (5.6.25)$$

где $\mathbf{H}\{\xi_1\} = \log\sqrt{2\pi e\sigma_0}$ — энтропия единичного гауссовского случайного символа с дисперсией σ_0.

В (5.6.25) первое слагаемое характеризует энтропию единичного (отдельного, изолированного) символа, а второе — зависимость символов в последовательности и выражается через нормированную спектральную плотность $s_\xi(\lambda)$, или, что эквивалентно, через нормированную корреляционную функцию.

Пример 4. Последовательность независимых гауссовских символов:

$$\mathbf{Cov}\{\xi_t, \xi_{t'}\} = \sigma_0\delta_{t,t'}.$$

При этом $\rho_\tau = \delta_{\tau,0}$, $s_\xi(\lambda) = 1 \equiv \text{const}$, $\lambda \in [-\pi, \pi]$. Тогда в силу (5.6.22) второе слагаемое — нулевое, и получаем установленный выше результат:

$$\mathbf{H}_d\{\Xi_n\} = n\mathbf{H}\{\xi_1\}, \quad h = \mathbf{H}\{\xi_1\}.$$

Пример 5. Пусть имеет место марковская корреляционная зависимость: $\sigma_\tau = \sigma_0 e^{-\alpha|\tau|}$, $\alpha > 0$. Тогда $D_n = \sigma_0^n(1-\rho^2)^{n-1}$, где $\rho = e^{-\alpha}$. Поэтому из (5.6.25) имеем

$$h = \mathbf{H}\{\xi_1\} + \log\sqrt{1-\rho^2} = \mathbf{H}\{\xi_1\} + \log\sqrt{1-e^{-2\alpha}}.$$

5.7. Оптимизация функционала энтропии на классе вероятностных распределений

Для криптологических применений важно исследовать случаи экстремальных значений функционала энтропии. Как видно из примеров 1—3 предыдущего параграфа, дифференциальная энтропия изменяется от $-\infty$ до $+\infty$. Для ИДС $\xi \in A$, как было установлено в п. 5.2, минимальное значение, равное нулю ($\mathbf{H}_{\min} = 0$), достигалось для вырожденного дискретного распределения вероятностей, а максимальное значение, равное $\mathbf{H}_{\min} = \log|A|$, — для дискретного равномерного распределения вероятностей. Как видим, имеются существенные различия дискретного и непрерывного случаев. Чтобы максимальные значения дифференциальной энтропии были конечны, будем осуществлять ее максимизацию на заданном ограниченном классе вероятностных распределений $\mathcal{P}$:

$$\mathbf{H}_d\{\xi\} = -\int_{-\infty}^{\infty} p(x)\log p(x)dx \to \max_{p(\ )\in\mathcal{P}}. \tag{5.7.1}$$

Исследуем три наиболее часто встречающихся в прикладных задачах класса абсолютно непрерывных вероятностных распределений.

Класс $\mathcal{P}_1(a,b)$:

$$\mathcal{P}_1(a,b) = \left\{ p(x),\ x \in \mathbb{R} : p(x) \ge 0,\ \int_{-\infty}^{\infty} \right.$$

$$\left. \int_{-\infty}^{\infty} p(x)dx = 1;\ p(x) = 0,\ x \notin [a,b] \right\} \tag{5.7.2}$$

есть семейство одномерных плотностей распределения с конечным носителем $[a,b]$, $-\infty < a < b < \infty$.

Класс $\mathcal{P}_2(a,\sigma^2)$:

$$\mathcal{P}_2(a,\sigma^2) = \left\{ p(x),\ x \in \mathbb{R}: p(x) \ge 0,\ \int_{-\infty}^{\infty} p(x)dx = 1, \right.$$

$$\left. \int_{-\infty}^{\infty} xp(x)dx = a,\ \int_{-\infty}^{\infty} (x-a)^2 p(x)dx \le \sigma^2 \right\} \tag{5.7.3}$$

есть семейство одномерных плотностей с конечными моментами первого и второго порядков: заданным математическим ожиданием (средним) $\mathbf{E}\{\xi\} = a$ и ограниченной дисперсией $\mathbf{D}\{\xi\} \le \sigma^2$.

Класс $\mathcal{P}_3(n,\mu,\Sigma)$:

$$\mathcal{P}_3(n,\mu,\Sigma)=\left\{p(x),\ x\in\mathbf{R}^n:\ p(x)\ge 0,\ \int\limits_{\mathbf{R}^n}p(x)dx=1,\right.$$

$$\left.\int\limits_{\mathbf{R}^n}xp(x)dx=\mu,\ \int\limits_{\mathbf{R}^n}(x-\mu)(x-\mu)^T p(x)dx=\Sigma\right\} \tag{5.7.4}$$

есть семейство n-мерных плотностей распределения с фиксированным n-вектором математического ожидания $\mu=(\mu_i)$ и невырожденной $(n\times n)$-ковариационной матрицей $\Sigma=(\sigma_{ij})$, $|\Sigma|\ne 0$.

Теорема 5.7.1. Для любого случайного символа $\xi\in\mathbf{R}$ с плотностью распределения $p(\cdot)\in\mathcal{P}_1(a,b)$ дифференциальная энтропия удовлетворяет неравенству

$$\mathsf{H}_d\{\xi\}\le\log(b-a), \tag{5.7.5}$$

причем верхняя граница, т.е. максимум дифференциальной энтропии по классу $\mathcal{P}_1(a,b)$, достигается в случае равномерного на $[a,b]$ распределения вероятностей $R[a,b]$ с плотностью

$$p^*(x)=\frac{1}{b-a}\mathbf{1}_{[a,b]}(x),\quad x\in\mathbf{R}, \tag{5.7.6}$$

где $\mathbf{1}_A(x)$ — индикаторная функция множества A.

Доказательство. Будем решать экстремальную задачу (5.7.1), (5.7.2) при $\mathcal{P}=\mathcal{P}_1(a,b)$. Без учета условия неотрицательности $p(\cdot)\ge 0$ в (5.7.2) эта задача эквивалентна следующей задаче вариационного исчисления с ограничением типа равенства:

$$\begin{aligned}\mathsf{H}_d\{\xi\}=&\ -\int\limits_a^b p(x)\log p(x)dx\to\max_{p(\cdot)},\\ &\int\limits_a^b p(x)dx=1.\end{aligned} \tag{5.7.7}$$

Для решения задачи (5.7.7) применим метод неопределенных множителей Лагранжа. Для этого составим функционал Лагранжа:

$$L(p(\cdot),\lambda)=\int\limits_a^b(-p(x)\log p(x)+\lambda p(x))dx,$$

где λ — неопределенный множитель Лагранжа; и запишем необходимое условие максимума:

$$\begin{cases} \delta L = \int\limits_a^b (-1 - \log p(x) + \lambda)\delta p(x)dx = 0, \\ \int\limits_a^b p(x)dx = 1, \end{cases} \tag{5.7.8}$$

где $\delta p(x)$ — вариация функции $p(\cdot)$, а δL — первая вариация функционала $L(\cdot)$. Так как $\delta p(x)$ — произвольная вариация, то система уравнений (5.7.8) примет эквивалентный вид

$$\begin{cases} -1 - \log p(x) + \lambda = 0, \\ \int\limits_a^b p(x)dx = 1. \end{cases} \tag{5.7.9}$$

Решим первое уравнение (5.7.9):

$$p(x) = 2^{\lambda - 1} = \text{const} = \mu, \quad x \in [a,b]$$

и, подставляя это решение во второе уравнение (5.7.9), получим единственное решение задачи (5.7.7) в виде (5.7.6). Заметим, что найденное решение удовлетворяет снятому ограничению $p(\cdot) \geq 0$.

Подставляя (5.7.6) в целевую функцию (5.7.7), получим

$$\max_{p(\ )} \mathsf{H}_d\{\xi\} = \log(b - a),$$

что и означает (5.7.5). □

Теорема 5.7.2. Для любого случайного символа $\xi \in \mathbb{R}$ с плотностью распределения $p(\cdot) \in \mathcal{P}_2(a, \sigma^2)$ дифференциальная энтропия удовлетворяет неравенству

$$\mathsf{H}_d\{\xi\} \leq \log\sqrt{2\pi e\sigma^2}, \tag{5.7.10}$$

причем верхняя граница в (5.7.10), т.е. максимум дифференциальной энтропии на классе $\mathcal{P}_2(a, \sigma^2)$, достигается в случае гауссовского (нормального) распределения вероятностей $\mathcal{N}_1(a, \sigma^2)$ с плотностью

$$p^*(x) = n_1(x \mid a, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R}. \tag{5.7.11}$$

Доказательство. Имеется два способа доказательства. Первый основан на решении задачи вариационного исчисления (5.7.1) при $\mathcal{P} = \mathcal{P}_2(a, \sigma^2)$ методом неопределенных множителей Лагранжа подобно доказательству предыдущей теоремы. Второй метод менее громоздкий и излагается ниже.

Получим сначала вспомогательное неравенство — оценку сверху для интеграла. Имеем

$$-\int_{-\infty}^{\infty} p(x)\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-a)^2}{2\sigma^2}}\right)dx = \log\sqrt{2\pi\sigma^2} + \frac{\log e}{2\sigma^2}\int_{-\infty}^{\infty}(x-a)^2 p(x)dx.$$

В силу (5.7.3) последний интеграл равен дисперсии $\mathbf{D}\{\xi\} \le \sigma^2$, поэтому

$$-\int_{-\infty}^{\infty} p(x)\log n_1(x\,|\,a,\sigma^2)\,dx \le \log\sqrt{2\pi e\sigma^2},\quad p(\cdot)\in\mathcal{P}_2(a,\sigma^2).$$

Из этого неравенства имеем оценку разности левой и правой частей (5.7.10):

$$\mathbf{H}_d\{\xi\} - \log\sqrt{2\pi e\sigma^2} \le -\int_{-\infty}^{\infty} p_X(x)\log p_X(x)xdx + \int_{-\infty}^{\infty} p_X(x)\log n_1(x\,|\,a,\sigma^2)dx =$$

$$= -\int_{-\infty}^{\infty} p_X(x)\log\frac{p_X(x)}{n_1(x\,|\,a,\sigma^2)}dx = -J(p(\cdot):n_1(\cdot)).$$

В силу леммы 5.6.1 $J(p(\cdot):n_1(\cdot)) \ge 0$, следовательно:

$$\mathbf{H}_d\{\xi\} - \log\sqrt{2\pi e\sigma^2} \le 0,$$

что эквивалентно (5.7.10).

Как установлено в п. 5.6, для гауссовского распределения (5.7.11) энтропия равна

$$\mathbf{H}_d\{\xi^*\} = \log\sqrt{2\pi e\sigma^2},$$

т.е. (5.7.10) обращается в равенство. □

Теорема 5.7.3. Для любой случайной n-символьной последовательности $\Xi_n = (\xi_i) \in \mathbb{R}^n$ с n-мерной плотностью распределения $p(\cdot) \in \mathcal{P}_3(n,\mu,\Sigma)$ дифференциальная энтропия удовлетворяет неравенству

$$\mathbf{H}_d\{\Xi_n\} \le \log\sqrt{(2\pi e)^n|\Sigma|}, \tag{5.7.12}$$

причем верхняя граница, т.е. максимум дифференциальной энтропии по классу $\mathcal{P}_3(n,\mu,\Sigma)$, достигается в случае n-мерного гауссовского (нормального) распределения вероятностей $\mathcal{N}_n(\mu,\Sigma)$ с плотностью

$$p^*(x) = n_n(x\,|\,\mu,\Sigma) = (2\pi)^{-n/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right). \tag{5.7.13}$$

Доказательство. Воспользуемся тем же методом доказательства, что и в теореме 5.7.2. Сначала получим вспомогательное равенство для n-кратного интеграла, используя свойства семейства $\mathcal{P}_3(n, \mu, \Sigma)$ и формулу (5.7.13):

$$Q = -\int_{\mathbf{R}^n} p_X(x)\log p^*(x)dx = \log\sqrt{(2\pi)^n|\Sigma|} + \frac{\log e}{2}\int_{\mathbf{R}^n}(x-\mu)^T\Sigma^{-1}(x-\mu)p_X(x)dx =$$

$$= \log\sqrt{(2\pi)^n|\Sigma|} + \frac{\log e}{2}\mathbf{E}\{\mathrm{tr}(\Sigma^{-1}(x-\mu)(x-\mu)^T)\} =$$

$$= \log\sqrt{(2\pi)^n|\Sigma|} + \frac{\log e}{2}\mathrm{tr}\left(\Sigma^{-1}\mathbf{E}\left\{(x-\mu)(x-\mu)^T\right\}\right) = \log\sqrt{(2\pi e)^n|\Sigma|}.$$

Оценим с учетом этого равенства разность левой и правой частей (5.7.12):

$$\mathbf{H}_d\{\xi\} - \log\sqrt{(2\pi e)^n|\Sigma|} = \mathbf{H}_d\{\xi\} - Q = -\int_{\mathbf{R}^n} p_X(X)\log\frac{p_X(x)}{p^*(x)}dx \le 0.$$

Здесь использована лемма 5.6.1. Как установлено в п. 5.6, для многомерного гауссовского распределения (5.7.13) дифференциальная энтропия равна

$$\mathbf{H}_d\{\Xi_n^*\} = \log\sqrt{(2\pi e)^n|\Sigma|},$$

т.е. (5.7.12) обращается в равенство. □

5.8. Задания

1. Пусть $\xi \in \{0, 1\}$ — двоичный случайный символ с распределением вероятностей Бернулли: $\mathbf{P}\{\xi = 1\} = p$, $\mathbf{P}\{\xi = 0\} = 1 - p$. Построить график зависимости энтропии $\mathbf{H}\{\xi\}$ от элементарной вероятности $p \in [0, 1]$, исследовать эту функцию на экстремум.

2. Пусть $\xi \in \{0, 1, \ldots, N-1\}$ — случайный символ, причем элементарная вероятность $p_0 = \mathbf{P}\{\xi = 0\} = \varepsilon$ фиксирована. Доказать, что при произвольных $p_i = \mathbf{P}\{\xi = i\}$ $(i = \overline{1, N-1})$:

$$\mathbf{H}\{\xi\} \le -\varepsilon\log\varepsilon - (1-\varepsilon)\log\frac{1-\varepsilon}{N-1},$$

причем равенство достигается тогда и только тогда, когда $p_i = (1-\varepsilon)/(N-1)$ $(i = \overline{1, N-1})$.

3. Доказать, что энтропия $\mathbf{H}\{\xi\} = -\sum_{i=1}^{N} p_i \log p_i$ — выпуклая функция в области D всех вероятностных векторов $p = (p_1, \ldots, p_N)$, задающих дискретные распределения вероятностей.

4. Пусть $\xi \in \{0, 1, \ldots, N-1\}$ — случайный символ с биномиальным распределением вероятностей ($\mathcal{L}\{\xi\} = \mathrm{Bi}(N-1, p)$): $p_i = \mathbf{P}\{\xi = i\} = C_{N-1}^{i} p^{i} (1-p)^{N-1-i}$ ($i = \overline{0, N-1}$). Вычислить энтропию $\mathbf{H}\{\xi\}$ и исследовать на экстремум ее зависимость от элементарной вероятности p.

5. Пусть $\xi \in \{0, 1, 2, \ldots\}$ — случайный символ с распределением вероятностей Пуассона ($\mathcal{L}\{\xi\} = \Pi(\lambda)$) : $p_i = \mathbf{P}\{\xi = i\} = e^{-\lambda}\lambda^{i}/i!$. Вычислить энтропию $\mathbf{H}\{\xi\}$ и исследовать на экстремум ее зависимость от параметра $\lambda > 0$.

6. Пусть $\xi \in \{0, 1, 2, \ldots\}$ — случайный символ с геометрическим распределением вероятностей ($\mathcal{L}\{\xi\} = G(p)$) : $p_i = \mathbf{P}\{\xi = i\} = p(1-p)^{i}$. Вычислить энтропию $\mathbf{H}\{\xi\}$ и исследовать на экстремум ее зависимость от параметра $p \in [0, 1]$.

7. Пусть $\xi \geq 0$ — случайная величина с экспоненциальным распределением ($\mathcal{L}\{\xi\} = E(\lambda)$), плотность которого равна $p_\xi(x) = \lambda e^{-\lambda x}$, $x \geq 0$. Вычислить дифференциальную энтропию $\mathbf{H}_d\{\xi\}$ и исследовать на экстремум ее зависимость от параметра $\lambda > 0$.

8. Исследовать семейство одномерных абсолютно непрерывных распределений вероятностей, заданных на полупрямой:

$$\mathcal{P}_4(\lambda) = \left\{ p_X(x),\ x \geq 0 : p_X(x) \geq 0,\ \int_0^\infty p_X(x)\,dx = 1,\right.$$

$$\left.\int_0^\infty x p_X(x)\,dx = \frac{1}{\lambda} \right\},\ \lambda > 0.$$

Найти максимум дифференциальной энтропии на классе $\mathcal{P}_4(\lambda)$. Показать, что этот максимум достигается для плотности экспоненциального распределения вероятностей $E(\lambda)$:

$$p^*(x) = \lambda e^{-\lambda x},\ x \geq 0.$$

9. Случайные символы ξ_1, ξ_2 в сообщении зависимы. Известно, что $\mathbf{H}\{\xi_1\} = 8$ битов, $\mathbf{H}\{\xi_2\} = 12$ битов. Какие значения может принимать условная энтропия $\mathbf{H}\{\xi_2 \mid \xi_1\}$, если $\mathbf{H}\{\xi_1 \mid \xi_2\}$ изменяется в максимально возможных пределах?

10. Пусть стационарный источник дискретных сообщений порождает случайную символьную последовательность $\Xi_n = (\xi_1, \ldots, \xi_n) \in A^n$, обладающую марковским свойством s-го порядка ($s \geq 0$):

$$\mathbf{P}\{\xi_n = a_n \mid \xi_{n-1} = a_{n-1}, \ldots, \xi_1 = a_1\} = \mathbf{P}\{\xi_n = a_n \mid \xi_{n-1} = a_{n-1}, \ldots, \xi_{n-s} = a_{n-s}\},$$

$$a_1, \ldots, a_n \in A,\ n > s.$$

Доказать, что условная энтропия

$$\mathsf{H}\{\xi_n \mid \Xi_{n-1}\} = \mathsf{H}\{\xi_s \mid \xi_{s-1}, \ldots, \xi_1\}, \quad n \geq s,$$

и удельная энтропия

$$h = \mathsf{H}\{\xi_s \mid \xi_{s-1}, \ldots, \xi_1\}.$$

11. Пусть $f_1(x)$, $f_2(x)$ — плотности распределения двух случайных символов сообщений ξ_1, ξ_2, отличающиеся только параметром сдвига. Доказать, что дифференциальные энтропии символов совпадают: $\mathsf{H}_d\{\xi_1\} = \mathsf{H}_d\{\xi_2\}$.

12. Пусть ξ — случайный символ сообщения с плотностью распределения вероятностей $f(x)$, а $\eta = a\xi + b$, где a, b — фиксированные неслучайные величины, причем $a \neq 0$. Доказать, что $\mathsf{H}_d\{\eta\} = \mathsf{H}_d\{\xi\} + \log|a|$.

SXJKN8EYHR3I4HR98
7YDWFUIHW9E8YR384
H5O34596EUF9O8WE7
6R83Y439846UHFE98
YDSFJH8374H4BR827
6RUBNWDF87Y6234BN
9WE872H3WRTJOJRTG
M905JRO9G7DFJBHVI
EURKDHFVIUDKFJBVK
DFVKJDBFKVJHB0954

Глава 6

МЕТОДЫ ТЕОРИИ ИНФОРМАЦИИ В КРИПТОЛОГИИ

SXJKN8EYHR3I4HR98
7YDWFUIHW9E8YR384
H5O34596EUF9O8WE7
6R83Y439846UHFE98
YDSFJH8374H4BR827
6RUBNWDF87Y6234BN
9WE872H3WRTJOJRTG
M905JRO9G7DFJBHVI
EURKDHFVIUDKFJBVK
DFVKJDBFKVJHB0954

6.1. Асимптотические свойства стационарного источника дискретных сообщений

Пусть рассматривается стационарный ИДС без памяти с алфавитом $A = V = \{0, 1\}$, порождающий случайную последовательность независимых в совокупности одинаково распределенных двоичных символов: $\Xi_n = (\xi_1, \ldots, \xi_n)$, $n = 1, 2, \ldots$, где $\xi_i \in \{0, 1\}$ — двоичный случайный символ с распределением вероятностей Бернулли ($i = \overline{1, n}$):

$$\mathbf{P}\{\xi_i = 1\} = p, \quad \mathbf{P}\{\xi_i = 0\} = q = 1 - p, \tag{6.1.1}$$

где $0 < p < 1/2$ — вероятность появления «1» (случай $p > 1/2$ сводится к этому случаю переобозначением символов $0 \leftrightarrow 1$).

Количество реализаций Ξ_n равно 2^n. В силу схемы независимых испытаний порождения символов распределение вероятностей Ξ_n задается соотношением

$$p_n(a) = p_n(a_1, \ldots, a_n) :: = \mathbf{P}\{\Xi_n = a\} = p^{\sum_{i=1}^{n} a_i} q^{n - \sum_{i=1}^{n} a_i} \tag{6.1.2}$$

или

$$\log p_n(a) = n \log q - \sum_{i=1}^{n} a_i \log \frac{q}{p} = n \log q - b \log \frac{q}{p},$$

где $a = (a_i) \in V_n$, $b = \sum_{i=1}^{n} a_i$ — суммарное количество «1» в двоичном n-векторе a.

Определим случайную величину

$$\eta_n = \sum_{i=1}^{n} \xi_i = |\Xi_n|^2. \tag{6.1.3}$$

Значения вероятностей (6.1.2) существенно изменяются при изменении величины b. Отношение наибольшей из этих вероятностей к наименьшей равно

$$\kappa_n = \frac{\max\limits_a p_n(a)}{\min\limits_a p_n(a)} = \frac{q^n}{p^n} = \left(\frac{q}{p}\right)^n > 1, \tag{6.1.4}$$

и эта величина экспоненциально растет с ростом n.

С учетом (6.1.1) и (6.1.3) вычислим моменты первого и второго порядка для случайной величины η_n:

$$\mathbf{E}\{\eta_n\} = n\mathbf{E}\{\xi_1\} = np < n/2, \quad \mathbf{D}\{\eta_n\} = n\mathbf{D}\{\xi_1\} = npq. \tag{6.1.5}$$

В силу (6.1.3) уклонение случайного числа единиц в Ξ_n от его среднего значения:

$$\zeta_n = \eta_n - np \tag{6.1.6}$$

имеет нулевое среднее и среднеквадратичное отклонение

$$\sigma_{\zeta_n} = \sqrt{\mathbf{D}\{\zeta_n\}} = \sqrt{npq},$$

а для относительного уклонения

$$\delta_n = \zeta_n / n \tag{6.1.7}$$

имеем

$$\mathbf{E}\{\delta_n\} = 0, \quad \sigma_{\delta n} = \sqrt{\mathbf{D}\{\delta_n\}} = \sqrt{\frac{pq}{n}}. \tag{6.1.8}$$

Как видно из (6.1.8), среднеквадратичное отклонение для случайной величины δ_n при $n \to \infty$ убывает как $1/\sqrt{n}$. Если $|b_n - np| = c\sqrt{npq}$, $c > 0$, то различие наибольшей и наименьшей вероятностей на этом подмножестве значений Ξ_n по-прежнему достаточно велико: $\kappa_n = (q/p)^{c\sqrt{npq}}$. Однако эта величина растет при $n \to \infty$ значительно медленнее, чем $(1/p)^{c\sqrt{npq}}$. Следовательно, справедливо неравенство

$$\log \frac{p_n(a)}{p_n(a')}\Bigg|_{|a|^2 = |a|^2 - c\sqrt{npq}} \ll \log \frac{1}{p_n(a)}.$$

Сформулируем это свойство в виде теоремы.

Теорема 6.1.1 (о высоковероятном подмножестве). Множество всех 2^n реализаций определенного выше двоичного случайного вектора $\Xi_n \in V_n$ можно разбить на два непересекающихся подмножества:

$$V_2^n = A_n \cup B_n, \; A_n \cap B_n = \varnothing, \tag{6.1.9}$$

так что при $n \to \infty$ выполняются свойства:

а) множество A_n имеет исчезающе малую вероятность

$$\mathbf{P}\{\Xi_n \in A_n\} = \sum_{a \in A_n} p_n(a) \to 0; \tag{6.1.10}$$

б) реализации из множества B_n становятся относительно равновероятными:

$$\left| \frac{\log p_n(a) - \log p_n(a')}{\log p_n(a)} \right| \to 0, \; a, a' \in B_n. \tag{6.1.11}$$

Доказательство. Воспользуемся неравенством Чебышёва (относительно дисперсий) для случайной величины η_n, определяемой (6.1.3) с учетом (6.1.5):

$$\forall \varepsilon > 0 : \mathbf{P}\{|\eta_n - np| \geq \varepsilon\} \leq \frac{\mathbf{D}\{\eta_n\}}{\varepsilon^2} = \frac{npq}{\varepsilon^2}.$$

Полагая $\varepsilon = n^{3/4}$, получим

$$\mathbf{P}\{|\eta_n - np| \geq \varepsilon^{3/4}\} \leq \frac{pq}{\sqrt{n}}. \tag{6.1.12}$$

Построим разбиение множества V_n следующим образом:

$$\begin{gathered} A_n = \left\{ a = (a_i) \in V_n : \left| \sum_{i=1}^{n} a_i - np \right| \geq n^{3/4} \right\}, \\ B_n = V_n \setminus A_n = \left\{ a = (a_i) \in V_n : \left| \sum_{i=1}^{n} a_i - np \right| < n^{3/4} \right\}. \end{gathered} \tag{6.1.13}$$

Тогда из (6.1.12) и (6.1.13) имеем

$$\mathbf{P}\{A_n\} = \mathbf{P}\{\Xi_n \in A_n\} \leq \frac{pq}{\sqrt{n}} \to 0,\ n \to \infty. \tag{6.1.14}$$

Следовательно, выполняется (6.1.10).

В силу (6.1.13), если $a \in B_n$, то

$$np - n^{3/4} < \sum_{i=1}^{n} a_i < np + n^{3/4},$$

поэтому, согласно (6.1.2),

$$n \log q - \log \frac{q}{p}(np + n^{3/4}) < \log p_n(a) < n \log q - \log \frac{q}{p}(np - n^{3/4}),$$

или

$$\begin{gathered} nq \log q + np \log p - n^{3/4} \log(q/p) < \log p_n(a) < \\ < nq \log q + np \log p + n^{3/4} \log(q/p),\ a \in B_n, \\ n(q \log q + p \log p) - n^{3/4} \log(q/p) < \log p_n(a) < \\ < n(q \log q + p \log p) + n^{3/4} \log(q/p). \end{gathered} \tag{6.1.15}$$

Отсюда имеем

$$\forall a, a' \in B_n : |\log p_n(a) - \log p_n(a')| < 2n^{3/4} \log \frac{q}{p}, \tag{6.1.16}$$

$$-\log p_n(a) = |\log p_n(a)| > -n(p \log p + q \log q) - n^{3/4} \log \frac{q}{p}. \tag{6.1.17}$$

Тогда из неравенств (6.1.16) и (6.1.17) следует:

$$\left|\frac{\log p_n(a) - \log p_n(a')}{\log p_n(a)}\right| < \frac{2n^{3/4}\log(q/p)}{-n(p\log p + q\log q) - n^{3/4}\log(q/p)} =$$

$$= \frac{2n^{-1/4}\log(q/p)}{-p\log p - q\log q - n^{-1/4}\log(q/p)} \to 0,\ n \to \infty,\ a,\ a' \in B_n$$

что совпадает с (6.1.11). □

Теорема 6.1.2. Пусть $B_n \subset V_n$ — «высоковероятное» подмножество реализаций, описанное в теореме (6.1.1). Тогда мощность этого подмножества $M_n = |B_n|$ при $n \to \infty$ удовлетворяет асимптотике

$$\frac{\log M_n}{n} \to \mathbf{H}\{\xi_1\} = -(p\log p + q\log q). \tag{6.1.18}$$

Доказательство. Согласно (6.1.2) и (6.1.13),

$$\mathbf{P}\{\Xi_n \in B_n\} = 1 - \mathbf{P}\{\Xi_n \in A_n\} = \sum_{a \in B_n} p_n(a). \tag{6.1.19}$$

В силу (6.1.14) из (6.1.19) имеем

$$1 \geq \mathbf{P}\{\Xi_n \in B_n\} \geq 1 - \frac{pq}{\sqrt{n}}. \tag{6.1.20}$$

С другой стороны, с учетом (6.1.15):

$$p_n(a) < 2^{n(p\log p + q\log q) + n^{3/4}\log(q/p)} = 2^{-n\mathbf{H}\{\xi_1\} + n^{3/4}\log(q/p)}.$$

Поэтому, используя второе равенство в (6.1.19), находим

$$\mathbf{P}\{\Xi_n \in B_n\} < M_n 2^{-n\mathbf{H}\{\xi_1\} + n^{3/4}\log(q/p)}. \tag{6.1.21}$$

Из (6.1.20) и (6.1.21) имеем

$$M_n 2^{-n\mathbf{H}\{\xi_1\} + n^{3/4}\log(q/p)} > 1 - \frac{pq}{\sqrt{n}}.$$

Следовательно,

$$M_n > \left(1 - \frac{pq}{\sqrt{n}}\right) 2^{n\mathbf{H}\{\xi_1\} - n^{3/4}\log(q/p)}. \tag{6.1.22}$$

Аналогично (6.1.22) получим оценку снизу для M_n. Воспользуемся в (6.1.15) и (6.1.20) левыми неравенствами:

$$1 \geq M_n 2^{-n\mathbf{H}\{\xi_1\} - n^{3/4}\log(q/p)}.$$

Отсюда имеем

$$M_n \leq 2^{n\mathbf{H}\{\xi_1\} + n^{3/4}\log(q/p)}. \tag{6.1.23}$$

Логарифмируя (6.1.22) и (6.1.23), деля на n и объединяя в совместное неравенство, получаем двухстороннюю оценку для $\log M_n/n$:

$$\mathsf{H}\{\xi_1\} - n^{-1/4}\log(q/p) < \frac{\log M_n}{n} \le \mathsf{H}\{\xi_1\} + n^{-1/4}\log(q/p).$$

Устремляя $n \to \infty$, приходим к (6.1.18). □

Теоремы 6.1.1 и 6.1.2 легко обобщаются для стационарного ИДС без памяти, у которого алфавит A имеет мощность $N = |A| > 2$. При этом $\Xi_n \in A^n$ принимает одно из N^n возможных различных значений. Согласно теоремам 6.1.1 и 6.1.2, внимания заслуживают лишь $M_n \approx 2^{n\mathsf{H}\{\xi_1\}}$ реализаций, которые можно считать равновероятными. Если распределение вероятностей $p_1(a_1)$, $a_1 \in A$ является равномерным: $p_1(a_1) = \mathsf{P}\{\xi_1 = a_1\} = 1/N$, $a_1 \in A$, то $\mathsf{H}\{\xi_1\} = \log N$ (формула Хартли) и $2^{n\mathsf{H}\{\xi_1\}} = N^n$, т.е. $M_n = |A|^n$. Однако если распределение вероятностей отлично от равномерного, то $\mathsf{H}\{\xi_1\} < \log N$ и доля заслуживающих внимания реализаций

$$u_n = \frac{|B_n|}{|A|^n} = \frac{2^{n\mathsf{H}\{\xi_1\}}}{N^n} = \left(\frac{2^{\mathsf{H}\{\xi_1\}}}{N}\right)^n = (2^{\mathsf{H}\{\xi_1\} - \log N})^n \tag{6.1.24}$$

неограниченно уменьшается с ростом n. Следовательно, подавляющее большинство реализаций при этом несущественные, и их можно отбросить. Этот факт лежит в основе теории кодирования сообщений и широко используется в криптологии.

Пример. Пусть рассматривается стационарный ИДС без памяти с двоичным алфавитом $A = V$ $(N = 2)$ и вероятностью появления единичного символа $0 \le p \le 1/2$. Тогда, согласно теореме 6.1.1, «высоковероятное множество» двоичных последовательностей имеет вид

$$B_n = \left\{ a = (a_1, \ldots, a_n) \in V_n : \left| \frac{1}{n}\sum_{i=1}^{n} a_i - p \right| < \frac{1}{n^{1/4}} \right\}.$$

Например, при $n = 10^4$ множество B_n состоит из двоичных последовательностей, для которых доля единиц заключена в следующих пределах:

$$p - 0.1 < \frac{1}{n}\sum_{i=1}^{n} a_i < p + 0.1.$$

При $p = 0.1$ количество единиц в таких двоичных последовательностях

$$0 \le \sum_{i=1}^{n} a_i \le 2000.$$

Доля этих «заслуживающих внимания» (по теореме 6.1.1) последовательностей составляет, согласно (6.1.24), величину

$$u_n = \left(\frac{2^{-(p\log p + (1-p)\log(1-p))}}{2} \right)^n = (F(p))^n,$$

где $F(p) = 2^{-(p\log p + (1-p)\log(1-p)+1)}$. Эта функция протабулирована ниже:

p	0.1	0.2	0.3	0.4	0.5
$F(p)$	0.692	0.825	0.921	0.980	1.000

Например, при $p = 0.1$ и $n = 100$ доля «заслуживающих внимания» реализаций составляет $u_n = 2^{-0.531n} \approx 10^{-16}$.

6.2. Энтропийная устойчивость случайных символьных последовательностей

Установленные в предыдущем параграфе асимптотические свойства разбиения множества реализаций для стационарных ИДС без памяти могут быть обобщены на класс зависимых случайных символьных последовательностей Ξ_n. При этом обобщении нам понадобится понятие *энтропийной устойчивости случайной последовательности.*

Случайная n-символьная последовательность $\Xi_n = (\xi_1, \ldots, \xi_n) \in A^n$ с n-мерным дискретным распределением вероятностей

$$p_n(a) = \mathbf{P}\{\Xi_n = a\},\ a = (a_i) \in A^n \tag{6.2.1}$$

и неубывающей по n энтропией $0 < \mathbf{H}\{X\} < \infty$:

$$\mathbf{H}\{X\} = \mathbf{E}\{-\log p_n(\Xi_n)\} = -\sum_{a \in A^n} p_n(a) \log p_n(a) \tag{6.2.2}$$

называется *энтропийно устойчивой*, если при увеличении числа символов ($n \to \infty$) имеет место сходимость по вероятности:

$$\frac{-\log p_n(\Xi_n)}{\mathbf{H}\{X\}} \xrightarrow{P} 1, \tag{6.2.3}$$

т.е. $\forall \varepsilon > 0\ \exists \bar{n} = \bar{n}(\varepsilon)$ такое, что при любом $n \geq \bar{n}(\varepsilon)$ выполняется неравенство

$$\mathbf{P}\left\{ \left| \frac{-\log p_n(\Xi_n)}{\mathbf{H}\{X\}} - 1 \right| \geq \varepsilon \right\} < \varepsilon; \tag{6.2.4}$$

при этом функция $\bar{n}(\varepsilon)$ — монотонно убывающая.

Отметим, что для реальных прикладных задач криптологии энтропия возрастает с ростом n: $\mathbf{H}\{X\} \to \infty$. Сформулируем важнейший результат, полученный в 1967 г. Р.П. Стратоновичем.

Теорема 6.2.1 (обобщенная теорема Стратоновича). Если $\Xi_n \in A^n$ — произвольная случайная n-символьная последовательность, удовлетворяющая свойству энтропийной устойчивости, то ее множество N^n реализаций A^n можно разбить на два непересекающихся подмножества A_n и B_n таким образом, что при $n \to \infty$ выполняются следующие асимптотические свойства:

а) суммарная вероятность реализаций подмножества A_n исчезающе мала:

$$\mathbf{P}\{\Xi_n \in A_n\} = \sum_{a \in A_n} p_n(a) \to 0; \tag{6.2.5}$$

б) реализации высоковероятного подмножества B_n становятся асимптотически относительно равновероятными в следующем смысле:

$$\left|\frac{\log p_n(a) - \log p_n(a')}{\log p_n(a)}\right| \to 0,\ a,\ a' \in B_n; \tag{6.2.6}$$

в) количество $M_n = |B_n|$ реализаций (мощность) множества B_n связано с энтропией (последовательности) $\mathbf{H}\{X\}$ асимптотическим соотношением

$$\frac{\log M_n}{\mathbf{H}\{X\}} \to 1. \tag{6.2.7}$$

Доказательство. Оно состоит из трех частей.

1. В силу свойства энтропийной устойчивости (6.2.3) имеем неравенство (6.2.4), в котором для некоторого натурального m положим $\varepsilon = 1/m$. Тогда для любых $m = 1, 2, \ldots, n$, $n \geq \bar{n}(1/m)$ из (6.2.4) имеем

$$\mathbf{P}\left\{\left|\frac{-\log p_n(\Xi_n)}{\mathbf{H}\{X\}} - 1\right| \geq \frac{1}{m}\right\} < \frac{1}{m}. \tag{6.2.8}$$

В силу указанного выше свойства монотонности $\bar{n}(\varepsilon)$ выберем $m = m_n \in \mathbb{N}$ так, чтобы $\bar{n}(1/m_n, 1/m_n) \leq n \leq \bar{n}(1/(m_n+1), 1/(m_n+1))$, и опеределим подмножества A_n и B_n, используя (6.2.8):

$$\begin{gathered} A_n = \{a = (a_i) \in A^n : |-\log p_n(\Xi_n)/\mathbf{H}\{X\} - 1| \geq 1/m_n\}, \\ B_n = A^n \setminus A_n. \end{gathered} \tag{6.2.9}$$

С учетом (6.2.8) и (6.2.9) получаем (6.2.5):

$$\mathbf{P}\{\Xi_n \in A_n\} < \frac{1}{m_n} \to 0,\ \ \mathbf{P}\{\Xi_n \in B_n\} > 1 - \frac{1}{m_n} \to 1,$$

так как очевидно, что $m_n \to \infty$ при $n \to \infty$.

2. В силу (6.2.9) $\forall a \in B_n$ имеем

$$\left(1 - \frac{1}{m_n}\right)\mathbf{H}\{X\} < -\log p_n(a) < \left(1 + \frac{1}{m_n}\right)\mathbf{H}\{X\}, \tag{6.2.10}$$

или (что эквивалентно)

$$1-\frac{1}{m_n}<\frac{-\log p_n(a)}{\mathsf{H}\{X\}}<1+\frac{1}{m_n}. \tag{6.2.11}$$

В силу (6.2.10) $\forall a, a' \in B_n$ имеем

$$\left|\frac{\log p_n(a)-\log p_n(a')}{-\log p_n(a)}\right| \le \frac{2\mathsf{H}\{X\}/m_n}{-\log p_n(a)}.$$

Согласно (6.2.11), $-H\{X\}/-\log p_n(a) < 1/(1-1/m_n)$, поэтому получаем неравенство

$$\left|\frac{\log p_n(a)-\log p_n(a')}{-\log p_n(a)}\right| < \frac{2/m_n}{1-1/m_n} = \frac{2}{m_n-1} \to 0,$$

что доказывает (6.2.5).

3. Для доказательства (6.2.7) запишем (6.2.10) $\forall a \in B_n$ в эквивалентном виде:

$$2^{-(1+1/m_n)\mathsf{H}\{X\}} < p_n(a) < 2^{-(1-1/m_n)\mathsf{H}\{X\}}. \tag{6.2.12}$$

Так как

$$1 \ge \mathsf{P}\{\Xi_n \in B_n\} = \sum_{a\in B_n} p_n(a) \ge 1-\frac{1}{m_n},$$

то с учетом (6.2.12) (как при доказательстве теоремы 6.1.2)

$$\left(1-\frac{1}{m_n}\right)2^{(1-1/m_n)\mathsf{H}\{X\}} < M_n < 2^{(1+1/m_n)\mathsf{H}\{X\}}.$$

Проводя логарифмирование и деление частей этого неравенства на $\mathsf{H}\{X\}$, получаем

$$1-\frac{1}{m_n}+\frac{\log(1-1/m_n)}{\mathsf{H}\{X\}} < \frac{\log M_n}{\mathsf{H}\{X\}} < 1+\frac{1}{m_n}.$$

Так как $\mathsf{H}\{X\}$ с ростом n не убывает, то из этих неравенств при $n \to \infty$ заключаем (6.2.7). □

Заметим, что проверка основного условия теоремы 6.2.1 — условия энтропийной устойчивости (6.2.3) — на практике затруднительна. В связи с этим сформулируем ряд легко проверяемых на практике достаточных условий, влекущих выполнение свойства (6.2.3).

Теорема 6.2.2. Если существует равный нулю предел

$$\lim_{n\to\infty}\frac{\mathsf{D}\{\log p_n(\Xi_n)\}}{(\mathsf{H}\{X\})^2} = 0, \tag{6.2.13}$$

то случайная символьная последовательность Ξ_n является энтропийно устойчивой.

Доказательство. Воспользуемся неравенством Чебышёва (относительно дисперсий) для произвольной случайной величины $\xi \in \mathbb{R}^1$ и произвольного $\varepsilon > 0$:

$$\mathbf{P}\{|\xi - \mathbf{E}\{\xi\}| \geq \varepsilon\} \leq \frac{\mathbf{D}\{\xi\}}{\varepsilon^2}.$$

Полагая $\xi = -\log p_n(\Xi_n)/\mathbf{H}\{X\}$, используя (6.2.13) и учитывая $E\{\xi\} = 1$, при $n \to \infty$ получаем

$$\mathbf{H}\left\{\left|\frac{-\log p_n(\Xi_n)}{\mathbf{H}\{X\}} - 1\right| \geq \varepsilon\right\} \to 0,$$

что по определению сходимости по вероятности и означает (6.2.3). □

Замечание 6.2.1. Фигурирующая в (6.2.13) величина

$$\delta_n^2 = \frac{\mathbf{D}\{\log p_n(\Xi_n)\}}{(\mathbf{H}\{X\})^2} = \frac{\mathbf{D}\{-\log p_n(\Xi_n)\}}{(\mathbf{E}\{-\log p_n(\Xi_n)\})^2} \geq 0$$

есть квадрат коэффициента вариации случайной величины $-\log p_n(\Xi_n)$.

Таким образом, условие (6.2.13) означает стремление к нулю коэффициента вариации: $\delta_n \to 0$.

Теорема 6.2.3. Если энтропия случайной символьной последовательности при $n \to \infty$ бесконечно возрастает ($\mathbf{H}\{X\} \to \infty$) и существует ограниченный верхний предел

$$\overline{\lim_{n\to\infty}} \frac{\mathbf{D}\{\log p_n(\Xi_n)\}}{\mathbf{H}\{X\}} \leq C < +\infty, \tag{6.2.14}$$

то такая символьная последовательность Ξ_n энтропийно устойчива.

Доказательство. Из (6.2.14) следует, что $\forall \varepsilon > 0$ найдется такой номер $\bar{n} = \bar{n}(\varepsilon)$, что $\forall n \geq \bar{n}$ справедлива оценка сверху:

$$\frac{\mathbf{D}\{\log p_n(\Xi_n)\}}{(\mathbf{H}\{X\})^2} \leq \frac{C + \varepsilon}{\mathbf{H}\{X\}}.$$

Так как по условию $\mathbf{H}\{X\} \to \infty$, то при $n \to \infty$ правая часть этого неравенства стремится к нулю, поэтому выполняется (6.2.13). Применяя теорему (6.2.1), получаем доказываемое. □

Теорема 6.2.4. Если случайная символьная последовательность Ξ_n такая, что для нее существует положительная удельная энтропия

$$h = \lim_{n\to\infty} \frac{\mathbf{H}\{X\}}{n} > 0 \tag{6.2.15}$$

и существует так называемая *удельная дисперсия*

$$d ::= \lim_{n\to\infty} \frac{\mathbf{D}\{\log p_n(\Xi_n)\}}{n} \geq 0,\ d < \infty, \tag{6.2.16}$$

то Ξ_n энтропийно устойчива.

Доказательство. Из (6.2.15) и (6.2.16) следуют асимптотические соотношения:

$$\mathbf{H}\{X\} = hn + o(n),\ \ \mathbf{D}\{\log p_n(\Xi_n)\} = dn + o(n).$$

Поэтому при $n \to \infty$ существует конечный предел:

$$\frac{\mathbf{D}\{\log p_n(\Xi_n)\}}{\mathbf{H}\{X\}} = \frac{d + o(1)}{h + o(1)} \to \frac{d}{h} \geq 0.$$

В силу (6.2.14) и теоремы 6.2.3 получаем доказываемое. □

Следствие 6.2.1. Если последовательность Ξ_n состоит из n независимых в совокупности одинаково распределенных невырожденных случайных символов (т.е. порождается стационарным ИДС без памяти), то она является энтропийно устойчивой.

Доказательство. Проверим третье достаточное условие энтропийной устойчивости, выражаемое теоремой 6.2.4. По условию следствия имеем:

$$\log p_n(a) = \sum_{i=1}^{n} \log p_1(a_i),$$

поэтому

$$\mathbf{H}\{X\} = \mathbf{E}\{-\log p_n(\Xi_n)\} = n\mathbf{H}\{\xi_1\},\ \ 0 < \mathbf{H}\{\xi_1\} < \infty,$$

$$\mathbf{D}\{\log p_n(\Xi_n)\} = n\mathbf{D}\{\log p_1(\xi_1)\},\ \ 0 < \mathbf{D}\{\log p_1(\xi_1)\} < \infty.$$

В результате, согласно (6.2.15) и (6.2.16),

$$h = \mathbf{H}\{\xi_1\} > 0,\ \ d = \mathbf{D}\{\log p_1(\xi_1)\} < \infty.$$ □

Замечание 6.2.2. Понятие энтропийной устойчивости можно ввести и для отдельной случайной величины ξ, если $\forall\varepsilon > 0$

$$\mathbf{P}\left\{\left|\frac{-\log p_1(\xi)}{\mathbf{H}\{\xi\}} - 1\right| < \varepsilon\right\} > 1 - \delta.$$

6.3. Количество информации по Шеннону и его свойства

Рассмотрим простейшую схему передачи дискретной информации (рис. 6.1). Пусть источник дискретных сообщений порождает случайную символьную последовательность — входной сигнал

$$X = (x_1, \ldots, x_n) \in \mathcal{A}^n,$$

где $\mathcal{A}$ — конечный «входной алфавит» символов, $|\mathcal{A}| = N < \infty$; n — количество входных символов. Канал связи, находящийся под воздействием «помех», преобразует входной сигнал X в выходную последовательность (выходной сигнал)

$$Y = (y_1, \ldots, y_{n'}) \in \mathcal{B}^{n'},$$

где $\mathcal{B}$ — конечный «выходной алфавит» символов, $|\mathcal{B}| = N' < \infty$; n' — количество выходных символов. Это преобразование может быть как детерминированным, так и стохастическим.

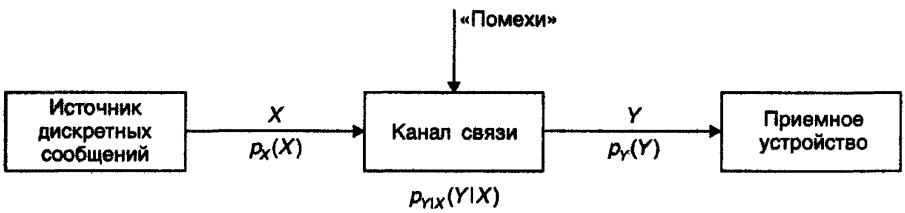


Рис. 6.1. Простейшая схема передачи информации

Обозначим: $p_X(X)$ — дискретное n-мерное распределение вероятностей входного сигнала; $p_Y(Y)$ — дискретное n'-мерное распределение вероятностей выходного сигнала; $p_{Y|X}(Y|X)$ — условное распределение вероятностей выходного сигнала Y при условии, что входной сигнал был X;

$$p_{X|Y}(X|Y) = \frac{p_{Y|X}(Y|X)p_X(X)}{p_Y(Y)} \tag{6.3.1}$$

есть условное распределение вероятностей входного сигнала X при условии, что выходной сигнал оказался Y.

Количеством информации по Шеннону, содержащейся в случайной символьной последовательности $Y \in \mathcal{B}^{n'}$ относительно входного сообщения $X \in \mathcal{A}^n$, называется разность безусловной и условной энтропий:

$$\mathsf{I}\{X, Y\} = \mathsf{H}\{X\} - \mathsf{H}\{X|Y\}, \tag{6.3.2}$$

где

$$\mathsf{H}\{X\} = -\sum_{X \in \mathcal{A}^n} p_X(X) \log p_X(X) \tag{6.3.3}$$

есть безусловная энтропия входного сообщения;

$$\mathsf{H}\{X|Y\} = -\sum_{\substack{X\in\mathcal{A}^n\\ Y\in\mathcal{B}^{n'}}} p_{X,Y}(X,Y)\log p_{X|Y}(X|Y) \tag{6.3.4}$$

есть условная энтропия входного сообщения X относительно выходного сообщения Y.

Данное определение количества информации как приращения энтропии введено К. Шенноном в 1948 г. Исследуем свойства шенноновского количества информации.

Свойство 6.3.1. Справедливы следующие эквивалентные выражения количества информации через энтропию:

$$\mathsf{I}\{X|Y\} = \mathsf{H}\{X\} + \mathsf{H}\{Y\} - \mathsf{H}\{X|Y\} = \mathsf{H}\{Y\} - \mathsf{H}\{Y|X\}. \tag{6.3.5}$$

Доказательство. Воспользуемся установленным ранее свойством иерархической аддитивности энтропии составной случайной последовательности $X||Y = (x_1, \ldots, x_n, y_1, \ldots, y_{n'}) \in \mathcal{A}^n \times \mathcal{B}^{n'}$:

$$\mathsf{H}\{X,Y\} = \mathsf{H}\{Y\} + \mathsf{H}\{X|Y\} = \mathsf{H}\{X\} + \mathsf{H}\{Y|X\}. \tag{6.3.6}$$

Из (6.3.6) имеем $\mathsf{H}\{X|Y\} = \mathsf{H}\{X,Y\} - \mathsf{H}\{Y\}$. Подставляя это в (6.3.2), получаем первое равенство в (6.3.5). Второе равенство в (6.3.5) получается подстановкой в его первую часть второго представления для $\mathsf{H}\{X,Y\}$ из (6.3.6). □

Свойство 6.3.2. Функционал шенноновского количества информации обладает свойством симметричности: $\mathsf{I}\{X,Y\} = \mathsf{I}\{Y,X\}$.

Доказательство. Согласно (6.3.5), имеем симметричное выражение

$$\mathsf{I}\{X,Y\} = \mathsf{H}\{X\} + \mathsf{H}\{Y\} - \mathsf{H}\{X,Y\} = \mathsf{I}\{Y,X\}.$$ □

Замечание 6.3.1. Это свойство показывает, что Y содержит такое же количество информации об X, что и X об Y.

Свойство 6.3.3. Справедлива следующая формула для вычисления шенноновского количества информации:

$$\mathsf{I}\{X,Y\} = -\sum_{\substack{X\in\mathcal{A}^n\\ Y\in\mathcal{B}^n}} p_{X,Y}(X,Y)\log\frac{p_{X,Y}(X,Y)}{p_X(X)p_Y(Y)}. \tag{6.3.7}$$

Доказательство. Вычислим правую часть (6.3.7), используя условие нормировки, свойство согласованности вероятностных распределений, а также (6.3.3) и (6.3.4):

$$B = -\mathsf{H}\{X,Y\} = -\sum_{X\in\mathcal{A}^n}\left(\sum_{X\in\mathcal{B}^{n'}} p_{X,Y}(X,Y)\log p_X(X)\right) - \sum_{X\in\mathcal{B}^{n'}}\left(\sum_{X\in\mathcal{A}^n} p_{X,Y}(X,Y)\log p_Y(Y)\right) =$$

$$= -\mathsf{H}\{X,Y\} + \mathsf{H}\{X\} + \mathsf{H}\{Y\},$$

что, согласно (6.3.5), совпадает с левой частью (6.3.7). □

Свойство 6.3.4. Количество информации, содержащейся в сообщении X о нем самом, равно энтропии сообщения X:

$$\mathsf{I}\{X,X\} = \mathsf{H}\{X\} = -\sum_{X\in\mathcal{A}^n} p_X(X)\log p_X(X). \tag{6.3.8}$$

Доказательство. Очевидно, что условное распределение X относительно X является вырожденным: $\mathsf{P}\{X = x \mid X = y\} = \delta_{xy}$. Поэтому $\mathsf{H}\{X \mid Y\} \equiv 0$ и из (6.3.2) следует (6.3.8). □

Замечание 6.3.2. Соотношение (6.3.8) означает, что энтропия, свойства которой исследованы в главе 5, может рассматриваться как частный случай функционала количества информации. Энтропию сообщения X поэтому иногда называют *собственной информацией* об X.

Свойство 6.3.5. Количество информации удовлетворяет неравенству

$$0 \le \mathsf{I}\{X,Y\} \le \min\{\mathsf{H}\{X\}, \mathsf{H}\{Y\}\}. \tag{6.3.9}$$

Доказательство. По свойствам условной энтропии:

$$0 \le \mathsf{H}\{X \mid Y\} \le \mathsf{H}\{X\},$$

поэтому из (6.3.2) и свойства 6.3.2 следует неравенство (6.3.9). □

Свойство 6.3.6. Количество информации по Шеннону $\mathsf{I}\{X,Y\}$ обращается в ноль тогда и только тогда, когда сообщения X, Y статистически независимы.

Доказательство. По свойству условной энтропии $\mathsf{H}\{X \mid Y\} = \mathsf{H}\{X\}$ тогда и только тогда, когда сообщения X,Y статистически независимы. Тогда из (6.3.2) получаем доказываемый результат.

Другой способ доказательства основан на использовании представления (6.3.8) и леммы 5.6.1. □

Свойство 6.3.7. При функциональных преобразованиях сообщений: $\tilde{X} = \varphi(X)$ или $\tilde{Y} = \psi(Y)$ количество информации не может возрасти:

$$\mathsf{I}\{X,Y\} \ge \mathsf{I}\{\varphi(X),Y\}, \tag{6.3.10}$$

$$\mathsf{I}\{X,Y\} \ge \mathsf{I}\{X,\psi(Y)\}, \tag{6.3.11}$$

причем равенства в (6.3.10) и (6.3.11) имеют место тогда и только тогда, когда $\varphi(\cdot)$, $\psi(\cdot)$ — биекции.

Доказательство. По определению (6.3.2) имеем

$$\mathsf{I}\{\tilde{X},Y\} = \mathsf{H}\{Y\} - \mathsf{H}\{Y|\tilde{X}\}.$$

По свойству условной энтропии $\mathsf{H}\{Y|\varphi(Y)\} \geq \mathsf{H}\{Y|X\}$, которое обращается в равенство лишь в случае, когда $\varphi(\cdot)$ — взаимно-однозначное функциональное преобразование. Поэтому справедливо неравенство (6.3.10). Неравенство (6.3.11) доказывается аналогично. □

Свойство 6.3.8. Если сообщения Y_1, Y_2 — независимы, то выполняется свойство аддитивности количества информации:

$$\mathsf{I}\{X,(Y_1,Y_2)\} = \mathsf{I}\{X,Y_1\} + \mathsf{I}\{X,Y_2\}. \tag{6.3.12}$$

Доказательство. Воспользуемся формулами (6.3.5), (6.3.2) и свойством аддитивности энтропии:

$$\mathsf{I}\{X,(Y_1,Y_2)\} = \mathsf{H}\{X,(Y_1,Y_2)\} - \mathsf{H}\{X\} - \mathsf{H}\{Y_1,Y_2\} =$$
$$= \mathsf{H}\{X,Y_1\} + \mathsf{H}\{X,Y_2\} - \mathsf{H}\{X\} - \mathsf{H}\{Y_1\} - \mathsf{H}\{Y_2\} = \mathsf{I}\{X,Y_1\} + \mathsf{I}\{X,Y_2\},$$

что совпадает с (6.3.12). □

Свойство 6.3.9. Справедливо неравенство

$$\mathsf{I}\{X,(Y_1,Y_2)\} \geq \max\{\mathsf{I}\{X,Y_1\}, \mathsf{I}\{X,Y_2\}\}.$$

Доказательство. В силу (6.3.5) и свойств энтропии

$$\mathsf{I}\{X,(Y_1,Y_2)\} = \mathsf{H}\{X\} - \mathsf{H}\{X|Y_1,Y_2\} \geq \mathsf{I}\{X,Y_i\},\ i = 1, 2. \qquad \square$$

Пример. Пусть двоичное сообщение описывается однородной цепью Маркова $\xi_1, \xi_2, \ldots \in \{0,1\}$ с дискретным временем, с начальным распределением

$$\pi = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

и матрицей вероятностей одношаговых переходов

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{pmatrix},\ 0 \leq \alpha \leq 1.$$

Пусть для некоторого момента времени $t = 1, 2, \ldots$ определены два соседних символа: $X ::= \xi_{t+1}$ (будущий символ = «пропущенный» символ), $Y ::= \xi_t$ (соседний наблюдаемый символ). Оценить количество информации о ξ_{t+1}, содержащейся в ξ_t.

Решение. По свойствам ОЦМ с дискретным временем имеем

$$\pi^* = \pi = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix},\ p_X(X) = p_Y(Y) \equiv \frac{1}{2},\ X, Y \in \{0,1\},$$

$$p_{X,Y}(X,Y)=\frac{1}{2}\begin{cases}1-\alpha, & \text{если } Y=X;\\ \alpha, & \text{если } Y\neq X.\end{cases}$$

Поэтому

$$\mathsf{I}\{\xi_{t+1},\xi_t\}=\sum_{X,Y=0}^{1}\frac{1}{2}((1-\alpha)\delta_{Y,X}+\alpha(1-\delta_{Y,X}))\log(2((1-\alpha)\delta_{Y,X}+\alpha(1-\delta_{Y,X})))=$$

$$=1+\alpha\log\alpha+(1-\alpha)\log(1-\alpha)=1+h(\alpha),$$

где $h(\alpha)=-\alpha\log\alpha-(1-\alpha)\log(1-\alpha)$ — функция, исследованная в примере из п. 5.5.

Рассмотрим систему передачи информации, в которой действует источник непрерывных сообщений $X=(x_1,\ldots,x_n)\in\mathbb{R}^n$, а принимаемое сообщение $Y=(y_1,\ldots,y_{n'})\in\mathbb{R}^{n'}$ также непрерывно.

Обозначим: $p_X(X)$ — плотность распределения вероятностей входного сообщения X; $p_Y(Y)$ — плотность распределения вероятностей выходного сообщения Y; $p_{Y|X}(Y|X)$ — условная плотность распределения выходного сообщения Y (при входном сообщении X);

$$p_{X|Y}(X|Y)=\frac{p_X(X)p_{Y|X}(Y|X)}{\int\limits_{\mathbb{R}^n}p_X(X')p_{Y|X}(Y|X')dX'}\tag{6.3.13}$$

есть условная плотность распределения входного сообщения X (при выходном сообщении Y); $p_{X,Y}(X,Y)$ — совместная плотность распределения;

$$\mathsf{H}_d\{X\}=-\int\limits_{\mathbb{R}^n}p_X(X)\log p_X(X)dX\tag{6.3.14}$$

есть безусловная дифференциальная энтропия входного случайного сообщения;

$$\mathsf{H}_d\{X|Y\}=-\int\limits_{\mathbb{R}^{n'}}\int\limits_{\mathbb{R}^n}p_{X,Y}(X,Y)\log p_{X|Y}(X|Y)\,dX\,dY\tag{6.3.15}$$

есть условная дифференциальная энтропия входного сигнала X относительно выходного сообщения Y.

Понятие количества информации при этом вводится аналогично дискретному случаю, рассмотренному выше.

Количеством информации по Шеннону, содержащейся в случайном выходном сигнале $Y\in\mathbb{R}^{n'}$, относительно случайного входного сообщения $X\in\mathbb{R}^n$ называется разность безусловной и условной дифференциальных энтропий (6.3.14) и (6.3.15):

$$\mathsf{I}\{X,Y\}=\mathsf{H}_d\{X\}-\mathsf{H}_d\{X|Y\}.\tag{6.3.16}$$

Свойства шенноновского количества информации (6.3.16) для дискретного случая сохраняют силу и для источника непрерывных сообщений. В частности,

аналогично свойству 6.3.3, удобна следующая формула для вычислений шенноновского количества информации:

$$\mathrm{I}\{X,Y\} = \int\limits_{\mathbb{R}^{n'}} \int\limits_{\mathbb{R}^{n}} p_{X,Y}(X,Y) \log \frac{p_{X,Y}(X,Y)}{p_X(X) p_Y(Y)} dX\, dY \geq 0. \qquad (6.3.17)$$

Пример. Пусть $X \in \mathbb{R}$, $Y \in \mathbb{R}$ — одномерные непрерывные сообщения (входной и выходной сигналы), имеющие совместное двумерное нормальное (гауссовское) распределение вероятностей:

$$\mathcal{L}\{X,Y\} = \mathcal{N}_2(\mu, \Sigma),$$

где $\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \in \mathbb{R}^2$ — вектор-столбец математических ожиданий, а

$$\Sigma = \begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{pmatrix}$$

есть ковариационная матрица входного и выходного сигналов.

Проводя вычисления согласно (6.3.17), получаем

$$\mathrm{I}\{X,Y\} = -\log\sqrt{1-\rho_{X,Y}^2}, \qquad (6.3.18)$$

где $\rho_{X,Y} = \sigma_{XY} / \sqrt{\sigma_{XX}\sigma_{YY}}$ — коэффициент корреляции входного и выходного сообщений.

Рассмотрим ситуацию, когда одно из сообщений X, Y непрерывно, а другое — дискретно. Пусть, например, имеется источник непрерывных сообщений $X \in \mathbb{R}^n$ с плотностью распределения $p_X(X)$, а канал связи является цифровым, поэтому $Y \in \mathcal{B}$ — дискретное выходное сообщение с дискретным распределением вероятностей $p_Y(Y)$, где $\mathcal{B}$ — дискретное множество. Обозначим через $p_{X|Y}(X|Y)$ условную плотность распределения случайного входного сигнала X при условии, что выходное сообщение — Y. Аналогично (6.3.17) получаем формулу для вычисления шенноновского количества информации:

$$\mathrm{I}\{X,Y\} = \sum_{Y \in \mathcal{B}} p_Y(Y) \int\limits_{\mathbb{R}^n} p_{X|Y}(X|Y) \log \frac{p_{X|Y}(X|Y)}{p_X(X)} dX. \qquad (6.3.19)$$

Свойства функционала (6.3.19) также совпадают со свойствами, доказанными ранее.

6.4. Шенноновские модели криптосистем

В двух заключительных параграфах данной главы рассмотрим применение Шенноновской теории информации к построению математических моделей криптосистем и оценке стойкости простейших симметричных криптосистем.

В настоящее время криптосистемы принято разделять на два класса: симметричные (одноключевые) и асимметричные (двухключевые). Общая схема симметричной криптосистемы приведена на рис. 6.2.

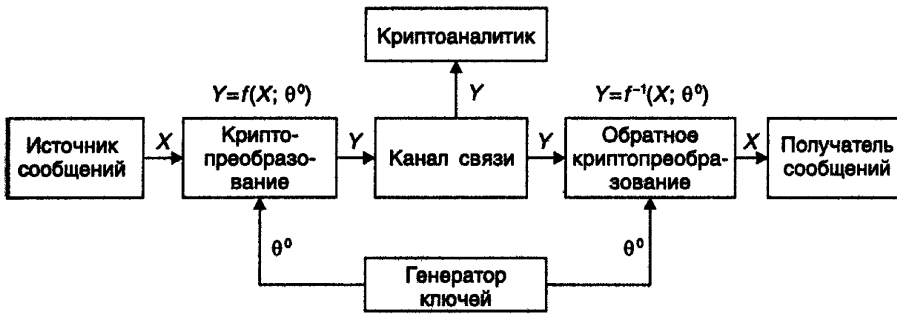


Рис. 6.2. Общая схема криптосистемы

Здесь приняты следующие обозначения: $X = (x_1, \ldots, x_n)$ — исходное сообщение (plaintext, message), генерируемое источником сообщений и представляющее собой последовательность n символов алфавита мощностью $\nu \geq 2$:

$$x_i \in A_\nu = \{0, 1, \ldots, \nu - 1\}, \quad i = \overline{1, n}, \quad X \in A_\nu^n;$$

$$\theta^0 = (\theta_1^0, \theta_2^0, \ldots, \theta_L^0) \in A_\mu^L$$

есть истинное значение ключа (key), L символов которого принимают значения из алфавита $A_\mu, \mu \geq 2$; $Y = (y_1, y_2, \ldots, y_n) \in A_\nu^n$ — шифртекст (криптограмма, выходная последовательность, зашифрованное сообщение, ciphertext), получающийся применением к X криптографического дискретного функционального преобразования (операции шифрования, cipher, encryption):

$$Y = f(X; \theta^0) : V_\nu^n \times V_\mu^L \to V_\nu^n, \tag{6.4.1}$$

или

$$y_i = f_i(x_1, \ldots, x_n; \theta^0), \quad i = \overline{1, n}.$$

Функция $f(\cdot)$ в (6.4.1) такова, что при любом фиксированном значении ключа θ^0 она является взаимооднозначным функциональным преобразованием (биекцией), и обратное преобразование (расшифровывание, дешифрование, decryption) единственно и восстанавливает исходный текст:

$$X = f^{-1}(Y; \theta^0). \tag{6.4.2}$$

Особенностью симметричных (одноключевых) систем является симметричное использование одного и того же ключа θ^0 отправителем и получателем (этот ключ «секретный» и поставляется абонентам специальным конфиденциальным способом). Отсюда и название данного класса криптосистем — одноключевые.

Приведем математические модели элементарных криптосистем, которые потребуются при оценке их стойкости, а также при решении задач криптоанализа и построения стандартных криптосистем.

1. Подстановка символов алфавита

Пусть определена подстановка на множестве $\{1, 2, \ldots, \nu\}$:

$$\begin{pmatrix} 1 & 2 & \ldots & \nu \\ s_1 & s_2 & \ldots & s_\nu \end{pmatrix} \tag{6.4.3}$$

и определен ν-вектор ($L = \nu = \mu$)

$$\theta^0 = (\theta_1^0, \ldots, \theta_\nu^0), \quad \theta_i^0 = s_i - 1, \quad i = \overline{1, \nu},$$

задающий перестановку символов алфавита (A_ν).

Тогда шифр простой подстановки — это криптографическое преобразование вида (6.4.1), осуществляемое поэлементно:

$$y_t = f_t(x_t; \theta^0) :: = \theta^0_{x_t + 1}, \quad t = 1, 2, \ldots \tag{6.4.4}$$

Обратное преобразование (6.4.2) при этом будет иметь вид

$$x_t = \overline{\theta}^0_{y_t + 1} = :: f_t^{-1}(y_t; \theta^0), \tag{6.4.5}$$

где $\overline{\theta}_i^0 = \overline{s}_i - 1$, а

$$\begin{pmatrix} 1 & 2 & \ldots & \nu \\ \overline{s}_1 & \overline{s}_2 & \ldots & \overline{s}_\nu \end{pmatrix}$$

есть подстановка, обратная (6.4.3).

2. Перестановка символов с периодом T

Пусть $T \in N$ — некоторый заданный период и определена некоторая подстановка на множестве $\{1, 2, \ldots, T\}$:

$$\begin{pmatrix} 1 & 2 & \ldots & T \\ s_1 & s_2 & \ldots & s_T \end{pmatrix}.$$

Как и в предыдущем преобразовании, с помощью этой подстановки определяется ключ $\theta^0 = (\theta_1^0, \ldots, \theta_T^0)$, $\theta_i^0 = s_i$, $i = \overline{1, T}$.

Криптопреобразование (6.4.1) осуществляется следующим образом. Исходное сообщение X разбивается на «блоки символов» длиной T, и внутри каждого блока производится перестановка символов в соответствии с заданным ключом θ^0. Для произвольного номера символа $t = (i-1)T + \tau$, где $i \in \{1, 2, \ldots\}$, $\tau \in \{1, 2, \ldots, T\}$, такое преобразование запишется в виде

$$y_t = x_{(i-1)T + \theta_\tau^0}. \tag{6.4.6}$$

Например, если $T = 5$, а $\theta^0 = (2\ 3\ 1\ 5\ 4)$, то сообщение

$$X = (x_1, x_2, x_3, x_4, x_5 \,|\, x_6, x_7, x_8, x_9, x_{10} \,|\ldots)$$

переходит в шифртекст

$$Y = (x_2, x_3, x_1, x_5, x_4 \,|\, x_7, x_8, x_6, x_{10}, x_9 \,|\ldots).$$

Легко убедиться, что обратное преобразование (по отношению к (6.4.6)) имеет вид

$$x_t = y_{(i-1)T+\bar{\theta}_\tau^0}, \tag{6.4.7}$$

где

$$\begin{pmatrix} 1 & 2 & \ldots & T \\ \bar{\theta}_1^0 & \bar{\theta}_2^0 & \ldots & \bar{\theta}_T^0 \end{pmatrix}$$

есть подстановка, обратная θ^0.

Для «усиления» стойкости криптопреобразования (6.4.6) к криптоанализу используют композицию нескольких перестановок с различными периодами. Если сделано $L \geq 2$ перестановок типа (6.4.6) с периодами $T_1, \ldots, T_L$, то составная перестановка, очевидно, будет иметь период

$$T = \text{НОК}(T_1, T_2, \ldots, T_L).$$

Следовательно, если периоды $\{T_1, \ldots, T_L\}$ — взаимно простые числа, то достигается максимальный период $T_{\max} = T_1 \cdots T_L$.

3. Шифр Виженера и его модификации

Как и в предыдущей криптосистеме, исходный текст X разбивается на блоки длиной T. Ключ θ^0 представляет собой фиксированный набор символов исходного алфавита A_ν ($\mu = \nu$):

$$\theta^0 = (\theta_1^0, \ldots, \theta_T^0), \quad \theta_i^0 \in V_\nu, \quad i = \overline{1, T}.$$

Криптофункция для произвольного номера $t = (i-1)T + \tau$, $i \in \{1, 2, \ldots\}$, $\tau \in \{1, 2, \ldots, T\}$ задается с помощью вычетов по модулю ν (см. п. 2.6):

$$y_t = (x_t + \theta_\tau^0) \bmod \nu. \tag{6.4.8}$$

Это криптопреобразование иногда называется преобразованием циклического сдвига с периодом T. Обратное преобразование по отношению к (6.4.8):

$$x_t = (y_t + \nu - \theta_\tau^0) \bmod \nu. \tag{6.4.9}$$

Пример шифртекста, построенного с помощью шифра Виженера с периодом $T = 6$ и ключом $\theta^0 = (4, 3, 2, 5, 1, 3)$, можно найти на с. 228 романа Ж. Верна «Жангада» [10].

Приведем ряд частных случаев криптопреобразования Виженера, известных с древних времен.

Шифр Цезаря — это частный случай преобразования Виженера с периодом $T = 1$ и ключом $\theta^0 \in A_\nu$:

$$y_t = (x_t + \theta^0) \bmod \nu, \quad t = 1, 2, \ldots \tag{6.4.10}$$

При этом, согласно (6.4.10), каждый символ (буква) исходного текста заменяется символом, циклически сдвинутым на фиксированное количество мест θ^0 по алфавиту A_ν.

В качестве примера приведем шифртекст длиной $n = 41$:

$$Y = \text{PELCGBYBTLVFPELCGBTENCULNAQPELCGBNANYLFVF},$$

полученный с помощью криптопреобразования Цезаря (6.4.10) при $\nu = 26$, $\theta^0 = 13$ из исходного отрывка английского текста:

$$X = \text{CRYPTOLOGYISCRYPTOGRAPHYANDCRYPTOANALYSIS}.$$

Иногда рассматривают обратный шифр Цезаря:

$$y_t = (\theta^0 + \nu - x_t) \bmod \nu. \tag{6.4.11}$$

Шифр Бофора — это модификация T-периодического шифра Виженера (6.4.8):

$$y_t = (\theta^0_\tau + \nu - x_t) \bmod \nu, \quad t = 1, 2, \ldots \tag{6.4.12}$$

Повторное применение $L \geq 2$ шифров Виженера называется составным шифром Виженера. Пусть определены L шифров Виженера, которые имеют периоды $T_1, \ldots, T_L$ и ключи $\theta^0_1 = (\theta^0_{11}, \ldots, \theta^0_{1T_1}), \ldots, \theta^0_L = (\theta^0_{L1}, \ldots, \theta^0_{LT_L})$. Если через $\{\theta^0_{lt} : t = 1, 2, \ldots\}$ обозначить ключ θ^0_τ, многократно периодически повторяемый с периодом T_l, то L-составной шифр Виженера имеет вид

$$y_t = (x_t + \theta^0_{1t} + \ldots + \theta^0_{Lt}) \bmod \nu, \quad t = 1, 2, \ldots \tag{6.4.13}$$

Легко показать, что L-составной шифр Виженера можно рассматривать как обычный шифр Виженера с периодом $T = \text{НОК}\{T_1, \ldots, T_L\}$.

4. Криптопреобразование Вернама (поточный шифр)

Криптопреобразование Вернама — специальный частный случай криптопреобразования Виженера (6.4.8), когда длина используемого ключа равна длине передаваемого сообщения n:

$$y_t = (x_t + \theta^0_t) \bmod \nu, \quad t = \overline{1, n}. \tag{6.4.14}$$

Обратное криптопреобразование имеет вид

$$x_t = (y_t + \nu - \theta^0_t) \bmod \nu, \quad t = \overline{1, n}.$$

В качестве ключа $\theta^0 = (\theta^0_1, \ldots, \theta^0_n) \in A^n_\nu$ используется реализация последовательности n независимых, одинаково распределенных на A_ν случайных величин либо некоторая реализация текста. Ключ θ^0 такого типа (используемого всего один раз) в литературе называется «бегущей строкой», «одноразовым блокнотом» (one-time pad) или гаммой. Иногда в качестве $\{\theta^0_t\}$ используется псевдослучайная последовательность, порождаемая специальным программным датчиком (см. главу 7). Более подробно модель поточного шифра рассмотрена в главе 10.

5. Биграммная (N-граммная) подстановка

Это криптопреобразование использует тот же принцип, что и простая подстановка (6.4.4). Однако в отличие от (6.4.4) в данном случае вместо подстановки одного символа: $A_\nu \leftrightarrow A_\nu$ осуществляется подстановка пар символов (биграмм): $A_\nu^2 \leftrightarrow A_\nu^2$ либо набора N соседних символов (N-грамм): $A_\nu^N \leftrightarrow A_\nu^N$. Ключ θ^0 в биграммной подстановке представляет собой $(\nu \times \nu)$-матрицу, (i, j)-й элемент которой — биграмма (i', j'), заменяющая биграмму (i, j).

В заключение отметим, что если имеются две произвольные криптосистемы T и R, то их часто можно комбинировать для получения новой криптосистемы

$$S = f(T, R).$$

Наиболее часто используются следующие два типа оператора комбинирования $f(\cdot)$.

1) *Произведение криптосистем:*

$$S = f_1(T, R) = TR, \quad S = f_2(T, R) = RT,$$

причем, вообще говоря, $TR \neq RT$. В существующих стандартных криптосистемах (см. главу 10) произведение шифров используется весьма часто. Например, после подстановки применяют транспозицию или после транспозиции — код Виженера.

2) *Взвешенная сумма криптосистем:*

$$S = pT + (1 - p)R, \quad p \in [0,1].$$

Выбор преобразования T осуществляется с вероятностью p, а преобразования R — с вероятностью $1 - p$. Эта функция комбинирования поясняется схемой на рис. 6.3.

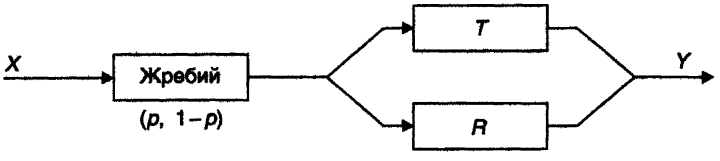


Рис. 6.3. Взвешенная сумма криптосистем

6.5. Теоретико-информационные оценки стойкости симметричных криптосистем

В этом параграфе исследуем общие вопросы устойчивости симметричных криптосистем к криптоанализу, используя теоретико-информационный подход Шеннона. Свойство устойчивости криптосистемы к криптоанализу принято называть криптостойкостью.

Прежде всего рассмотрим вопросы, связанные с потенциальной (т.е. максимально возможной) криптостойкостью. Насколько устойчива криптосистема, если криптоаналитик не ограничен временем и вычислительными средствами для анализа шифртекстов? Имеет ли шифртекст Y единственное решение (т.е. однозначно ли определяет ключ θ^0)? Какой должна быть минимальная длина шифртекста $n_{\min}$, чтобы решение стало единственным? Существуют ли криптосистемы, в которых нельзя найти единственное решение независимо от длины n исследуемого шифртекста? Существуют ли криптосистемы, в которых криптоаналитик не получает никакой информации, сколько бы он ни перехватил шифртекстов? Для решения этих проблем К. Шеннон применил аппарат математической теории информации. Будем рассматривать симметричные криптосистемы, описанные в п. 6.4.

Основное модельное предположение К. Шеннона об исходном сообщении X состоит в том, что «язык источника сообщений может рассматриваться как некоторый вероятностный процесс, порождающий дискретную последовательность символов в соответствии с некоторой системой вероятностей» (модели сообщений рассмотрены в главе 5). Поэтому исходное сообщение X К. Шеннон предполагал случайным n-вектором с дискретным распределением вероятностей:

$$\mathbf{P}\{X = X^{(i)}\} = q_i, \quad i = \overline{1, \nu^n}, \qquad q_1 + q_2 + \ldots + q_{\nu^n} = 1, \tag{6.5.1}$$

где $X^{(i)} \in A_\nu^n$ — i-й возможный вариант исходного n-символьного сообщения из алфавита A_ν. Ключ $\theta^0 = (\theta_1^0, \ldots, \theta_m^0)$ также предполагается (генерируется) случайным вектором, не зависящим от X, с дискретным распределением вероятностей:

$$\mathbf{P}\{\theta^0 = \theta^{(j)}\} = p_j, \quad j = \overline{1, \mu^m}, \quad \sum_{j=1}^{\mu^m} p_j = 1, \tag{6.5.2}$$

где $\theta^{(j)} \in A_\mu^m$ — j-й возможный вариант ключевой m-символьной последовательности в алфавите A_μ; p_j — *априорная вероятность ключа* $\theta^{(j)}$.

Симметричная криптосистема называется *совершенно криптостойкой*, если апостериорное распределение вероятностей исходного случайного сообщения X при регистрации случайного шифртекста $Y = f(X; \theta^0)$ совпадает с априорным распределением вероятностей:

$$\mathbf{P}\{X = X^{(i)} \mid Y = Y^{(l)}\} = \mathbf{P}\{X = X^{(i)}\} = q_i, \quad i, l = \overline{1, \nu^n}. \tag{6.5.3}$$

Смысл условия (6.5.3) в том, что хотя криптоаналитик и имеет шифртекст, он не добавляет ему информации о переданном сообщении.

Теорема 6.5.1. Необходимое и достаточное условие совершенной криптостойкости состоит в том, что условное распределение вероятностей шифртекста $Y \in A_\nu^n$ при фиксированном сообщении $X \in A_\nu^n$ не зависит от X:

$$\mathbf{P}\{Y = Y^{(l)} \mid X = X^{(i)}\} = \mathbf{P}\{Y = Y^{(l)}\}, \quad i, l = \overline{1, \nu^n}. \tag{6.5.4}$$

Доказательство. По формуле Байеса имеем

$$\mathbf{P}\{X = X^{(i)} \mid Y = Y^{(l)}\} = \frac{q_i\,\mathbf{P}\{Y = Y^{(l)} \mid X = X^{(i)}\}}{\mathbf{P}\{Y = Y^{(l)}\}}.$$

Видно, что (6.5.3) выполняется тогда и только тогда, когда

$$\frac{\mathbf{P}\{Y = Y^{(l)} \mid X = X^{(i)}\}}{\mathbf{P}\{Y = Y^{(l)}\}} = 1,$$

что совпадает с (6.5.4). □

Следствие 6.5.1. Если выполняется условие совершенной криптостойкости, то количество информации по Шеннону, содержащейся в шифртексте Y об исходном сообщении X, равно нулю:

$$\mathbf{I}\{X, Y\} = \mathbf{I}\{Y, X\} = 0. \tag{6.5.5}$$

Доказательство. Вычислим энтропию исходного сообщения X и условную энтропию X относительно шифртекста Y с учетом (6.5.3):

$$\mathbf{H}\{X\} = -\sum_{i=1}^{\nu^n} q_i \log q_i,$$

$$\mathbf{H}\{X \mid Y\} = -\sum_{i=1}^{\nu^n}\sum_{l=1}^{\nu^n} \mathbf{P}\{X = X^{(i)}, Y = Y^{(l)}\} \cdot \log \mathbf{P}\{X = X^{(i)} \mid Y = Y^{(l)}\} =$$

$$= -\sum_{i,l=1}^{\nu^n} \mathbf{P}\{X = X^{(i)}, Y = Y^{(l)}\} \cdot \log \mathbf{P}\{X = X^{(i)}\} = -\sum_{i=1}^{\nu^n} q_i \log q_i = \mathbf{H}\{X\}.$$

Тогда по определению количества информации имеем

$$\mathbf{I}\{X, Y\} = \mathbf{H}\{X\} - \mathbf{H}\{X \mid Y\} = 0.$$

Второе равенство в (6.5.5) вытекает из свойства симметричности функционала количества информации. □

Следствие 6.5.2. Пусть $\{Y^{(l)} : l = \overline{1, \nu^n}\} = A_\nu^n$ — множество всевозможных шифртекстов, порождаемых криптофункцией $Y = f(X; \theta)$;

$$\mathcal{J}_{il} = \{j : f(X^{(i)}, \theta^{(j)}) = Y^{(l)}\} \subset N \tag{6.5.6}$$

есть подмножество номеров ключей, переводящих исходный текст $X^{(i)}$ в один и тот же шифртекст $Y^{(l)}$. Чтобы $f(\cdot)$ удовлетворяла свойству совершенной криптостойкости, необходимо и достаточно выполнения свойства

$$\sum_{j \in \mathcal{J}_{il}} p_j = a_l = \underset{1 \le l \le \nu^n}{\text{invar}}. \tag{6.5.7}$$

Доказательство. Прежде всего отметим, что множество индексов $\mathcal{J}_{ll}$ не пусто в силу биективности $f(\cdot)$. Вычислим и сравним левую и правую части (6.5.4) с учетом (6.5.6) и свойства независимости θ^0 и X:

$$\mathbf{P}\{Y = Y^{(l)} \mid X = X^{(l)}\} = \mathbf{P}\{f(X^{(l)};\theta^0) = Y^{(l)} \mid X = X^{(l)}\} =$$

$$= \sum_{j\in\mathcal{J}_{ll}} \mathbf{P}\{\theta^0 = \theta^{(j)} \mid X = X^{(l)}\} = \sum_{j\in\mathcal{J}_{ll}} \mathbf{P}\{\theta = \theta^{(j)}\} = \sum_{j\in\mathcal{J}_{ll}} p_j; \qquad (6.5.8)$$

$$\mathbf{P}\{Y = Y^{(l)}\} = \mathbf{P}\{f(X;\theta^0) = Y^{(l)}\} = \sum_{i,j} q_i p_j \delta_{f(X^{(i)};\theta^{(j)}),Y^{(l)}} = \sum_{i=1}^{\nu^n} q_i \sum_{j\in\mathcal{J}_{ll}} p_j. \qquad (6.5.9)$$

Сравнивая (6.5.8) и (6.5.9), заключаем справедливость (6.5.7). □

Теорема 6.5.2. Необходимым условием выполнения свойства совершенной криптостойкости является справедливость следующих неравенств энтропий:

$$\mathbf{H}\{\theta^0\} \geq \mathbf{H}\{X\}, \qquad (6.5.10)$$

$$\mathbf{H}\{\theta^0\} \geq \mathbf{H}\{Y\}. \qquad (6.5.11)$$

Доказательство. Как было установлено при доказательстве теоремы 6.5.1, справедливо равенство

$$\mathbf{H}\{X\} = \mathbf{H}\{X \mid Y\}. \qquad (6.5.12)$$

Воспользуемся свойством иерархической аддитивности энтропии (см. главу 5):

$$\mathbf{H}\{Y, X, \theta^0\} = \mathbf{H}\{Y\} + \mathbf{H}\{X \mid Y\} + \mathbf{H}\{\theta^0 \mid X, Y\} =$$

$$= \mathbf{H}\{Y, \theta^0, X\} = \mathbf{H}\{Y\} + \mathbf{H}\{\theta^0 \mid Y\} + \mathbf{H}\{X \mid \theta^0, Y\}. \qquad (6.5.13)$$

Так как при фиксированном шифртексте Y и ключе θ^0 исходный текст $X = f^{-1}(Y;\theta^0)$ не случаен, то $\mathbf{H}\{X \mid \theta^0, Y\} = 0$. Поэтому из уравнения (6.5.13) находим $\mathbf{H}\{X \mid Y\} = \mathbf{H}\{\theta^0 \mid Y\} - \mathbf{H}\{\theta^0 \mid X, Y\}$. По свойствам энтропии отсюда следует:

$$\mathbf{H}\{X \mid Y\} \leq \mathbf{H}\{\theta^0 \mid Y\} \leq \mathbf{H}\{\theta^0\}.$$

Используя это в (6.5.12), приходим к (6.5.10). Неравенство (6.5.11) доказывается аналогично. □

Теорему 6.5.2 в криптологии называют «пессимистическим утверждением К. Шеннона», так как она требует, чтобы энтропия (неопределенность) ключа θ^0 была не меньше энтропии исходного текста X (или шифртекста Y). Поскольку распределение $\{q_i\}$ исходного текста X может быть произвольным, а $\max_{\{q_i\}} \mathbf{H}\{X\} = \log \nu^n$, то неравенство (6.5.10) примет вид

$$\mathbf{H}\{\theta^0\} \geq \log \nu^n.$$

Для его выполнения в случае $\mu = \nu$ требуется, чтобы длина ключа m была не меньше длины шифруемого текста: $m \geq n$. Для практики наиболее интересен случай самых коротких ключей: $m = n$.

Теорема 6.5.3. Если $\mu = \nu$, $m = n$ и $\forall i, l \in \{1, 2, \ldots, \nu^n\}$ уравнение

$$f(X^{(i)}; \theta^{(j)}) = Y^{(l)} \tag{6.5.14}$$

имеет единственное решение $j = j_{il}$ (т.е. $\mathcal{J}_{il} = \{j_{il}\}$ — одноточечное множество), то необходимым и достаточным условием совершенной криптостойкости является равновероятность используемых ключей:

$$p_j \equiv \text{const} = \frac{1}{\nu^n}, \quad j = \overline{1, \nu^n}. \tag{6.5.15}$$

Доказательство. Согласно (6.5.7),

$$p_{j_{il}} = a_l \quad \forall i, j, l. \tag{6.5.16}$$

Очевидно (в силу биекции), что j_{il} зависит от i и l так, что при изменении $i \in \{1, 2, \ldots, \nu^n\}$ индекс j_{il} пробегает ν^n всевозможных значений. Поэтому (6.5.16) невозможно без выполнения (6.5.15). □

Следствие 6.5.3. Криптопреобразование Вернама (6.4.14) при условии равновероятности ключей (6.5.12) обладает свойством совершенной криптостойкости.

Доказательство. В силу (6.4.14) уравнение (6.5.14) имеет единственное решение для любых $X^{(i)}, Y^{(l)}$: $\theta^{(i)} = (Y^{(l)} + \nu - X^{(i)}) \bmod \nu$, где вычеты вычисляются покомпонентно. Поэтому из (6.5.12) и теоремы 6.5.3 получаем доказываемый результат. □

Данное следствие объясняет, почему для передачи и защиты наиболее важной информации широко используются поточные криптосистемы, базирующиеся на криптопреобразовании Вернама. Следствие объясняет также, почему к качеству генерации ключевой последовательности $\theta^0 = (\theta_1^0, \theta_2^0, \ldots, \theta_n^0)$ (гаммы) предъявляются столь высокие требования (см. главу 7).

Рассмотрим еще одну важную характеристику криптосистем (связанную с криптостойкостью), введенную К. Шенноном, — расстояние единственности U. Для этого воспользуемся следующими соотношениями:

$\mathsf{H}\{\theta^0, X\} = \mathsf{H}\{\theta^0\} + \mathsf{H}\{X\}$ (по условию независимости θ^0, X);

$\mathsf{H}\{\theta^0 | X\} = \mathsf{H}\{Y | X\}$ (по свойствам криптосистемы);

$\mathsf{H}\{\theta^0, X\} = \mathsf{H}\{Y, X\}$ (вытекает из предыдущего);

$\mathsf{H}\{Y | X\} = \mathsf{H}\{Y\} + \mathsf{H}\{X | Y\}$ (по свойству энтропии).

Отсюда получаем выражение для условной энтропии исходного текста X относительно наблюдаемого шифртекста Y:

$$\mathsf{H}\{X\,|Y\} = \mathsf{H}\{X\} - (\mathsf{H}\{Y\} - \mathsf{H}\{\theta^0\}).$$

Так как условная энтропия неотрицательна, получаем неравенство

$$\mathsf{H}\{X\,|Y\} = \mathsf{H}\{X\} - (\mathsf{H}\{Y\} - \mathsf{H}\{\theta^0\}) \geq 0. \qquad (6.5.17)$$

Расстояние единственности — такая минимальная длина шифртекста Y (и исходного текста), при которой исчезает неопределенность в исходном тексте X при наблюдении шифртекста Y:

$$U = \min\{n : \mathsf{H}\{X\,|Y\} = 0\}. \qquad (6.5.18)$$

Следуя М. Хеллману (см. п. 6.1, 6.2), построим оценки энтропий, входящих в (6.5.17) и (6.5.18):

$$\mathsf{H}\{X\} = n\log\nu_X, \quad \mathsf{H}\{Y\} = n\log\nu_Y, \quad \mathsf{H}\{\theta^0\} = \log|\Theta|, \qquad (6.5.19)$$

где ν_X (ν_Y) — число, подбираемое так, что приближенно ν_X^n (ν_Y^n) реализаций исходных текстов X (шифртекстов Y) имеют вероятности, значительно отличающиеся от нуля, а остальные реализации имеют пренебрежимо малую вероятность; $|\Theta|$ — мощность пространства используемых ключей; $\nu_X \leq \nu_Y < \nu$.

Подставляя (6.5.19) в (6.5.17) и (6.5.18), находим приближенное выражение для расстояния единственности:

$$U = \frac{\log|\Theta|}{\log(\nu_Y/\nu X)}. \qquad (6.5.20)$$

Для реальных криптосистем обычно оказывается, что шифртекст Y имеет распределение, близкое к равномерному, поэтому $\nu_Y \approx \nu$. Тогда формула (6.5.20) принимает следующий вид:

$$U = \frac{\log|\Theta|}{k}, \quad k = \log\frac{\nu}{\nu_X} > 0,$$

где k — коэффициент, характеризующий избыточность языка. Например, для текстов на английском, немецком и французском языках этот коэффициент приближенно одинаков и равен $k \approx 0.53$.

6.6. Задания

1. Для трех дискретных случайных символов ξ_1, ξ_2, ξ_3 энтропии одинаковы: $\mathsf{H}\{\xi_1\} = \mathsf{H}\{\xi_2\} = \mathsf{H}\{\xi_3\} = h$. Вычислить количество информации $\mathsf{I}\{\xi_1, \xi_2, \xi_3\}$, если:
   а) $\mathsf{H}\{\xi_1, \xi_2, \xi_3\} = 3h$;
   б) $\mathsf{H}\{\xi_1, \xi_2, \xi_3\} = h$.

2. Алфавит состоит из 8 согласных и 8 гласных букв. В сообщении все буквы алфавита равновероятны и независимы. После прохождения через канал связи согласные всегда принимаются безошибочно, а гласные — только в половине случаев; в другой половине имеют место ошибки «случайного перепутывания гласных букв». Какое количество информации по Шеннону содержится в принятом символе о переданном?

3. ИДС генерирует независимую последовательность из алфавита, состоящего из 16 равновероятных символов. При передаче по каналу связи символы искажаются так, что четверть всех символов алфавита принимается неправильно, причем все ошибки равновероятны. Вычислить шенноновское количество информации в принятом символе о переданном.

4. Студент может получить зачет с вероятностью 0.3, не проработав весь учебный материал, и с вероятностью 0.9, проработав его полностью. Какое количество информации о степени проработки учебного материала можно получить по результату сдачи зачета, если в среднем 90 % студентов полностью проработали учебный материал?

5. Имеется случайная последовательность двоичных символов $\xi_1, \ldots, \xi_n \in \{0,1\}$. Известно, что возможно появление только таких реализаций, для которых $\xi_1 + \ldots + \xi_n = 0$, причем все они равновероятны. Вычислить:

$$I\{\xi_1, \xi_2\}, I\{(\xi_1, \xi_2), \xi_3\}, I\{(\xi_1, \ldots, \xi_{n-1}), \xi_n\}.$$

6. Статистика прогнозирования дождя в некотором городе характеризуется следующей таблицей вероятностей событий (ξ, η):

Прогноз η	Истинное состояние ξ	
	Дождь	Нет дождя
Дождь	2/16	3/16
Нет дождя	1/16	10/16

Студент заметил, что бюро прогнозов не ошибается лишь в 12/16 случаев. Если всегда давать прогноз «Нет дождя», то доля правильных решений окажется больше: 13/16. Студент предложил этот способ прогнозирования с просьбой о гонораре. Однако начальник бюро прогнозов как специалист по теории информации посчитал проект студента необоснованным. Почему?

7. Доказать, что $I\{\xi, f(\eta)\} \le I\{\xi, \eta\}$, каким бы ни было функциональное преобразование $f(\cdot)$.

8. Известно, что энтропия, приходящаяся на одну букву русскоязычного текста, составляет приблизительно 1.2 бита. Каково минимальное среднее

количество десятичных символов, необходимых для передачи информации, содержащейся в телеграмме из 100 букв?

9. На входе канала связи имеется последовательность двоичных независимых равновероятных символов. Ошибки в канале приводят к изменению значений некоторых символов на обратные, причем вероятность ошибки в k-м символе зависит лишь от наличия ошибки в предыдущем $(k-1)$-м символе: она равна 0.2 при наличии ошибки в $(k-1)$-м символе и 0.05 при ее отсутствии. Найти среднее количество передаваемой информации в расчете на символ.

10. Пусть $\xi, \eta \in \mathbb{R}^N$ — два независимых случайных вектора, причем η — случайный вектор с независимыми компонентами: $\mathbf{Cov}\{\eta, \eta\} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_N\}$. Пусть $\zeta = \xi + \eta$. Доказать, что количество информации по Шеннону удовлетворяет неравенству

$$\mathbf{I}\{\xi, \zeta\} \le \frac{1}{2}\sum_{i=1}^{N} \log\left(1 + \frac{\mathbf{D}\{\xi_i\}}{\lambda_i}\right),$$

которое обращается в равенство, если $\xi_1, \ldots, \xi_N$ — независимые гауссовские случайные величины.

Пусть $\xi_1, \ldots, \xi_m$ — максимальная невырожденная подсистема системы $n > m$ случайных символов $\xi_1, \ldots, \xi_n$. Доказать, что

$$\mathbf{I}\{\xi_1, \ldots, \xi_n, \eta\} = \mathbf{I}\{\xi_1, \ldots, \xi_m, \eta\}.$$

11. Используя понятия энтропии и информации по Шеннону, оценить, сколько вопросов надо задать, чтобы отгадать задуманное собеседником натуральное число, не превосходящее N, если он дает лишь двоичные ответы: «да» или «нет».

12. Имеется N монет одного достоинства; одна из них — фальшивая: легче или тяжелее остальных. Используя понятия энтропии и информации по Шеннону, оценить наименьшее число k взвешиваний на чашечных весах без гирь, которые позволяют найти фальшивую монету и определить, легче она или тяжелее остальных.

SXJKN8EYHR3I4HR
987YDWFUIHW9E8Y
R384H5O34596EUF
9O8WE76R83Y4398
46UHFE98YDSFJH8
374H4BR8276RUBN
WDF87Y6234BN9WE
872H3WRTJOJRTGM
905JRO9G70954TO
IRF7ERFH2347FQ7
WERF9UJ1349YUR7
WEF90UQ3I4OUJHR
P8937407140098E7

Глава 7

СТАТИСТИЧЕСКОЕ ТЕСТИРОВАНИЕ СЛУЧАЙНЫХ И ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

SXJKN8EYHR3I4HR
987YDWFUIHW9E8Y
R384H5O34596EUF
9O8WE76R83Y4398
46UHFE98YDSFJH8
374H4BR8276RUBN
WDF87Y6234BN9WE
872H3WRTJOJRTGM
905JRO9G70954TO
IRF7ERFH2347FQ7
WERF9UJ1349YUR7
WEF90UQ3I4OUJHR
P8937407140098E7

7.1. Равномерно распределенная случайная последовательность и ее свойства

Случайные числа и их генераторы являются неотъемлемыми элементами современных криптосистем. Приведем конкретные примеры использования случайных чисел в криптологии:

1) сеансовые и другие ключи для симметричных криптосистем, таких как DES, ГОСТ 28147—89, Blowfish;

2) стартовые значения для программ генерации ряда математических величин в асимметричных криптосистемах, например, «больших простых чисел» в криптосистемах RSA, ElGamal;

3) случайные слова, комбинируемые с парольными словами для нарушения «атаки угадывания» пароля криптоаналитиком;

4) вектор инициализации для блочных криптосистем, работающих в режиме обратной связи;

5) случайные значения параметров для многих систем электронной цифровой подписи, например DSA или белорусского стандарта СТБ 1176.2—99;

6) случайные выборы в протоколах аутентификации, например в протоколе Цербер (Kerberos);

7) случайные параметры протоколов для обеспечения уникальности различных реализаций одного и того же протокола, например в протоколах SET и SSL.

Отметим, что для некоторых из этих криптографических применений необходимы огромные массивы случайных чисел, которые по своему назначению требуют конфиденциального использования. Например, в протоколе Цербер сетевой сервер генерирует тысячи сессионных ключей ежечасно. К сожалению, компьютеры по своей конструкции предназначены быть детерминированными системами, поэтому на современных компьютерах генерация случайных чисел весьма затруднительна.

Известно, что проблема генерации случайной последовательности с произвольным законом распределения вероятностей сводится к проблеме генерации так называемой *равномерно распределенной случайной последовательности*

(РРСП), или, как ее часто называют в криптографических приложениях, «чисто случайной» последовательности. РРСП — это случайная последовательность $x_1, x_2, \ldots, x_t, x_{t+1}, \ldots$ со значениями в дискретном множестве $\mathcal{A} = \{0, 1, \ldots, N-1\}$, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ и удовлетворяющая двум свойствам — C_1 и C_2.

Свойство C_1. Для любого $n \in \mathbb{N}$ и произвольных значений индексов $1 \le t_1 < \ldots < t_n$ случайные величины $x_{t_1}, \ldots, x_{t_n} \in \mathcal{A}$ независимы в совокупности.

Свойство C_2. Для любого номера $t \in \mathbb{N}$ случайная величина x_t имеет дискретное равномерное на $\mathcal{A}$ распределение вероятностей:

$$\mathbf{P}\{x_t = i\} = \frac{1}{N}, \quad i \in \mathcal{A}.$$

Из базовых свойств C_1, C_2 вытекают следующие дополнительные свойства, используемые при генерации случайных чисел.

Свойство C_3. Если $\{x_t\}$ — РРСП, то для любого $n \in \mathbb{N}$ и любой фиксированной последовательности индексов $1 \le t_1 < \ldots < t_n$ n-мерное дискретное распределение вероятностей вектора (слова) $(x_{t_1}, x_{t_2}, \ldots, x_{t_n})$ является равномерным:

$$\mathbf{P}\{x_{t_1} = i_1, \ldots, x_{t_n} = i_n\} = \frac{1}{N^n}, \quad i_1, \ldots, i_n \in \mathcal{A}.$$

Свойство C_4. Если $x_t \in \mathcal{A}$ — элемент РРСП, то $\forall k \in \mathbb{N}$ справедливы следующие выражения его начального и центрального моментов k-го порядка:

$$\alpha_k = \mathbf{E}\{x_t^k\} = \frac{1}{N(k+1)} \sum_{l=0}^{k} \binom{k+1}{l} B_l N^{k+1-l};$$

$$\mu_k = \mathbf{E}\{(x_t - \alpha_1)^k\} = k! \sum_{l=0}^{k} \frac{B_l}{l!} \sum_{s=1}^{k+1-l} \frac{N^{s-1}}{s!} \frac{(-(N-1)/2)^{k+1-l-s}}{(k+1-l-s)!},$$

где $\{B_l\}$ — числа Бернулли [14].

Свойство C_5. Для ковариационной функции и спектральной плотности РРСП $\{x_t\}$ справедливы следующие выражения:

$$r(\tau) = \mathbf{E}\{(x_t - \alpha_1)(x_{t+\tau} - \alpha_1)\} = \frac{N^2 - 1}{12} \delta_{\tau,0}, \quad \tau \in \mathbb{Z};$$

$$S(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{+\infty} r(\tau) \cos(\lambda\tau) = \frac{N^2 - 1}{24\pi}, \quad \lambda \in [-\pi, +\pi].$$

Свойство C_6 (воспроизводимость при прореживании). Для любой фиксированной последовательности моментов времени $1 \le t_1 < \ldots < t_n < t_{n+1} < \ldots$ при «прореживании» РРСП $\{x_t\}$ возникает подпоследовательность

$$y_1 = x_{t_1}, \ldots, y_n = x_{t_n}, y_{n+1} = x_{t_{n+1}}, \ldots,$$

которая также является РРСП.

Свойство C_7 (воспроизводимость при суммировании). Если $x_t \in \mathcal{A}$ — РРСП, а $\xi_t \in \mathcal{A}$ — произвольная неслучайная либо случайная последовательность, не зависящая от $\{x_t\}$, то случайная последовательность

$$y_t = (x_t + \xi_t) \bmod N$$

также является РРСП.

Свойство C_8. Если $\{x_t\}$ — РРСП, то $\forall n \in \mathbb{N}$ количество информации по Шеннону, содержащейся в отрезке последовательности $X_n = (x_1, \ldots, x_n) \in \mathcal{A}^n$, о будущем элементе x_{n+1} равно нулю:

$$\mathrm{I}\{x_{n+1}, X_n\} = 0,$$

поэтому для любого алгоритма прогнозирования $\hat{x}_{n+1} = f(X_n) : \mathcal{A}^n \to \mathcal{A}$ вероятность ошибки прогнозирования не может быть меньше, чем для «угадывания по жребию»:

$$\mathbf{P}\{\hat{x}_{n+1} \ne x_{n+1}\} \ge 1 = \frac{1}{N}.$$

Свойство C_9. Если $\{x_t\}$ — РРСП, то для любого $k \in \mathbb{N}$ и произвольной интегрируемой борелевской функции k переменных $y = f(z_1, \ldots, z_k), z_1, \ldots, z_k \in \mathbb{R}$, при $n \to \infty$ имеет место сходимость «почти наверное»:

$$\frac{1}{n}\sum_{\tau=1}^{n} f(x_{(\tau-1)k+1}, \ldots, x_{\tau k}) \xrightarrow{\text{п.н.}} \frac{1}{N^k} \sum_{i_1, \ldots, i_k \in \mathcal{A}} f(t_{i_1}, \ldots, i_{i_k}).$$

Свойство C_{10}. Если $\{x_t\}$ — равномерно распределенная последовательность порядка $k = \infty$ в смысле Г. Вейля [35], то $\{x_t\}$ — РРСП.

С учетом свойств C_1 — C_{10} определим понятия генератора случайной последовательности и его типов.

Генератор РРСП — устройство, позволяющее по запросу получить реализацию равномерно распределенной случайной последовательности $x_1, \ldots, x_n \in \mathcal{A}$ длиной $n \in \mathbb{N}$; элементы $x_1, \ldots, x_n$ этой реализации принято называть случайными числами. Существует три типа генераторов РРСП: 1) *табличный*, 2) *физический*, 3) *программный*. В этой главе будем рассматривать наиболее распространенные программные генераторы.

Программный генератор РРСП — программа имитации на компьютере реализации РРСП. Имитируемая последовательность $\{x_t\}$ называется *псевдослучайной*, так как она вычисляется на компьютере по известному детерминированному (обычно рекуррентному) соотношению, и в то же время ее статистические свойства «близки» (по определенным критериям) к свойствам РРСП.

7.2. Универсальный алгоритм статистического тестирования случайных и псевдослучайных последовательностей

Как отмечалось в п. 7.1, в современных криптосистемах интенсивно используются равномерно распределенные случайные последовательности или имитирующие их псевдослучайные последовательности. Оказывается, что стойкость криптосистем существенно зависит от того, насколько точно соответствует модели РРСП используемая последовательность $x_t \in \mathcal{A}\ (t = 1, 2, \ldots)$. Проверка близости $\{x_t\}$ к модели РРСП осуществляется методами статистического тестирования и состоит в проверке гипотез выполнения базовых свойств C_1, C_2 РРСП (см. п. 7.1). Иногда вместо базовых свойств C_1, C_2 проверяют выполнение свойств C_3— C_{10} РРСП, следующих из C_1 и C_2.

Сформулируем *универсальный алгоритм статистического тестирования* случайных и псевдослучайных последовательностей. Пусть имеется программный (физический или табличный) генератор, порождающий псевдослучайную (случайную) последовательность

$$x_1, x_2, \ldots \in \mathcal{A} = \{0, 1, 2, \ldots, N-1\}$$

необходимой длины. Пусть определены две гипотезы: H_{0n} — о том, что наблюдаемая последовательность $x_1, x_2, \ldots, x_n \in \mathcal{A}$ заданной длины $n \in \mathbb{N}$ является РРСП и имеет n-мерное дискретное равномерное распределение вероятностей (см. свойство C_3 п. 7.1); $H_{1n} = \overline{H_{0n}}$ — альтернатива общего вида, утверждающая, что свойство C_3 нарушено и $\{x_t\}$ не соответствует модели РРСП. Согласно принятой в математической статистике классификации, гипотеза H_{0n} является простой, а H_{1n} — сложной.

Пусть для $m \in \mathbb{N}$ построена некоторая m-мерная векторная статистика

$$t = \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} = t(x_1, \ldots, x_n) = \begin{pmatrix} t_1(x_1, \ldots, x_n) \\ \vdots \\ t_m(x_1, \ldots, x_n) \end{pmatrix} \in \mathbb{R}^m, \tag{7.2.1}$$

удовлетворяющая двум условиям:

1) найдено дискретное распределение вероятностей

$$p_{0j} = \mathbf{P}_{H_{0n}}\{t(x_1, \ldots, x_n) = u_j\} > 0, \quad j = \overline{1, L}, \tag{7.2.2}$$

статистики (7.2.1) при верной гипотезе H_{0n}, где $U = \{u_1, \ldots, u_L\} \subset \mathbb{R}^m$ — конечное множество L упорядоченных (в лексикографическом порядке) возможных значений статистики (7.2.1);

2) при верной альтернативе H_{1n} дискретное распределение вероятностей

$$p_{1j} = \mathsf{P}_{H_{1n}}\{t(x_1, \ldots, x_n) = u_j\}, \quad j = \overline{1, L}, \tag{7.2.3}$$

статистики (7.2.1) отличается от распределения (7.2.2).

Универсальный алгоритм тестирования основан на выявлении различия дискретных вероятностных распределений (7.2.2) и (7.2.3) по наблюдаемой выборке с помощью χ^2-критерия согласия Пирсона и состоит из пяти этапов.

1. Генерируем подпоследовательность длиной $n_+ = Mn$ и разбиваем ее на M фрагментов:

$$X = (x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}, \ldots, x_{Mn}) = (X_1 \,||\, X_2 \,||\, \ldots \,||\, X_M),$$

где $X_i = (x_{(i-1)n+1}, x_{(i-1)n+2}, \ldots, x_{in}) \in \mathcal{A}^n$ — i-й фрагмент $i \in \{1, 2, \ldots, M\}$ длиной n; M — количество фрагментов.

2. Для i-го ($i = \overline{1, M}$) фрагмента X_i вычисляем значение m-мерной статистики (7.2.1):

$$t_i = t(x_{(i-1)n+1}, \ldots, x_{in}) \in U.$$

3. По сформированной таким образом выборке объема M вычисляем частоты встречаемости всех возможных значений статистики $t(\cdot)$:

$$\nu_j = \sum_{i=1}^{M} \delta_{t_i, u_j}, \quad j = \overline{1, L}, \tag{7.2.4}$$

где δ_{kl} — символ Кронекера.

4. Вычисляем χ^2-статистику Пирсона:

$$\chi^2 = \sum_{j=1}^{L} \frac{(\nu_j - Mp_{0j})^2}{Mp_{0j}} \geq 0. \tag{7.2.5}$$

5. Выносим решение по правилу:

$$\text{принимается гипотеза} \begin{cases} H_{0n}, & \text{если } P > \varepsilon, \\ H_{1n}, & \text{если } P \leq \varepsilon, \end{cases} \tag{7.2.6}$$

где $P = 1 - G_{L-1}(\chi^2)$ — так называемое *P-значение* для статистики (7.2.5); $G_f(\cdot)$-функция χ^2-распределения с f степенями свободы [4]; $\varepsilon \in (0, 1)$ — задаваемый уровень значимости теста.

Замечание 7.2.1. При увеличении количества фрагментов ($M \to \infty$) асимптотический размер теста (7.2.4)—(7.2.6), т.е. вероятность ошибки первого рода, совпадает с заданным уровнем значимости ε:

$$\lim_{M \to \infty} \mathsf{P}_{H_{0n}}\{\text{принимается } H_{1n}\} = \varepsilon,$$

и тест является состоятельным, т.е. его мощность стремится к единице (вероятность ошибки второго рода стремится к нулю):

$$\lim_{M \to \infty} \mathbf{P}_{H_{1n}}\{\text{принимается } H_{1n}\} = 1.$$

Замечание 7.2.2. Для практического применения теста (7.2.4)—(7.2.6) необходимо выбирать количество фрагментов M так, чтобы $M \min_j p_{0j} > 5$.

Замечание 7.2.3. Возможна модификация описанного теста, состоящая в проверке гипотезы о том, что статистика (7.2.5) имеет χ^2-распределение с $L-1$ степенями свободы. Для проверки такой гипотезы целесообразно s-кратно повторить выборку и использовать критерий Колмогорова [5], основанный на эмпирической функции распределения.

7.3. Тест n-серий

Тест n-серий является классическим тестом n-мерной равномерности. Он входит в *«батарею тестов» Кнута* [14] и является частным случаем универсального алгоритма тестирования (7.2.4)—(7.2.6), описанного в п. 7.2:

$$U = \mathcal{A}^n, \quad L = N^n; \quad t_k(x_1, \ldots, x_n) = x_k, \quad k = \overline{1,n};$$
$$m = n; \quad p_{0j} = \frac{1}{N^n}, \quad j = \overline{1,L}. \tag{7.3.1}$$

При этом вычисляемая, согласно (7.2.4), статистика ν_j — это частота встречаемости j-й возможной n-серии $(i_1, i_2, \ldots, i_n) \in \mathcal{A}^n$:

$$j = 1 + \sum_{k=1}^{n} i_k \cdot N^{k-1}.$$

Отметим, что серии, отличающиеся только перестановкой символов, считаются различными, поэтому общее количество серий $L = N^n$.

Отметим, что с учетом (7.3.1) и замечания 7.2.2 для применения теста (7.2.4)—(7.2.6) требуется, чтобы количество M фрагментов исследуемой последовательности $\{x_t\}$ было не менее критической величины

$$M_{\min} = 5 \cdot N^n.$$

Таким образом, при увеличении исследуемого порядка равномерности n требуемое количество фрагментов растет экспоненциально быстро. В этом заключается главный недостаток теста n-серий.

7.4. Тест интервалов

Тест интервалов — это еще один классический тест n-мерной равномерности, входящий в «батарею тестов» Кнута [14]. В этом тесте исследуются частоты случайных событий, состоящих в попадании элементов наблюдаемой последовательности $\{x_t\}$ в заданный числовой промежуток $[\alpha,\beta)$, где α, $\beta \in \mathcal{A}$, $\alpha < \beta$. При этом оценивается случайная длина серии t непопаданий в $[\alpha,\beta)$ до первого попадания. Тест интервалов является частным случаем общего алгоритма тестирования (см. п. 7.2) при следующих значениях определяющих его характеристик:

$$n = \infty, \quad m = 1, \quad U = \mathbb{N}_0 = \{0, 1, 2, \ldots\}, \quad L = \infty,$$

$$t(x_1, x_2, \ldots) = \begin{cases} 0, & x_1 \in [\alpha,\beta), \\ 1, & x_1 \notin [\alpha,\beta),\ x_2 \in [\alpha,\beta), \\ \ldots & \\ t, & x_1, \ldots, x_t \notin [\alpha,\beta),\ x_{t+1} \in [\alpha,\beta), \\ \ldots & \end{cases} \tag{7.4.1}$$

$$p_{0j} = p(1-p)^j, \quad j \in \mathbb{N}_0, \quad p = \frac{\beta - \alpha}{N}.$$

Отличительной особенностью данного теста является отсутствие заранее фиксированного размера фрагмента и, следовательно, заранее фиксированной длины n_+ генерируемой последовательности $\{x_t\}$; величина n_+ определяется как суммарная длина M фрагментов:

$$n_+ = \sum_{k=1}^{M} (t^{(k)} + 1), \tag{7.4.2}$$

где $t^{(k)} \in \mathbb{N}_0$ — зарегистрированная длина серии непопаданий в k-м фрагменте.

Заметим, что в силу (7.4.1) и (7.4.2) и свойств геометрического распределения вероятностей $\{p_{0j}\}$ средняя длина последовательности, необходимой для построения теста интервалов (7.2.4)—(7.2.6), (7.4.1), равна

$$\bar{n}_+ = \mathbf{E}\{n_+\} = \frac{MN}{\beta - \alpha},$$

а среднеквадратичное отклонение

$$\sigma_{n_+} = \sqrt{\mathbf{D}\{n_+\}} = \frac{N}{\beta - \alpha}\sqrt{M\left(1 - \frac{\beta - \alpha}{N}\right)}.$$

7.5. Обобщенный покер-тест

Обобщенный покер-тест и его специальный случай — *классический покер-тест* — также относятся к классическим тестам n-мерной равномерности, входящим в «батарею тестов» Кнута [14]. В этом тесте вида (7.2.4)—(7.2.6) ис-

следуются частоты встречаемости некоторых комбинаций символов в n-сериях $x = (x_1, \ldots, x_n) \in \mathcal{A}^n$. Статистика $t = t(x)$, на которой основан обобщенный покер-тест, определяется как количество различных символов в n-серии x ($n \geq 2$). При этом $L = \min(n, N)$,

$$p_{0j} = \frac{N(N-1)\cdots(N-j+1)}{N^n}\left\{ {n \atop j} \right\}, \quad j \in U = \{1, 2, \ldots, L\},$$

где $\left\{ {n \atop j} \right\}$ — *число Стирлинга второго рода* [28]. Комбинаторные числа Стирлинга определяются равенствами

$$x^n = \sum_{j=0}^{n} \left\{ {n \atop j} \right\} x(x-1)\cdots(x-j+1), \quad \left\{ {n \atop j} \right\} = \frac{1}{j!}\sum_{l=0}^{j}(-1)^{j-l}\binom{j}{l} l^n$$

и могут быть вычислены при помощи рекуррентных соотношений

$$\left\{ {n \atop j} \right\} = \left\{ {n-1 \atop j-1} \right\} + j\left\{ {n-1 \atop j} \right\}, \quad n \geq 1.$$

с начальными условиями $\left\{ {0 \atop 0} \right\} = 1$, $\left\{ {0 \atop j} \right\} = 0$, $j \neq 0$.

В классическом покер-тесте рассматриваются серии длиной $n = 5$, $n \leq N$, а статистика $t = t(x)$ определяет количество различных элементов, содержащихся в серии x:

$$t(x) = \begin{cases} 1, & \text{если } x \text{ состоит из 5 одинаковых элементов;} \\ 2, & \text{если } x \text{ содержит 2 различных элемента;} \\ 3, & \text{если } x \text{ содержит 3 различных элемента;} \\ 4, & \text{если } x \text{ содержит 4 различных элемента;} \\ 5, & \text{если все элементы в } x \text{ различны.} \end{cases}$$

Распределение вероятностей этой статистики при гипотезе H_{0n} легко находится с помощью комбинаторного анализа и приведено в работе [33].

7.6. Тест «собирателя купонов»

Тест «собирателя купонов» также является частным случаем теста (7.2.4)—(7.2.6), относится к классическим тестам Кнута [14] и основывается на исследовании случайной длины фрагмента последовательности $\{x_t\}$, обладающего заданным характерным свойством. Этот тест обычно применяется при $N > 2$. Если исходная тестируемая последовательность — двоичная, то можно

увеличить мощность алфавита до значения $N' = 2^r$, рассматривая «укрупненную последовательность» фрагментов размера $r > 1$:

$$x'_1 = \sum_{i=1}^{r} x_i 2^{i-1}, \quad x'_2 = \sum_{i=1}^{r} x_{i+r} 2^{i-1}, \ldots$$

Для проверки гипотезы n-мерной равномерности целесообразно брать $r = n$, тогда $\mathcal{A}' = \{0, 1, \ldots, 2^n - 1\}$, $N' = 2^n$.

В тесте «собирателя купонов» находится такой фрагмент $\{x'_1, x'_2, \ldots, x'_t\}$ минимальной длины t в «укрупненной последовательности», который «впервые соберет полный комплект N' купонов», т.е. содержит все N' символов из $\mathcal{A}'$.

Тест «собирателя купонов» имеет вид (7.2.4)—(7.2.6) при следующих значениях определяющих его характеристик:

$$m = 1, \quad U = \{N', N' + 1, \ldots\}, \quad L = \infty,$$

$$t(x'_1, x'_2, \ldots) = \min\{\tau \in \mathbb{N}_0 : \mathcal{A}' \subseteq \{x'_1, x'_2, \ldots, x'_\tau\}\}, \tag{7.6.1}$$

$$p_{0j} = \frac{N'!}{(N')^j} \left\{ \begin{matrix} j-1 \\ N'-1 \end{matrix} \right\}, \quad j \in U.$$

7.7. Тест перестановок

Тест перестановок — еще один из классических тестов Кнута [14], являющийся частным случаем теста общего вида (7.2.4)—(7.2.6).

Тест перестановок можно рассматривать как развитие покер-теста, основанного на частоте встречаемости различных перестановок символов «укрупненной последовательности»

$$x'_i \in \mathcal{A}' = \{0, 1, \ldots, N' - 1\}, \quad N' = 2^r,$$

формируемой как и в п. 7.6.

Для построения теста выберем произвольное натуральное число (параметр теста) $2 \le n' \le N'$ и рассмотрим вариационный ряд

$$0 \le x'_{(1)} \le x'_{(2)} \le \ldots \le x'_{(n')} \le N' - 1$$

«укрупненной последовательности» $x'_1, x'_2, \ldots, x'_{n'}$. Статистику $t(\cdot)$ теста определим следующим комбинаторным соотношением:

$$t = t(x'_1, \ldots, x'_{n'}) = \begin{cases} j, & \text{если } x'_1, \ldots, x'_{n'} \text{ попарно различны,} \\ n'! + 1 & \text{в противном случае,} \end{cases} \tag{7.7.1}$$

где $j = j(x'_1, \ldots, x'_{n'})$ — номер перестановки, переводящей упорядоченный набор $(x'_{(1)}, \ldots, x'_{(n')})$ в наблюдаемый набор $(x'_1, \ldots, x'_{n'})$.

При этом в обозначениях п. 7.2

$$U = \{1, 2, \ldots, n'!, n'!+1\}, \quad L = n'!+1,$$

$$p_{0j} = \begin{cases} \dfrac{1}{n'!}\prod\limits_{i=1}^{n'-1}\left(1-\dfrac{i}{N'}\right), & j \in \{1, 2, \ldots, n'!\}, \\ 1-\prod\limits_{i=0}^{n'-1}\left(1-\dfrac{i}{N'}\right), & j = n'!+1. \end{cases}$$

7.8. Тест пересекающихся n-грамм

Этот тест входит в *«батарею статистических тестов» Дж. Марсальи*, предложенную им в 1985 г. и предъявляющую «более жесткие требования» к исследуемой последовательности $\{x_t\}$, чем «батарея тестов» Кнута. Данный тест похож на использованный в заданиях 1, 5. Отличие состоит в том, что для построения теста используются частоты всевозможных комбинаций, вычисленные по пересекающимся n-граммам. Это усложняет нахождение распределения тестовой статистики при гипотезе H_{0n}, но существенно уменьшает длину n_+ требуемой для тестирования случайной последовательности $\{x_t\}$.

Тест пересекающихся n-грамм состоит в следующем. Подвергаемая тестированию последовательность $(x_1, x_2, \ldots, x_{n_+}) \in \mathcal{A}^{n_+}$ длиной n_+ $(n_+ > n)$ «зацикливается» добавлением в конце $n-1$ ее начальных символов:

$$X' = (x_1, x_2, \ldots, x_{n-1}, x_n, \ldots, x_{n_+}, x_1, x_2, \ldots, x_{n-1}).$$

Затем анализируются n_+ «пересекающихся n-грамм» (со сдвигом на одну позицию):

$$(x_1, x_2, \ldots, x_n), (x_2, x_3, \ldots, x_{n+1}), \ldots, (x_{n_+}, x_1, \ldots, x_{n-1}) \in \mathcal{A}^n,$$

и для $i_1, i_2, \ldots, i_n \in \mathcal{A} = \{0, 1, \ldots, N-1\}$ вычисляются частоты встречаемости N^n всевозможных n-грамм:

$$\nu_{i_1, \ldots, i_n} = \sum_{t=1}^{n_+} \mathbf{I}\{x_t = i_1, x_{t+1} = i_2, \ldots, x_{t+n-1} = i_n\}.$$

Аналогично вычисляются частоты встречаемости N^{n-1} всевозможных $(n-1)$-грамм $\{\nu_{i_1, \ldots, i_{n-1}}\}$.

С использованием вычисленных частот строится статистика

$$Q = \sum_{i_1, \ldots, i_n \in \mathcal{A}} \frac{(\nu_{i_1, \ldots, i_n} - n_+ N^{-n})^2}{n_+ N^{-n}} - \sum_{i_1, \ldots, i_{n-1} \in \mathcal{A}} \frac{(\nu_{i_1, \ldots, i_{n-1}} - n_+ N^{-n+1})^2}{n_+ N^{-n+1}} \geq 0, \quad (7.8.1)$$

Доказано [128], что при $n_+ \to \infty$ для истинной гипотезы H_{0n} распределение статистики Q сходится к χ^2-распределению с $f = N^n - N^{n-1} = N^{n-1}(N-1)$ степенями свободы:

$$\mathcal{L}\{Q \mid H_{0n}\} \to \chi^2_f.$$

Данный факт позволяет по аналогии с п. 7.2 построить решающее правило, основанное на P-значении $P = 1 - G_f(Q)$:

$$\text{принимается гипотеза } \begin{cases} H_{0n}, & \text{если } P > \varepsilon, \\ H_{1n}, & \text{если } P \le \varepsilon, \end{cases}$$

где $\varepsilon \in (0, 1)$ — заданный уровень значимости теста; $G_f(\cdot)$ — стандартная функция χ^2-распределения с f степенями свободы.

7.9. Тест, основанный на рангах двоичных матриц

Этот тест, входящий в «батарею статистических тестов» Дж. Марсальи, является частным случаем теста (7.2.4)—(7.2.6) и основан на вероятностных свойствах двоичной матрицы, которая строится по исследуемому фрагменту двоичной последовательности $x_1, x_2, \ldots, x_n \in \mathcal{A}$.

Пусть длина n исследуемого фрагмента последовательности представима в виде произведения двух чисел:

$$n = kl, \tag{7.9.1}$$

где k и l — некоторые натуральные числа, $2 \le k \le l$. Представим исследуемый фрагмент $x_1, x_2, \ldots, x_{kl} \in \mathcal{A}$ в виде двоичной $(k \times l)$-матрицы

$$X = (x_{ij}) ::= \begin{pmatrix} x_1 & x_2 & \ldots & x_l \\ x_{l+1} & x_{l+2} & \ldots & x_{2l} \\ \ldots & & & \\ x_{(k-1)l+1} & x_{(k-1)l+2} & \ldots & x_{kl} \end{pmatrix}. \tag{7.9.2}$$

Для теста (7.2.4)—(7.2.6) в качестве тестовой статистики $t(\cdot)$ используется ранг двоичной матрицы (7.9.2):

$$t = t(X) = \operatorname{rank}(X) \in U = \{0, 1, \ldots, k\}. \tag{7.9.3}$$

Вероятностные свойства статистики (7.9.3), необходимые для построения теста (7.2.4)—(7.2.6), выражаются следующим утверждением [175, 128].

Теорема 7.9.1. Если верна гипотеза n-мерной равномерности H_{0n}, то распределение вероятностей статистики (7.9.3) имеет вид

$$\mathsf{P}_{H_{0n}}\{\operatorname{rank}(X) = j\} = p_{0j} = 2^{j(k+l-j)-kl} \prod_{i=0}^{j-1} \frac{(1-2^{i-k})(1-2^{i-l})}{1-2^{i-j}}, \quad j \in U. \tag{7.9.4}$$

Отметим, что случай $j = 0$ соответствует нулевой матрице X; при этом

$$p_{00} = 2^{-kl}. \tag{7.9.5}$$

Таким образом, *тест, основанный на рангах двоичных матриц*, имеет вид (7.2.4)—(7.2.6) с учетом обозначений (7.9.1)—(7.9.5).

В заключение отметим, что вычисление ранга (7.9.3) осуществляется методом Гаусса — приведением матрицы X к матрице верхнего треугольного вида с помощью элементарных преобразований ее строк.

7.10. Спектральные тесты

Спектральные тесты — это тесты псевдослучайных (случайных) последовательностей, основанные на *преобразовании Фурье* (спектре Фурье) исследуемых последовательностей.

Рассмотрим спектральный тест для выявления скрытой периодичности в исследуемой последовательности $x_1, x_2, \ldots, x_n \in \mathbb{R}$. Определим гипотезу H_0: $x_t = \xi_t$ — РРСП с нулевым средним значением и дисперсией $\sigma^2 = \mathbf{D}\{x_t\}$ и альтернативу H_1 о наличии скрытой периодичности с некоторым неизвестным периодом $1 \le T_0 < n/2$ и соответствующей частотой $\omega_0 = 2\pi/T_0$:

$$x_t = \xi_t + f(t), \quad f(t) = \sum_{k=1}^{K}(a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t)), \tag{7.10.1}$$

где $f(t)$ — *периодический тренд* с некоторыми неизвестными коэффициентами $\{a_k\}, \{b_k\}$; $K \ge 1$ — параметр тренда, определяющий количество гармоник с частотами, кратными основной исследуемой частоте ω_0.

Если исследуется «чисто гармонический» тренд с единственной частотой ω_0, то полагаем $K = 1$. Будем предполагать, что n кратно T_0. Это предположение не является критическим, если n достаточно велико.

Определим конечное множество $L = [n/(2K)]$ частот

$$\Omega = \{\omega_j:\ \omega_j = 2\pi j/n,\ j = \overline{1, L}\}$$

и *дискретное преобразование Фурье* временного ряда $\{x_t\}$ на частоте $\omega \in [0, \pi]$ (i — мнимая единица):

$$X(\omega) = \frac{2}{n}\left(\sum_{t=1}^{n} x_t \cos(\omega t) + i\sum_{t=1}^{n} x_t \sin(\omega t)\right). \tag{7.10.2}$$

Для обнаружения скрытой периодичности строится специальная статистика — *гармограмма*:

$$H(\omega_j) = \frac{n}{2\sigma^2}\sum_{k=1}^{K}|X(k\omega_j)|^2 \ge 0, \quad j = \overline{1, L}. \tag{7.10.3}$$

Оказывается, если верна гипотеза H_0, то при $n \to \infty$ и известной дисперсии σ^2 статистика (7.10.3) имеет асимптотическое χ^2-распределение с $2K$ степенями свободы ($j = \overline{1, L}$) [2]:

$$\mathcal{L}\{H(\omega_j)\} \to \chi^2_{2K}. \tag{7.10.4}$$

Соотношения (7.10.1)—(7.10.4) позволяют построить следующий алгоритм статистического обнаружения периодичности и оценивания периода T_0.

1. Вычисление максимума гармограммы:

$$M_H = \max_{1 \le j \le L} H(\omega_j), \quad j_0 = \arg \max_{1 \le j \le L} H(\omega_j).$$

2. Статистическая проверка гипотез H_0, H_1 с помощью решающего правила:

$$\text{принимается} \begin{cases} H_0, & \text{если } P_1 > \varepsilon, \\ H_1, & \text{если } P_1 \le \varepsilon, \end{cases}$$

где $P_1 = 1 - (G_{2K}(M_H))^L$ — P-значение; $G_f(z)$ — функция χ^2-распределения с f степенями свободы; ε — заданный уровень значимости.

3. Если принята гипотеза H_1, то вычисляются статистические оценки частоты и периода:

$$\hat{\omega}_0 = \omega_{j_0}, \quad \hat{T}_0 = \frac{n}{j_0}.$$

Построим семейство спектральных тестов, основанных на асимптотическом свойстве (7.10.4) при $K = 1$. Согласно (7.10.2) и (7.10.3) построим последовательность значений выборочной спектральной плотности (периодограммы):

$$h_j = \frac{n}{2\sigma^2} \left| X\left(\frac{2\pi j}{n}\right) \right|^2 \ge 0, \quad j = 1, 2, \ldots, n/2; \tag{7.10.5}$$

здесь предполагается, что n — четное число.

Известно [2], что если верна гипотеза H_0, то определяемые (7.10.5) случайные величины $\{h_j\}$ при $n \to \infty$ асимптотически независимы и одинаково распределены по закону χ^2 с двумя степенями свободы:

$$\mathcal{L}\{h_j\} \to \chi^2_2, \quad F_{h_j}(z) \to G_2(z), \quad z \ge 0. \tag{7.10.6}$$

Зададим некоторое положительное число $c > 0$ и с помощью (7.10.5) построим последовательность двоичных случайных величин ($j = 1, 2, \ldots, n/2$):

$$\xi_j = \begin{cases} 1, & \text{если } h_j < c, \\ 0, & \text{если } h_j \ge c. \end{cases} \tag{7.10.7}$$

С учетом (7.10.5)—(7.10.7) гипотеза H_0 трансформируется в гипотезу H_0': $\{\xi_j\}$ — случайная выборка объема $n/2$ из распределения Бернулли: $Bi(1, p)$, $p = G_2(c)$.

Для проверки гипотезы H_0' против альтернативы общего вида $H_1' = \overline{H_0'}$ построим тест, основанный на статистике

$$m_n = \sum_{j=1}^{n/2} \xi_j \in \{0, 1, \ldots, n/2\}, \tag{7.10.8}$$

представляющей собой частоту появления «1» в выборке $\{\xi_j\}$. В силу теоремы Муавра — Лапласа при $n \to \infty$ и истинной гипотезе H_0 нормированная статистика

$$\eta_n = \frac{m_n - np/2}{\sqrt{np(1-p)/2}} \in \mathbb{R} \tag{7.10.9}$$

распределена асимптотически нормально:

$$\mathcal{L}\{\eta_n\} \to \mathcal{N}_1(0, 1).$$

На основании этого факта можно построить тест, асимптотический размер которого равен заданному значению $\varepsilon \in [0, 1]$:

$$\text{принимается} \begin{cases} H_0, & \text{если } P_2 > \varepsilon, \\ H_1, & \text{если } P_2 \le \varepsilon, \end{cases} \tag{7.10.10}$$

где $P_2 = 2(1 - \Phi(\eta_n))$, $\Phi(\cdot)$ — функция распределения стандартного нормального закона.

Отметим, что $c > 0$ — параметр построенного семейства решающих правил (7.10.5)—(7.10.10). В работе [175] используется частный случай этого решающего правила, когда параметр c выбран из условия

$$p = G_2(c) = 0.95, \quad c = G_2^{-1}(0.95) \approx 6,$$

а тестируемая последовательность $\{x_t\}$ является двоичной и для достижения нулевого значения математического ожидания преобразуется следующим образом:

$$x_t ::= (-1)^{x_t} \in \{-1, 1\}.$$

При этом для гипотезы H_0 дисперсия $\sigma^2 = 1$, а формула (7.10.7) принимает упрощенный вид

$$\xi_j = \begin{cases} 1, & \text{если } |X(2\pi j/n)| < \sqrt{12/n}, \\ 0, & \text{если } |X(2\pi j/n)| \ge \sqrt{12/n}. \end{cases}$$

Построим еще один спектральный тест, также основанный на асимптотическом (при $n \to \infty$) поведении статистик $\{h_j\}$, определяемых (7.10.5). Для этого по выборке $\{h_j\}$ построим *выборочную (эмпирическую) функцию распределения*

$$F_n(z) = \frac{2}{n} \sum_{j=1}^{n/2} \mathbf{1}(z - h_j), \quad z \ge 0. \tag{7.10.11}$$

Затем вычислим *расстояние Колмогорова* между выборочной функцией распределения (7.10.11) и гипотетической функцией распределения $G_2(z)$:

$$D_n = \sqrt{\frac{n}{2}} \max_{z \geq 0} \left| F_n(z) - G_2(z) \right|. \tag{7.10.12}$$

Тест Колмогорова, основанный на статистике (7.10.12), имеет следующий вид [38]:

$$\text{принимается} \begin{cases} H_0, \text{ если } D_n < \Delta_\varepsilon, \\ H_1, \text{ если } D_n \geq \Delta_\varepsilon, \end{cases} \tag{7.10.13}$$

где $\Delta_\varepsilon = \mathcal{K}^{-1}(1-\varepsilon)$ — квантиль уровня $1-\varepsilon$ распределения Колмогорова. Приведем некоторые значения функции Δ_ε:

Уровень значимости ε	0.01	0.05	0.10	0.20
Квантиль Δ_ε	1.63	1.36	1.22	1.07

7.11. Тесты случайного блуждания

Предположим, что исследуемая последовательность $\{x_t\}$ двоичная: $x_t \in \mathcal{A}$. С помощью преобразования $y_t = 2x_t - 1$, $t = \overline{1,T}$ перейдем к последовательности $\{y_t\}$, $y_t \in \{-1, 1\}$. Обозначим через S_n частичную сумму элементов последовательности $\{y_t\}$: $S_n = \sum_{i=1}^{n} y_i$, $n > 0$, $S_0 = 0$. *Тесты случайного блуждания* основаны на вероятностных свойствах *траектории* S_n, $n = 1, 2, \ldots$

Если $\{y_t\}$ — РРСП (гипотеза H_0), то для распределения частичных сумм $\{S_0, S_1, \ldots, S_n\}$ известны следующие результаты [33]:

$$p_{n,r} = \mathbf{P}\{S_n = r\} = \binom{n}{\frac{n+r}{2}} 2^{-n}, \quad p_{0,r} = \delta_{0,r}, \quad r = \overline{0,n}, \quad u_{2n} = \mathbf{P}\{S_{2n} = 0\} = \binom{2n}{n} 2^{-2n}.$$

Определим статистики [33], на основании которых построим тесты случайного блуждания.

Количество возвратов $\mathrm{RN}_{2L} = \left|\{2k : S_{2k} = 0,\ 0 < k \leq L\}\right|$. Распределение вероятностей статистики RN_{2L} при условии, что верна H_0, имеет вид

$$\mathbf{P}\{\mathrm{RN}_{2L} = r\} = p_{2L-r,r}, \quad 0 \leq r \leq L. \tag{7.11.1}$$

Количество изменений знаков $\mathrm{CH}_{2L+1} = \left|\{2k : S_{2k-1}S_{2k+1} < 0,\ 0 < k \leq L\}\right|$. Распределение вероятностей CH_{2L+1} при истинности H_0:

$$\mathbf{P}\{\mathrm{CH}_{2L+1} = r\} = 2p_{2L+1,2r+1}, \quad 0 \leq r \leq L. \tag{7.11.2}$$

Вес Хэмминга $\mathrm{HW}_{2L} = \sum_{k=1}^{2L} x_t = S_{2L}/2 + L$. Распределение вероятностей HW_{2L} при истинности H_0:

$$\mathbf{P}\{\mathrm{HW}_{2L} = r\} = \binom{2L}{r} 2^{-2L}, \quad 0 \le r \le 2L. \tag{7.11.3}$$

Исходная последовательность $\{x_t\}$ длиной n разбивается на отрезки длиной $2L$ (для статистики CH_{2L+1} — длиной $2L+1$) $\{X_1, \ldots, X_M\}$. На каждом из них вычисляются ряды частичных сумм $\{S_{(1)}, \ldots, S_{(M)}\}$, $S_{(j)} = (S_j^0, \ldots, S_j^M)$, $j = \overline{1, M}$.

Для каждого ряда $S_{(j)}$, $j = \overline{1, M}$, вычисляются значения используемой статистики (одной из трех описанных ранее; условимся статистику обозначать $\gamma_{2L}^j = \gamma_{2L}(S_{(j)})$). В результате получаем последовательность значений статистики

$$\{\gamma_{2L}^j\}, \ j = \overline{1, M}.$$

Данную последовательность M значений статистики разбивают на J групп $\{\Gamma_{(1)}, \ldots, \Gamma_{(J)}\}$ одинаковой длины $M_J = M/J$, $\Gamma_{(i)} = \{\gamma_{2L}^{(i-1)M_J+1}, \ldots, \gamma_{2L}^{iM_J}\}$. По каждой группе $\Gamma_{(i)}$, $i = \overline{1, J}$, находится эмпирическое распределение вероятностей статистики γ_{2L}:

$$p_m^j = \left|\{l : \gamma_{2L}^l = m, \ l = \overline{(j-1)M_J + 1, jM_J}\}\right|, \quad m = \overline{0, 2L}, \quad j = \overline{1, J}.$$

На основании j-го эмпирического распределения вероятностей вычисляем χ^2-статистику:

$$q_j = \sum_{m=0}^{K} \frac{(\tilde{p}_m^j - M_J \mu_{2L,m})^2}{M_J \mu_{2L,m}},$$

где $K = \max\{n : M_J \mathbf{P}\{\gamma_{2L} = n\} > 10\}$; $\mu_{2L,k} = \mathbf{P}\{\gamma_{2L} = k\}$; $\tilde{p}_k^j = p_k^j$, $k < K$; $\mu_{2L,K} = \sum_{k=K}^{2L} \mathbf{P}\{\gamma_{2L} = k\}$; $\tilde{p}_K^j = \sum_{k=K}^{2L} \tilde{p}_k^j$.

В результате получаем выборку $\{q_j\}$, $j = \overline{1, J}$, значений χ^2-статистики для групп $\{\Gamma_{(1)}, \ldots, \Gamma_{(J)}\}$. Далее находим выборочную функцию распределения $F_{J(z)}$ по выборке $\{q_j\}$, $j = \overline{1, J}$:

$$F_J(z) = \frac{1}{J} \sum_{j=1}^{J} \mathbf{1}\{z - q_j\}, \quad z \ge 0.$$

С помощью этой выборочной функции распределения вычисляется расстояние Колмогорова до функции χ^2-распределения $G_{J-1}(z)$ с $J-1$ степенями свободы:

$$D_J = \sqrt{J} \max_{z \ge 0} \left| F_J(z) - G_{(J-1)}(z) \right| \ge 0.$$

На основе этой статистики строится тест Колмогорова (аналогично (7.10.13)).

7.12. Универсальный статистический тест Маурера

Универсальный статистический тест Маурера основан на следующей идее: последовательность случайна, когда ее невозможно значительно упаковать или сжать. Для этого вводится мера, которая оценивает степень сжимаемости. Впервые этот метод был предложен Зивом [196]; в данном параграфе используется мера Элиса и Вильемса [90, 193] для *универсального кодирующего алгоритма.*

Пусть $x_1, \ldots, x_{n_+}$ — наблюдаемая последовательность над алфавитом $\mathcal{A} = \{0, 1\}$, $n_+ = (K + Q)L$, где K, Q и L — фиксированные натуральные числа. Параметр L — длина фрагмента, Q — количество начальных данных теста, K — количество основных шагов теста. Пусть $Y_i = (x_{L(i-1)+1}, \ldots, x_{Li})$ — фрагмент (слово) длиной L и $i = \overline{1, Q + K}$.

Для $Q + 1 \le i \le Q + K$ определим величину

$$A_i = \begin{cases} i, \text{ если не существует } 1 \le j < i, \text{ что } Y_i = Y_{i-j}; \\ \min\{j : j \ge 1, Y_i = Y_{i-j}\} \text{ в противном случае.} \end{cases} \tag{7.12.1}$$

Универсальный статистический тест Маурера основан на следующей статистике (мере сжимаемости):

$$F_{\text{uni}} = \frac{1}{K} \sum_{i=Q+1}^{Q+K} \log_2 A_i. \tag{7.12.2}$$

Теорема 7.12.1 (Маурер [139]). Если выполняется гипотеза H_0 и $Q \to \infty$, то для статистик (7.12.1) и (7.12.2) моменты имеют вид

$$\mu = \mathbf{E}\{F_{\text{uni}}\} = \mathbf{E}\{\log_2 A_i\} = 2^{-L} \sum_{i=1}^{\infty} (1 - 2^{-L})^{i-1} \log_2 i, \tag{7.12.3}$$

$$\mathbf{D}\{\log_2 A_i\} = 2^{-L} \sum_{i=1}^{\infty} (1 - 2^{-L})^{i-1} (\log_2 i)^2 - (\mathbf{E}\{F_{\text{uni}}\})^2. \tag{7.12.4}$$

Дисперсия статистики F_{uni} вычисляется по формуле

$$\sigma^2 = \mathbf{D}\{F_{\text{uni}}\} = c(L, K)^2 \frac{\mathbf{D}\{\log_2 A_i\}}{K}, \tag{7.12.5}$$

где для $L \ge 3$ $c(L, 2^L)$ близко к 0.8, и если $K \gg 2^L$, то $c(L, K)$ близко к 0.5, 0.6 и 0.65 соответственно для $L = 4$, $L = 8$ и $L = 12$. Для $K \ge 2^L$ достаточна точна аппроксимация [139]:

$$c(L, K) \approx 0.7 - \frac{0.8}{L} + \frac{(4 + 32/L)K^{-3/L}}{15}. \tag{7.12.6}$$

На практике L рекомендуется выбирать от 6 до 16, $Q \ge 10 \cdot 2^L$ и K «как можно больше» (например, $K = 1000 \cdot 2^L$).

Универсальный статистический тест Маурера, таким образом, имеет вид.

1. Выбираются параметры Q, K, L, и регистрируется последовательность $x_1, x_2, \ldots, x_{(Q+K)L} \in \mathcal{A}$.
2. На основании последовательности $\{x_i\}$ строится статистика F_{uni}
3. Принимается гипотеза

$$\begin{cases} H_0, & \text{если } t_1 < F_{\text{uni}} < t_2, \\ H_1 & \text{в противном случае,} \end{cases} \tag{7.12.7}$$

где $t_1 = \mu + \sigma\Phi^{-1}(\varepsilon/2)$ и $t_2 = \mu - \sigma\Phi^{-1}(\varepsilon/2)$; ε — уровень значимости критерия; $\Phi^{-1}(\varepsilon/2)$ — квантиль уровня $\varepsilon/2$ стандартного нормального распределения.

7.13. Тесты на основе приращений энтропии

Пинкус и Сингер [156] для оценки «степени» случайности последовательности ввели специальную характеристику, основанную на *приращениях энтропии*. Эта характеристика вводится в предположении, что «похожие» фрагменты остаются такими при последовательном увеличении количества членов во фрагментах при некоторых нарушениях гипотезы о РРСП.

Пусть $x_1, \ldots, x_n \in \mathcal{A} = \{0, 1, \ldots, N-1\}$ — наблюдаемая последовательность, $N \geq 2$. Определена следующая гипотеза H_0: $\{x_t\}$ — РРСП (равномерно распределенная случайная последовательность) и альтернатива $H_1 = \overline{H_0}$. Обозначим через

$$Y_i(m) = (x_i, x_{i+1}, \ldots, x_{i+m-1}) \in \mathcal{A}^m, \quad i = \overline{1, n-m+1},$$

слово из m элементов, начинающееся i-м элементом последовательности. Вычислим относительную частоту встречаемости фрагмента $Y_i(m)$ в последовательности

$$A_m^i = \frac{1}{n-m+1} \left| \{ j : 1 \leq j \leq n-m+1,\ Y_j(m) = Y_i(m) \} \right|. \tag{7.13.1}$$

Пусть

$$F^{(m)} = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} \log A_m^i \tag{7.13.2}$$

есть *выборочная энтропия* фрагментов длиной m. Определим *статистику приращения энтропии* для фрагментов длиной m и $m+1$:

$$g(0) = -F^{(1)}, \quad g(m) = F^{(m)} - F^{(m+1)}, \quad m = 1, 2, \ldots, n-1. \tag{7.13.3}$$

Заметим, что малым значениям $g(m)$ соответствует последовательность с «сильной регулярностью». Напротив, большие значения $g(m)$ соответствуют чисто случайным последовательностям.

Асимптотическое поведение статистики $g(m)$ при увеличении объема регистрируемых данных n определяется следующей теоремой [175].

Теорема 7.13.1. Для фиксированного m при $n \to \infty$ и выполнении гипотезы H_0 имеет место сходимость к χ^2-распределению:

$$\mathcal{L}\{2n(\log_2 N - g(m))\} \to \chi^2_f, \quad f = N^{m+1} - N^m.$$

С помощью этой теоремы построим тест, основанный на приращении энтропии при фиксированном m.

1. Регистрируется последовательность $x_1, x_2, \ldots, x_n \in \mathcal{A}$ и задается параметр m.
2. На основании последовательности $\{x_i\}$ вычисляется статистика приращения энтропии $g(m)$ согласно (7.13.1)—(7.13.3).
3. Вычисляется статистика $\chi^2 = 2n|\log N - g(m)|$.
4. Выносится решение с помощью решающего правила:

$$\text{принимается} \begin{cases} H_0, & \text{если } P > \varepsilon, \\ H_1, & \text{если } P \le \varepsilon; \end{cases} \tag{7.13.4}$$

$$P = 1 - G_f(\chi^2), \quad f = N^{m+1} - N^m,$$

где ε — уровень значимости критерия; $G_f(z)$ — функция χ^2-распределения с f степенями свободы.

Отметим, что тест (7.13.4) применим лишь тогда, когда m достаточно мало, так что $n/N^m \to \infty$. С целью снизить требуемую длину последовательности n рассмотрим «менее жесткую» асимптотику: $n/N^m \to \lambda$, $0 < \lambda < \infty$.

Для этого рассмотрим следующую функцию N аргументов:

$$f(u_1, \ldots, u_N) = -u_1 \log\frac{u_1}{u_1 + \ldots + u_N} - \ldots - u_N \log\frac{u_N}{u_1 + \ldots + u_N}.$$

Пусть $\Pi_1, \ldots, \Pi_N$ — N независимых случайных величин, распределенных по закону Пуассона с параметром λ. Вычислим моменты функции $f(\cdot)$:

$$\mu_n = \frac{1}{N\lambda}\mathbf{E}\{f(\Pi_1, \ldots, \Pi_N)\};$$

$$\nu = \frac{1}{N\lambda}\mathbf{Cov}\{f(\Pi_1, \ldots, \Pi_N), \Pi_1 + \ldots + \Pi_N\};$$

$$\sigma_n^2 = N^m(\mathbf{D}\{f(\Pi_1, \ldots, \Pi_N) - N\lambda\nu^2)\}.$$

Асимптотическое поведение статистики $g(m)$ при совместном увеличении параметра m и объема регистрируемых данных n определяется следующей теоремой [175].

Теорема 7.13.2. Если $n, m \to \infty$, $n/N^m \to \lambda > 0$ и верна гипотеза H_0, то

$$\mathbf{P}\left\{ n\frac{g(m) - \mu_n}{\sigma_n} < x \right\} \to \Phi(x),$$

где $\Phi(x)$ — функция распределения стандартного нормального закона.

На основе этой теоремы построен следующий тест.

1. Выбирается параметр m и регистрируется последовательность $x_1, x_2, \ldots, x_n \in \mathcal{A}$.
2. На основе последовательности $\{x_i\}$ строится статистика $g(m)$ согласно (7.13.3).
3. Выносится решение с помощью решающего правила:

$$\begin{cases} H_0, \text{ если } t_1 < g(m) < t_2, \\ H_1 \text{ в противном случае,} \end{cases} \tag{7.13.5}$$

где $t_1 = \mu_n + \sigma_n \Phi^{-1}(\varepsilon/2)/n$ и $t_2 = \mu_n - \sigma_n \Phi^{-1}(\varepsilon/2)/n$; ε — уровень значимости критерия; $\Phi^{-1}(\varepsilon/2)$ — квантиль уровня $\varepsilon/2$ стандартного нормального распределения.

7.14. Тест, основанный на алгоритме сжатия Лемпеля — Зива

Этот тест основан на следующей идее. Подлежащая тестированию двоичная ($N = 2$) последовательность $x_1, x_2, \ldots, x_n \in \mathcal{A}$ подвергается сжатию с помощью *алгоритма Лемпеля — Зива* [196]. Если результаты сжатия статистически значимо отличаются от теоретических результатов для РРСП, то тестируемая последовательность признается отличной от РРСП.

Алгоритм Лемпеля — Зива состоит из трех шагов.

Шаг 1. Разбиваем исходную последовательность $x_1, \ldots, x_n \in \{0, 1\}$ на двоичные слова (цепочки) так, чтобы каждое последующее слово было кратчайшим словом, не встречавшимся ранее.

Шаг 2. Нумеруем все полученные слова последовательно в двоичном алфавите.

Шаг 3. Каждое слово новой («сжатой») двоичной последовательности $y_1, y_2, \ldots, y_m$ состоит из префикса и суффикса: префикс — номер предыдущего слова, которое отличается от данного слова лишь последним символом (окончанием); суффикс — окончание.

В данном тесте в качестве статистики используется статистика W_n — число слов, построенных на шагах 1, 2 алгоритма сжатия. Например, если задана исходная двоичная последовательность $\{x_t\}$ длиной $n = 11$: 01100111010, то после шага 1 получаем $W_n = 6$ двоичных слов: 0; 1; 10; 01; 11; 010. При этом сжатая последовательность будет иметь длину $m = 21$:

0; 001 1; 010 0; 001 1; 010 1; 100 0.

Отметим, что при малых значениях n (как в данном числовом примере) длина m сжатой последовательности может оказаться больше n.

Тест основан на следующем асимптотическом свойстве статистики W_n. Если верна гипотеза H_0 и $n \to \infty$, то W_n распределена асимптотически нормально:

$$\mathcal{L}\left\{\frac{W_n - \mu_n}{\sigma_n}\right\} \to \mathcal{N}_1(0,1), \tag{7.14.1}$$

причем для математического ожидания μ_n и среднеквадратичного отклонения σ_n справедливы асимптотические приближения:

$$\mu_n \approx \frac{n}{\log_2 n}, \quad \sigma_n \approx \sqrt{\frac{0.266n}{(\log_2 n)^3}}. \tag{7.14.2}$$

Для фиксированных значений n величины μ_n и σ_n можно вычислить экспериментально методом Монте-Карло с помощью некоторого «истинно случайного» генератора. Например, генератора BBS (Blum — Blum — Shub) (см. п. 8.8). Таким способом для $n = 10^6$ было найдено

$$\mu_n \approx 69586.25, \quad \sigma_n \approx 70.448718. \tag{7.14.3}$$

Тест, основанный на статистике

$$z_n = \frac{W_n - \mu_n}{\sigma_n}, \tag{7.14.4}$$

имеет вид, аналогичный (7.10.10):

$$\text{принимается} \begin{cases} H_0, & \text{если } P > \varepsilon, \\ H_1, & \text{если } P \le \varepsilon; \end{cases} \tag{7.14.5}$$

$$P = 2(1 - \Phi(z_n)),$$

где $\Phi(\cdot)$ — функция распределения стандартного нормального закона; $\varepsilon \in (0,1)$ — заданный асимптотический уровень значимости.

7.15. Тест, основанный на линейной сложности

Линейной сложностью наблюдаемой двоичной последовательности $X_n = (x_1, x_2, \ldots, x_n)$ называется минимальный порядок

$$L_n = L(X_n) \in \{1, 2, \ldots, n\}$$

линейной рекурренты над $\mathbb{F}_2$:

$$x_t = (\theta_1 x_{t-1} + \theta_2 x_{t-2} + \ldots + \theta_{L_n} x_{t-L_n}) \bmod 2, \quad t = L_n + 1, L_n + 2, \ldots,$$

где $\theta_1, \ldots, \theta_{L_n} \in \mathbb{F}_2$ — коэффициенты линейной рекуррентной последовательности, порождающей наблюдаемую последовательность X_n длиной n. Для вычисления линейной сложности L_n может быть использован *алгоритм Берлекэм-*

па — Месси [143]. Оказывается, если верна гипотеза H_0, то математическое ожидание статистики L_n равно [175]:

$$\mu_n = \mathbf{E}_{H_0}\{L_n\} = \frac{n}{2} + \frac{9+(-1)^{n+1}}{36} - \frac{1}{2^n}\left(\frac{n}{3}+\frac{2}{9}\right). \qquad (7.15.1)$$

При этом для $n \to \infty$ дискретное распределение вероятностей статистики

$$T_n = (-1)^n(L_n - \mu_n) + \frac{2}{9}$$

сходится к следующему предельному распределению:

$$\mathbf{P}_{H_0}\{T_n = l\} \to \begin{cases} 2^{-(2|l|+1)}, & \text{если } l = 0, -1, -2, \ldots, \\ 2^{-2l}, & \text{если } l = 1, 2, \ldots \end{cases} \qquad (7.15.2)$$

Из-за быстрого убывания этих вероятностей при $|l| \to \infty$ удобно разбить область значений T_n на $K = 7$ ячеек:

$$\Gamma_1 = \{T_n \le -2.5\}, \quad \Gamma_2 = \{-2.5 < T_n \le -1.5\},$$
$$\Gamma_3 = \{-1.5 < T_n \le -0.5\}, \quad \Gamma_4 = \{-0.5 < T_n \le 0.5\}, \qquad (7.15.3)$$
$$\Gamma_5 = \{0.5 < T_n \le 1.5\}, \quad \Gamma_6 = \{1.5 < T_n \le 2.5\}, \quad \Gamma_7 = \{L_n > 2.5\}.$$

Согласно (7.15.2) и (7.15.3) находятся теоретические вероятности попадания T_n в эти ячейки:

$$\pi_1 = 0.01047, \quad \pi_2 = 0.03125, \quad \pi_3 = 0.12500, \quad \pi_4 = 0.50000,$$
$$\pi_5 = 0.2500, \quad \pi_6 = 0.06250, \quad \pi_7 = 0.02078.$$

Алгоритм тестирования при этом имеет следующий вид.

1. Исходная n-битовая последовательность X_n разбивается на M блоков по N битов ($500 < N < 5000$):

$$X_n = X_N^{(1)} \,||\, X_N^{(2)} \,||\, \ldots \,||\, X_N^{(M)}.$$

2. С помощью алгоритма Берлекэмпа — Месси для i-го блока ($i = \overline{1, M}$) вычисляется линейная сложность $L_N^{(i)} = L(X_N^{(i)})$.

3. Согласно (7.15.1) находится теоретическое математическое ожидание $\mu_N = \mathbf{E}_{H_0}\{L_N^{(i)}\}$.

4. Для каждого $i = \overline{1, M}$ вычисляется статистика

$$T_i = (-1)^N(L_N^{(i)} - \mu_N) + \frac{2}{9}.$$

5. Вычисляются частоты попадания $\{T_i\}$ в ячейки $\{\Gamma_k\}$:

$$\nu_k = \sum_{i=1}^{M} \mathbf{1}_{\Gamma_k}(T_i), \quad k = \overline{1, 7},$$

где $\mathbf{1}_\Gamma(\cdot)$ — индикаторная функция множества Γ ($\nu_1 + \ldots + \nu_7 \equiv M$).

6. Вычисляется χ^2-статистика:

$$\chi^2 = \sum_{k=1}^{7} \frac{(\nu_k - M\pi_k)^2}{M\pi_k} \geq 0.$$

7. Выносится решение по правилу:

$$\text{принимается} \begin{cases} H_0, \text{ если } P > \varepsilon, \\ H_1, \text{ если } P \leq \varepsilon; \end{cases}$$

$$P = 1 - G_6(\chi^2),$$

где $G_f(\cdot)$ — функция χ^2-распределения с f степенями свободы.

7.16. Тест на основе экстремальной статистики скалярного произведения

Этот тест предложен в работе [37]. Пусть $x_1, x_2, \ldots, x_n \in \mathcal{A} = \{0, 1\}$ — наблюдаемая двоичная последовательность. Определены гипотеза H_0: $\{x_i\}$ — независимые в совокупности одинаково распределенные случайные величины Бернулли с равномерным распределением $\mathsf{P}\{x_i = 0\} = \mathsf{P}\{x_i = 1\} = \frac{1}{2}$ и альтернатива $H_1 = \overline{H}_0$.

Для заданного $N \in \mathbb{N}$ разобъем последовательность $x_1, x_2, \ldots, x_n$ длиной $n = mN$ на m последовательных непересекающихся фрагментов длиной N: $X_1, X_2, \ldots, X_m \in \mathcal{A}^N$, где

$$X_i = \begin{pmatrix} x_{(i-1)N+1} \\ x_{(i-1)N+2} \\ \vdots \\ x_{iN} \end{pmatrix} \in \mathcal{A}^N, \quad i = \overline{1,m}. \tag{7.16.1}$$

Определим на $X_1, X_2, \ldots, X_m$ статистики

$$Y_j = X_1^T X_j = \sum_{k=1}^{N} x_k x_{(j-1)N+k}, \quad j = \overline{2,m}. \tag{7.16.2}$$

Статистика Y_j характеризует «степень похожести» первого и j-го фрагментов. Статистики $Y_2, \ldots, Y_m$ при выполнении гипотезы H_0 имеют следующее распределение вероятностей.

Лемма 7.16.1. Если верна гипотеза H_0, то распределение статистики Y_j имеет вид

$$\mathsf{P}_{H_0}\{Y_j = y\} = 2^{-2N} 3^{N-y} \binom{N}{y},$$

где $y \in \{0, 1, \ldots, N\}$, $j = \overline{2, m}$.

Лемма 7.16.2. Если верна гипотеза H_0, то

$$\mathbf{E}_{H_0}\{Y_j\} = \frac{N}{4}, \quad \mathbf{D}_{H_0}\{Y_j\} = \frac{3N}{16}, \quad \mathbf{Corr}_{H_0}\{Y_i, Y_j\} = \frac{1}{3},$$

для $i, j = \overline{2,m}$, $i \neq j$.

Лемма 7.16.3. Совместное распределение статистик $Y_2, \ldots, Y_m$ при H_0 имеет вид

$$p_m(y_2, \ldots, y_m; N) = 2^{-N} \sum_{\substack{l = \max\{y_j\} \\ 2 \le j \le m}}^{N} 2^{-(m-1)l} \binom{N}{l} \prod_{j=2}^{m} \binom{l}{y_j},$$

$$0 \le y_2, y_3, \ldots, y_m \le N.$$

На статистиках $Y_2, \ldots, Y_m$ построим *статистику максимального скалярного произведения*:

$$Y_{\max} = \max\{Y_2, \ldots, Y_m\}. \tag{7.16.3}$$

Теорема 7.16.1. Если верна гипотеза H_0, то статистика $Y_{\max}$ имеет следующее распределение:

$$q_0 = \mathbf{P}\{Y_{\max} = 0\} = 2^{-N}(1 + 2^{1-m})^N,$$

$$q_k = \mathbf{P}\{Y_{\max} = k\} = 2^{-N} \sum_{l=0}^{N} \binom{N}{l} \left(\sum_{y=0}^{k} 2^{-l} \binom{l}{y} \right)^{m-1} - 2^{-N} \sum_{l=0}^{N} \binom{N}{l} \left(\sum_{y=0}^{k-1} 2^{-l} \binom{l}{y} \right)^{m-1}, \; k > 0.$$

Следствие 7.16.1. При $m \to \infty$, если верна гипотеза H_0, то асимптотическое распределение статистики $Y_{\max}$ имеет вид

$$\lim_{m \to \infty} q_k = 2^{-N} \binom{N}{k}, \quad k = \overline{0,N}.$$

Замечание 7.16.1. Вместо (7.16.2) для расчета «степени похожести» можно применять следующие формулы:

$$Y_j = \sum_{k=1}^{N} (1 - x_k) x_{(j-1)N+k}, \quad j = \overline{2,m},$$

$$Y_j = \sum_{k=1}^{N} (1 - x_k)(1 - x_{(j-1)N+k}), \quad j = \overline{2,m},$$

$$Y_j = \sum_{k=1}^{N} x_k (1 - x_{(j-1)N+k}), \quad j = \overline{2,m}.$$

При этом распределение статистик $\{Y_j\}$ при выполнении H_0 не изменится.

При помощи теоремы 7.16.1 построим *тест максимального скалярного произведения* на основе χ^2-статистики.

1. Регистрируется последовательность $x_1, x_2, \ldots, x_T \in \mathcal{A}$ длиной $T = LmN = Ln$, где $m, N \in \mathbb{N}$, $m \geq 2$, задаются пользователем, а $L \in \mathbb{N}$ выбирается так, чтобы обеспечить условие применимости χ^2-критерия.

2. Последовательность длиной T разбивается на L подпоследовательностей длиной $n = mN$, а каждая из этих подпоследовательностей разбивается на m фрагментов длиной N согласно (7.16.1):

$$X_i^{(l)} = (x_{(l-1)n+(i-1)N+1}, \ldots, x_{(l-1)n+iN})^T \in \mathcal{A}, \quad i = \overline{1,m}, \quad l = \overline{1,L}.$$

По l-й ($l = \overline{1,L}$) подпоследовательности $X^{(l)}$, согласно (7.16.1), вычисляется $(m-1)$ статистик:

$$Y_j^{(l)} = X_1^{(l)T} X_j^{(l)} = \sum_{k=1}^{N} x_{(l-1)n+k}\, x_{(l-1)n+(j-1)N+k}, \quad j = \overline{2,m}.$$

Для l-й подпоследовательности по статистикам $Y_2^{(l)}, \ldots, Y_m^{(l)}$ вычисляется экстремальная статистика $Y_{\max}^{(l)}$.

3. По выборке экстремальных статистик $\{Y_{\max}^{(l)} : l = \overline{1,L}\}$ вычисляется χ^2-статистика:

$$\chi^2 = \sum_{k=0}^{N} \frac{(f_k - Lq_k)^2}{Lq_k}, \quad f_k = \sum_{l=1}^{L} \mathrm{I}\{Y_{\max}^{(l)} = k\}.$$

4. Выносим решение с помощью статистического правила:

$$\text{принимается гипотеза} \begin{cases} H_0, & \text{если } P > \varepsilon, \\ H_1, & \text{если } P \leq \varepsilon, \end{cases} \tag{7.16.4}$$

где $P = 1 - G_N(\chi^2)$ — P-значение; ε — уровень значимости.

Тест максимального скалярного произведения на основе распределения Колмогорова получается из предыдущего теста повторением шагов 1—3 с последующим применением теста Колмогорова к полученной выборке статистики χ^2. Приведем описание этого теста.

1. Регистрируется последовательность $x_1, x_2, \ldots, x_T \in \mathcal{A}$ длиной $T = LSmN = LSn$, где $m, N \in \mathbb{N}$, $m \geq 2$, задаются пользователем, а $L \in \mathbb{N}$ и $S \in \mathbb{N}$ выбираются так, чтобы обеспечить выполнение соответствующих критериев.

2. Последовательность длиной T разбивается на S подпоследовательностей длиной Ln, которые в свою очередь разбиваются на L подпоследовательностей длиной $n = mN$, а каждая из этих подпоследовательностей разбивается на m фрагментов длиной N согласно (7.16.1):

$$X_i^{(l)(s)} = (x_{(s-1)Ln+(l-1)n+(i-1)N+1}, \ldots, x_{(s-1)Ln+(l-1)n+iN})^T,$$

где $i = \overline{1,m}$, $l = \overline{1,L}$, $s = \overline{1,S}$.

По l-й ($l=\overline{1,L}$) подпоследовательности $X^{(l)(s)}$, согласно (7.16.1), вычисляется $(m-1)$ статистик:

$$Y_i^{(l)(s)} = X_1^{(l)(s)T} X_j^{(l)(s)} = \sum_{k=1}^{N} x_{(s-1)Ln+(l-1)n+k}\, x_{(s-1)Ln+(l-1)n+(j-1)N+k}, \quad j=\overline{2,m}.$$

Для l-й подпоследовательности по статистикам $Y_2^{(l)(s)}, \ldots, Y_m^{(l)(s)}$ вычисляется экстремальная статистика $Y_{\max}^{(l)(s)}$.

3. По выборке экстремальной статистики $\{Y_{\max}^{(l)(s)} : l=\overline{1,L}\}$ вычисляется χ^2-статистика:

$$\chi_s^2 = \sum_{k=0}^{N} \frac{(f_k^{(s)} - Lq_k)^2}{Lq_k}, \quad f_k^{(s)} = \sum_{l=1}^{L} \delta_{Y_{\max}^{(l)(s)},k}, \quad s=\overline{1,S}.$$

4. Над выборкой статистики $\{\chi_s^2 : s=\overline{1,S}\}$ проведем следующее преобразование: $t_s = G_N(\chi_s^2)$, где $G_N(x)$ — функция χ^2-распределения с N степенями свободы. Если выполняется гипотеза H_0, то статистика t_s асимптотически равномерно распределена на отрезке [0, 1]. Пусть $t_{(1)}, \ldots, t_{(S)}$ — значения $t_1, \ldots, t_S$, упорядоченные по возрастанию. Статистика

$$D_S = \sqrt{S} \max_{1\le j\le S} \left| \frac{j}{S} - t_{(j)} \right|$$

имеет распределение Колмогорова.

5. Выносим решение с помощью статистического правила:

$$\text{принимается гипотеза} \begin{cases} H_0, & D_S < D(1-\varepsilon), \\ H_1, & D_S \ge D(1-\varepsilon), \end{cases} \tag{7.16.5}$$

где ε — уровень значимости критерия; $D(1-\varepsilon)$ — квантиль уровня $(1-\varepsilon)$ распределения Колмогорова.

7.17. Тест на основе экстремальной статистики дельта-произведения

Этот тест описан в работе [37]. Пусть $x_1, x_2, \ldots, x_n \in \mathcal{A} = \{0,1\}$ — наблюдаемая двоичная последовательность. Определены гипотеза H_0: $\{x_i\}$ — независимые в совокупности одинаково распределенные случайные величины Бернулли с равномерным распределением $\mathbf{P}\{x_i=0\} = \mathbf{P}\{x_i=1\} = \frac{1}{2}$, и альтернатива $H_1 = \overline{H_0}$.

Для заданного $N \in \mathbb{N}$ разобьем последовательность $x_1, x_2, \ldots, x_n$ длиной $n = mN$ на m последовательных непересекающихся фрагментов длиной N: $X_1, X_2, \ldots, X_m \in \mathcal{A}^m$, где

$$X_i = \begin{pmatrix} x_{(i-1)N+1} \\ x_{(i-1)N+2} \\ \vdots \\ x_{iN} \end{pmatrix} \in \mathcal{A}^N, \quad i = \overline{1,m}. \tag{7.17.1}$$

Определим на $X_1, X_2, \ldots, X_m$ статистики

$$Y_j^* = \sum_{k=1}^{N} \mathrm{I}\{x_k = x_{(j-1)N+k}\}, \quad j = \overline{2,m}. \tag{7.17.2}$$

Статистика Y_j^* характеризует «степень похожести» первого и j-го фрагментов. Статистики $Y_2^*, \ldots, Y_m^*$ при выполнении гипотезы H_0 имеют следующее распределение.

Лемма 7.17.1. Если верна гипотеза H_0, то статистики $\{Y_j^*\}$ распределены по биномиальному закону распределения вероятностей:

$$\mathbf{P}_{H_0}\{Y_j^* = y\} = 2^{-N}\binom{N}{y}, \quad 0 \le y \le N, \quad j = \overline{2,m}.$$

Теорема 7.17.1. Совместное распределение $Y_2^*, \ldots, Y_m^*$ при H_0 имеет вид

$$p_m^*(y_2, \ldots, y_m; N) = \prod_{j=2}^{m} 2^{-N}\binom{N}{y_j}, \quad 0 \le y_2, \ldots, y_m \le N.$$

Следствие 7.17.1. Если верна гипотеза H_0, то статистики $Y_2^*, \ldots, Y_m^*$ независимы в совокупности.

На статистиках $Y_2^*, \ldots, Y_m^*$ построим *статистику максимального дельта-произведения*:

$$Y_{\max}^* = \max\{Y_2^*, \ldots, Y_m^*\}.$$

Следствие 7.17.2. Функция распределения статистики $Y_{\max}^*$ при выполнении гипотезы H_0 имеет вид

$$F_{Y_{\max}^*}(k) = \left(\sum_{y=0}^{k} 2^{-N}\binom{N}{y}\right)^{m-1}, \quad k = \overline{0,N}.$$

Следствие 7.17.3. При $m \to \infty$, если верна гипотеза H_0, то асимптотическое распределение статистики $Y_{\max}^*$ имеет вид

$$\lim_{m\to\infty} q_k^* = \begin{cases} 0, & k < N, \\ 1, & k = N. \end{cases}$$

Замечание 7.17.1. Вместо формулы (7.17.2) можно применять следующую:

$$Y_j^* = \sum_{k=1}^{N} \mathrm{I}\{x_k \neq x_{(j-1)N+k}\}, \quad j = \overline{2,m}.$$

При этом распределение статистик $\{Y_j^*\}$ не изменится.

Пользуясь теоремой 7.17.1, построим *тест максимального дельта-произведения* на основе χ^2-статистики.

1. Регистрируется последовательность $x_1, x_2, \ldots, x_T \in \mathcal{A}^T, \mathcal{A} = \{0,1\}$, длиной $T = LmN = Ln$, где $m, N \in \mathbb{N}$, $m \geq 2$, задаются пользователем, а $L \in \mathbb{N}$ выбирается так, чтобы обеспечить выполнение условий применимости χ^2-критерия.

2. Последовательность длиной T разбивается на L подпоследовательностей длиной $n = mN$, а каждая из этих подпоследовательностей на m фрагментов длиной N согласно (7.16.1):

$$X_i^{(l)} = (x_{(l-1)n+(i-1)N+2}, \ldots, x_{(l-1)n+iN})^T \in \mathcal{A}^N, \quad i = \overline{1,m}, \quad l = \overline{1,L}.$$

По l-й ($l = \overline{1,L}$) подпоследовательности $X^{(l)}$, согласно (7.16.1), вычисляется $(m-1)$ статистик:

$$Y_j^{*(l)} = \sum_{k=1}^{N} \mathrm{I}\{x_{(l-1)n+k} = x_{(l-1)n+(j-1)N+k}\}, \quad j = \overline{2,m}.$$

Для l-й подпоследовательности по статистикам $Y_2^{*(l)}, \ldots, Y_m^{*(l)}$ вычисляется экстремальная статистика $Y_{\max}^{*(l)}$.

3. По выборке экстремальной статистики $\{Y_{\max}^{*(l)} : l = \overline{1,L}\}$ вычисляется χ^2-статистика:

$$\chi^2 = \sum_{k=0}^{N} \frac{(f_k - Lq_k)^2}{Lq_k}, \quad f_k = \sum_{l=1}^{L} \mathrm{I}\{Y_{\max}^{*(l)} = k\}.$$

4. Выносим решение с помощью статистического правила (7.16.4).

Тест максимального дельта-произведения на основе распределения Колмогорова получается из предыдущего теста повторением шагов 1—3 с последующим применением теста Колмогорова к полученной выборке статистики χ^2 аналогично тому, как это сделано в п. 7.16.

7.18. Алгоритмическое определение случайности

Все статистические тесты РРСП, представленные в предыдущих параграфах, базируются на вероятностной модели случайности [15]. Однако в последние годы интенсивно развивается алгоритмическая концепция случайности, берущая начало от работ А.Н. Колмогорова и В.А. Успенского [15], П. Мар-

тин-Лёфа, Г.Дж. Чайтина, К.П. Шнорра. Как отмечается в [15], «хотя алгоритмы и случайность могут показаться диаметрально противоположными, но это только на первый взгляд. Рассматривая наши представления о случайности и об алгоритмах более пристально, можно наблюсти связи между ними... Алгоритмы могут быть использованы для определения случайности, а случайность — для построения нового класса алгоритмов...» Имея в виду потенциальную возможность построения нового семейства тестов, базирующихся на алгоритмическом определении случайности, изложим основные понятия, связанные с этим определением.

Примем обозначения: $V = \{0, 1\}$ — двоичный алфавит; Ξ — множество всех двоичных цепочек (слов, strings) $x = (x_1, \ldots, x_n)$ произвольной конечной длины $n = |x| \in \mathbb{N}$; $x \in V^{|x|}$; $\Omega = V^{\infty}$ — множество всех бесконечных двоичных последовательностей; $\mu(\omega)$, $\omega \in \Omega$ — равномерная бернуллиева мера на Ω, $\mu(\Omega) = 1$; $R \in \Omega$ — подлежащее построению подмножество «случайных в некотором разумном неформальном смысле» последовательностей.

Случайные последовательности обладают тремя фундаментальными свойствами: типичность, хаотичность, устойчивость частот, каждое из которых может быть взято за основу некоторого определения случайности.

Класс $T \in \Omega$ «типических последовательностей» введен Мартин-Лёфом, так что $T \in R$. Свойство «типичности означает принадлежность последовательности к каждому разумному их большинству». На первый взгляд кажется, что в качестве T можно взять пересечение всех подмножеств множества Ω, имеющих меру 1. Однако, как показал Мартин-Лёф, такое пересечение есть пустое множество $\varnothing$.

Чтобы построить класс T, понадобится ряд вспомогательных понятий. Каждой конечной цепочке $X \in \Xi$ длиной $n = |x| < +\infty$ поставим в соответствие шар

$$\Gamma_x = \{(x \,||\, x_{n+1}, x_{n+2}, \ldots) \in \Omega : x_{n+1}, x_{n+2}, \ldots \in V\} \subset \Omega,$$

состоящий из всевозможных бесконечных продолжений цепочки x, причем

$$\mu(\Gamma_x) = \frac{1}{2^n}.$$

Известно, что $\mu(M) = 0$, т.е. подмножество $M \subset \Omega$ является *пренебрежимым* тогда и только тогда, когда для любого рационального $\varepsilon > 0$ существует такая последовательность конечных цепочек $X_k \in \Xi$, $k \in \mathbb{N}$, что соответствующая последовательность шаров удовлетворяет двум условиям:

1) $M \subset \cup_k \Gamma_{X_k}$;
2) $\Sigma_k \mu(\Gamma_{X_k}) < \varepsilon$.

Если к тому же последовательность цепочек X_k, $k \in \mathbb{N}$ является вычислимой, т.е. существует алгоритм для вычисления X_k по заданному k, реализуемый конечной программой $\langle X_k \rangle$, то принято говорить, что *множество M имеет эффективно меру ноль*:

$$\mu(M) \underset{eff}{=} 0,$$

или *является эффективно пренебрежимым*. При этом множество $L \subset \Omega$ имеет *эффективно меру единица*:

$$\mu(L) \underset{eff}{=} 1,$$

если для его дополнения имеем $\mu(\Omega \setminus L) \underset{eff}{=} 0$.

Теорема 7.18.1 (Мартин-Лёф). Пересечение T всех множеств эффективной меры 1 не только не пусто, но и само имеет эффективную меру 1.

Указанное в этой теореме пересечение T является наименьшим множеством эффективной меры 1 и называется *конструктивным носителем* (constructive support) меры. Это множество и принимается в качестве класса всех типических последовательностей, $\mu(\mathrm{T}) \underset{eff}{=} 1$.

Последовательность $X \in \mathrm{T}$, принадлежащая конструктивному носителю меры, называется *случайной по Мартин-Лёфу*.

Заметим, что многие предельные теоремы теории вероятностей (например, закон повторного логарифма), в которых указывается сходимость «почти наверное», на самом деле верны для класса T, т.е. по эффективной мере 1.

Класс $\mathrm{И} \subset \Omega$ *«хаотических последовательностей»* введен А.Н. Колмогоровым ($\mathrm{R} \subset \mathrm{C}$): «Всякая случайная последовательность — хаотическая в том смысле, что для нее отсутствует какой-либо простой закон, управляющий чередованием ее членов». Таким образом, хаотичность означает «сложность строения» последовательности.

Пусть $k = K(x) : \Xi \to \mathbb{N}$ — мера сложности конечной двоичной цепочки x. В качестве меры сложности $K(\cdot)$ целесообразно, оказывается, использовать монотонную энтропию $L(\cdot)$, предложенную А.К. Звонкиным и К. Шнорром. Алгоритм вычисления монотонной энтропии — довольно сложный и использует язык программирования для одноголовочной машины Тьюринга.

Последовательность $(x_1, x_2, \ldots, x_n) \in \Omega$ принадлежит *классу* $\mathrm{C} \subset \Omega$ *хаотических последовательностей* тогда и только тогда, когда для любого $n \in \mathbb{N}$

$$K(x_1, x_2, \ldots, x_n) \geq n - c,$$

где $c \geq 0$ — некоторая константа, не зависящая от n, т.е. когда энтропия растет достаточно быстро при $n \to \infty$. Последовательность $X \in \mathrm{C}$ при этом называется *случайной по Колмогорову*.

Теорема 7.18.2 (Левин — Шнорр). Классы C и T совпадают: $\mathrm{T} = \mathrm{C}$.

Классом случайных последовательностей в связи с этим называют $\mathrm{R} = \mathrm{T} = \mathrm{C}$.

Класс $\mathrm{S} \subset \Omega$ *«стохастических последовательностей»* содержит такие последовательности $X \in \Omega$, для которых выполняется *свойство устойчивости частот*:

$$\frac{1}{s}\sum_{t=1}^{s} x_t \underset{s\to\infty}{\to} \frac{1}{2};$$

причем это свойство сохраняется и для подпоследовательности $(x_{\gamma_1}, x_{\gamma_2}, \ldots) \in \Omega$, при «любом допустимом» выборе индексов $\gamma_1, \gamma_2, \ldots$ Существует два понятия допустимости индексных последовательностей $\gamma_1, \gamma_2, \ldots$

Класс стохастических по Чёрчу последовательностей $\mathrm{CS} \subset \Omega$ допускает следующий способ построения индексных последовательностей:

$$\gamma_i = f(\gamma_1, \ldots, \gamma_{i-1}, x_1, \ldots, x_i), \quad \gamma_1 < \gamma_2 < \gamma_3 < \ldots$$

Класс стохастических по Колмогорову последовательностей $\mathrm{KS} \subset \Omega$ отличается тем, что условие упорядоченности индексов $\{\gamma_i\}$ снимается. При этом

$$\mathrm{R} \subset \mathrm{KS} \subset \mathrm{CS}, \quad \mathrm{KS} \neq \mathrm{CS}.$$

Открытой пока остается проблема: $\mathrm{KS} \stackrel{?}{=} \mathrm{R}$.

В заключение кратко остановимся на алгоритмическом определении случайности для конечных цепочек. Как отмечает А.Н. Колмогоров [15], «кажется естественным назвать цепочку $x \in \Xi$ случайной, если она не может быть записана в более сжатом виде, т.е. если кратчайшая программа ее порождения $\langle x \rangle$ столь же длинна, как она сама. Правильный вопрос состоит не в том, является ли данная цепочка случайной, а в том, насколько она случайна». Для этого вводится понятие «дефект случайности».

Пусть $M \subset \Xi$ — произвольное фиксированное множество конечных двоичных цепочек и $x \in M$. Тогда дефект случайности элемента x относительно M есть

$$d = d(x \mid M) = \log_2 |M| - H\{x \mid M\},$$

где $H\{x \mid M\} \geq 0$ — условная сложность описания элемента x. Иначе говоря, d — разность в длине между двумя описаниями элемента x средствами множества M: между стандартным (указать номер элемента в алфавите M) и кратчайшим описаниями. При этом элемент $x \in M$ называется *δ-случайным* (*δ-стохастическим*) для некоторого фиксированного $\delta > 0$, если

$$d(x \mid M) \leq \delta.$$

7.19. Задания

1. (п. 7.2). Для $N = 2$; $t_k(x_1, \ldots, x_n) = x_k$, $k = \overline{1,n}$; $m = n$; $L = 2^n$, $U = \mathcal{A}^n$; $\varepsilon = 0.01$ реализовать на компьютере тест (7.2.4)—(7.2.6) n-мерной дискретной равномерности двоичной последовательности $\{x_t\}$. Методом статистического моделирования на компьютере для заданных значений n оценить вероятности ошибок первого и второго рода. Оценить сложность алгоритма тестирования аналитически и проверить согласованность полученной оценки с реальным быстродействием алгоритма. В качестве тестируемых последовательностей использовать `File00`—`File06` из Архива дискретных последовательностей (АДП) (прил. Б).

2. (п. 7.2). Реализовать на компьютере тест (7.2.4)—(7.2.6) при $N=2$; $N=8$, 16, 32, 64; $m=8$; $U=\mathcal{A}^8$;

$$t_1(x_1, \ldots, x_n) = x_1 \oplus x_2 \oplus \ldots \oplus x_n,$$

$$t_2 = (x_1, \ldots, x_n) = \sum_{1 \le i_1 < i_2 \le n} x_{i_1} x_{i_2},$$

$$t_8 = (x_1, \ldots, x_n) = \sum_{1 \le i_1 < i_2 < \ldots < i_8 \le n} x_{i_1} x_{i_2} \cdots x_{i_8},$$

где суммирование выполняется по mod 2.
Методом статистического моделирования на компьютере оценить вероятности ошибок первого и второго рода. Оценить сложность алгоритма тестирования аналитически и проверить согласованность полученной оценки с реальным быстродействием алгоритма. Для компьютерных экспериментов использовать те же последовательности, что и в задании 1.

3. (п. 7.2). В задании 2 исследовать влияние выбора m и функций $\{t_k(\cdot)\}$ на мощность теста и быстродействие алгоритма тестирования.

4. (п. 7.2). Аналогично заданию 2 рассмотреть еще один универсальный способ задания статистик $\{t_k(\cdot)\}$:

$$t_k(x_1, \ldots, x_n) = \mathbf{1}_{\Gamma_{k-1}}(x_1, \ldots, x_n), \quad k = \overline{1,m},$$

где $\mathbf{1}_\Gamma(\cdot)$ — индикаторная функция множества $\Gamma \subseteq \mathcal{A}^n$; $\Gamma_0, \ldots, \Gamma_{m-1} \subset \mathcal{A}^n$ — разбиение пространства $\mathcal{A}^n$ на m «ячеек»:

$$\mathcal{A}^n = \bigcup_{k=0}^{m-1} \Gamma_k, \quad \Gamma_k \cap \Gamma_l = \varnothing, \quad l \ne k.$$

Пример такого разбиения при $m=8$, $n=16$:

$$\Gamma_{i_1+2i_2+4i_3} = \{(i_1, i_2, i_3, j_1, \ldots, j_{13}) \colon j_1, \ldots, j_{13} \in \mathcal{A}\}, \quad i_1, i_2, i_3 \in \mathcal{A}.$$

При таком разбиении все ячейки имеют одинаковую мощность:

$$|\Gamma_0| = |\Gamma_1| = \ldots = |\Gamma_7| = 2^{13}.$$

5. (п. 7.2). Выполнить задание 1 для $N > 2$, $m = n$,

$$t_k(x_1, \ldots, x_n) = x_k, \quad k = \overline{1,n},$$

$U = \mathcal{A}^n$, $L = N^n$. Для компьютерных экспериментов использовать последовательности `File07`, `File08` из АДП при $N=8$.

6. (п. 7.2). Разработать модифицированный алгоритм тестирования n-мерной равномерности на основе подхода, указанного в замечании 7.2.3. Выпол-

нить с модифицированным алгоритмом эксперименты, указанные в задании 1. Провести сравнительный анализ модифицированного и базового алгоритма (7.2.4)—(7.2.6) по быстродействию и мощности.

7. (п. 7.3). Выполнить задания 1, 5.

8. (п. 7.2). Оптимизировать по быстродействию программу, реализующую тест n-серий.

9. (п. 7.4). Выполнить задания 1, 5 при заданных значениях α, β.

10. (п. 7.4). Модифицировать тест интервалов так, чтобы обеспечить выполнение условия, указанного в замечании 7.2.2. Для этого вместо статистики (7.4.1) рассмотреть «усеченную статистику»:

$$\tilde{t}(x_1, x_2, \ldots) = \begin{cases} t(x_1, x_2, \ldots), & \text{если } t(x_1, x_2, \ldots) < t_+, \\ t_+, & \text{если } t(x_1, x_2, \ldots) \geq t_+, \end{cases} \tag{7.19.1}$$

где $t_+ \in \mathbb{N}$ — заданная верхняя граница такая, что

$$Mp(1-p)^{t_+-1} > 5.$$

11. (п. 7.4). Исследовать вопрос выбора параметров α, β с целью увеличить мощность теста интервалов в рамках задания 1 для класса альтернатив H_{1n}, определяемых однородной цепью Маркова (прил. Б).

12. (п. 7.4). Выполнить задание 11 для класса альтернатив H_{1n}, определяемых двоичной авторегрессионной последовательностью BAR(p) (прил. Б).

13. (п. 7.4). Методом статистического моделирования модифицированного теста, основанного на статистике (7.19.1), исследовать зависимость мощности теста от «параметра усечения» t_+ на основе данных задания 11.

14. (п. 7.5). Выполнить задания 1, 5 с помощью классического покер-теста. Исследовать влияние n на вычислительную сложность и мощность теста.

15. (п. 7.5). Выполнить задания 1, 5 с помощью обобщенного покер-теста. Исследовать влияние n на вычислительную сложность и мощность теста.

16. (п. 7.5). Осуществить модификацию классического покер-теста и обобщенного покер-теста для двоичных последовательностей ($N = 2$), переходя от битов к фрагментам последовательности — полубайтам $N' = 16$, байтам ($N'' = 256$), n-кам ($N''' = 2^n$). Для проверки n-мерной равномерности третий вариант более предпочтителен.

17. (п. 7.5). Провести сравнительный анализ классического и обобщенного покер-теста по вычислительной сложности и мощности.

18. (п. 7.6). Выполнить задания 1, 5 с помощью теста «собирателя купонов». Аналитически и методом компьютерного моделирования исследовать влияние «параметра укрупнения» r на мощность и вычислительную сложность теста.

19. (п. 7.6). Аналитически и методом компьютерного моделирования оценить математическое ожидание и дисперсию статистики (7.6.1) при истинности гипотезы H_{0n}. Использовать эти результаты для уточнения оценки вычислительной сложности теста.

20. (п. 7.6). Разработать обобщение теста «собирателя купонов» (в целях снижения его вычислительной сложности), в котором осуществляется «сбор неполного комплекта из $N'_0 \le N'$ купонов».

21. (п. 7.6). Выполнить задания 1, 5 с помощью обобщенного теста «собирателя купонов». Исследовать влияние параметра N'_0 на мощность теста и вычислительную сложность реализующего его алгоритма.

22. (п. 7.7). Выполнить задание 1 с помощью теста перестановок.

23. (п. 7.7). В условиях задания 1 аналитически и методом компьютерного моделирования исследовать влияние параметра n' на мощность и вычислительную сложность теста.

24. (п. 7.7). Выполнить задание 22 для обобщенного теста перестановок, учитывающего раздельно комбинации с совпадающими компонентами. Сравнить этот тест по мощности и вычислительной сложности с тестом, основанным на статистике (7.7.1).

25. (п. 7.8). Выполнить задание 1 с помощью теста пересекающихся n-грамм.

26. (п. 7.8). Сравнить тест пересекающихся n-грамм и ранее рассмотренный (задание 1) тест непересекающихся n-грамм при одних и тех же значениях n, n_+, ε по мощности и вычислительной сложности.

27. (п. 7.8). Выполнить задание 5 с помощью теста пересекающихся n-грамм.

28. (п. 7.8). Разработать и исследовать (методом компьютерного моделирования) модификацию теста пересекающихся n-грамм, заключающуюся в том, что в (7.8.1) вместо частот $(n-1)$-грамм используются частоты $(n-2)$-грамм.

29. (п. 7.9). Реализовать на компьютере алгоритм вычисления ранга двоичной матрицы.

30. (п. 7.9). Выполнить задание 1 с помощью теста, основанного на рангах двоичных матриц, для заданных значений k и l в (7.9.1).

31. (п. 7.9). Методом компьютерного моделирования для последовательностей **File01–File06** из АДП (прил. Б) исследовать зависимость мощности теста и его вычислительной сложности от количества строк k и столбцов l в (7.9.1) при фиксированном n. Дать рекомендации по использованию произвола при выборе k, l для достижения максимально возможной мощности теста.

32. (п. 7.9). Модифицировать данный тест на основе использования иных, чем (7.9.3), функций от двоичной матрицы X.

33. (п. 7.10). Выполнить задание 1 с помощью алгоритма обнаружения периодичности и оценивания периода (для $K = 1$; 2; 5; 10). Методом компьютерного моделирования для последовательностей **File01–File06** из АДП исследовать зависимость мощности теста и его вычислительной сложности от основных параметров алгоритма.

34. (п. 7.10). Выполнить задание 33 с помощью семейства спектральных тестов (7.10.5)—(7.10.10). Исследовать вопрос оптимального выбора параметра c из условия максимума мощности теста.

35. (п. 7.10). Выполнить задание 33 с помощью спектрального теста (7.10.11)—(7.10.13).

36. (п. 7.11). Выполнить задание 1 с помощью теста случайного блуждания, основанного на статистике «количество возвратов» (7.11.1). Методом компьютерного моделирования для последовательностей **File01–File06** из АДП исследовать зависимость мощности теста и его вычислительной сложности от основных параметров.

37. (п. 7.11). Выполнить задание 36 для теста случайного блуждания, основанного на статистике «количество изменений знаков» (7.11.2).

38. (п. 7.11). Выполнить задание 36 для теста случайного блуждания, основанного на статистике «вес Хэмминга» (7.11.3).

39. (п. 7.11). Сравнить рассмотренные тесты случайного блуждания по мощности и вычислительной сложности.

40. (п. 7.11). Аналогично (7.11.1)—(7.11.3) предложить статистику и построить новый тест случайного блуждания.

41. (п. 7.12). Выполнить задание 1 с помощью универсального статистического теста Маурера (7.12.1)—(7.12.7) для заданных параметров Q, K, L.

42. (п. 7.12). Методом компьютерного моделирования для последовательностей **File01–File06** из АДП исследовать зависимость мощности теста и его вычислительной сложности от параметров Q, K, L.

43. (п. 7.12). Решить задачу оптимального выбора параметров теста Маурера по критерию максимума мощности теста при фиксированной длине n_+ наблюдаемой двоичной последовательности.

44. (п. 7.12). Указать класс альтернатив H_1, для которых тест Маурера:
1) наиболее чувствителен (имеет наибольшую мощность);
2) нечувствителен (не отличает эти альтернативы от H_0).

45. (п. 7.13). Выполнить задание 1 с помощью теста (7.13.4) для заданных значений m.

46. (п. 7.13). Методом компьютерного моделирования для последовательностей `File01`–`File06` из АДП (прил. Б) исследовать зависимость мощности теста (7.13.4) и его вычислительной сложности от параметров n, N, m.

47. (п. 7.13). Выполнить задание 45 для теста (7.13.5).

48. (п. 7.13). Выполнить задание 46 для теста (7.13.5).

49. (п. 7.13). Для $N = 2$ разработать программную реализацию тестов (7.13.4) и (7.13.5) с максимальным быстродействием.

50. (п. 7.14). Выполнить задание 1 с помощью теста (7.14.1), (7.14.2), (7.14.4), (7.14.5). Для $n = 10^6$ выполнить это же задание с помощью теста (7.14.1), (7.14.3)—(7.14.5). Сравнить уровни значимости этих двух тестов.

51. (п. 7.14). Методом компьютерного моделирования для последовательностей `File01`–`File06` из АДП исследовать мощность и вычислительную сложность тестов из задания 50.

52. (п. 7.14). Используя «надежный» генератор псевдослучайной двоичной последовательности (предложенный преподавателем), разработать программу оценки параметров μ_n, σ_n в (7.14.1) методом Монте-Карло.

53. (п. 7.14). Разработать модификацию данного теста, основанную на обработке M реализаций длиной n каждая и на проверке согласия выборочной функции распределения P-значений (7.14.5) с функцией распределения стандартного равномерного закона $R[0, 1]$ с помощью критерия Колмогорова.

54. (п. 7.15). Выполнить задание 1 с помощью алгоритма тестирования, основанного на статистике линейной сложности.

55. (п. 7.15). Методом компьютерного моделирования для последовательностей `File01`–`File06` из АДП (прил. Б) исследовать мощность и вычислительную сложность алгоритма тестирования.

56. (п. 7.15). Разработать модификацию теста, используя произвол в задании ячеек (7.15.3) и их количества.

57. (п. 7.16). Выполнить задание 1 с помощью теста (7.16.2), (7.16.3), (7.16.4). Методом компьютерного моделирования для последовательностей `File01–File06` из АДП исследовать зависимость мощности теста и его вычислительной сложности от основных параметров.

58. (п. 7.16). Выполнить задание 57 для теста (7.16.2), (7.16.3), (7.16.5).

59. (п. 7.16). Выполнить задание 57 с учетом замечания 7.16.1.

60. (п. 7.16). Выполнить задание 58 с учетом замечания 7.16.1.

61. (п. 7.17). Выполнить задания 57—60 с использованием теста на основе экстремальной статистики дельта-произведения.

kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654ughet
r873yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr98374h34yr
98e76d9w8yedyt8354753griewgfr87we68r73yt4irg384687efiuweghr834rh
iufh8e7r63gh4fiye86r8re7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4r
hw9e76r98e6r98y3i4u5h39kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986
659374uhreifhei654ughetr873yf5u3eiry9rewyejhrf98794358yoihoir7f8
4uy5ihjliehjr98374h34yr9 8e76d9w8yedyt8354753grie
wgfr87we68r73yt4irg38468 7efiuweghr834rhiufh8e7r6
3gh4fiye86r8re7yrfiuehri **Часть 2** t763549y43iurhieuyf98e4i
5uh3i4rhw9e76r98e6r98y3i 4u5h394kdfjhgjdhrgjkldfk
vjsiertuesrlojeoiruyt986659374uhreifhei654ughetr873yf5u3eiry9rew
yejhrf98794358yoihoir7f84uy5ihjliehjr98374h34yr98e76d9w8yedyt835
4753griewgfr87we68r73yt4irg384687efiuweghr834rhiufh8e7r63gh4fiye
86r8re7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i
4u5h 394k
dfjh **МАТЕМАТИЧЕСКИЕ И КОМПЬЮТЕРНЫЕ** gjdh
rgjk **МЕТОДЫ КРИПТОЛОГИИ** ldfk
vjsi ertu
esrl ojeo
iruyt986659374uhreifhei654ughetr873yf5u3eiry9rewyejhrf98794358yo
ihoir7f84uy5ihjliehjr98374h34yr98e76d9w8yedyt8354753griewgfr87we
68r73yt4irg384687efiuweghr834rhiufh8e7r63gh4fiye8r8re7yrfiuehrit
763549y43iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h394kdfjhgjdhr
gjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654ughetr873yf5u3e
iry9rewyejhrf987943griewgfr87we68r73yt4irg384687efiuweghr834rhiu
fh8e7r63gh4fiye86r8re7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4rhw
9e76r98e6r98y3i4u5h394kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt9866
59374uhreifhei654ughetr873yf5u3eiry9rewyejhrf98794358yoihoir7f84
uy5ihjliehjr98374h34yr98e76d9w8yedyt8354753griewgfr87we68r73yt4i
rg384687efiuweghrreubgnerugfiuerghfqoewihoqiewhfquywegqwieughiwq
ugefwgeue834rhiufh8e7r63gh4fiye86r8re7yrfiuehrit763549y43iurhieu
yf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h394kdfjhgjdhrgjkldfkvjsiertue
srlojeoiruyt986659374uhreifhei654ughetr873yf5u3eiry9rewyejhrf987
94358yoihoir7f84uy5ihjliehjr98374h34yr98e76d9w8yedyt8354753griew
gfr87we68r73yt4irg384687efiuweghr834rhiufh8e7r63gh4fiye86r8re7yr
fiuehrit763549y43iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h394kd
fjhgjdhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhei654ughetr8
73yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr98374h34yr98
e76d9w8yedyt8354753griewgfr87we68r73yt4irg384687efiuweghr834rhiu
fh8e7r63gh4fiye86r8re7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4rhw
9e76r98e6r98y3i4u5h394kdfjhgjdhrgjkldfkvjsiertuesrlojeoiruyt9866
59374uhreifhei654ughetr873yf5u3eiry9rewyejhrf98794358yoihoir7f84
uy5ihjliehjr98374h34yr98e76d9w8yedyt8354753griewgfr87we68r73yt4i
rg384687efiuweghr834rhiufh8e7r63gh4fiye86r8re7yrfiuehrit763549y4
3iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4u5h394kdfjhgjdhrgjkldfkv
jsiertuesrlojeoiruyt98665937hjliehjr98374h34yr98e76d9w8yedyt8354
753griewgfr87we68r73yt4irg384687efiuweghr834rhiufh8e7r63gh4fiye8
6r8re7yrfiuehrit763549y43iurhieuyf98e4i5uh3i4rhw9e76r98e6r98y3i4
u5h394klerkdnfwep[weleryjgh74356ghn8o6y567546ue534123err4oedfjhg
jdhrgjkldfkvjsiertuesrlojeoiruyt986659374uhreifhrthrei654ughetr8
73yf5u3eiry9rewyejhrf98794358yoihoir7f84uy5ihjliehjr98374h34yr98

SXJKN8EYHR3I4HR987
YDWFUIHW9E8YR384H5
O34596EUF9O8WE76R8
3Y439846UHFE98YDSF
JH8374H4BR8276RUBN
WDLSKFNO34H34O4FHO
3IHF3843H4F8347F3N
FDKFDOF34KJ9DS809S
JDFJ093489U3J4OPI2
93J4JDO9S90DUFIOJ4
FO9SD90FUSDO9JFO93
4709374UO47DF87Y62
34BN9WE872H3RO98E7

Глава 8

АЛГОРИТМЫ ГЕНЕРАЦИИ ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

SXJKN8EYHR3I4HR987
YDWFUIHW9E8YR384H5
O34596EUF9O8WE76R8
3Y439846UHFE98YDSF
JH8374H4BR8276RUBN
WDLSKFNO34H34O4FHO
3IHF3843H4F8347F3N
FDKFDOF34KJ9DS809S
JDFJ093489U3J4OPI2
93J4JDO9S90DUFIOJ4
FO9SD90FUSDO9JFO93
4709374UO47DF87Y62
34BN9WE872H3RO98E7

8.1. Классификация алгоритмов генерации

Согласно п. 7.1, *программный генератор РРСП* — это программа для имитации на компьютере реализации РРСП с помощью псевдослучайной последовательности $\{x_t\}$, которая: 1) вычисляется на компьютере по известному детерминированному рекуррентному соотношению; 2) «по своим статистическим свойствам не отличима» от РРСП.

Классификация существующих алгоритмов генерации псевдослучайных последовательностей представлена на рис. 8.1. Выделяются три основных подхода к построению алгоритмов генерации: 1) *прямые методы* построения элементарных рекуррентных последовательностей:

$$x_t = f(x_{t-1}, x_{t-2}, \ldots, x_{t-m}) : \mathcal{A}^m \to \mathcal{A},$$

«проходящих» статистические тесты из главы 7; 2) *методы «улучшения элементарных последовательностей»*, заключающиеся в специальных функциональных преобразованиях этих последовательностей для уменьшения отклонения их статистических свойств от свойств РРСП; 3) *комбинирование алгоритмов генерации*, построенных с помощью первого или второго подхода.

8.2. Линейные и мультипликативные конгруэнтные генераторы

Линейным конгруэнтным генератором (ЛКГ) с параметрами (x_0, a, c, N) называется программный генератор РРСП, порождающий псевдослучайную последовательность $x_1, x_2, \ldots \in \mathcal{A}$, $\mathcal{A} = \{0, 1, \ldots, N-1\}$, с помощью рекуррентного соотношения

$$x_{t+1} = (ax_t + c) \bmod N, \quad t = 0, 1, \ldots \tag{8.2.1}$$

Параметры этого генератора (8.2.1) имеют следующий смысл: $x_0 \in \mathcal{A}$ — *начальное*, или *стартовое*, *значение* (seed); $a \in \mathcal{A} \setminus \{0\}$ — ненулевой *множитель*; $c \in \mathcal{A}$ — *приращение*; $N \in \mathbb{N}$ — *модуль*, равный мощности алфавита $\mathcal{A}$.

Если приращение $c = 0$, то генератор (8.2.1) называется *мультипликативным конгруэнтным генератором* (МКГ), а если $c \neq 0$, то *смешанным конгруэнтным генератором* (СКГ).

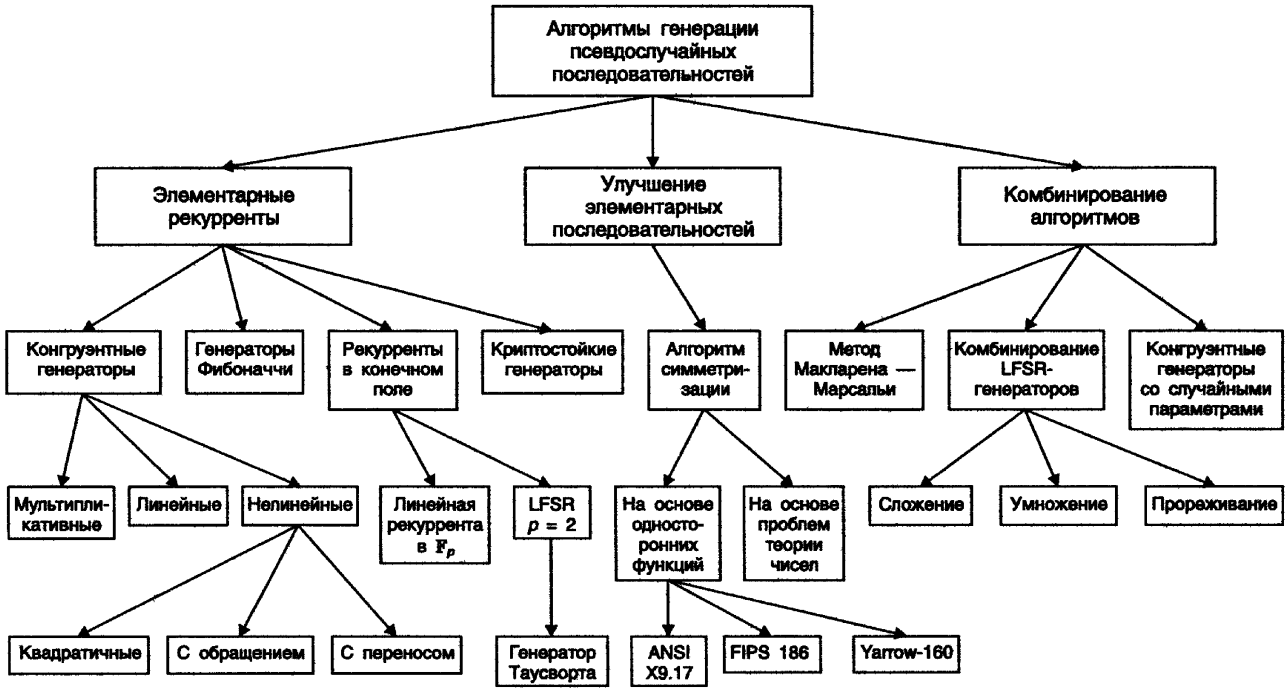


Рис. 8.1. Классификация алгоритмов генерации псевдослучайных последовательностей

Перечислим свойства псевдослучайной последовательности, порождаемой ЛКГ.

Свойство C_1. Для общего члена последовательности (8.2.1) справедлива формула

$$x_t = \left(a^t x_0 + \frac{a^t - 1}{a - 1} c \right) \bmod N, \quad t \geq 1. \tag{8.2.2}$$

Свойство C_2. Найдется номер $\tau \in \mathcal{A}$, начиная с которого последовательность (8.2.1) «зацикливается» с периодом $T \leq N - \tau$.

Свойство C_3. Для любого $k \geq 2$ подпоследовательность $x_0, x_k, x_{2k}, x_{3k}, \ldots \in \mathcal{A}$, полученная из псевдослучайной последовательности (8.2.1) удалением всех членов, не кратных k, оказывается псевдослучайной последовательностью, порожденной ЛКГ (8.2.1), с параметрами $(x_0, \tilde{a}, \tilde{c}, N)$, где

$$\tilde{a} = a^k \bmod N, \quad \tilde{c} = (c(a^k - 1)/(a - 1)) \bmod N.$$

Свойство C_4. Псевдослучайная последовательность (8.2.1), порождаемая ЛКГ, достигает максимального значения периода $T_{\max} = N(\tau = 0)$ тогда и только тогда, когда выполнены следующие три условия:

а) c и N — взаимно простые, т.е. НОД $(c, N) = 1$;

б) число $b = a - 1$ кратно p для любого простого числа $p < N$, являющегося делителем N;

в) число b кратно 4, если N кратно 4.

Свойство C_5. Для МКГ ($c = 0$), если x_0 и N — взаимно простые, a — первообразный элемент по модулю N, а $\varphi(N)$ — максимально возможный порядок по модулю N, то псевдослучайная последовательность имеет максимальный период $T_{\max} = \varphi(N)$.

Свойство C_6. Для МКГ, если $N = 10^q$, $q \geq 5$ и x_0 не кратно 2 или 5, то максимально возможное значение периода $T_{\max} = 5 \cdot 10^{q-2} = N/20$ псевдослучайной последовательности достигается тогда и только тогда, когда вычет $a \bmod 200$ принимает одно из 32 следующих «магических» значений:

$$3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83, 91, 109, 117, 123,$$
$$131, 133, 139, 141, 147, 163, 171, 173, 179, 181, 187, 189, 197.$$

Свойство C_7. Для МКГ, если $N = 2^q$, $q \geq 4$, то максимально возможное значение периода $T_{\max} 2^{q-2} = N/4$ псевдослучайной последовательности достигается, если $x_0 \geq 1$ —нечетно и вычет $a \bmod 8 \in \{3, 5\}$.

Свойство C_8. Для МКГ, если $x_0 \neq 0$, N — простое число и справедливо разложение на множители: $N-1 = p_1^{m_1} \cdots p_s^{m_s}$, где $p_1, \ldots, p_s$ — простые числа, а $m_1, \ldots, m_s \in \mathbb{N}$, то максимально возможное значение периода $T_{\max} = N-1$ псевдослучайной последовательности достигается для случая, когда

$$a^{(N-1)/p_j} \neq 1 \pmod N, \quad j = \overline{1,s}.$$

Свойство C_9. «Слабость» ЛКГ и МКГ заключается в том, что если рассматривать последовательные биграммы $(z_1^{(t)}, z_2^{(t)})$: $z_1^{(t)} = x_t$, $z_2^{(t)} = x_{t-1}$, то точки $z^t = (z_1^{(t)}, z_2^{(t)})$, $t = 1, 2, \ldots$, на плоскости $\mathbb{R}^2$ будут лежать на прямых из семейства $z_2 = az_1 + c - kN$, $k = 0, 1, \ldots$

8.3. Нелинейные конгруэнтные генераторы

Свойство C_9 линейного и мультипликативного конгруэнтных генераторов псевдослучайных последовательностей (см. п. 8.2) представляет «слабость» этих генераторов и может активно использоваться для построения криптоатак в целях оценки параметров a, c, x_0. Для устранения этого недостатка используют нелинейные конгруэнтные генераторы псевдослучайных последовательностей, среди которых известны: квадратичный конгруэнтный генератор; генератор Эйхенауэра — Лена [88] с обращением; конгруэнтный генератор, использующий умножение с переносом; квадратичный генератор Ковэю; генератор «середины квадрата». Наибольшее распространение получили первые три подхода, описание которых приводится ниже.

8.3.1. Квадратичный конгруэнтный генератор

Этот алгоритм генерации псевдослучайной последовательности $x_t \in \mathcal{A} = \{0, 1, \ldots, N-1\}$ определяется квадратичным рекуррентным соотношением

$$x_{t+1} = (dx_t^2 + ax_t + c) \bmod N, \quad t = 0, 1, \ldots, \tag{8.3.1}$$

где $x_0, a, c, d \in \mathcal{A}$ — параметры генератора. Выбор этих параметров осуществляется на основе следующих двух свойств последовательности (8.3.1).

Свойство C_1. Квадратичная конгруэнтная последовательность (8.3.1) имеет наибольший период $T_{\max} = N$ тогда и только тогда, когда выполнены следующие условия:

1) c, N — взаимно простые числа;
2) d, $a-1$ — кратны p, где p — любой нечетный простой делитель N;
3) d — четное число, причем

$$d = \begin{cases} (a-1) \bmod 4, & \text{если } N \text{ кратно } 4, \\ (a-1) \bmod 2, & \text{если } N \text{ кратно } 2; \end{cases}$$

4) если N кратно 9, то либо $d \bmod 9 = 0$, либо $d \bmod 9 = 1$ и $cd \bmod 9 = 6$.

Свойство C_2. Если $N = 2^q$, $q \geq 2$, то наибольший период $T_{\max} = 2^q$ тогда и только тогда, когда c — нечетно, d — четно, a — нечетное число, удовлетворяющее соотношению

$$a = (d+1) \bmod 4.$$

8.3.2. Генератор Эйхенауэра — Лена с обращением

Псевдослучайная нелинейная конгруэнтная последовательность Эйхенауэра — Лена с обращением [88] определяется следующим нелинейным рекуррентным соотношением ($t = 0, 1, \ldots$):

$$x_{t+1} = \begin{cases} (ax_t^{-1} + c) \bmod N, & \text{если } x_t \geq 1, \\ c, & \text{если } x_t = 0, \end{cases} \tag{8.3.2}$$

где $x_t^{-1} \in \mathcal{A}$ — обратный к x_t элемент по модулю N, т.е. $x_t x_t^{-1} \equiv 1 (\bmod N)$; $x_0, a, c, \in \mathcal{A}$ — параметры генератора.

Свойство C_3. Если $N = 2^q$, a, x_0 — нечетны, c — четно, то генератор (8.3.2) имеет максимально возможный период $T_{\max} = 2^{q-1}$ тогда и только тогда, когда

$$a \equiv 1 \pmod 4, \quad c = 2 \pmod 4.$$

8.3.3. Конгруэнтный генератор, использующий умножение с переносом

При этом нелинейная конгруэнтная псевдослучайная последовательность определяется рекуррентным соотношением

$$x_{t+1} = (ax_t + c_t) \bmod N, \quad t = 0, 1, \ldots, \tag{8.3.3}$$

где, в отличие от (8.3.1), «приращение» $c_t = c(x_{t-1}, x_{t-2}, \ldots, x_0)$ изменяется во времени и зависит от указанных аргументов нелинейно:

$$c_t = \left\lfloor \frac{ax_{t-1} + c_{t-1}}{N} \right\rfloor. \tag{8.3.4}$$

Параметрами нелинейного конгруэнтного генератора (8.3.3), (8.3.4) являются x_0, c_0, a, N. Рекомендации по выбору этих параметров даны в работе [35, 130].

8.4. Рекурренты в конечном поле

Обобщением мультипликативной конгруэнтной последовательности (см. п. 8.1) является *линейная рекуррентная последовательность* порядка $k \geq 1$ над конечным полем $\mathbb{F}_p$ (p — простое число):

$$x_{t+1} = (a_1 x_t + a_2 x_{t-1} + \ldots + a_k x_{t-k+1}) \bmod p, \quad t = 0, 1, \ldots, \tag{8.4.1}$$

где $a_1, a_2, \ldots, a_k \in \mathcal{A} = \{0, 1, \ldots, p-1\}$ — коэффициенты рекурренты, а $x_0, x_{-1}, \ldots, x_{-k+1} \in \mathcal{A}$ — начальные значения рекурренты.

Параметры генератора псевдослучайной последовательности (8.4.1): p; k; $a_1, \ldots, a_k$; $x_0, x_{-1}, \ldots, x_{-k+1}$. Начальные значения $x_0, x_{-1}, \ldots, x_{-k+1} \in \mathcal{A}$ выбираются произвольно так, чтобы не обращались в ноль одновременно. Коэффициенты рекурренты $a_1, \ldots, a_k \in \mathcal{A}$ выбираются таким образом, чтобы порождающий полином

$$f(x) = x^k - a_1 x^{k-1} - \ldots - a_{k-1}x - a_k \tag{8.4.2}$$

являлся примитивным многочленом по модулю p, т.е. многочлен (8.4.2) имел корень x_*, являющийся первообразным элементом поля $\mathbb{F}_{p^k}$. При таком выборе параметров достигается максимально возможный период $T_{\max} = p^k - 1$ псевдослучайной последовательности (8.4.1).

В качестве примеров эффективно реализуемых на компьютере алгоритмов генерации вида (8.4.1) приведем четыре генератора псевдослучайных последовательностей с периодом $T_{\max} \approx 2^{96}$ [128]:

$$\begin{aligned} x_{t+1} &= (1176x_t + 1476x_{t-1} + 1776x_{t-2}) \bmod (2^{32} - 5); \\ x_{t+1} &= 2^{13}(x_t + x_{t-1} + x_{t-2}) \bmod (2^{32} - 5); \\ x_{t+1} &= (1995x_t + 1998x_{t-1} + 2001x_{t-2}) \bmod (2^{35} - 849); \\ x_{t+1} &= 2^{19}(x_t + x_{t-1} + x_{t-2}) \bmod (2^{35} - 1629). \end{aligned} \tag{8.4.3}$$

Для «повышения качества псевдослучайных последовательностей» в линейную рекурренту (8.4.1) вводится нелинейное «приращение» c_t (как и в конгруэнтном генераторе, использующем умножение с переносом, см. п. 8.3):

$$x_{t+1} = (a_1 x_t + a_2 x_{t-1} + \ldots + a_k x_{t-k+1} + c_t) \bmod p, \tag{8.4.4}$$

где «приращение-перенос»

$$c_t = c(x_{t-1}, x_{t-2}, \ldots, x_0, x_{-1}, \ldots, x_{-k+1}) \in \mathcal{A}$$

зависит от своих аргументов нелинейно согласно рекуррентному соотношению

$$c_t = \lfloor (a_1 x_t + a_2 x_{t-1} + \ldots + a_k x_{t-k+1} + c_{t-1})/p \rfloor. \tag{8.4.5}$$

Примером рекуррентного генератора с переносом (8.4.4), (8.4.5) является генератор с периодом $T_{\max} \approx 2^{158}$ [128]:

$$x_{t+1} = (5115x_t + 1776x_{t-1} + 1492x_{t-2} + 2111111111x_{t-3} + c_t) \bmod 2^{32}. \tag{8.4.6}$$

8.5. Последовательности, порождаемые линейными регистрами сдвига с обратной связью (LFSR)

Линейным регистром сдвига с обратной связью (Linear Feedback Shift Register, сокращенно LFSR) называется логическое устройство, схема которого изображена на рис. 8.2 [97].

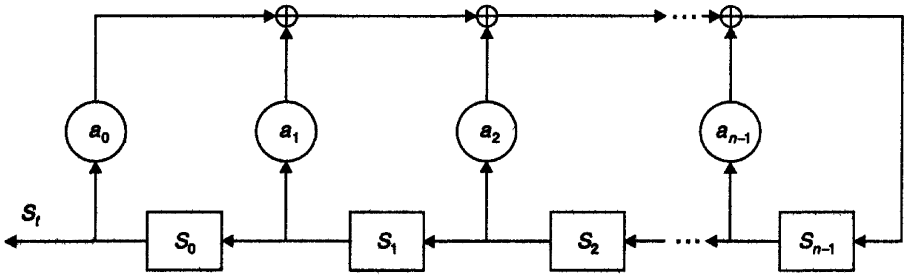


Рис. 8.2. Блок-схема LFSR

LFSR состоит из n ячеек памяти, двоичные состояния которых в момент времени $t = 0, 1, \ldots$характеризуются значениями $S_0(t), S_1(t), \ldots, S_{n-1}(t) \in \mathcal{A} = \{0, 1\}$. Выходы ячеек памяти связаны не только последовательно друг с другом, но и с сумматорами $\oplus$ в соответствии с коэффициентами передачи $a_0, a_1, \ldots, a_n \in \mathcal{A}$: если $a_i = 1$, то значение $S_i(t)$ i-й ячейки передается на один из входов i-го сумматора; если же $a_i = 0$, то такая передача отсутствует. Полагается $a_0 \equiv 1$. Состояние LFSR в текущий момент времени t задается двоичным n-вектор-столбцом $S(t) = (S_{n-1}(t), S_{n-2}(t), \ldots, S_0(t))'$.

Содержание ячеек LFSR с течением времени изменяется следующим образом, определяя тем самым динамику состояний LFSR:

$$S_i(t+1) = \begin{cases} S_{i+1}(t), & \text{если } i \in \{0, 1, \ldots, n-2\}, \\ \sum_{j=0}^{n-1} a_j S_j(t), & \text{если } i = n-1, \end{cases} \tag{8.5.1}$$

или в матричном виде:

$$S(t+1) = AS(t), \quad A = \begin{pmatrix} a_0 \; a_1 \; a_2 \ldots & \vdots & a_{n-1} \\ & \vdots & 0 \\ I_{n-1} & \vdots & \vdots \\ & \vdots & 0 \end{pmatrix}, \tag{8.5.2}$$

где I_{n-1} — единичная $(n-1) \times (n-1)$-матрица, $t = 0, 1, \ldots$

Текущие значения нулевой ячейки регистра используются в качестве элементов порождаемой LFSR двоичной псевдослучайной последовательности $s_t = S_0(t)$ (см. рис. 8.2), которая с учетом (8.5.1) удовлетворяет линейному рекуррентному соотношению

$$s_{t+n} = \sum_{j=0}^{n-1} a_j s_{t+j}, \quad t = 0, 1, \ldots \tag{8.5.3}$$

Модель (8.5.3) является частным случаем модели (8.4.1) линейной рекурренты над полем $\mathbb{F}_2$, поэтому коэффициенты $\{a_j\}$ выбираются согласно методике, приведенной в п. 8.4.

Опишем генератор Таусворта, основанный на LFSR. Он аналогичен (8.5.1)—(8.5.3) и определяется следующими соотношениями:

$$A = (I_n + (L')^p)(I_n + L^q),$$
$$S(t+1) = AS(t), \quad s_t = S_0(t), \quad t = 0, 1, \ldots, \tag{8.5.4}$$

где p, q — некоторые заданные натуральные числа (параметры алгоритма); L — $(n \times n)$-двоичная матрица, все единицы которой расположены лишь под главной диагональю. Преобразование $\tilde{S} = L^q S$ осуществляет сдвиг всех компонентов вектора S «вниз» на q позиций, заполняя освободившиеся позиции нулями; точно так же $\tilde{S} = (L')^p S'$ осуществляет сдвиг «вверх» на p позиций.

Выбор параметров p, q осуществляется таким образом, чтобы матрица A имела порядок $2^n - 1$ в группе невырожденных $(n \times n)$-матриц; при этом последовательность двоичных векторов $\{S(t)\}$, определяемых (8.5.4), имеет максимально возможный период $T_{\max} = 2^n - 1$. Это условие выполняется, например, если:

$$n = 31, \quad p = 13, \quad q = 18 \quad \text{или} \quad p = 18, \quad q = 13;$$
$$n = 32, \quad p = 15, \quad q = 17 \quad \text{или} \quad p = 17, \quad q = 15. \tag{8.5.5}$$

8.6. Генераторы Фибоначчи

Общий вид рекуррентного соотношения, определяющего *генератор Фибоначчи*, задается уравнением

$$x_t = x_{t-r} \Diamond x_{t-s}, \quad t = r, r+1, r+2, \ldots, \tag{8.6.1}$$

где $r, s \in \mathbb{N}$ $(r > s)$ — параметры генератора; $\Diamond$ — символ бинарного отношения:

$$\Diamond \in \{+, -, \times, \oplus\}.$$

В случае «+» или «–» $\{x_t\}$ — целые числа $(\bmod 2^k)$ для некоторого заданного $k \in \mathbb{N}$; в случае «×» $\{x_t\}$ — нечетные числа $(\bmod 2^k)$; в случае «⊕» элемент $x_t \in V_k$ представляет собой двоичный k-вектор и действие $\oplus$ выполняется покомпонентно.

Свойства псевдослучайной последовательности Фибоначчи (8.6.1) и методика выбора параметров r, s, k излагаются в [35, 130].

8.7. Криптостойкие генераторы на основе односторонних функций

Для повышения стойкости алгоритмов генерации псевдослучайных последовательностей к криптоанализу в последнее время предлагается [143] синтезировать алгоритмы на основе известных в криптографии односторонних функций (см. главу 13). Характерное свойство *односторонних (one-way) функций* состоит в том, что для вычисления значения функции по заданному значению

аргумента существует полиномиально-сложный алгоритм, в то время как для вычисления аргумента по заданному значению функции полиномиально-сложного алгоритма не существует (или он не известен). Доказательство свойства односторонности функции является трудной математической задачей, поэтому в настоящее время в криптосистемах часто используются «кандидаты в односторонние функции», для которых показано лишь, что *в настоящее время не известны* полиномиально-сложные алгоритмы вычисления обратной функции. Примерами таких «кандидатов» являются некоторые известные криптоалгоритмы (например, DES) и хэш-функции (например, SHA-1).

8.7.1. Генератор ANSI X9.17

Он представляет собой национальный стандартный алгоритм США для генерации двоичной псевдослучайной последовательности, входящей в FIPS (USA Federal Information Processing Standard). В этом генераторе в качестве «кандидата» односторонней функции используется так называемый «тройной DES» с двумя ключами $K_1, K_2 \in V_{64}$ — это алгоритм шифрования вида

$$F_K = E_{K_1} D_{K_2} E_{K_1},$$

где $K = (K_1 \,||\, K_2) \in V_{128}$ — составной 128-битовый ключ; E_{K_1} — алгоритм шифрования DES с ключом K_1; D_{K_2} — алгоритм расшифровывания DES с ключом $K_2 \neq K_1$.

Входными данными алгоритма ANSI X9.17 являются: некоторое случайное (и конфиденциальное) 64-битовое стартовое слово $s \in V_{64}$; 128-битовый составной ключ $K \in V_{128}$; m — количество 64-битовых двоичных слов, которые необходимо получить на выходе генератора.

Выходными данными являются m 64-битовых двоичных слов $x_1, x_2, \ldots, x_m \in V_{64}$.

Алгоритм генерации состоит из пяти шагов.

Шаг 1. Фиксируется 64-битовое представление $d \in V_{64}$ даты и времени при обращении к программе генерации и вычисляется вспомогательное 64-битовое двоичное слово

$$I = F_K(d).$$

Шаг 2. Для $i = \overline{1,m}$ повторяются шаги 3, 4.

Шаг 3. Вычисляется значение i-го выходного слова:

$$x_i = F_K(I \oplus s).$$

Шаг 4. Вычисляется новое значение параметра s:

$$s = F_K(x_i \oplus I).$$

Шаг 5. Формируется выходная последовательность m слов $x_1, x_2, \ldots, x_m \in V_{64}$, либо двоичная последовательность из $64m$ битов:

$$X = (x_1 \,||\, x_2 \,||\, \ldots \,||\, x_m) \in V_{64m}.$$

8.7.2. Генераторы FIPS-186

Они также являются национальным стандартом США и предназначены для генерации конфиденциальных параметров и ключевой информации для национального стандартного алгоритма электронной цифровой подписи DSA. В качестве «кандидата» односторонней функции используется алгоритм шифрования DES или алгоритм хэширования SHA-1; таким образом, имеются два варианта односторонней функции G_1 и G_2 (алгоритмы их вычисления приведены ниже).

Входными данными алгоритма генерации FIPS-186 являются целое число m и 160-битовое простое число q.

Выходные данные: последовательность m псевдослучайных чисел $x_1, x_2, \ldots, x_m \in \{0, 1, \ldots, q-1\}$, которые рекомендовано использовать как ключи в стандарте электронной цифровой подписи DSA.

Алгоритм генерации состоит из девяти шагов.

Шаг 1. Если в алгоритме используется односторонняя функция G_1, то задать $b = 160$; если же G_2, то задать произвольное число b: $160 \le b \le 512$.

Шаг 2. Сгенерировать неким образом случайное (и конфиденциальное) b-битовое стартовое значение $s \in V_b$.

Шаг 3. Задать 160-битовое «магическое» вспомогательное слово (в шестнадцатеричной записи):

$$t = 67452301\ \text{EFCDAB89}\ \text{98BADCFE}\ \text{10325476}\ \text{C3D2E1F0}_{16}.$$

Шаг 4. Для $i = \overline{1,m}$ выполнить шаги 5—8.

Шаг 5. По усмотрению пользователя: либо произвольно задать b-битовое число $y_i \in V_b$, либо положить его равным нулю ($y_i = \mathbf{0} \in V_b$).

Шаг 6. Вычислить $z_i = (s + y_i) \bmod 2^b \in V_b$.

Шаг 7. Вычислить $x_i = G(t, z_i) \bmod q \in \mathbb{F}_q$, где G — используемый вариант односторонней функции, т.е. $G = G_1$ либо $G = G_2$; алгоритмы вычисления значений G_1, G_2 приведены ниже.

Шаг 8. Вычислить $s = (1 + s + x_i) \bmod 2^b \in V_b$.

Шаг 9. Формируется псевдослучайная последовательность

$$x_1, x_2, \ldots, x_m \in \mathbb{F}_q.$$

Опишем односторонние функции G_1 и G_2.

Алгоритм вычисления значений функции G_1. *Входные данные*: два 160-битовых слова t, $c \in V_{160}$. *Выходные данные*: 160-битовое слово $G_1(t, c) \in V_{160}$. Алгоритм вычисления значений функции G_1 состоит из одиннадцати шагов.

Шаг 1. Разбить слово t на 5 блоков по 32 бита: $t = t_0 \,||\, t_1 \,||\, t_2 \,||\, t_3 \,||\, t_4$.

Шаг 2. Разбить аналогично слово: $c = c_0 \,||\, c_1 \,||\, c_2 \,||\, c_3 \,||\, c_4$.

Шаг 3. Для $i = \overline{0, 4}$ вычислить $u_i = t_i \oplus c_i$.

Шаг 4. Для $i = \overline{0, 4}$ выполнить шаги 5—9.

Шаг 5. Вычислить $b_1 = c_{(i+4) \bmod 5}$, $b_2 = c_{(i+3) \bmod 5}$.

Шаг 6. Вычислить $a_1 = u_2$, $a_2 = u_{(i+1) \bmod 5} \oplus u_{(i+4) \bmod 5}$.

Шаг 7. Вычислить $A = (a_1 \,||\, a_2)$, $B = (\tilde{b}_1 \,||\, b_2)$, где $\tilde{b}_1 \in V_{24}$ обозначает 24 младших бита 32-битового слова b_1.

Шаг 8. Осуществить DES-шифрование слова A с помощью 56-битового ключа B: $y_i = E_B(A)$.

Шаг 9. Разбить полученное после шифрования слово $y_i \in V_{64}$ на два 32-битовых слова: $y_i = L_i \,||\, R_i$.

Шаг 10. Для $i = \overline{0, 4}$ вычислить $z_i + L_i \oplus R_{(i+2) \bmod 5} \oplus L_{(i+3) \bmod 5}$.

Шаг 11. Выходное 160-битовое слово является конкатенацией:

$$G(t, c) = z_0 \,||\, z_1 \,||\, z_2 \,||\, z_3 \,||\, z_4 \in V_{160}.$$

Алгоритм вычисления значений функции G_2. *Входные данные*: 160-битовое слово t и b-битовое слово c, где $160 \le b \le 512$. *Выходные данные*: 160-битовое слово $G_2(t, c) \in V_{160}$. Алгоритм вычисления значений функции G_2 состоит из пяти шагов.

Шаг 1. Разбить слово t на пять 32-битовых слов:

$$t = H_1 \,||\, H_2 \,||\, H_3 \,||\, H_4 \,||\, H_5.$$

Шаг 2. Дополнить $512 - b$ нулями слово c, чтобы получить 512-битовое слово: $U = c \,||\, \mathbf{0}^{512-b}$.

Шаг 3. Разбить слово U на шестнадцать 32-битовых слов: $U = u_0 \,||\, u_1 \,||\, \ldots \,||\, u_{15}$ и задать $m = 1$.

Шаг 4. Выполнить шаг 4 алгоритма SHA-1, изменяющий $\{H_i\}$.

Шаг 5. Выходное слово является конкатенацией:

$$G_2(t, c) = H_1 \,||\, H_2 \,||\, H_3 \,||\, H_4 \,||\, H_5.$$

8.7.3. Генератор Yarrow-160

Генератор псевдослучайных двоичных слов Yarrow-160 (по-английски yarrow означает тысячелистник — название растения) предложен Келси, Шнайером и Фергюсоном [109] как «криптографический генератор псевдослучайных чисел». Он использует «тройной DES» с «тройным ключом» и американский стандарт хэш-функции SHA-1.

Очередное псевдослучайное число $x \in V_{64}$ определяется по правилам

$$C \leftarrow (C + 1) \bmod 2^{64}, \quad x \leftarrow E_{K_3} E_{K_2}^{-1} E_{K_1}(C),$$

где C — 64-битовый счетчик; E_{K_i} — алгоритм DES-шифрования с ключом $K_i \in V_{56}$.

В алгоритме Yarrow предусмотрена специальная схема обновления ключа K и значения счетчика C с использованием хэш-функции SHA-1 и некоторых вспомогательных «аккумуляторов случайности».

8.8. Криптостойкие генераторы, основанные на проблемах теории чисел

Стойкость генераторов псевдослучайных последовательностей, рассматриваемых в данном параграфе, основывается на неразрешимости с полиномиальной сложностью (на данный момент) некоторых известных *проблем теории чисел*: *факторизации больших чисел* и *дискретного логарифмирования*.

8.8.1. RSA-алгоритм генерации псевдослучайных последовательностей

Этот алгоритм основывается на том факте, что в настоящее время криптоанализ RSA не может быть осуществлен с полиномиальной сложностью.

RSA-алгоритм генерации двоичной псевдослучайной последовательности $x_1, x_2, \ldots, x_n \in \mathcal{A}$ состоит из шести шагов.

Шаг 1. Как и в RSA-криптосистеме, генерируются достаточно большие различные простые числа p, q и вычисляются числа $N = pq$, $\varphi = (p-1)(q-1)$. Выбирается случайное целое число k, $1 < k < \varphi$, взаимно простое с φ, так, что НОД $(k, \varphi) = 1$.

Шаг 2. Выбрать случайное целое — стартовое значение — $u_0 \in \{1, 2, \ldots, N-1\}$.

Шаг 3. Для $i = \overline{1, n}$ выполнить шаги 4, 5.

Шаг 4. Вычислить $u_i = u_{i-1}^k \bmod N \in \{0, 1, \ldots, N-1\}$.

Шаг 5. Вычислить $x_i \in \mathcal{A} = \{0, 1\}$ — самый младший бит числа u_i в двоичном представлении.

Шаг 6. Сформировать выходную последовательность $x_1, x_2, \ldots, x_n$.

Недостатком RSA-алгоритма является его невысокое быстродействие при реализации на универсальных компьютерах, вызванное большими затратами машинного времени при выполнении модулярного умножения на шаге 4. Для повышения быстродействия на шаге 5 можно выделять сразу несколько младших битов. Эта идея используется в представленной ниже *модификации Микали — Шнорра RSA-алгоритма*.

Шаг 1. Как и в RSA-криптосистеме, генерируются достаточно большие различные простые числа p, q и вычисляются $N = pq$, $\varphi = (p-1)(q-1)$. Вычисляется битовая длина $m = \lceil \log_2 N \rceil$ числа N. Выбирается целое k, $1 < k < \varphi$, так, что НОД $(k, \varphi) = 1$, $80k \le m$. Вычисляются $K = \lfloor m(1 - 2/k) \rfloor$, $r = m - K$.

Шаг 2. Выбрать стартовое число u_0, содержащее в двоичном представлении r значимых битов: $u_0 \in \{0, 1, \ldots, 2^r - 1\}$.

Шаг 3. Генерируем псевдослучайную последовательность длиной Kn. Для $i = \overline{1,n}$ выполнить шаги 4—6.

Шаг 4. Вычислить $y_i = u_{i-1}^k \bmod N$.

Шаг 5. Вычислить u_i как число, двоичное представление которого образовано r старшими битами двоичного представления числа y_i.

Шаг 6. Вычислить $x_i \in V_K$ как двоичное слово, образованное K младшими битами двоичного представления числа y_i.

Шаг 7. Сформировать выходную двоичную последовательность $x_1, x_2, \ldots, x_n \in V_K$.

Отметим, что условие криптостойкости алгоритма Микали — Шнорра является более жестким, чем для RSA-алгоритма, и состоит в следующем: вероятностное распределение случайной величины $y = x^k \bmod N$ для r-битовой случайной величины x должно быть неотличимо полиномиально-сложными статистическими тестами от равномерного распределения на множестве $\{0, 1, \ldots, N-1\}$.

8.8.2. BBS (Blum — Blum — Shub)-алгоритм генерации псевдослучайных последовательностей

Этот алгоритм генерации двоичной псевдослучайной последовательности $x_1, x_2, \ldots, x_n \in \mathcal{A} = \{0, 1\}$ основывается на проблеме факторизации больших чисел. Он используется в криптосистеме Блюма — Голдвассера с открытым ключом и состоит из шести шагов.

Шаг 1. Генерируются два достаточно больших секретных случайных (и различных) простых числа $p, q \equiv 3 (\bmod 4)$ и вычисляется $N = pq$.

Шаг 2. Выбрать случайное стартовое целое число $s \in \{1, 2, \ldots, N-1\}$ так, что НОД $(s, N) = 1$. Вычислить $u_0 = s^2 \bmod N$.

Шаг 3. Для $i = \overline{1,n}$ выполнить шаги 4, 5.

Шаг 4. Вычислить $u_i = u_{i-1}^2 \bmod N$.

Шаг 5. Вычислить $x_i \in \mathcal{A}$ как самый младший бит двоичного представления числа u_i.

Шаг 6. Сформировать выходную последовательность $x_1, x_2, \ldots x_n \in \mathcal{A}$.

Отметим, что BBS-алгоритм допускает модификацию, подобную модификации Микали — Шнорра.

8.9. Методы «улучшения» элементарных псевдослучайных последовательностей

Пусть $\xi_1, \xi_2, \ldots \in \mathcal{A}$ — некоторая двоичная псевдослучайная последовательность, сгенерированная одним из простейших методов и называемая поэтому (в данном параграфе) элементарной. Для того чтобы построить псевдослучай-

ную последовательность со свойствами, более близкими к свойствам РРСП, чем элементарная последовательность, осуществим функциональное преобразование $\{\xi_i\}$:

$$x_1 = f_1(\xi_1, \xi_2, \ldots), \quad x_2 = f_2(\xi_1, \xi_2, \ldots), \ldots, \tag{8.9.1}$$

где $f_1(\cdot)$, $f_2(\cdot)$, ... — некоторые функционалы, задающие отображение $\mathcal{A}^\infty \to \mathcal{A}$.

Функционалы $f_1(\cdot)$, $f_2(\cdot)$, ... следует подбирать так, чтобы преобразованная последовательность $\{x_k\}$ имела вероятностное распределение, более близкое к распределению РРСП, чем распределение $\{\xi_i\}$.

Выбирая различные функционалы $f_1(\cdot)$, $f_2(\cdot)$, ... и метрики в пространстве вероятностных распределений, можно разработать множество методов и алгоритмов «улучшения» элементарных псевдослучайных последовательностей.

Проиллюстрируем эту идею на простейшем примере (алгоритм симметризации псевдослучайных последовательностей). Пусть в элементарной последовательности $\{\xi_i\}$ нарушено условие равновероятности символов «0» и «1»:

$$\mathbf{P}\{\xi_i = 1\} = p, \quad \mathbf{P}\{\xi_i = 0\} = q = 1 - p, \quad p \neq \frac{1}{2}. \tag{8.9.2}$$

В целях преодоления этого отклонения от свойств РРСП осуществим следующее функциональное преобразование вида (8.9.1):

$$\eta_j = \begin{cases} 0, & \text{если } \xi_{2j-1} = 0, \xi_{2j} = 1, \\ 1, & \text{если } \xi_{2j-1} = 1, \xi_{2j} = 0, \quad j = 1, 2, \ldots, \\ 2 & \text{в противном случае.} \end{cases} \tag{8.9.3}$$

Затем из этой последовательности $\{\eta_j\}$ выходная двоичная последовательность $\{x_k\}$ получается прореживанием: отбрасываются значения $\eta_j \notin \mathcal{A} = \{0, 1\}$, т.е. те значения, которые совпадают с символом «2».

Из (8.9.2) и (8.9.3) следует, что

$$\begin{gathered} \pi_0 = \mathbf{P}\{\eta_j = 0\} = \mathbf{P}\{\xi_{2j-1} = 0, \xi_{2j} = 1\} = p(1-p); \\ \pi_1 = \mathbf{P}\{\eta_j = 1\} = \mathbf{P}\{\xi_{2j-1} = 1, \xi_{2j} = 0\} = p(1-p); \\ \mathbf{P}\{\eta_j = 2\} = 1 - 2p(1-p). \end{gathered} \tag{8.9.4}$$

Из (8.9.4) видно, что $\pi_0 = \pi_1$, поэтому после прореживания получаем двоичную псевдослучайную последовательность с равновероятными значениями:

$$\mathbf{P}\{x_k = 0\} = \mathbf{P}\{x_k = 1\} = \frac{1}{2}. \tag{8.9.5}$$

Заметим, что при получении (8.9.4), (8.9.5) использовалось предположение о попарной независимости элементов исходной последовательности $\{\xi_i\}$.

Отметим, что доля (в процентах) отброшенных значений в последовательности $\{\eta_j\}$ равна $k = (1 - 2p(1-p)) \cdot 100\,\%$ и тем больше, чем больше p отличается от 1/2. Это позволяет вычислить средний коэффициент использования исходной последовательности для получения одного элемента выходной последовательности.

8.10. Комбинирование алгоритмов генерации методом Макларена — Марсальи

Пусть имеется два простейших генератора псевдослучайных последовательностей: G_1 и G_2. Генератор G_1 порождает «элементарную» последовательность над алфавитом мощности N:

$$x_0, x_1, x_2, \ldots \in \mathcal{A}(N) = \{0, 1, \ldots, N-1\},$$

а генератор G_2 — над алфавитом мощности K:

$$y_0, y_1, y_2, \ldots \in \mathcal{A}(K) = \{0, 1, \ldots, K-1\}.$$

Пусть имеется вспомогательная таблица

$$T = \{T(0), T(1), \ldots, T(K-1)\}$$

из K целых чисел (память из K ячеек).

Метод Макларена — Марсальи комбинирования последовательностей $\{x_i\}$, $\{y_i\}$ для получения выходной псевдослучайной последовательности $\{z_k\}$ состоит в следующем. Сначала T-таблица заполняется K первыми членами последовательности $\{x_i\}$:

$$T(i) \leftarrow x_i, \quad i = \overline{0, K-1}.$$

Элементы выходной последовательности вычисляются следующим образом:

$$s \leftarrow y_k, \quad z_k = T(s), \quad T(s) \leftarrow x_{K+k}, \quad k = 0, 1, \ldots$$

Таким образом, генератор G_2 осуществляет «случайный» выбор из T-таблицы, а также ее «случайное» заполнение «случайными» числами, порождаемыми генератором G_1.

Метод комбинирования Макларена — Марсальи позволяет ослабить зависимость между членами $\{z_k\}$ и увеличить период псевдослучайной последовательности.

8.11. Комбинирование LFSR-генераторов

LFSR-генераторы часто используются в качестве генераторов элементарных псевдослучайных последовательностей и применяются для комбинирования генераторов [174]. Прежде всего отметим, что LFSR-генераторы можно использовать в качестве G_1, G_2 в генераторе Макларена — Марсальи. Рассмотрим еще ряд способов комбинирования LFSR-генераторов.

8.11.1. Полиномиальное комбинирование элементарных последовательностей

Рассмотрим общую модель комбинирования $M \geq 2$ элементарных псевдослучайных последовательностей, порожденных M генераторами LFSR (см. п. 8.5): LFSR_1, ..., LFSR_M (рис. 8.3). Здесь $x_{it} \in \mathcal{A} = \{0, 1\}$ — элемент двоичной

псевдослучайной последовательности, порождаемой l-м LFSR-генератором в момент времени $t = 1, 2, \ldots$ ($l = \overline{1, M}$), $y = F(x_1, \ldots, x_M) : \mathcal{A}^M \to \mathcal{A}$ — заданная двоичная функция M двоичных переменных, $y_t \in \mathcal{A}$ — элемент выходной двоичной псевдослучайной последовательности в момент времени t.

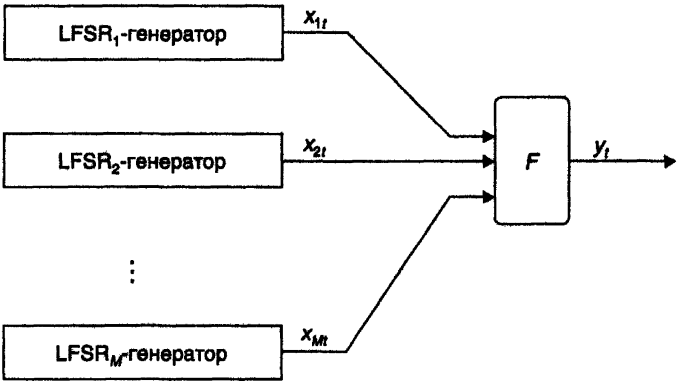


Рис. 8.3. Общая модель комбинирования LFSR-генераторов

Будем рассматривать функцию $F(\cdot)$ общего полиномиального вида:

$$y = F(x) = \left(a_0 + \sum_{1 \le i \le M} a_i x_i + \sum_{1 \le i < j \le M} a_{ij} x_i x_j + \ldots + a_{12\ldots M} x_1 \cdots x_M \right) \bmod 2, \quad (8.11.1)$$

где $a_0, \{a_i\}, \{a_{ij}\}, \ldots, a_{12\ldots M} \in \mathcal{A}$ — заданные двоичные коэффициенты.

Определим аналогично (8.11.1) полином степени M над полем действительных чисел $\mathbb{R}$:

$$F^*(x) = a_0 + \sum_{1 \le i \le M} a_i x_i + \sum_{1 \le i < j \le M} a_{ij} x_i x_j + \ldots + a_{12\ldots M} x_1 \cdots x_M. \quad (8.11.2)$$

Исследуем влияние функции комбинирования $F(\cdot)$ на свойства выходной последовательности. Так как каждая из последовательностей $\{x_{1t}\}$, $\{x_{2t}\}$, ..., $\{x_{Mt}\}$, порождаемая LFSR-генератором, периодична, то и выходная двоичная последовательность

$$y_t = F(x_{1t}, x_{2t}, \ldots, x_{Mt}), \quad t = 1, 2, \ldots \quad (8.11.3)$$

также периодична. Известно, что всякую периодическую последовательность $\{y_t\}$ можно представить как порождаемую некоторым LFSR-генератором. Наименьший порядок этого LFSR принято называть *линейной сложностью* последовательности $\{y_t\}$ и обозначать $L\{y_t\}$. Справедливо следующее утверждение, связывающее линейные сложности входных последовательностей $\{x_{1t}\}, \ldots, \{x_{Mt}\}$ и выходной последовательности $\{y_t\}$ [174].

Теорема 8.11.1. Пусть M двоичных LFSR-последовательностей максимальной длины (соответствующих примитивным многочленам) комбинируются согласно (8.11.1). Если все стартовые значения LFSR — ненулевые, все M порядков примитивных многочленов $L_1, \ldots, L_M$ различны и $L_j > 2\,(j = \overline{1, M})$, то выходная последовательность (8.11.3) имеет линейную сложность

$$L\{y_t\} = F^*(L_1, L_2, \ldots, L_M), \tag{8.11.4}$$

где $F^*(\cdot)$ определяется полиномом (8.11.2).

Следствие 8.11.1. Если комбинирование LFSR-генераторов осуществляется сложением по модулю 2 их выходных последовательностей, то

$$L\{y_t\} = \sum_{i=1}^{M} L_i.$$

Чем выше линейная сложность $L\{y_t\}$, тем более криптостойким является генератор псевдослучайной последовательности. Отметим, однако, что при подборе функции комбинирования (8.11.1) необходимо учитывать еще требование сбалансированности функции $F(\cdot)$:

$$W(F) = \sum_{x_1, \ldots, x_M \in \{0, 1\}} F(x_1, \ldots, x_M) = 2^{M-1}. \tag{8.11.5}$$

Например, свойству сбалансированности (8.11.5) при $M = 4$ удовлетворяет многочлен

$$F(x) = (x_1 + x_2 + x_3 + x_4 + x_1x_2 + x_3x_4 + x_1x_2x_3) \bmod 2.$$

В силу отмеченных фактов подбор «оптимальной» функции комбинирования $F_0(\cdot)$ сводится к решению экстремальной задачи:

$$F^*(L_1, L_2, \ldots, L_M) \to \max_{\{a_i\}, \{a_{ij}\}, \ldots, \{a_{12\ldots M}\}} \tag{8.11.6}$$

при ограничении (8.11.5).

8.11.2. Комбинирование с помощью псевдослучайного прореживания

Рассмотрим еще один способ комбинирования двух LFSR-генераторов G_1, G_2. Пусть LFSR-генератор G_1 порождает «элементарную» двоичную последовательность $\{a_t\}$, а LFSR-генератор G_2 — двоичную «селектирующую» последовательность $\{s_t\}$. С помощью этих двух последовательностей $\{a_t\}$, $\{s_t\}$ строится выходная последовательность $\{x_t\}$, включающая те биты x_t, для которых соответствующее значение селектора $s_t = 1$; если $s_t = 0$, то значение a_t игнорируется (отбрасывается). Такой генератор двоичной псевдослучайной последовательности предложен в работе [72] и назван *SG-генератором* (*Shrinking Generator*).

Свойства SG-генератора выражаются следующими двумя утверждениями.

Теорема 8.11.2. Пусть T_a, T_s — соответственно периоды последовательностей $\{a_t\}$ и $\{s_t\}$. Если генераторы G_1, G_2 используют примитивные порождающие многочлены степеней L_1 и L_2 соответственно, а периоды T_a, T_s — взаимно простые числа, то выходная последовательность $\{x_t\}$ имеет период

$$T = (2^{L_1} - 1)2^{L_2 - 1}.$$

Теорема 8.11.3. В условиях теоремы 8.11.2 линейная сложность выходной последовательности $\{x_t\}$ удовлетворяет неравенствам

$$L_1 2^{L_2 - 2} \le L\{x_t\} \le L_1 2^{L_2 - 1}.$$

8.12. Конгруэнтный генератор со случайными параметрами

Еще один способ комбинирования двух генераторов G_1, G_2 заключается в том, что G_2 изменяет параметры генератора G_1 с течением времени. Проиллюстрируем это в случае, когда G_1 — линейный конгруэнтный генератор:

$$x_t = (a_t x_{t-1} + b_t) \bmod N, \quad t = 1, 2, \ldots, \tag{8.12.1}$$

где $x_0 \in A = \{0, 1, \ldots, N-1\}$ — некоторое стартовое значение, а

$$A_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix} \in B, \quad t = 1, 2, \ldots \tag{8.12.2}$$

есть некоторая псевдослучайная последовательность векторов, равномерно распределенных в B.

В работе [66] доказано, что если $|B| = 3$, то наибольшее приближение распределения $\{x_t\}$ к равномерному достигается, если множество параметров B имеет следующий вид:

$$B = \left\{ \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\}. \tag{8.12.3}$$

8.13. Задания

1. (п. 8.2). Разработать программные реализации всех генераторов псевдослучайных последовательностей, представленных в данном параграфе, для значений параметров, указанных преподавателем.

2. (п. 8.2). Аналитически и с помощью компьютерного моделирования оценить вычислительную сложность (быстродействие) алгоритмов генерации, представленных в данном параграфе.

3. (п. 8.2). Осуществить статистическое тестирование алгоритмов генерации с помощью батареи тестов, разработанных при выполнении заданий главы 7.

4. (п. 8.2). Подобрать параметры алгоритмов генерации из условия минимума уклонения статистических характеристик генерируемой последовательности от характеристик РРСП.

5. (п. 8.2). Выявить «слабости» генераторов, представленных в данном параграфе, которые могут использоваться для криптоанализа генераторов.

6. (п. 8.2). Разработать алгоритм статистической (или алгебраической) оценки параметров генератора по наблюдаемой выходной последовательности; оценить точность и вычислительную сложность этого алгоритма.

7. (п. 8.2). Провести сравнительный анализ генераторов, представленных в данном параграфе, и сформулировать рекомендации по условиям их применения.

8. (п. 8.2). Проиллюстрировать графически свойство C_9.

9. (п. 8.3). Выполнить задания 1—7 применительно к нелинейным конгруэнтным генераторам (8.3.1)—(8.3.4).

10. (п. 8.3). Разработать и исследовать модификацию квадратичного конгруэнтного генератора, основанную на введении «приращения» c_t аналогично (8.3.3), (8.3.4).

11. (п. 8.3). Разработать и исследовать модификацию генератора Эйхенауэра — Лена, основанную на введении «приращения» c_t аналогично (8.3.3), (8.3.4).

12. (п. 8.4). Выполнить задания 1—7 применительно к генераторам псевдослучайных последовательностей (8.4.1)—(8.4.6).

13. (п. 8.4). Экспериментально исследовать влияние параметра c_0 в генераторе (8.4.6) на его свойство периодичности.

14. (п. 8.5). Выполнить задания 1—7 применительно к LFSR и генератору Таусворта (8.5.1)—(8.5.5).

15. (п. 8.5). Разработать и исследовать модификацию генератора Таусворта:

$$A = \prod_{k=1}^{m} (I_n + (L')^{p_k})(I_n + L^{q_k}).$$

16. (п. 8.5). Разработать и исследовать модификацию генератора Таусворта:

$$s_t = b' S(t),$$

где b — некоторый заданный двоичный n-вектор (параметр генератора).

17. (п. 8.6). Выполнить задания 1—7 применительно к генераторам Фибоначчи вида (8.6.1).

18. (п. 8.6). Установить связь между генераторами Фибоначчи и рекуррентами в конечном поле.

19. (п. 8.7). Выполнить задания 1—7 применительно к генератору ANSI X9.17.

20. (п. 8.7). Выполнить задания 1—7 применительно к генератору FIPS-186 с использованием функции G_1.

21. (п. 8.7). Выполнить задания 1—7 применительно к генератору FIPS-186 с использованием функции G_2.

22. (п. 8.7). Осуществить программную реализацию генератора Yarrow-160 и исследовать его свойства.

23. (п. 8.8). Выполнить задания 1—7 применительно к RSA-генератору.

24. (п. 8.8). Выполнить задания 1—7 применительно к генератору Микали — Шнорра.

25. (п. 8.8). Выполнить задания 1—7 применительно к BBS-генератору.

26. (п. 8.8). Аналогично подп. 8.8.1 построить модификацию BBS-алгоритма, используя подход Микали — Шнорра.

27. (п. 8.9). Выполнить задания 1—7 применительно к алгоритму симметризации элементарной двоичной последовательности. В качестве элементарных последовательностей использовать данные из АДП (прил. Б).

28. (п. 8.9). Оценить коэффициент использования к [14, 38] для алгоритма симметризации (8.9.1)—(8.9.5).

29. (п. 8.9). Предполагая, что элементарная последовательность $\{\xi_i\}$ является двоичной однородной цепью Маркова, исследовать изменение результатов (8.9.4), (8.9.5).

30. (п. 8.9). В предположениях задания 29 подобрать преобразования (8.9.1) так, чтобы $\{x_k\}$ стала РРСП.

31. (п. 8.9). Построить обобщение алгоритма в задании 30 для случая, когда $\{\xi_i\}$ — s-связная однородная цепь Маркова.

32. (п. 8.10). Выполнить задания 1—7 применительно к генератору Маклорена — Марсальи, используя в качестве G_1, G_2 мультипликативные конгруэнтные генераторы.

33. (п. 8.10). Исследовать влияние параметра K на качество генерируемой псевдослучайной последовательности.

34. (п. 8.10). Теоретически исследовать период для генератора Маклорена — Марсальи.

35. (п. 8.11). Выполнить задания 1—7 применительно к генератору (8.11.1)—(8.11.3) с полиномиальным комбинированием элементарных последовательностей.

36. (п. 8.11). Разработать алгоритм нахождения сбалансированных полиномов $F(\cdot)$, удовлетворяющих условию (8.11.5).

37. (п. 8.11). Разработать алгоритм нахождения «оптимальной» функции комбинирования $F_0(\cdot)$ согласно (8.11.5), (8.11.6).

38. (п. 8.11). Выполнить задания 1—7 применительно к SG-генератору, основанному на псевдослучайном прореживании LFSR-последовательности.

39. (п. 8.12). Выполнить задания 1—7 применительно к линейному конгруэнтному генератору со случайными параметрами (8.12.1)—(8.12.3).

40. (п. 8.12). Разработать и исследовать алгоритм генерации со случайными параметрами, в котором в качестве G_1, G_2 используются LFSR-генераторы.

SXJKN8EYHR3I4HR987Y
DWFUIHW9E8YR384H5O3
4596EUF9O8WE76R83Y4
39846UHFE98YDSFJH83
74H4BR8276RUBNWDF87
Y6234BN9WE872H3WRTJ
OJRTGM905JRO9G70954

Глава 9
ПОТОЧНЫЕ КРИПТОСИСТЕМЫ

SXJKN8EYHR3I4HR987Y
DWFUIHW9E8YR384H5O3
4596EUF9O8WE76R83Y4
39846UHFE98YDSFJH83
74H4BR8276RUBNWDF87
Y6234BN9WE872H3WRTJ
OJRTGM905JRO9G70954

9.1. Основные понятия

Поточной криптосистемой будем называть совокупность $G = \{G_\theta : \theta \in \Theta\}$ последовательностей бесконечной длины над алфавитом $\mathcal{A}$. Пусть на $\mathcal{A}$ задана групповая операция $\odot$, a^{-1} — обратный к $a \in \mathcal{A}$ элемент группы. Шифрование открытого текста $X = (x_0, \ldots, x_{T-1}) \in \mathcal{A}^T$ на ключе θ выполняется по правилу $y_t = x_t \odot g_t$, расшифровывание — $x_t = y_t \odot g_t^{-1}$, $t = 0, \ldots, T-1$. Здесь $g_0, \ldots, g_{T-1}$ — начальный отрезок последовательности G_θ.

Если символы $g_0, \ldots, g_{T-1}$ выбираются случайно, независимо и равновероятно, то шифртекст $Y = (y_0, \ldots, y_{T-1})$ не содержит информации об открытом тексте X и криптосистема G является *совершенно криптостойкой* по Шеннону (см. п. 6.5). Как уже отмечалось, обеспечение условия идеальной криптостойкости предполагает распространение одноразовых ключей, длина которых совпадает с длиной обрабатываемых данных, что зачастую трудно осуществимо. На практике *генератор* последовательностей G_θ криптосистемы G состоит из одного или нескольких генераторов рекуррентных последовательностей, начальные значения которых определяются ключом θ, а выходные символы комбинируются для получения символов G_θ [173].

В отличие от блочных криптосистем, слова $X_t = (x_{n(t-1)}, \ldots, x_{nt-1})$ открытого текста обрабатываются, вообще говоря, различными для разных $t = 1, 2, \ldots$ подстановками $F_{\theta,t} \in \mathcal{S}(\mathcal{A}^n)$. Использование режимов поточного шифрования (см. 10.1.3) является простым способом преобразования блочной системы в поточную.

Атаки при известных, выбранном и выбираемом открытом текстах для поточных криптосистем эквивалентны. В условиях этих атак криптоаналитику становится известен отрезок $g_0, \ldots, g_{T-1}$ последовательности G_θ. Требуется решить одну из следующих задач:

S1) определить ключ θ;

S2) не определяя θ, построить алгоритм вычисления последующих значений $g_T, g_{T+1}, \ldots$;

S3) удостовериться, что наблюдается отрезок последовательности из G.

9.2. Рекуррентные последовательности

Пусть задана функция $f(x_1, \ldots, x_n)\colon \mathcal{A}^n \to \mathcal{A}$, которая порождает *рекуррентную последовательность s порядка n* с элементами

$$s_{t+n} = f(s_{t+n-1}, \ldots, s_t), \quad t = 0, 1, \ldots$$

при заданных начальных значениях $s_0, \ldots, s_{n-1}$.

Всякая рекуррентная последовательность является периодической — существуют числа $t_0 \in \mathbb{N}_0$ и $r \in \mathbb{N}$ такие, что $s_{t+r} = s_t$ для всех $t \geq t_0$. Минимальное r с данным свойством называется *минимальным периодом*, а минимальное t_0 — *предпериодом*. Если $t_0 = 0$, то s — *чисто периодическая* последовательность.

Очевидно, что $r \leq |\mathcal{A}|^n$. Если $r = |\mathcal{A}|^n$, то s называется *последовательностью де Брейна*. *Векторы состояния* $S_0 = (s_0, \ldots, s_{n-1})$, $S_1 = (s_1, \ldots, s_n)$, $\ldots$, $S_{r-1} = (s_{r-1}, \ldots, s_{r+n-2})$ такой последовательности пробегают все множество $\mathcal{A}^n$, $S_r = S_0$.

Далее в главе будем рассматривать бинарный алфавит $\mathcal{A} = \mathbb{F}_2$ и считать, что $\odot = \oplus$, $f \in \mathcal{F}_n$.

9.3. Линейные рекуррентные последовательности

Линейная булева функция $f \in \mathcal{F}_n$ порождает *линейную рекуррентную последовательность* (ЛРП), в которой

$$s_{t+n} = \alpha_{n-1} s_{t+n-1} \oplus \ldots \oplus \alpha_0 s_t, \quad t = 0, 1, \ldots$$

или

$$s_t = S_{t,1}, \quad S_{t+1} = S_t A,$$

где $S_{t,1}$ — первая компонента вектора $S_t \in V_n$,

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & \alpha_0 \\ 1 & 0 & \ldots & 0 & \alpha_1 \\ 0 & 1 & \ldots & 0 & \alpha_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & \alpha_{n-1} \end{pmatrix}.$$

Далее полагаем, что вектор S_0 начального состояния ЛРП является ненулевым.

Характеристический многочлен

$$p(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_0$$

матрицы A называется *характеристическим многочленом* ЛРП $s_0, s_1, \ldots$; A при этом называется *сопровождающей матрицей* $p(\lambda)$.

Последовательность $s_0, s_1, \ldots$ удовлетворяет целому семейству линейных рекуррентных соотношений, например, всякому соотношению, заданному характеристическим многочленом $p(\lambda)\,q(\lambda)$, $q(\lambda) \in \mathbb{F}_2[\lambda]$, $q(0) = 1$. При этом однозначно определен *минимальный многочлен* $m(\lambda)$ такой, что всякий характеристический многочлен последовательности $s_0, s_1, \ldots$ делится на $m(\lambda)$. Известно, что минимальный период линейной рекуррентной последовательности совпадает с порядком ее минимального многочлена.

Для криптографических приложений важность представляют ЛРП с примитивными характеристическими многочленами, называемые *последовательностями максимального периода*, или *m-последовательностями*. В табл. 9.1 приведен список всех примитивных триномов степени $n \le 100$.

Таблица 9.1

Примитивные триномы $\lambda^n + \lambda^m + 1$, $1 \le n \le 100$, $1 \le m \le n/2$

n	m	n	m	n	m	n	m
2	1	20	3	41	3, 20	73	25, 28, 31
3	1	21	2	47	5, 14, 20, 21	79	9, 19
4	1	22	1	49	9, 12, 15, 22	81	4, 16, 35
5	2	23	5, 9	52	3, 7, 19, 21	84	13
6	1	25	3, 7	55	24	87	13
7	1, 3	28	3, 9, 13	57	7, 22	89	38
9	4	29	2	58	19	93	2
10	3	31	3, 6, 7, 13	60	1, 11	94	21
11	2	33	13	63	1, 5, 31	95	11, 17
15	1, 4, 7	35	2	65	18, 32	97	6, 12, 33, 34
17	3, 5, 6	36	11	68	9, 33	98	11, 27
18	3, 7, 9	39	4, 8, 14	71	6, 9, 18, 20, 35	100	37

Использование триномов удобно при генерации ЛРП, так как для вычисления нового элемента последовательности требуется выполнить всего одно сложение $\oplus$. В табл. 9.1 приведен один элемент пары «многочлен $p(\lambda)$ — *возвратный многочлен* $p^*(\lambda) = \lambda^n p(1/\lambda)$», поскольку $p(\lambda)$ и $p^*(\lambda)$ являются примитивными одновременно.

На рис. 9.1, *а* изображена схема устройства для вычисления элементов линейной рекуррентной последовательности с характеристическим многочленом $\lambda^8 + \lambda^4 + \lambda^3 + \lambda^2 + 1$, называемого *регистром сдвига с обратной связью*. В *триггерах* 0, 1, ..., 7 такого регистра в момент времени t содержатся соответственно компоненты $s_t, s_{t+1}, \ldots, s_{t+7}$ вектора состояния S_t.

При $\alpha_0 \neq 0$ характеристический многочлен $\tilde{p}(\lambda)$ матрицы

$$\tilde{A} = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{n-1} & \alpha_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

является возвратным к характеристическому многочлену $p(\lambda)$ матрицы A. Регистр сдвига, вырабатывающий последовательность

$$\tilde{s}_t = \tilde{S}_{t,1}, \quad \tilde{S}_{t+1} = \tilde{S}_t \tilde{A}, \quad t = 0, 1, \dots,$$

при выборе $\tilde{p}(\lambda) = \lambda^8 + \lambda^5 + \lambda^4 + \lambda^3 + 1$, представлен на рис. 9.1, *б*. Программная реализация такого регистра может оказаться эффективнее стандартной реализации (рис. 9.1, *а*).

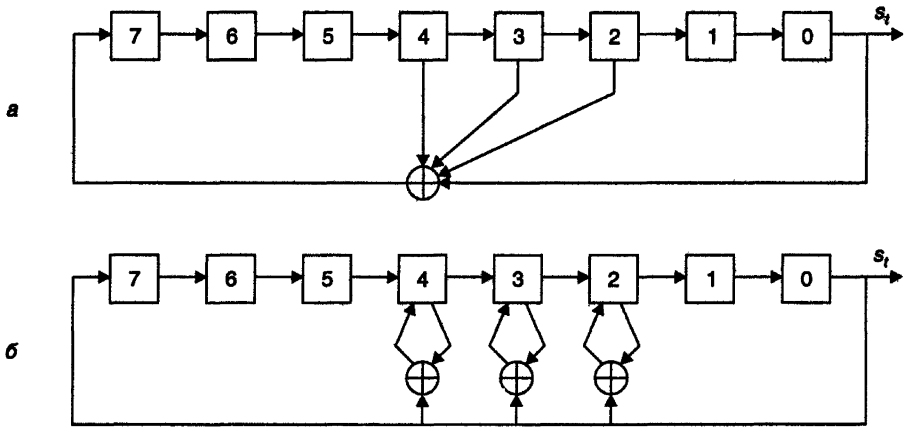


Рис. 9.1. Линейные регистры сдвига

9.4. Оценка параметров и распознавание ЛРП

Если $s_0, s_1, \dots$ — линейная рекуррентная последовательность с минимальным многочленом $m(\lambda)$ степени n, то *ганкелев определитель*

$$D_t^{(k)} = \begin{vmatrix} s_t & s_{t+1} & s_{t+k-1} \\ s_{t+1} & s_{t+2} & s_{t+k} \\ \dots & \dots & \dots \\ s_{t+k-1} & s_{t+k} & s_{t+2k-2} \end{vmatrix}$$

равен 0 для всех $k \geq n+1$ и всех $t \geq 0$. С другой стороны, если $s_0, s_1, \ldots$ — последовательность независимых случайных величин, $0 < \mathbf{P}\{s_t = 0\} < 1$, то

$$\mathbf{P}\{D_0^{(k)} = \ldots = D_{T-1}^{(k)} = 0\} \to 0$$

при $T \to \infty$ для любого фиксированного натурального k. Данное свойство можно использовать для построения алгоритмов распознавания ЛРП.

Алгоритм Берлекэмпа — Месси позволяет по $2k$-отрезку $s_0, \ldots, s_{2k-1}$ линейной рекуррентной последовательности определить минимальный многочлен $m(\lambda)$, если степень последнего не превосходит k.

При выполнении алгоритма формируются последовательности многочленов $m_0^*(\lambda) = 1$, $m_1^*(\lambda), \ldots, m_{2k}^*(\lambda)$ и целых чисел $l_0 = 0, l_1, \ldots, l_{2k}$, а именно для $t = 0, \ldots, 2k-1$:

1) вычисляется невязка

$$e_t = s_t \oplus \alpha_{t,1} s_{t-1} \oplus \ldots \oplus \alpha_{t,n_t} s_{t-l_t},$$

где $\alpha_{t,i}$ — коэффициент при степени λ^i в $m_t^*(\lambda)$;

2) вычисляется многочлен

$$m_{t+1}^*(\lambda) = \begin{cases} m_t^*(\lambda) + m_\tau^*(\lambda)\lambda^{t-\tau}, & \text{если } e_t = 1, \\ m_t^*(\lambda) & \text{в противном случае,} \end{cases}$$

где τ — максимальное число из множества $\{-1, 0, 1, \ldots, t-1\}$ такое, что $l_\tau < l_t$ (l_{-1} считается равным -1, $m_{-1}^*(\lambda) = 1$);

3) вычисляется

$$l_{t+1} = \begin{cases} t + 1 - l_t, & \text{если } e_t = 1 \text{ и } l_t \leq t/2, \\ l_t & \text{в противном случае.} \end{cases}$$

Искомый минимальный многочлен

$$m(\lambda) = \lambda^{l_{2k}} m_{2k}^*(1/\lambda).$$

9.5. Линейная сложность

Пусть s — произвольная последовательность $s_0, s_1, \ldots$ элементов поля $\mathbb{F}_2$. *Линейной сложностью* $l_t(s)$ начального t-отрезка последовательности s называется минимальное целое n такое, что $s_0, \ldots, s_{t-1}$ — первые t элементов линейной рекуррентной последовательности порядка n (если $s_0 = \ldots = s_{t-1} = 0$, то $l_t(s) = 0$). Линейная сложность $l(s)$ всей последовательности s определяется как $l_\infty(s)$. Если линейная сложность $l(G_\theta)$ последовательностей G_θ криптосистемы G невелика, то, применив алгоритм Берлекэмпа — Месси, можно эффективно решить задачу S2 криптоанализа G.

Известно [173], что для последовательности $s_0, s_1, \ldots$ независимых случайных величин, принимающих значения 0 и 1 с равной вероятностью, выполняется

$$\mathbf{E}\{l_t(s)\} = \frac{t}{2} + \frac{4 + b_t}{18} - \frac{1}{2^t}\left(\frac{1}{3}t + \frac{2}{9}\right),$$

$$\mathbf{D}\{l_t(s)\} = \frac{86}{81} - \frac{1}{2^t}\left(\frac{14 - b_t}{27}t + \frac{82 - 2b_t}{81}\right) - \frac{1}{2^{2t}}\left(\frac{1}{9}t^2 + \frac{4}{27}t + \frac{4}{81}\right),$$

где $b_t = t \bmod 2$. Значения моментов статистик $l_t(s)$ используются в тесте для проверки гипотезы о независимости и равновероятности символов $s_0, s_1, \ldots$ (см. п. 7.15).

Последовательность $l_1(s), l_2(s), \ldots$ называется *профилем линейной сложности s*. Такой профиль совпадает с формируемой при выполнении алгоритма Берлекэмпа — Месси последовательностью $l_1, l_2, \ldots$

9.6. Определение начального состояния ЛРП

Пусть ЛРП $s_0, s_1, \ldots$ с характеристическим многочленом

$$p(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_1\lambda + 1,$$

содержащим $(q + 1)$ ненулевых коэффициентов, непосредственно не наблюдается. Криптоаналитику становится известной *искаженная последовательность* [182]

$$z_t = s_t \oplus \xi_t, \quad t = 0, \ldots, T - 1.$$

Здесь $\xi_0, \ldots, \xi_{T-1}$ — реализация схемы независимых испытаний, в которой событие $\{\xi_0 = 0\}$ наступает с вероятностью $(1 + \varepsilon)/2$, $0 < \varepsilon < 1$. Требуется определить вектор начального состояния $S_0 \in V_n$.

Криптоаналитик выбирает всевозможные векторы начального состояния $\hat{S}_0 \in V_n \setminus \{0\}$ и строит ЛРП $\hat{s}_0, \hat{s}_1, \ldots, \hat{s}_{T-1}$. В качестве оценки S_0 выбирается вектор $\hat{S}_0$, доставляющий минимум сумме

$$\sum_{t=0}^{T-1} 1\{\hat{s}_t = z_t\}.$$

Вероятность ошибочного решения близка к нулю, если

$$\frac{n}{T} < 1 - h_2\left(\frac{1 + \varepsilon}{2}\right),$$

где $h_2(\cdot)$ — *двоичная энтропия* (см. главу 5);

$$h_2(x) = -x\log_2(x) - (1 - x)\log_2(1 - x), \quad 0 \le x \le 1, \quad 0\log_2(0) = 0.$$

Эффективность атаки можно повысить, если q мало. Используя задание 11, для чисел t, $k = 0, 1, \ldots, t + 2^k n < T$, запишем линейные равенства

$$s_t \oplus \alpha_1 s_{t+2^k} \oplus \ldots \oplus \alpha_{n-1} s_{t+2^k(n-1)} \oplus s_{t+2^k n} = 0,$$

связывающие элементы ЛРП. Если $z_t = s_t$, т.е. $\xi_t = 0$, то соотношение

$$z_t \oplus \alpha_1 z_{t+2^k} \oplus \ldots \oplus \alpha_{n-1} z_{t+2^k(n-1)} \oplus z_{t+2^k n} = 0$$

выполняется с вероятностью

$$\mathbf{P}\{\alpha_1 \xi_{t+2^k} \oplus \ldots \oplus \alpha_{n-1} \xi_{t+2^k(n-1)} \oplus \xi_{t+2^k} = 0\} = \frac{1}{2}(1 + \varepsilon^q).$$

Если же $z_t \neq s_t$, то такое соотношение выполняется с вероятностью $(1 - \varepsilon^q)/2$.

Пусть наблюдение z_t входит в R соотношений, r из которых выполняются. При таком условии вероятность β_t совпадения символов z_t и s_t равна

$$\beta_t = \frac{(1+\varepsilon)(1+\varepsilon^q)^r(1-\varepsilon^q)^{R-r}}{(1+\varepsilon)(1+\varepsilon^q)^r(1-\varepsilon^q)^{R-r} + (1-\varepsilon)(1-\varepsilon^q)^r(1+\varepsilon^q)^{R-r}}.$$

В ходе описанной в работе [141] атаки выбираются элементы $z_{t_1}, \ldots, z_{t_m}$, $m \geq n$, наблюдаемой последовательности, соответствующие максимальным значениям вероятностей $\beta_{t_1}, \ldots, \beta_{t_m}$. Задается порог Δ, $0 < \Delta < 1$ определяются всевозможные векторы невязок $(e_1, \ldots, e_m)$, для которых

$$\prod_{i=1}^{m} \beta_{t_i}^{1-e_i}(1-\beta_{t_i})^{e_i} \geq \Delta,$$

и составляется система линейных уравнений

$$\hat{s}_{t_i} = \hat{s}_{t_i}(\hat{S}_0) = z_{t_i} \oplus e_i, \quad i = 1, \ldots, m,$$

относительно неизвестного вектора $\hat{S}_0$. Искомая оценка начального состояния S_0 выбирается среди решений получаемых систем.

9.7. Комбинирование последовательностей

Опишем основные способы комбинирования выходных рекуррентных последовательностей $s_0^{(i)}, s_1^{(i)}, \ldots$генераторов $G^{(i)}$, $i = 1, \ldots, d$, для получения последовательности $G_\theta = (g_0, g_1, \ldots)$ поточной криптосистемы G. Стартовые значения $S_0^{(i)}$ генераторов $G^{(i)}$ определяются ключом θ крипосистемы.

При *параллельном подключении* (рис. 9.2) выходные символы $G^{(i)}$ обрабатываются функцией $g \in \mathcal{F}_d$:

$$g_t = g(s_t^{(1)}, \ldots, s_t^{(d)}), \quad t = 0, 1, \ldots$$

Например, для *генератора Геффе* [95] $d = 3$ и

$$g(x_1, x_2, x_3) = x_1 x_2 \oplus (x_1 \oplus 1) x_3.$$

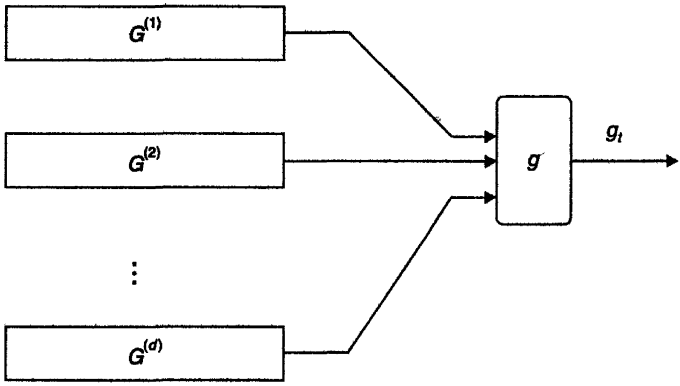


Рис. 9.2. Параллельное подключение генераторов

Выходной символ $s_t^{(1)}$ генератора $G^{(1)}$ определяет выбор между выходными символами $s_t^{(2)}$ и $s_t^{(3)}$ двух других генераторов.

Если $G^{(i)}$ — генераторы m-последовательностей порядка n_i, $i = 1, \ldots, d$, то линейную сложность последовательности $g_0, g_1, \ldots$ можно определить, вычислив значение многочлена Жегалкина $g(x_1, \ldots, x_d)$ при значениях n_i переменных x_i, выполняя умножение и сложение в кольце $\mathbb{Z}$, а не в поле $\mathbb{F}_2$. Например, линейная сложность выходных последовательностей генератора Геффе равна $n_1 n_2 + n_1 n_3 + n_3$.

В *пороговом генераторе* [63] $d = 2k + 1$ — нечетное число. Используется *пороговая функция* g: $g(x) = 1$, только если $\mathrm{w}(x) \geq k + 1$, где $\mathrm{w}(x)$ — вес Хэмминга (количество ненулевых координат) вектора x. Таким образом, $g_t = 1$, только если среди выходных символов $s_t^{(1)}, \ldots, s_t^{(d)}$ единиц больше, чем нулей.

При *последовательном подключении* выходные символы генератора $G^{(i)}$ управляют выполнением преобразований на генераторе $G^{(i+1)}$. Он может либо выполнять стандартное рекуррентное преобразование (умножение вектора состояния $S_t^{(i+1)}$ на сопровождающую матрицу характеристического многочлена $A^{(i+1)}$, *шаг*), либо простаивать ($S_{t+1}^{(i)} = S_t^{(i)}$, *стоп*). Например, в *каскаде Голлманна* [65] (рис. 9.3) генератор $G^{(1)}$ вырабатывает обычную линейную рекуррентную последовательность. Функционирование генераторов $G^{(i)}$, $i = 2, \ldots, d$ осуществляется по правилам

$$s_t^{(i)} = s_t^{(i-1)} \oplus S_{t,1}^{(i)}, \quad S_{t+1}^{(i)} = \begin{cases} S_t^{(i)} A^{(i)}, & \text{если } s_t^{(i-1)} = 1, \\ S_t^{(i)}, & \text{если } s_t^{(i-1)} = 0 \end{cases}$$

и $g_t = s_t^{(d)}$, $t = 0, 1, \ldots$

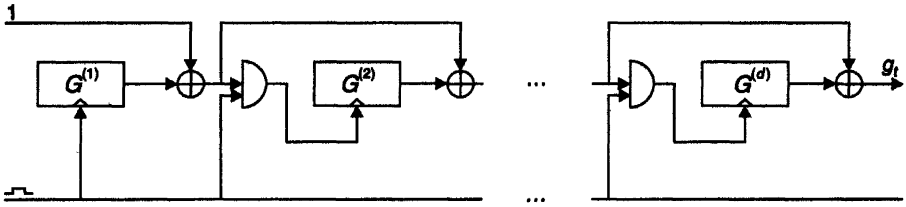


Рис. 9.3. Каскад Голлманна

Переключающий шаг-стоп генератор [98] (рис. 9.4) использует параллельное и последовательные подключения. Здесь $d = 3$,

$$S_{t+1}^{(2)} = \begin{cases} S_t^{(2)} A^{(2)}, & s_t^{(1)} = 1, \\ S_t^{(2)}, & s_t^{(1)} = 0, \end{cases} \quad S_{t+1}^{(3)} = \begin{cases} S_t^{(3)}, & s_t^{(1)} = 1, \\ S_t^{(3)} A^{(3)}, & s_t^{(1)} = 0 \end{cases}$$

и $g_t = s_t^{(2)} \oplus s_t^{(3)}$.

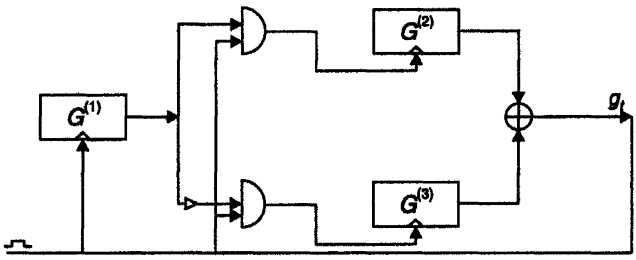


Рис. 9.4. Переключающий шаг-стоп генератор

Другой способ последовательного подключения — выходная последовательность одного генератора управляет выбором выходных символов другой. Например, для *прореживающего генератора* [72] $d = 2$,

$$g_t = s_{\tau_t}^{(2)},$$

где τ_t — номер t-й единицы в последовательности $s_0^{(1)}, s_1^{(1)}, \ldots$

9.8. Корреляционный криптоанализ

Пусть в поточной криптосистеме G используется d генераторов $G^{(1)}, \ldots, G^{(d)}$. Если при случайном выборе ключа θ символы $g_0, g_1, \ldots$ наблюдаемой последовательности G_θ коррелируют с символами ненаблюдаемой выходной последо-

вательности $s_0^{(1)}, s_1^{(1)}, \ldots$ генератора $G^{(1)}$, то возможно определение части ключа θ — вектора $S_0^{(1)}$ начального состояния генератора $G^{(1)}$ [182]. Действительно, пусть для простоты $G^{(1)}$ — генератор ЛРП и

$$\mathbf{P}\{\xi_t = 0\} = \frac{1+\varepsilon}{2}, \quad \text{где } \xi_t = g_t \oplus s_t^{(1)}, \quad 0 < \varepsilon < 1.$$

Применяя описанные в п. 9.6 методы, определяем вектор $S_0^{(1)}$ начального состояния искаженной ЛРП $s_0^{(1)}, s_1^{(1)}, \ldots$ Если

$$\mathbf{P}\{g_t = s_t^{(2)}\} \neq \frac{1}{2},$$

то можно определить вектор $S_0^{(2)}$ начального состояния генератора $G^{(2)}$ и т.д. Такая процедура иллюстрирует важный криптоаналитический принцип — *«разделяй и властвуй»*.

Рассмотрим применение корреляционных методов для криптоанализа *генератора Беса — Пайпера* [51] (рис. 9.5), в котором:

а) $d = 3$;

б) $G^{(1)}, G^{(3)}$ — генераторы m-последовательностей с известными примитивными характеристическими многочленами $p_1(\lambda)$ и $p_3(\lambda)$;

в) многочлен $p_3(\lambda)$ содержит нечетное количество ненулевых коэффициентов;

г) $s_t^{(2)} = S_{t,1}^{(2)}$, $S_{t+1}^{(2)} = S_t^{(2)} A^{(2)}$, если $s_t^{(1)} = 1$; $S_{t+1}^{(2)} = S_t^{(2)}$, если $s_t^{(1)} = 0$;

д) $A^{(2)}$ — сопровождающая матрица примитивного многочлена $p_2(\lambda)$;

е) $g_t = s_t^{(2)} \oplus s_t^{(3)}$.

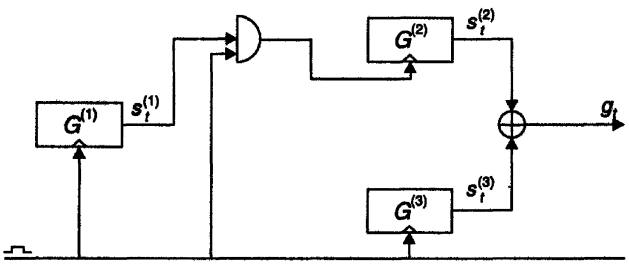


Рис. 9.5. Генератор Беса — Пайпера

Описанная в работе [195] корреляционная атака позволяет по последовательности $g_0, g_1, \ldots$ генератора Беса — Пайпера определить векторы $S_0^{(1)}$, $S_0^{(2)}$, $S_0^{(3)}$ и многочлен $p_2(\lambda)$. Атака проводится следующим образом (для упрощения выкладок считаем, что в m-последовательностях символы 0 и 1 встречаются одинаково часто).

1. По наблюдениям $g_0 \oplus g_1$, $g_1 \oplus g_2$, ... определяется линейная рекуррентная последовательность $s_0^{(3)} \oplus s_1^{(3)}$, $s_1^{(3)} \oplus s_2^{(3)}$, ... Это возможно, так как

$$\mathbf{P}\{s_t^{(3)} \oplus s_{t+1}^{(3)} = g_t \oplus g_{t+1}\} = \mathbf{P}\{s_t^{(2)} = s_{t+1}^{(2)}\} = \frac{3}{4} \neq \frac{1}{2}.$$

2. По последовательности $s_0^{(3)} \oplus s_1^{(3)}$, $s_1^{(3)} \oplus s_2^{(3)}$, ... определяется последовательность $s_0^{(3)}$, $s_1^{(3)}$, ... (см. задание 1б).

3. Определяется последовательность $s_t^{(2)} = g_t \oplus s_t^{(3)}$, $t = 0, 1, \ldots$

4. Строится последовательность

$$\hat{s}_t^{(1)} = g_t \oplus s_t^{(3)} \oplus g_{t+1} \oplus s_{t+1}^{(3)}, \quad t = 0, 1, \ldots$$

5. По последовательности $\hat{s}_0^{(1)}$, $\hat{s}_1^{(1)}$, ...определяется линейная рекуррентная последовательность $s_0^{(1)}$, $s_1^{(1)}$, ...генератора $G^{(1)}$. Это возможно, поскольку

$$\mathbf{P}\{s_t^{(1)} \neq \hat{s}_t^{(1)}\} = \mathbf{P}\{s_t^{(1)} \neq s_t^{(2)} \oplus s_{t+1}^{(2)}\} =$$

$$= \mathbf{P}\{s_t^{(1)} = 1, s_t^{(2)} = s_{t+1}^{(2)}\} = \frac{1}{4} \neq \frac{1}{2}.$$

6. В последовательности $s_0^{(2)}$, $s_1^{(2)}$, ... удаляются повторяющиеся при $s_t^{(1)} = 0$ символы $s_{t+1}^{(2)}$. К полученной последовательности применяется алгоритм Берлекэмпа — Месси и определяется многочлен $p_2(\lambda)$.

9.9. Задания

1. (п. 9.2). Пусть функция $f(x_1, \ldots, x_n)$ порождает последовательность де Брейна. Доказать:
   а) $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_{n-1}) \oplus x_n$, где $g \in \mathcal{F}_{n-1}$, причем $g(0, \ldots, 0) = g(1, \ldots, 1) = 1$;
   б) функции

$$g(x_{n-1}, \ldots, x_1) \oplus x_n,$$
$$g(x_1 \oplus 1, \ldots, x_{n-1} \oplus 1) \oplus x_n,$$
$$g(x_1, \ldots, x_{n-1}) \oplus x_1 \cdots x_{n-2} \oplus x_2 \cdots x_{n-1} \oplus x_n,$$
$$g(x_1, \ldots, x_{n-1}) \oplus (x_1 \oplus x_{n-1})(x_2 \oplus 1) \cdots (x_{n-2} \oplus 1) \oplus x_n$$

   также порождают последовательности де Брейна;
   в) функция $f(x_1, \ldots, x_n) \oplus (x_1 \oplus 1) \cdots (x_{n-1} \oplus 1)$ порождает последовательность, векторы состояния которой пробегают все множество $V_n \setminus \{0\}$.

2. (п. 9.2). Найти все функции $f \in \mathcal{F}_n$, порождающие последовательности де Брейна порядка $n = 2, 3, 4$. Проверить, что количество таких функций равно $2^{2^{n-1}-n}$.

3. (п. 9.3). Доказать, что при $\alpha_0 \neq 0$:
а) ЛРП $s_0, s_1, \ldots$ является чисто периодической последовательностью;
б) минимальный период ЛРП $s_0, s_1, \ldots$ делит порядок матрицы A как элемента *полной линейной группы* $GL_n(\mathbb{F}_2)$ (мультипликативной группы всех невырожденных $n \times n$ матриц над полем $\mathbb{F}_2$);
в) если $n > 1$ и $s_0 = \ldots = s_{n-2} = 0$, $s_{n-1} = 1$, то минимальный период ЛРП $s_0, s_1, \ldots$ совпадает с порядком матрицы A в группе $GL_n(\mathbb{F}_2)$.

4. (п. 9.3). Доказать, что:
а) многочлен $p(\lambda) \in \mathbb{F}_2[\lambda]$ делит многочлен $\lambda^r - 1$ тогда и только тогда, когда $\operatorname{ord} p(\lambda) \mid r$;
б) если $p(\lambda)$ — неприводимый многочлен степени n, то $\operatorname{ord} p(\lambda) \mid (2^n - 1)$ (просуммировать периоды всех ЛРП с характеристическим многочленом $p(\lambda)$);
в) если $2^n - 1$ — простое число Мерсенна, то всякий неприводимый многочлен $p(\lambda)$ степени n является примитивным.

5. (п. 9.3). Написать программу, которая по заданным характеристическому многочлену $p(\lambda)$ и вектору начального состояния S_0 строит ЛРП заданной длины T.

6. (п. 9.4). Разработать алгоритм и написать программу оценки снизу порядка заданной линейной рекуррентной последовательности. Использовать ганкелевы определители.

7. (п. 9.4). Написать программу определения минимального многочлена заданной линейной рекуррентной последовательности.

8. (п. 9.5). Доказать:
а) $l_t(s) \leq t$, причем $l_t(s) = t$, только если s_{t-1} — единственный ненулевой элемент в начальном t-отрезке s;
б) если s — периодическая последовательность с периодом r, то $l(s) \leq r$;
в) $l_{t+1}(s) > l_t(s)$, только если $l_t(s) \leq t/2$, при этом $l_{t+1}(s) + l_t(s) = t + 1$.

9. (п. 9.5). Написать программу вычисления профиля линейной сложности заданной последовательности s. Вычислить профиль последовательности

$$s_t = \begin{cases} 1, & \text{если } t = 0, 1, 2^2 - 1, 2^3 - 1, \ldots, \\ 0 & \text{в противном случае.} \end{cases}$$

Сравнить с теоретическим результатом [173]: $l_t(s) = \lfloor (t+1)/2 \rfloor$.

10. (п. 9.5). Известно [94], что если $r = 2^n$ — период s, $n > 0$,

$$L(s) = (s_0, \ldots, s_{r/2-1}, s_0, \ldots), \quad R(s) = (s_{r/2}, \ldots, s_{r-1}, s_{r/2+1}, \ldots),$$

то

$$l(s)=\begin{cases} l(L(s)), & L(s)=R(s),\\ 2^{n-1}+l(L(s)\oplus R(s)), & L(s)\neq R(s).\end{cases}$$

Используя данный факт, разработать алгоритм и программу вычисления линейной сложности последовательностей де Брейна порядка n. Найти последовательности де Брейна с максимальной и минимальной линейной сложностью для $n = 2, 3, 4$. Проверить, что при $n \geq 3$ справедлива оценка

$$2^{n-1}+n\leq l(s)\leq 2^n-1.$$

11. (п. 9.6). Доказать, что $(p(\lambda))^{2^k}=p(\lambda^{2^k})$ и

$$s_{t+2^k n}=\alpha_{n-1}s_{t+2^k(n-1)}\oplus\ldots\oplus\alpha_1 s_{t+2^k}\oplus\alpha_0 s_t$$

для любых t, $k = 0, 1, 2, \ldots$

12. (п. 9.6). Написать программу определения начального состояния искаженной ЛРП. Положить $n = 20$, $p(\lambda)=\lambda^{20}+\lambda^3+1$, $T = 2000$, $\varepsilon=\frac{3}{8}$.

13. (п. 9.7). Пороговая функция $g(x_1,\ldots,x_d)$ *симметрическая*, ее значение не изменяется при перестановке значений переменных. Найти вид многочлена Жегалкина для g при $d = 3, 5, 7$. Доказать, что

$$\mathbf{P}\{g(x_1,\ldots,x_d)=x_1\}=\frac{1}{2}+\frac{1}{2^d}\binom{d-1}{\frac{d-1}{2}}$$

для случайного вектора $(x_1,\ldots,x_d)$ с равномерным на V_d распределением.

14. (п. 9.7). Пусть в G используется d генераторов $G^{(i)}$ m-последовательностей порядка n_i с минимальным периодом $r_i=2^{n_i}-1$, $i=1,\ldots,d$. Пусть r — минимальный период внутреннего состояния G, т.е. минимальное натуральное число, для которого

$$(S_r^{(1)},\ldots,S_r^{(d)})=(S_0^{(1)},\ldots,S_0^{(d)}).$$

Доказать:

а) $r=\text{НОК}(r_1,\ldots,r_d)$ при параллельном подключении генераторов;

б) $r=r_1\,\text{НОК}(r_2, r_3/(2^m-1))$, $m=\text{НОД}(n_1-1, n_3)$ для переключающего шаг-стоп генератора;

в) $r=\text{НОК}((r_1+1)/2, r_2)$ для прореживающего генератора.

15. (п. 9.7). Разработать и реализовать программно пороговую криптосистему G, используя комбинацию из пяти генераторов ЛРП. Оценить период и линейную сложность последовательностей G.

16. (п. 9.8). Доказать, что на втором этапе атаки на генератор Беса — Пайпера последовательность $s_0^{(3)}, s_1^{(3)}, \ldots$ определяется однозначно. Использовать тот факт, что многочлен $p_3(\lambda)$ содержит нечетное количество q ненулевых коэффициентов. Предложить модификацию атаки для случая четного q.

17. (п. 9.8). Используя результат задания 13, провести корреляционную атаку на разработанную в задании 15 пороговую криптосистему.

Глава 10

МАТЕМАТИЧЕСКИЕ МОДЕЛИ СТАНДАРТНЫХ БЛОЧНЫХ КРИПТОСИСТЕМ

SXJKN8EYHR3I4HR987YD
WFUIHW9E8YR384H5O345
96EUF9O8WE76R83Y4398
46UHFE98YDSFJH8374H4
BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7E
RFH2347FQ7WERF9UJ134
9YUR7DFKJHGJDFBVDGFJ
VBDSKFJHKJFVKBDFKJVJ
KGFKJSBDFJBVKJDSGFJK
VBKSDJBFKVBDGFVAKDLK
WEF90UQ3I4OUJHRP8937

10.1. Криптосистема DES и ее свойства

В настоящее время разработано большое количество *симметричных криптосистем*, многие из которых являются национальными или ведомственными стандартами. Наибольшую известность приобрели системы DES, ГОСТ 28147—89, IDEA, Rijndael и некоторые другие. Они находятся под пристальным вниманием современных криптологов, основной задачей которых является поиск «слабостей» этих и других криптосистем, а также изучение их свойств. Этим вопросам посвящены многие журнальные публикации и книги (см. список литературы). Ниже дано достаточно полное математическое описание перечисленных стандартов, обоснована корректность их работы.

Стандарт DES впервые был опубликован в США в 1975 г. Он разработан корпорацией IBM путем модификации более ранней версии, известной под названием LUCIFER. В настоящее время появились дальнейшие его модификации. Это объясняется тем, что DES, как никакой другой стандарт, не изучался столь пристально криптоаналитиками.

10.1.1. Описание стандарта DES

В криптосистеме DES используется блочный принцип шифрования двоичного текста. Длина блока шифрования составляет 64 бита. Размер ключа — также 64 бита. При этом каждый восьмой бит является служебным и в шифровании не участвует. Каждый такой бит является двоичной суммой семи предыдущих и служит лишь для обнаружения ошибок при передаче ключа по каналу связи.

Процесс криптопреобразования включает следующие три основных этапа.

1. Биты исходного сообщения x подвергаются начальной подстановке IP в соответствии с табл. 10.1.

Таблица 10.1

Подстановка *IP*

IP	58	50	42	34	26	18	10	2	60	52	44	36	28	20	12	4
	62	54	46	38	30	22	14	6	64	56	48	40	32	24	16	8
	57	49	41	33	25	17	9	1	59	51	43	35	27	19	11	3
	61	53	45	37	29	21	13	5	63	55	47	39	31	23	15	7

Это значит, что 58-й бит становится первым, 50-й — вторым и т.д. Затем полученный вектор $x_0 = IP(x)$ представляется в виде $x_0 = L_0R_0$, где L_0 — левая половина из 32 битов, а R_0 — правая половина из 32 битов.

2. Сообщение L_0R_0 преобразуется далее 16 раз по так называемой схеме Фейстеля:

$$L_i = R_{i-1}, \quad R_i = L_{i-1} \oplus f(R_{i-1}, K_i), \quad i = 1, 2, \ldots, 16,$$

где функция f и расписание ключей $K_1, K_2, \ldots, K_{16}$ будут описаны отдельно (рис. 10.1).

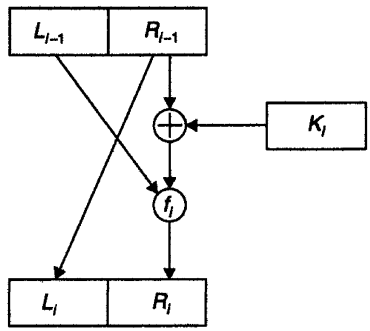


Рис. 10.1. Криптопреобразование Фейстеля

3. Сообщение $L_{16}R_{16}$ перемешивается подстановкой IP^{-1}:

$$y = IP^{-1}(R_{16}L_{16}),$$

где y — зашифрованное сообщение.

Шифрование осуществляется по схеме, приведенной на рис. 10.2.

Функция f. Она имеет два аргумента: A и B. Первый состоит из 32 битов, а второй — из 48 битов. Результат состоит из 32 битов.

1. Аргумент A, имеющий 32 бита, преобразуется в 48-битовый вектор $P(A)$ путем перестановки с повторениями исходного вектора A. Эта процедура одинакова для всех тактов. Она задана в табл. 10.2.

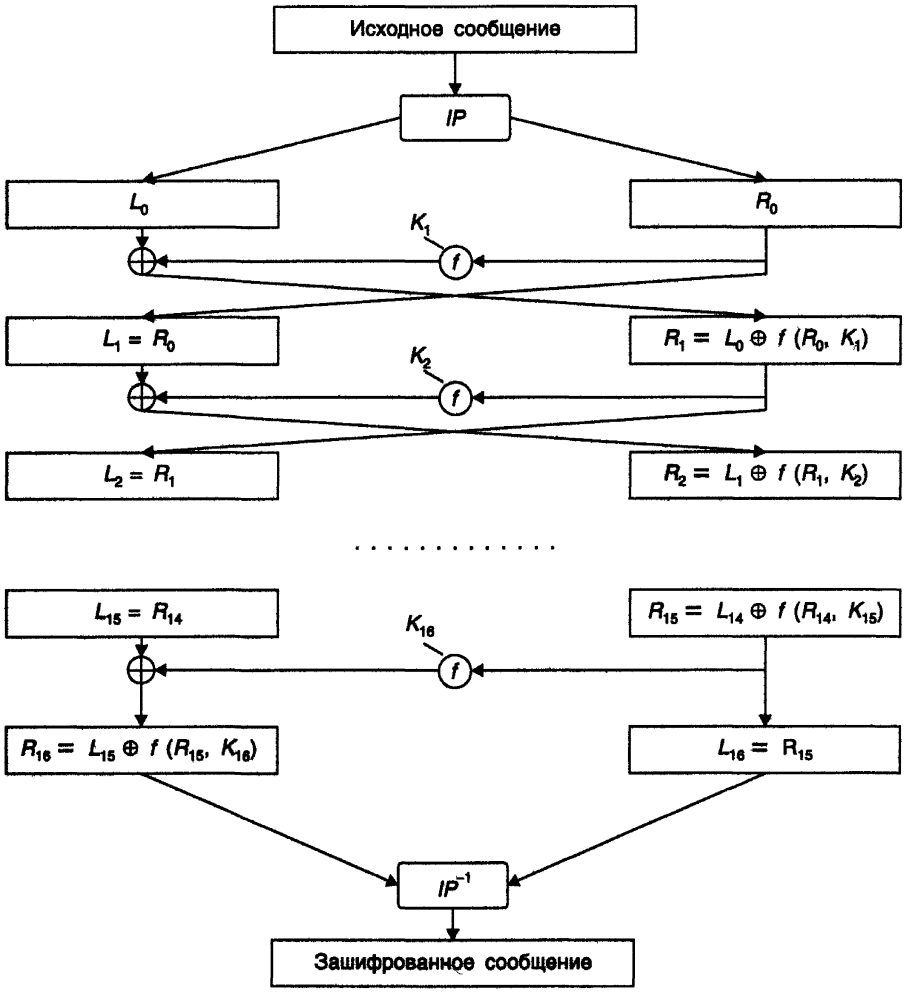


Рис. 10.2. Схема криптопреобразования DES

Таблица 10.2

Подстановка P_1

P_1	32	1	2	3	4	5	4	5	6	7	8	9	8	9	10	11
	12	13	12	13	14	15	16	17	16	17	18	19	20	21	20	21
	22	23	24	25	24	25	26	27	28	29	28	29	30	31	30	1

2. Далее вычисляется сумма $P(A) \oplus B$ и записывается в виде конкатенации восьми 6-битовых слов: $P(A) \oplus B = B_1 B_2 B_3 B_4 B_5 B_6 B_7 B_8$.

3. На этом этапе каждое слово B_i поступает на соответствующий S-блок S_i. Блок S_i преобразует 6-битовый вход B_i в 4-битовый выход C_i. S-блок есть 4×16 матрица с целыми элементами в диапазоне от 0 до 16.

Два первых бита слова B_i, если их рассматривать как двоичную запись числа, определяют номер строки матрицы S-блока. Четыре последних бита определяет некоторый столбец. Тем самым найден некоторый элемент матрицы. Его двоичная запись и является выходом.

В табл. 10.3—10.10 представлены все 8 блоков DES.

Таблица 10.3

Блок S_1

S_1	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13

Таблица 10.4

Блок S_2

S_2	13	1	8	14	6	11	3	4	9	7	2	13	12	0	5	10
	3	13	4	7	15	2	8	14	12	0	1	10	6	9	11	5
	0	14	7	11	10	4	13	1	5	8	12	6	9	3	2	15
	13	8	10	1	3	15	4	2	11	6	7	12	0	5	14	9

Таблица 10.5

Блок S_3

S_3	10	0	9	14	6	3	15	5	1	13	12	7	11	4	2	8
	13	7	0	9	3	4	6	10	2	8	5	14	12	11	15	1
	13	6	4	9	8	15	3	0	11	1	2	12	5	10	14	7
	1	10	13	0	6	9	8	7	4	15	14	3	11	5	2	12

Таблица 10.6

Блок S_4

S_4	7	13	14	3	0	6	9	10	1	2	8	5	11	12	4	15
	13	8	11	5	6	15	0	3	4	7	2	12	1	10	14	9
	10	6	9	0	12	11	7	13	15	1	3	14	5	2	8	4
	3	15	0	6	10	1	13	8	9	4	5	11	12	7	2	14

Таблица 10.7

Блок S_5

S_5	2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9
	14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6
	4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14
	11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3

Таблица 10.8

Блок S_6

S_6	12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
	10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
	9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
	4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13

Таблица 10.9

Блок S_7

S_7	4	11	2	14	15	0	8	13	3	12	9	7	5	10	6	1
	13	0	11	7	4	9	1	10	14	3	5	12	2	15	8	6
	1	4	11	13	12	3	7	14	10	15	6	8	0	5	9	2
	6	11	13	8	1	4	10	7	9	5	0	15	14	2	3	12

Таблица 10.10

Блок S_8

S_8	13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7
	1	15	13	8	10	3	7	4	12	5	6	11	0	14	9	2
	7	11	4	1	9	12	14	2	0	6	10	13	15	3	5	8
	2	1	14	7	4	10	8	13	15	12	9	0	3	5	6	11

4. Выход $C = C_1C_2 \cdots C_8$ перемешивается фиксированной подстановкой P_2 (табл. 10.11).

Таблица 10.11

Подстановка P_2

P_2	16	7	20	21	29	12	28	17	1	15	23	26	5	18	31	10
	2	8	24	14	32	27	3	9	19	13	30	6	22	11	4	25

Расписание ключей

1. В 64-битовом ключе K устраняются биты 8, 16, ..., 64. Оставшиеся биты перемешиваются подстановкой P_3 (табл.10.12).

Таблица 10.12

Подстановка P_3

P_3	57	49	41	33	25	17	9	1	58	50	42	34	26	18
	10	2	59	51	43	35	27	19	11	3	60	52	44	36
	53	55	47	39	31	23	15	7	62	54	46	38	30	22
	14	6	61	53	45	37	29	21	13	5	28	20	12	4

Выход $P_3(K)$ представляется в виде $P_3(K) = C_0D_0$, где C_0 — левая половина, D_0 — правая.

2. Очередное значение C_i, D_i вычисляется по схеме

$$C_i = L_i(C_{i-1}), \quad D_i = L_i(D_{i-1}),$$

где L_i — циклический сдвиг влево на одну позицию, если $i = 1, 2, 9, 16$. В остальных случаях L_i — сдвиг влево на две позиции.

3. На этом этапе выход перемешивается подстановкой P_4 (табл. 10.13).

Таблица 10.13

Подстановка P_4

P_4	14	17	11	24	1	5	3	28	15	6	21	10
	23	19	12	4	26	8	16	7	27	20	13	2
	41	52	31	37	47	55	30	40	51	45	33	48
	44	49	39	56	34	53	46	42	50	36	29	32

10.1.2. Расшифровывание в стандарте DES

В процессе шифрования последовательно используются преобразования $IP, \Phi_1, \Phi_2, \ldots, \Phi_{16}, T, IP^{-1}$. Каждое преобразование Фейстеля является композицией двух преобразований. Первое из них — транспозиция L и R: $T(L, R) = (R, L)$. Второе, фактически, является преобразованием Вернама 32-битовых слов, т.е. $V(L, R) = (L, R \oplus C)$, где C зависит от ключа и L.

Таким образом, для любого i

$$\Phi_i = V_iT.$$

В свою очередь, легко убедиться, что $V_i^2 = T^2 = id$ — тождественное преобразование. Такие преобразования (элементы группы) принято называть *инволюциями*. Для них, в частности, верно $V_i^{-1} = V_i$, $T^{-1} = T$.

По правилу обращения произведения элементов неабелевой группы имеем

$$D = E^{-1} = ((IP)^{-1} TV_{16} TV_{15} T \cdots V_1 T(IP))^{-1} =$$

$$= (IP)^{-1} TV_1 TV_2 \cdots V_{16} T(IP).$$

Это означает, что расшифровывание осуществляется теми же алгоритмом и ключом, но в расписании ключей надо поменять на обратный порядок генерации ключей.

10.1.3. Применение алгоритма DES для поточного шифрования

Поточное шифрование осуществляется побитным суммированием сообщения M произвольной длины с ключом K такой же длины. Расшифровывание производится аналогично, так как $(M \oplus K) \oplus K = M \oplus (K \oplus K) = M$.

Пусть, например, надо зашифровать сообщение $X = x_1 x_2 x_3, \ldots,$ где x_i, $i \geq 1$ — 64-битовый блок. Выберем вектор инициализации IV и ключ k стандарта DES. Построим ключ $K = k_1 k_2 k_3 \ldots$ нужной длины:

$$k_1 = E_K(IV), \;\; k_2 = E_K(k_1), \;\; k_3 = E_K(k_2), \; \ldots$$

Шифровать и расшифровывать будем по правилам

$$Y = X \oplus K, \quad X = Y \oplus K.$$

При таком подходе поточный ключ зависит лишь от ключа K и вектора инициализации.

Второй способ состоит в следующем. Генерация поточного ключа $K = k_1 k_2 k_3 \ldots$ производится по правилу

$$k_1 = E_K(IV), \;\; k_2 = E_K(x_1 + k_1), \;\; k_3 = E_K(x_2 + k_3), \; \ldots$$

Если к зашифрованной последовательности $y_1 = x_1 \oplus k_1$, $y_2 = x_2 \oplus k_2, \ldots$ применить то же правило, то получим исходную последовательность $X = x_1 x_2 x_3 \ldots$

10.2. Криптосистема IDEA

10.2.1. Модель IDEA

Как и DES, стандарт IDEA шифрует 64-битовые блоки. Для расшифровывания применяется прежний алгоритм. Меняется лишь расписание ключей. Шифрование состоит из восьми тактов. Исходное сообщение X разбивается на четыре 16-битовых блока: X_1, X_2, X_3, X_4. В шифровании используются следующие операции: XOR (обозначается знаком $\oplus$), сложение по модулю 2^{16} (обозначается $\boxplus$), умножение по модулю простого числа $2^{16} + 1$ (обозначается $\odot$).

В каждом такте выполняются следующие операции:

1) X_1 умножается на первый подключ;

2) X_2 складывается (⊞) со вторым подключом;

3) X_3 складывается (⊞) с третьим подключом;

4) X_4 умножается на четвертый подключ;

5) складываются (⊕) результаты (1) и (3);

6) складываются (⊕) результаты (2) и (4);

7) результат (5) умножается на пятый подключ;

8) складываются (⊞) результаты (6) и (7);

9) результат (8) умножается на шестой подключ;

10) складываются (⊞) результаты (7) и (9);

11) складываются (⊕) результаты (1) и (9);

12) складываются (⊕) результаты (3) и (9);

13) складываются (⊕) результаты (2) и (10);

14) складываются (⊕) результаты (4) и (10);

15) результаты (11), (13), (12), (14) подаются на вход следующего такта;

16) после восьми тактов на вход финального преобразования подаются результаты (11), (12), (13), (14).

Финальное преобразование (FT):

1) X_1 умножается на первый подключ;

2) X_2 складывается (⊞) со вторым подключом;

3) X_3 складывается (⊞) с третьим подключом;

4) X_4 умножается на четвертый подключ.

Конкатенация результатов этих четырех преобразований и есть зашифрованный текст

10.2.2. Расписание ключей и расшифровывание

В алгоритме используется 52 подключа, так как в каждом из восьми тактов используется по шесть ключей и еще четыре необходимы для финального преобразования. Эти ключи генерируются следующим образом.

Сначала исходный 128-битовый ключ делится на восемь 16-битовых подключей. Первые шесть из них используются в первом такте, а два последних — во втором. Затем исходный ключ сдвигается циклически влево на 25 позиций и опять делится на восемь подключей. Первые четыре из них предназначены для второго такта, а последние четыре — для третьего. Этот процесс продолжается до окончания алгоритма.

Расшифровывание осуществляется аналогично. Порядок использования ключей обращается, а сами они мультипликативно или аддитивно обратны ключам шифрования (табл. 10.14, 10.15).

При мультипликативном обращении нулевого ключа по модулю $2^{16}+1$ его полагают равным $2^{16}=-1$ и, таким образом, $0^{-1}=0$.

Таблица 10.14

Ключи для шифрования

Такт	Ключи для шифрования					
1	$Z_1^{(1)}$	$Z_2^{(1)}$	$Z_3^{(1)}$	$Z_4^{(1)}$	$Z_5^{(1)}$	$Z_6^{(1)}$
2	$Z_1^{(2)}$	$Z_2^{(2)}$	$Z_3^{(2)}$	$Z_4^{(2)}$	$Z_5^{(2)}$	$Z_6^{(2)}$
3	$Z_1^{(3)}$	$Z_2^{(3)}$	$Z_3^{(3)}$	$Z_4^{(3)}$	$Z_5^{(3)}$	$Z_6^{(3)}$
4	$Z_1^{(4)}$	$Z_2^{(4)}$	$Z_3^{(4)}$	$Z_4^{(4)}$	$Z_5^{(4)}$	$Z_6^{(4)}$
5	$Z_1^{(5)}$	$Z_2^{(5)}$	$Z_3^{(5)}$	$Z_4^{(5)}$	$Z_5^{(5)}$	$Z_6^{(5)}$
6	$Z_1^{(6)}$	$Z_2^{(6)}$	$Z_3^{(6)}$	$Z_4^{(6)}$	$Z_5^{(6)}$	$Z_6^{(6)}$
7	$Z_1^{(7)}$	$Z_2^{(7)}$	$Z_3^{(7)}$	$Z_4^{(7)}$	$Z_5^{(7)}$	$Z_6^{(7)}$
8	$Z_1^{(8)}$	$Z_2^{(8)}$	$Z_3^{(8)}$	$Z_4^{(8)}$	$Z_5^{(8)}$	$Z_6^{(8)}$
FT	$Z_1^{(9)}$	$Z_2^{(9)}$	$Z_3^{(9)}$	$Z_4^{(9)}$		

Таблица 10.15

Ключи для расшифровывания

Такт	Ключи для расшифровывания					
1	$Z_1^{(9)-1}$	$-Z_2^{(9)}$	$-Z_3^{(9)}$	$Z_4^{(9)-1}$	$Z_5^{(8)}$	$Z_6^{(8)}$
2	$Z_1^{(8)-1}$	$-Z_3^{(8)}$	$-Z_2^{(8)}$	$Z_4^{(8)-1}$	$Z_5^{(7)}$	$Z_6^{(7)}$
3	$Z_1^{(7)-1}$	$-Z_3^{(7)}$	$-Z_2^{(7)}$	$Z_4^{(7)-1}$	$Z_5^{(6)}$	$Z_6^{(6)}$
4	$Z_1^{(6)-1}$	$-Z_3^{(6)}$	$-Z_2^{(6)}$	$Z_4^{(6)-1}$	$Z_5^{(5)}$	$Z_6^{(5)}$
5	$Z_1^{(5)-1}$	$-Z_3^{(5)}$	$-Z_2^{(5)}$	$Z_4^{(5)-1}$	$Z_5^{(4)}$	$Z_6^{(4)}$
6	$Z_1^{(4)-1}$	$-Z_3^{(4)}$	$-Z_2^{(4)}$	$Z_4^{(4)-1}$	$Z_5^{(3)}$	$Z_6^{(3)}$
7	$Z_1^{(3)-1}$	$-Z_3^{(3)}$	$-Z_2^{(3)}$	$Z_4^{(3)-1}$	$Z_5^{(2)}$	$Z_6^{(2)}$
8	$Z_1^{(2)-1}$	$-Z_3^{(2)}$	$-Z_2^{(2)}$	$Z_4^{(2)-1}$	$Z_5^{(1)}$	$Z_6^{(1)}$
FT	$Z_1^{(1)-1}$	$-Z_2^{(1)}$	$-Z_3^{(1)}$	$Z_4^{(1)-1}$		

10.2.3. Корректность расшифровывания

Процесс шифрования можно описать следующим образом. Рассмотрим преобразования пространства V_{64}:

$$H(X) = (Z_1X_1,\ Z_2 \boxplus X_2,\ Z_3 \boxplus X_3,\ Z_4X_4),\quad F(X) = (Y_1, Y_2, Y_3, Y_4),$$

$$Y_1 = X_1 \oplus \varphi(X_1 \oplus X_3,\ X_2 \oplus X_4,\ Z_5, Z_6),$$

$$Y_2 = X_2 \oplus \psi(X_1 \oplus X_3,\ X_2 \oplus X_4,\ Z_5, Z_6),$$

$$Y_3 = X_3 \oplus \varphi(X_1 \oplus X_3,\ X_2 \oplus X_4,\ Z_5, Z_6),$$

$$Y_4 = X_4 \oplus \psi(X_1 \oplus X_3,\ X_2 \oplus X_4,\ Z_5, Z_6),$$

где

$$\varphi(U_1, U_2, U_3, U_4) = ((U_1, U_3) \boxplus Y_2)Y_4,$$

$$\psi(U_1, U_2, U_3, U_4) = f(U_1, U_2, U_3, U_4) \boxplus (U_1U_3).$$

Рассмотрим функцию $T(X_1, X_2, X_3, X_4) = T(X_1, X_3, X_2, X_4)$. Функции H, F, T — обратимы. В самом деле,

$$H^{-1}(X) = (Z_1^{-1}X, (-Z_2) \boxplus X_2, (-Z_3) \boxplus X_3, Z_4^{-1}X_4).$$

Обратимость нуля в операции умножения уже оговорена. Обратимость остальных элементов вытекает из простоты числа $2^{16} + 1$. Обратимость функций F и T вытекает из их инволютивности:

$$F^2 = T^2 = id \Rightarrow F^{-1} = F,\ T^{-1} = T.$$

Докажем инволютивность F. С этой целью вычислим первый компонент вектора $F^2(X)$:

$$Y_1 \oplus \varphi(Y_1 \oplus Y_3,\ Y_2 \oplus Y_4,\ Z_5, Z_6) =$$

$$= X_1 \oplus \varphi(X_1 \oplus X_3,\ X_2 \oplus X_4,\ Z_5, Z_6) + \varphi(Y_1 \oplus Y_3, Y_2 \oplus Y_4,\ Z_5, Z_6).$$

Последнее выражение равно X_1, поскольку $X_1 \oplus X_3 = Y_1 \oplus Y_3$, $X_2 \oplus X_4 = Y_2 \oplus Y_4$. Заметим, что хотя преобразование H и не инволютивно, обратное преобразование H^{-1} относится к такому же типу.

Преобразование каждого такта, за исключением последнего, имеет вид TF_iH_i. В последнем такте отсутствует T.

Работа алгоритма завершается преобразованием H_9. Поэтому любое преобразование стандарта можно записать в виде

$$E = H_9F_8H_8\ TF_7H_7 \cdots TF_2H_2\ TF_1H_1.$$

По правилу обращения произведения в неабелевой группе и в силу инволютивности T и F_i:

$$E^{-1} = H_1^{-1}F_1T\ H_2^{-1}F_2T \cdots H_7^{-1}F_7T\ H_8^{-1}F_8H_9^{-1}.$$

Полученное преобразование отличается от преобразований стандарта лишь расстановкой преобразований T. Это обстоятельство можно игнорировать за счет изменения расписания ключей. В самом деле,

$$E^{-1} = H_1^{-1} F_1 (TH_2^{-1}T) TF_2 \cdots (TH_7^{-1}T) TF_7 (TH_8^{-1}T) TF_8 H_9^{-1}.$$

Теперь преобразования T стоят на нужных местах. Остается показать, что преобразования $T H_i^{-1} T$, $i = 2, 3, \ldots, 8$ относятся к преобразованиям типа H. Это достигается прямым вычислением:

$$TH^{-1}T = (Z_1^{-1} X_1, -Z_3 \boxplus X_2, -Z_2 \boxplus X_3, Z_4^{-1} X_4).$$

Последняя формула и вид E^{-1} определяют указанное расписание ключей для расшифровывания.

10.3. Криптосистема ГОСТ 28147—89

ГОСТ 28147—89 является 64-битовым 32-тактовым алгоритмом с 256-битовым ключом. Имеющиеся в алгоритме S-блоки также можно использовать как ключ. При шифровании сообщение представляется в виде конкатенации двух блоков LR одинаковой длины. Один такт алгоритма осуществляет одно преобразование Фейстеля (рис. 10.3):

$$L_i = R_{i-1}, \quad R_i = L_{i-1} \oplus f(R_{i-1}, K_i).$$

Отличие от стандарта DES — лишь в конструкции функции f. Работа алгоритма завершается транспозицией $LR \to_1 RL$.

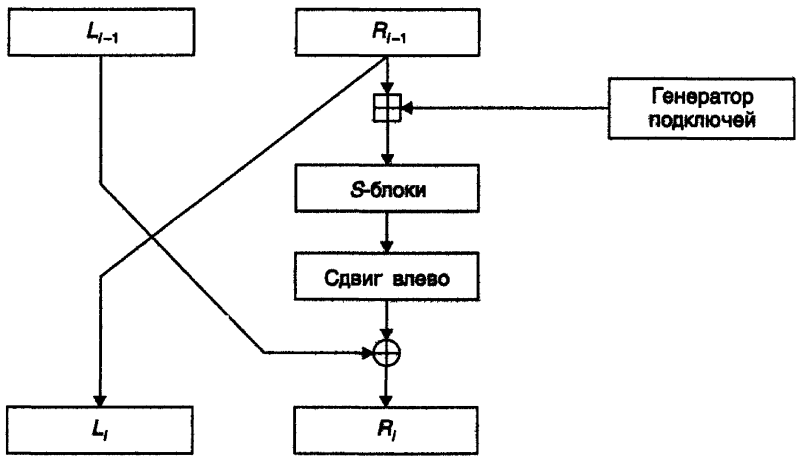


Рис. 10.3. Один такт алгоритма ГОСТ 28147—89

Описание функции f. Сначала правый блок R_i складывается по модулю 2^{32} с подключом K_i. Полученное 32-битовое сообщение делится на восемь 4-битовых частей. Каждое из этих 4-битовых чисел преобразуется соответствующим S-блоком в другое 4-битовое число. Поэтому любой S-блок определяется некоторой 16-битовой перестановкой на множестве из 16 элементов 0, 1, ..., 15. Реально использовались, например, следующие S-блоки:

$S_1 = (4\ 10\ 9\ 2\ 13\ 8\ 0\ 14\ 6\ 11\ 1\ 12\ 7\ 15\ 5\ 3)$,
$S_2 = (14\ 11\ 4\ 12\ 6\ 13\ 15\ 10\ 2\ 3\ 8\ 1\ 0\ 7\ 5\ 9)$,
$S_3 = (5\ 8\ 1\ 13\ 10\ 3\ 4\ 2\ 14\ 15\ 12\ 7\ 6\ 0\ 9\ 11)$,
$S_4 = (7\ 13\ 10\ 1\ 0\ 8\ 9\ 15\ 14\ 4\ 6\ 12\ 11\ 2\ 5\ 3)$,
$S_5 = (6\ 12\ 7\ 1\ 5\ 15\ 13\ 8\ 4\ 10\ 9\ 14\ 0\ 3\ 11\ 2)$,
$S_6 = (4\ 11\ 10\ 0\ 7\ 2\ 1\ 13\ 3\ 6\ 8\ 5\ 9\ 12\ 15\ 14)$,
$S_7 = (13\ 11\ 4\ 1\ 3\ 15\ 5\ 9\ 0\ 10\ 14\ 7\ 6\ 8\ 2\ 12)$,
$S_8 = (1\ 15\ 13\ 0\ 5\ 7\ 10\ 4\ 9\ 2\ 3\ 14\ 6\ 11\ 8\ 12)$.

После преобразования S-блоками полученное 32-битовое сообщение сдвигается влево на 11 позиций.

Ключевое расписание. Исходный 256-битовый ключ делится на восемь 32-битовых подключей $k_1, k_2, \ldots, k_8$; они используются в 32 тактах в следующем порядке: 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 8 7 6 5 4 3 2 1. При дешифровании порядок использования подключей меняется на противоположный.

10.4. Общая структура алгоритма Rijndael

Предполагая знакомство читателя с формальным описанием алгоритма, отметим некоторые его алгебраические свойства.

В алгоритме Rijndael к исходному блоку открытого текста последовательно применяются следующие преобразования:

1) K_0 — первоначальное добавление ключа;
2) несколько тактовых преобразований;
3) заключительное тактовое преобразование.

Каждое тактовое преобразование включает:

1) B — преобразование байтовой подстановки;
2) S — циклический сдвиг байтовых блоков;
3) M — преобразование аффинного типа;
4) K_i — добавление тактового ключа.

В заключительном такте нет преобразования M. Алгоритм в принципе может иметь любое количество единичных тактов. В простейшем случае зашифрованный блок имеет вид $K_2SBK_1MSBK_0(m)$. Пусть $T = K_2SBK_1MSBK_0$, тогда

$$T^1 = (K_2SBK_1MSBK_0)^{-1} = K_0^{-1}B^{-1}S^{-1}M^{-1}K_1^{-1}B^{-1}S^{-1}K_2^{-1}.$$

Это произведение можно преобразовать с учетом свойств элементарных преобразований.

Во-первых, $K^{-1} = K$, так как добавление ключа — это просто преобразование второго порядка $x \to x \oplus k$. Преобразование S^{-1} так же, как и преобразование S, является циклическим сдвигом байтовых блоков. Преобразование B является аналогом S-блока, в его основе лежит подстановка на конечном поле $\mathbb{F}_{2^8}$ вида $x \to x^{-1}(0 \to 0)$. Таким образом, множители преобразования T^{-1} несколько иные, чем преобразования T, и их порядок иной. Это обстоятельство, по мнению авторов алгоритма, делает практически невозможным выполнение условия $T^{-1} = T$, т.е. $T^2 = id$. Тем самым, по их мнению, в алгоритме нет полуслабых ключей. В связи с этим отметим, что структуры множителей у преобразований T и T^{-1} на самом деле отличаются не столь сильно. Во-вторых, $SB = BS \Rightarrow S^{-1}B^{-1} = B^{-1}S^{-1}$, так как произведение сдвига и подстановки коммутирует.

Рассмотрим произведение MK. С одной стороны,

$$MK(x) = M(x + k) = M(x) + M(k),$$

а с другой стороны,

$$KM(x) = M(x) + k.$$

Таким образом, множители K и M, а значит K^{-1} и M^{-1}, также можно менять местами при внесении в расписание ключей изменений $k \to M(k)$. Следовательно,

$$T^{-1} = K_0 S^{-1} B^{-1} K_1 M^{-1} S^{-1} B^{-1} K_2.$$

Отметим, наконец, что алгоритм не использует схему Фейстеля, а влияние ключа (в отличие от DES) происходит в самом начале его работы. Небольшое количество тактов в алгоритме, по-видимому, компенсируется большим, чем в DES, количеством элементарных преобразований в одном такте и их разнообразием.

10.5. Использование алгебры полиномов в алгоритме Rijndael

Алгоритм Rijndael в существенной степени использует умножение и сложение элементов полей колец и полиномов над ними. В первую очередь это относится к полю $\mathbb{F}_{2^8}$. Оно интерпретируется как фактор-кольцо

$$\mathbb{F}_2[x]/(x^8 + x^4 + x^3 + x + 1)\mathbb{F}_2[x],$$

а его элементы отождествляются с байтами. Это поле используется для построения S-блока. В основу конструкции положено преобразование $x \to x^{-1}$, где $x \in \mathbb{F}_{2^8}^*$, $0 \to 0$. При этом достигаются две цели. Во-первых, это преобразование имеет над $\mathbb{F}_{2^8}$ высокую интерполяционную степень, так как $x^{256} = x \Rightarrow x^{255} = x^{-1}$, $x \neq 0$. Во-вторых, если преобразование $x \to x^{-1}$ записать как преобразование $V_8 \to V_8$, то оно оказывается нелинейным. Чтобы устра-

нить очевидный недостаток $0 \to 0$, конструкция S-блока дополняется еще одним преобразованием. Оно определено над полем $\mathbb{F}_2$ и имеет вид $Y = AX + B$, где B — некоторый сбалансированный вектор. Такое преобразование принято называть аффинным. И в этом случае авторы не придерживаются традиции. Преобразование $Y = AX + B$ не является инвалютивным, хотя за счет выбора матриц A и B это легко достижимо. В самом деле, если $A^2 = E$, а B является собственным вектором матрицы A, отвечающим собственному значению $\lambda = 1$, то

$$A(AX + B) + B = A^2X + AB + B = EX + B + B = X.$$

Эта возможность не использована, так как для матрицы A из алгоритма $A^2 \neq E$. В предполагаемом алгоритме еще раз используется техника умножения многочленов, чтобы реализовать операцию сдвига. Но более важную роль операция умножения многочленов играет в конструкции перемешивающего преобразования MixColumn. Оно действует на сообщение (шифруемый блок) следующим образом. Сначала само сообщение представляется в виде

$$\begin{pmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,5} \\ a_{1,0} & a_{1,1} & \dots & a_{1,5} \\ & & \dots & \\ a_{3,0} & a_{3,1} & \dots & a_{3,5} \end{pmatrix}, \quad \text{где } a_{i,j} \in \mathbb{F}_{2^8}.$$

Любой его столбец, например первый, отождествляется с многочленом над полем $\mathbb{F}_{2^8}$, т.е.

$$(a_{0,0}, a_{1,0}, a_{2,0}, a_{3,0})^T = a_{3,0}x^3 + a_{2,0}x^2 + a_{1,0}x + a_{0,0} = a(x),$$

где $a(x) \in \mathbb{F}_{2^8}[x]$. Затем этот многочлен преобразуется в многочлен $b(x)$ по правилу

$$b(x) = a(x)c(x) \pmod{x^4 + 1},$$

где $c(x) = 03x^3 + 01x^2 + 01x + 02$. Поясним принятые обозначения коэффициентов. В исходном виде это шестнадцатеричная запись некоторого числа. Это число следует записать в двоичной системе счисления. Полученный байт даст набор коэффициентов многочлена седьмой степени, который затем отождествляется с элементом фактор-кольца $\mathbb{F}_2[x]/(m(x))$. Например:

$$03_{10} = 3_{10} = 2 + 1 = (0,0,0,0,0,0,1,1) = x + 1 \in \mathbb{F}_{2^8}.$$

В этом случае имеем дело с операцией умножения не в поле, а в кольце с делителями нуля, так как многочлен $x^4 + 1$ приводим над любым полем характеристики 2:

$$x^4 + 1 = (x + 1)^4.$$

Условие обратимости требует, чтобы многочлен $c(x)$ был взаимно прост с многочленом $x^4 + 1$. Для фактического обращения этой операции предлагается

использовать алгоритм Евклида. Выбор многочлена $x^4 + 1$, по-видимому, обусловлен двумя причинами. Во-первых, в операции перемешивания оказывается задействована операция из нового объекта: кольца с делителями нуля. Во-вторых, редукция $\mathrm{mod}(x^4 + 1)$ осуществляется очень просто с вычислительной точки зрения, так как $x^j \pmod{x^4 + 1} \equiv x^{j \pmod 4}$.

В заключение сделаем некоторые выводы.

Предложенное представление элементов поля $\mathbb{F}_{2^8}$, как элементов фактор-кольца по модулю идеала, порожденного многочленом $m(x) = x^8 + x^4 + x^3 + x + 1$, авторами, по существу, не обосновано. Это, по-видимому, связано с тем, что выбор любого другого неприводимого многочлена восьмой степени приводит к изоморфному полю.

Выбор поля $\mathbb{F}_{2^8}$ связан с тем, что его элементы являются байтами, а это облегчает его реализацию на 8-битовом процессоре.

Конструкция S-блока включает аффинное преобразование $Y = AX + B$ над полем $\mathbb{F}_2$. Не вполне ясно, почему авторы отказались от возможности сделать его инволютивным, что упростило бы операцию его обращения.

Операция MixColumn, основанная на умножении элементов кольца с делителями нуля $\mathbb{F}_{2^8}[x]/(x^4 + 1)$, ранее в криптоалгоритмах не применялась. Она быстро реализуется и обратима за счет выбора многочлена $c(x)$.

10.6. Задания

1. Выписать в явном виде подстановку IP^{-1}. Найти ее разложение в произведение независимых циклов и вычислить ее порядок.

2. Выписать в явном виде ключи всех 16 тактов стандарта DES.

3. Доказать, что преобразования T, V, IP, IP^{-1}, а потому и все преобразования стандарта DES в целом, являются четными подстановками на пространстве V_{64}.

4. Убедиться, что все S-блоки стандарта DES являются нелинейными преобразованиями из V_6 в V_8.

5. Установить свойство дополнительности DES: $E_{\bar{k}}(\bar{x}) = \overline{E_k(x)}$, где черта означает взятие дополнительного вектора, т.е. $x + \bar{x} = 0$.

6. Пусть $T(L, R) = (R, L)$, $V_1(L, R) = (L, R \oplus C_1)$. Вычислить порядок преобразования $\Phi_1 = TV_1$. Вычислить порядки композиции $\Phi_1\Phi_2$, $\Phi_1\Phi_2\Phi_3$.

7. Показать, что преобразования вида H в стандарте IDEA образуют группу относительно их композиции.

8. Изобразить блок-схему алгоритма IDEA.

9. Доказать, что всякое преобразование стандарта ГОСТ 28147—89 обратимо и является четной подстановкой на множестве всех 64-битовых слов.

10. Предложите какие-нибудь алгоритмы генерации псевдослучайных последовательностей при помощи алгоритма ГОСТ 28147—89. Оцените величину их периода.

11. Предложите возможные криптоатаки на двух- или трехтактовый ГОСТ 28147—89.

SXJKN8EYHR3I4HR9
87YDWFUIHW9E8YR3
84H5O34596EUF9O8
WE76R83Y439846UH
FE98YDSFJH8374H4
BR8276RUBNWDF87Y
6234BN9WE872H3WR
TJOJRTGM905JRO9G
70954TOIRF7ERFH2
347FQ7WERF9UJ134
9YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFV
KBDFKJVJKGFKJSBD

Глава 11

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОАНАЛИЗА СИММЕТРИЧНЫХ СИСТЕМ

SXJKN8EYHR3I4HR9
87YDWFUIHW9E8YR3
84H5O34596EUF9O8
WE76R83Y439846UH
FE98YDSFJH8374H4
BR8276RUBNWDF87Y
6234BN9WE872H3WR
TJOJRTGM905JRO9G
70954TOIRF7ERFH2
347FQ7WERF9UJ134
9YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFV
KBDFKJVJKGFKJSBD

11.1. Задачи и принципы криптоанализа

Любая математическая теория наряду с прямыми рассматривает обратные задачи, которые часто оказываются вычислительно более сложными. *Криптоанализ* занимается задачами, обратными задачам криптографии. По определению известного криптолога В. Фридмена, «криптоанализ включает определение используемого языка, типа криптосистемы, ключа и исходного текста; обычно именно в этом порядке».

Следует сразу заметить, что основная цель криптоанализа состоит не в проникновении в компьютерные сети противника для получения конфиденциальной информации, а в оценке стойкости (надежности) используемых и разрабатываемых криптосистем [11]. Методы криптоанализа позволяют оценивать стойкость (уровень безопасности) криптосистемы количественно в виде количества W компьютерных операций, необходимых криптоаналитику для «вскрытия» ключа (или исходного текста). Для пользователя криптосистемы важно, чтобы W было настолько велико, что при реализации криптоатаки на современной вычислительной технике требовалось бы машинное время $t_{\text{Оащ}}$ (например, несколько лет), в течение которого переданная в зашифрованном виде информация теряла бы свою ценность.

Приведем пять основных принципов, «выстраданных криптологами — теоретиками и практиками», которые используются при решении задач криптоанализа [46].

1. Не следует недооценивать противника.

2. (*Принцип Керкхоффа*). Только криптоаналитик может судить о безопасности криптосистемы.

3. (*Принцип Керкхоффа — Шеннона*). Противник знает используемую криптосистему с точностью до ключевой информации.

4. (*Принцип Живерже*). Поверхностные усложнения криптосистемы могут быть иллюзорны, так как порождают ложные оценки ее стойкости.

5. При оценке стойкости некоторой криптосистемы необходимо учитывать криптографические ошибки и другие нарушения дисциплины безопасности.

Согласно принципу Керкхоффа — Шеннона, основной задачей криптоаналитика является оценка ключевой информации при условии, что сама исполь-

зуемая криптосистема известна. Алгоритм оценки ключевой информации принято называть *криптоатакой*. Как отмечалось в главе 1, в зависимости от условий взаимодействия криптоаналитика с криптосистемой различают следующие четыре типа криптоатак:

A_1) криптоатака с использованием только шифртекстов;

A_2) криптоатака с использованием открытых текстов и соответствующих им шифртекстов;

A_3) криптоатака с использованием выбираемых криптоаналитиком открытых текстов и соответствующих им шифртекстов;

A_4) криптоатака с использованием активного аппаратного воздействия на криптосистему.

Эти типы криптоатак упорядочены в порядке возрастания степени активности криптоаналитика при взаимодействии с криптосистемой. В частности, четвертый тип криптоатак A_4, появившийся в последние годы [21, 22], предполагает, что криптоаналитик имеет физический доступ к чипу, реализующему криптосистему. При этом предполагается, что доступ к области памяти, содержащей ключевые параметры и алгоритм шифрования, не может быть осуществлен без стирания содержащейся в ней информации. Аппаратное воздействие криптоаналитика состоит в том, что задается специальный открытый текст и анализируются соответствующие ему шифртексты в двух ситуациях: а) чип работает в обычных условиях; б) чип подвержен воздействию специального поля, порождающего сбои в шифртексте. Отметим, что этот новый вид криптоатак возник в связи с массовым использованием интеллектуальных электронных карточек (smart cards), в чипе которых обязательно используется криптоалгоритм.

История возникновения и развития криптоанализа (от докомпьютерной эры до нашего времени) достаточно подробно изложена в [46, 79, 21, 12]. В качестве первой монографии, посвященной криптоанализу, следует упомянуть книгу А. Синкова [186]. Не останавливаясь на истории криптоанализа, заметим лишь, что пакеты прикладных программ по криптоанализу появились уже и на свободном рынке программных продуктов [46]. Так, например, американская фирма Access Data Recovery распространяет программный продукт для криптоанализа шифрований, используемых в программных системах MS Word, Word Perfect, MS Excel, Lotus 1-2-3, Symphony, Paradox, Quattro Pro.

Хотя криптоанализ возник давно и интенсивно развивается на современном этапе, единая математическая теория криптоанализа еще не разработана. Результаты исследований по криптоанализу сосредоточены в основном в зарубежных публикациях и посвящены частным задачам реализации определенных криптоатак для конкретных криптосистем. Анализ зарубежных исследований по криптоанализу позволяет выделить следующие четыре основных класса *методов криптоанализа*:

1) метод «опробования»;

2) методы криптоанализа с использованием теории статистических решений;

3) линейный криптоанализ;

4) разностный криптоанализ.

Методы первого класса применяются для криптоанализа как симметричных, так и асимметричных криптосистем; методы остальных трех классов применяются для криптоанализа симметричных криптосистем. В связи с этим в последующих параграфах описываются в основном симметричные криптосистемы. Вопросы криптоанализа криптосистем с открытым ключом рассматриваются в главах 12, 14, 15.

11.2. Метод «опробования» и его вычислительная сложность

Рассмотрим произвольную симметричную криптосистему, общая математическая модель которой описана в п. 6.4. Пусть $X = (x_1, \ldots, x_n) \in \mathcal{A}_N^n$ — открытый текст, подлежащий шифрованию; $\theta^0 = (\theta_1^0, \ldots, \theta_L^0) \in \Theta \subseteq \mathcal{A}_m^L$ — истинное значение ключа, использованное в данном сеансе шифрования; $f(\cdot)$ — криптографическое преобразование; $Y = (y_1, \ldots, y_n) \in V_N^n$ — шифртекст (криптограмма):

$$Y = f(X; \theta^0). \tag{11.2.1}$$

Напомним, что обязательным условием любой криптосистемы является биективность преобразования (11.2.1) при фиксированном значении ключа (параметра) θ^0 так, что выполняется

$$f^{-1}(Y; \theta^0) = X. \tag{11.2.2}$$

Соотношение (11.2.2) фактически определяет алгоритм расшифровывания зарегистрированного шифртекста Y санкционированным получателем информации, который знает ключ θ^0.

Метод «опробования» (полного перебора, «brute force attack») базируется на соотношении (11.2.2) и заключается в следующем:

1) на основе имеющегося шифртекста $Y \in \mathcal{A}_N^n$ и соответствующего ей открытого текста $X \in \mathcal{A}_N^n$ составляется система уравнений относительно $\theta = (\theta_1, \ldots, \theta_L)$:

$$f_i^{-1}(Y; \theta) = x_i, \quad i = \overline{1, n} \tag{11.2.3}$$

или в векторном виде:

$$f^{-1}(Y; \theta) = X;$$

2) полным перебором всех возможных значений ключевого параметра $\theta \in \Theta \subseteq \mathcal{A}_m^L$ находится подмножество $\Theta_0 = \{\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(l)}\}$ l решений этой системы;

3) если количество найденных решений $l = 1$, то в силу (11.2.2) определено истинное значение параметра $\hat{\theta} = \theta^0$; в противном случае Θ_0 не является одноточечным множеством и определено подмножество l значений искомого параметра, одно из которых в силу (11.2.2) совпадает с истинным значеним θ^0.

В случае $l > 1$ целесообразно увеличить количество уравнений в системе (11.2.3). Для этого необходимо увеличить количество символов n в исходном сообщении X либо получить и использовать в (11.2.3) $q > 1$ пар «сообщение — шифртекст» $(X^{(1)}, Y^{(1)}), \ldots, (X^{(q)}, Y^{(q)})$:

$$f^{-1}(Y^{(j)}; \theta) = X^{(j)}, \quad j = \overline{1, q}. \tag{11.2.4}$$

Увеличивая n или q, можно достичь случая $l = 1$ и, следовательно, безошибочного оценивания ключа: $\hat{\theta} = \theta^0$.

Как видно из описания, в данном виде метод «опробования» используется в криптоатаках типа A_2 (т.е. при наличии открытых текстов и соответствующих им шифртекстов). Метод «опробования» иногда применяется и в криптоатаках типа A_1 (при наличии только шифртекстов); при этом вместо равенства (11.2.4) для всех $\theta \in \Theta$ проверяется, «имеет ли текст $f^{-1}(Y^{(j)}; \theta)$ смысловое значение». Однако такое применение возможно лишь в случае, если априорно известно, что передаваемые сообщения представляют собой «имеющий смысл текст».

Вычислительная сложность метода «опробования» характеризуется количеством компьютерных операций, необходимых для реализации этого метода:

$$W = W(n, q, |\Theta|) = |\Theta| \, W_1(n, q), \tag{11.2.5}$$

где $|\Theta|$ — количество всевозможных допустимых значений параметра (ключа) θ^0, $W_1(n, q)$ — количество компьютерных операций, затрачиваемых на «опробование» (проверку (11.2.4)) для одного значения θ. Обычно $W_1(n, q)$ зависит от q линейно:

$$W_1(n, q) = w(n) \, q, \tag{11.2.6}$$

где $w(n)$ — количество компьютерных операций, затрачиваемых на проверку (11.2.3) применительно к n-символьным сообщениям для единичного значения $\theta \in \Theta$; $w(n)$ фактически совпадает с количеством компьютерных операций при расшифровывании шифртекста (вычисление обратной функции $f^{-1}(\cdot)$) при известном ключе. Заметим, что величина $w(n)$ является известной характеристикой для каждой реальной криптосистемы, причем чем выше быстродействие криптосистемы, тем эта величина меньше.

Из (11.2.5) и (11.2.6) следует, что основным фактором, определяющим вычислительную сложность метода «опробования» и одновременно криптостойкость алгоритма шифрования по отношению к рассматриваемому методу криптоатаки, является мощность пространства ключей $|\Theta|$. С учетом этого оценим вычислительную сложность метода «опробования» для криптоанализа известных криптосистем, описанных в п. 6.4, 10.1—10.3.

1. Подстановка символов алфавита. Для этой криптосистемы Θ есть множество всевозможных подстановок символов алфавита V_N, за исключением тождественной подстановки, поэтому

$$|\Theta| = N! - 1, \quad W = (N! - 1) \, q w(n). \tag{11.2.7}$$

Иногда, чтобы усилить «перепутывание» символов, задают Θ как множество моноциклических подстановок, которые имеют единственный цикл максимального порядка N. При этом очевидно:

$$|\Theta| = (N-1)!, \quad W = (N-1)!\,qw(n). \tag{11.2.8}$$

2. Перестановка символов текста с периодом T. При этом Θ есть множество всевозможных подстановок индексов $\{1, 2, \ldots, T\}$ переставляемых символов текста, за исключением тождественной подстановки, поэтому

$$|\Theta| = T!-1, \quad W = (T!-1)\,qw(n). \tag{11.2.9}$$

Как и в (11.2.8), в случае моноциклических перестановок Θ имеем

$$|\Theta| = (T-1)!, \quad W = (T-1)!\,qw(n). \tag{11.2.10}$$

Отметим, что, как видно из (11.2.9) и (11.2.10), при шифровании с помощью перестановки символов текста вычислительная сложность криптоатаки методом «опробования» не зависит от мощности алфавита N (11.2.7) и (11.2.8), а определяется величиной периода T.

3. Шифр Виженера и его модификации. В этом случае

$$\Theta = \{\theta = (\theta_1, \ldots, \theta_T) : \theta_i \in V_N, i = \overline{1, T}\} \setminus \{\mathbf{0}_T\},$$

поэтому

$$|\Theta| = N^T - 1, \quad W = (N^T - 1)\,qw(n). \tag{11.2.11}$$

При $T = 1$ имеем шифр Цезаря и Бофора, для которых:

$$|\Theta| = N - 1, \quad W = (N-1)\,qw(n). \tag{11.2.12}$$

4. Шифр Вернама (поточный шифр). Его можно рассматривать как частный случай шифра Виженера, когда $T = n$, т.е. период совпадает с длиной шифруемого текста:

$$|\Theta| = N^n - 1, \quad W = (N^n - 1)\,qw(n). \tag{11.2.13}$$

5. Блочно-итерационный шифр DES. В этом случае двоичный текст ($N = 2$) шифруется блоками размера $n = 64$ и используется 56-битовый ключ:

$$\Theta = \{\theta = (\theta_1, \ldots, \theta_{56}) : \theta_i \in V_2, i = \overline{1, 56}\},$$

поэтому

$$|\Theta| = 2^{56}, \quad W = 2^{56}\,qw(64). \tag{11.2.14}$$

Отметим, что используя некоторые специальные свойства DES, вычислительную сложность W можно уменьшить вдвое (см. п. 11.4).

6. Блочно-итерационный шифр IDEA. В отличие от DES, размер ключа увеличен до 128 битов, поэтому

$$|\Theta| = 2^{128}, \quad W = 2^{128}\,qw(64). \tag{11.2.15}$$

7. Блочно-итерационный шифр ГОСТ 28147—89. В отличие от DES и IDEA размер ключа увеличен до 256 битов, поэтому

$$|\Theta| = 2^{256}, \quad W = 2^{256} qw(64). \tag{11.2.16}$$

В табл. 11.1 приведены для сравнения вычисленные с помощью (11.2.7)—(11.2.16) и таблиц из приложения А значения $|\Theta|$, а также значения среднего времени перебора $|\Theta|$ с помощью суперкомпьютера INTEL ASCI RED (в порядке возрастания) для некоторых часто используемых значений параметров N, T, n. Для оценки значимости приведенных чисел отметим, что количество секунд в 70-летней человеческой жизни составляет примерно $2.21 \cdot 10^9$, а количество секунд с момента Рождества Христова до сегодняшнего дня — приблизительно $6.31 \cdot 10^{10}$.

Таблица 11.1

Мощностные характеристики пространства ключей

Криптосистема	$\lvert\Theta\rvert$	Среднее время перебора для INTEL ASCI RED
Цезаря, $N = 26$	25	$1.2 \cdot 10^{-11}$ с
DES	$7.21 \cdot 10^{16}$	9.4 ч
Моноциклическая подстановка символов, $N = 26$	$1.55 \cdot 10^{25}$	$2.3 \cdot 10^6$ лет
Подстановка символов, $N = 26$	$4.03 \cdot 10^{26}$	$6.0 \cdot 10^7$ лет
IDEA	$3.40 \cdot 10^{38}$	$1.3 \cdot 10^{21}$ лет
Виженера, $N = 26$, $T = 32$	$1.90 \cdot 10^{45}$	$7.3 \cdot 10^{27}$ лет
ГОСТ 28147—89	$1.16 \cdot 10^{77}$	$1.7 \cdot 10^{58}$ лет
T-моноциклическая перестановка символов текста, $T = 64$	$1.99 \cdot 10^{88}$	$2.9 \cdot 10^{69}$ лет
T-перестановка символов текста, $T = 64$	$1.27 \cdot 10^{90}$	$1.9 \cdot 10^{71}$ лет
Вернама (поточный шифр), $N = 26$, $n = 128$	$1.30 \cdot 10^{181}$	$2.0 \cdot 10^{162}$ лет

11.3. Методы криптоанализа с использованием теории статистических решений

Рассмотрим симметричную криптосистему, описываемую математической моделью (11.2.1), (11.2.2):

$$\begin{gathered} Y = f(X;\theta^0), \quad X = f^{-1}(Y;\theta^0), \quad X = (x_1, \ldots, x_n) \in \mathcal{A}_N^n, \\ Y = (y_1, \ldots, y_n) \in \mathcal{A}_N^n, \quad \theta^0 = (\theta_1^0, \ldots, \theta_L^0) \in \Theta \subseteq \mathcal{A}_m^L. \end{gathered} \tag{11.3.1}$$

Будем рассматривать криптоатаку типа A_1, когда криптоаналитик априорно знает криптографическую функцию $f(\cdot)$ с точностью до ключа θ^0 и наблюдает шифртекст Y длиной n. Задача криптоанализа заключается в оценке неизвестного параметра $\theta^0 \in \Theta$ криптографической функции по наблюдению Y.

Методы *криптоанализа с использованием теории статистических решений* основаны на следующих дополнительных модельных предположениях:

U_1) исходный текст X представляет собой случайную последовательность с некоторым заданным дискретным распределением вероятностей:

$$p_n(a) = p_n(a_1, \ldots, a_n) = \mathbf{P}\{X = a\} = \mathbf{P}\{x_1 = a_1, \ldots, x_n = a_n\},$$
$$0 \le p(a) \le 1, \quad a = (a_1, \ldots, a_n) \in V_N^n, \quad \sum_{a \in A_N^n} p_n(a) = 1; \tag{11.3.2}$$

U_2) параметр θ_0 также является случайным вектором, не зависящим от X и имеющим дискретное распределение вероятностей:

$$\pi(\theta) = \pi(\theta_1, \ldots, \theta_L) = \mathbf{P}\{\theta_0 = \theta\} = \mathbf{P}\{\theta_1^0 = \theta_1, \ldots, \theta_L^0 = \theta_L\},$$
$$0 \le \pi(\theta) \le 1, \quad \theta = (\theta_1, \ldots, \theta_L) \in \Theta \subseteq V_m^L, \quad \sum_{\theta \in \Theta} \pi(\theta) = 1. \tag{11.3.3}$$

Среди методов, использующих теорию статистических решений, рассмотрим два: метод максимального правдоподобия (частотного анализа) и байесовский метод [186, 46, 119].

Метод максимального правдоподобия для решения сформулированной выше задачи криптоанализа использует только предположение U_1; при этом параметр θ^0 считается детерминированной неизвестной величиной. Данный метод базируется на *функции правдоподобия*:

$$L(\theta) = q_n(y_1, \ldots, y_n \mid \theta), \quad \theta \in \Theta, \tag{11.3.4}$$

где

$$q_n(b \mid \theta) = \mathbf{P}\{Y = b \mid \theta\}, \quad b = (b_1, \ldots, b_n) \in \mathcal{A}_N^n \tag{11.3.5}$$

есть дискретное распределение вероятностей шифртекста Y в ситуации, когда $\theta^0 = \theta$. Это распределение вероятностей $q_n(b \mid \theta)$ легко находится в силу U_1, (11.3.2), (11.3.5) и функционального преобразования $Y = f(X; \theta)$:

$$q_n(b \mid \theta) = \mathbf{P}\{f(X; \theta) = b\} = \mathbf{P}\{X = f^{-1}(b; \theta)\} = p_n(f^{-1}(b; \theta)),$$
$$b \in V_N^n, \quad \theta \in \Theta. \tag{11.3.6}$$

Подставляя (11.3.6) в (11.3.4), находим явный вид функции правдоподобия:

$$L(\theta) = p_n(f^{-1}(Y; \theta)), \quad \theta \in \Theta. \tag{11.3.7}$$

Отметим, что при вычислениях вместо функции $L(\theta)$ удобно использовать монотонную функцию от нее — логарифм (например, по основанию 2):

$$l = l(\theta) = \log L(\theta) = \log p_n(f^{-1}(Y; \theta)). \tag{11.3.8}$$

Функция (11.3.8) называется *логарифмической функцией правдоподобия*. Отметим, что функция $L(\theta)$ (и $l(\theta)$) характеризует, насколько правдоподобно получить шифртекст Y в предположении, что параметр был равен θ.

Суть криптоанализа методом максимального правдоподобия заключается в том, что в качестве оценки параметра θ^0 принимается «наиболее правдоподобное значение» $\hat{\theta} \in \Theta$, которое максимизирует функцию правдоподобия (11.3.7), или (что эквивалентно) логарифмическую функцию правдоподобия (11.3.8):

$$\hat{\theta} = \arg\max_{\theta \in \Theta} L(\theta) = \arg\max_{\theta \in \Theta} l(\theta). \qquad (11.3.9)$$

Конкретизируем алгоритм криптоанализа (11.3.9) для двух наиболее распространенных на практике моделей источников сообщений X, рассмотренных в главе 4: стационарного источника дискретных сообщений без памяти (X — последовательность независимых в совокупности одинаково распределенных символов алфавита V_N) и марковского источника дискретных сообщений (X — однородная цепь Маркова с пространством состояний V_N).

Примем обозначения:

$$X(\theta) = (x_1(\theta), \ldots, x_n(\theta)) = f^{-1}(Y;\theta) = (f_1^{-1}(Y;\theta), \ldots, f_n^{-1}(Y;\theta)) \qquad (11.3.10)$$

есть вариант расшифровывания шифртекста Y в предположении, что ключевой параметр — θ; при этом если $\theta = \theta^0$, то получаем истинный исходный текст $X(\theta^0) = X$; δ_{jk} — символ Кронекера;

$$\bar{\nu}_j(Y;\theta) = \sum_{i=1}^{n} \delta_{j,x_i(\theta)} = \sum_{i=1}^{n} \delta_{j,f_i^{-1}(Y;\theta)} \in \{0, 1, \ldots, n\} \qquad (11.3.11)$$

есть частота встречаемости символа $j \in V_N$ в последовательности $X(\theta)$, определяемой (11.3.10); при этом выполняется условие нормировки:

$$\sum_{j=0}^{N-1} \bar{\nu}_j(Y;\theta) \equiv n. \qquad (11.3.12)$$

Теорема 11.3.1. Если исходный текст X есть последовательность независимых в совокупности символов $x_1, \ldots, x_n \in \mathcal{A}_n$, имеющих одинаковое распределение вероятностей

$$p_1(u) = \mathbf{P}\{x_i = u\}, \quad u \in \mathcal{A}_N, \quad i = \overline{1, n}, \qquad (11.3.13)$$

то оценка максимального правдоподобия для параметра $\theta^0 \in \Theta$ криптосистемы (11.3.1) по наблюдаемому шифртексту $Y = (y_1, \ldots, y_n) \in \mathcal{A}_N^n$ имеет вид

$$\hat{\theta} = \hat{\theta}(Y) = \arg\max_{\theta \in \Theta} \sum_{j=0}^{N-1} \bar{\nu}_j(Y;\theta) \log p_1(j). \qquad (11.3.14)$$

Доказательство. По условию теоремы, используя (11.3.13), найдем совместное n-мерное дискретное распределение вероятностей (11.3.2) исходного текста X:

$$p_n(a) = p_n(a_1, \ldots, a_n) = \prod_{i=1}^{n} d_i(a_i), \quad a \in \mathcal{A}_N^n.$$

Поэтому в силу (11.3.8) и (11.3.10) логарифмическая функция правдоподобия примет вид

$$l(\theta) = \log \prod_{i=1}^{n} p_1(f_1^{-1}(Y;\theta)) = \sum_{i=1}^{n} \log p_1(f_i^{-1}(Y;\Theta)).$$

Приводя в последней сумме подобные члены и используя обозначение (11.3.11), получим:

$$l(\theta) = \sum_{j=0}^{N-1} \bar{\nu}_1(Y;\theta) \log p_1(j), \quad \theta \in \Theta. \tag{11.3.15}$$

Подставляя выражение (11.3.15) логарифмической функции правдоподобия в (11.3.9), находим оценку $\hat{\theta}$ в виде (11.3.14). □

С помощью теоремы 11.3.1 можно построить криптоатаки типа A_1 для криптосистем, описанных в п. 6.5. Для иллюстрации рассмотрим простейшую криптосистему Цезаря.

Обозначим:

$$\nu_j(Y) = \sum_{i=1}^{n} \delta_{j,y_i}, \quad j \in \mathcal{A}_N, \quad Y \in \mathcal{A}_N^n, \quad \theta \in \Theta \tag{11.3.16}$$

есть частота встречаемости символа j в наблюдаемом шифртексте Y; при этом, как и в случае (11.3.12), выполняется условие нормировки:

$$\sum_{j=0}^{N-1} \nu_j(Y) \equiv n.$$

Следствие 11.3.1. Если используется криптосистема Цезаря:

$$y_i = (x_i + \theta^0) \bmod N, \quad i = \overline{1,n}, \tag{11.3.17}$$

где $\theta^0 \in \Theta = \{1, 2, \ldots, N-1\}$ — неизвестный ключ, то алгоритм криптоанализа по методу максимального правдоподобия имеет вид

$$\hat{\theta} = \hat{\theta}(Y) = \arg\max_{\theta \in \Theta} l(\theta), \quad l(\theta) = \sum_{j=0}^{N-1} \nu_{(j+\theta) \bmod N}(Y) \log p_1(j). \tag{11.3.18}$$

Доказательство. Из (11.3.17) имеем обратную криптофункцию

$$x_i(\theta) = f_i^{-1}(Y;\theta) = (u_i - \theta) \bmod N, \quad i = \overline{1, n},$$

поэтому из (11.3.11) находим частоты $\{\bar{\nu}_j\}$:

$$\bar{\nu}_j(Y;\theta) = \sum_{i=1}^{n} \delta_{j,(y_i-\theta) \bmod N}.$$

Так как $(y_i - \theta) \bmod N = j$ тогда и только тогда, когда $y_i = (j + \theta) \bmod N$, то с учетом (11.3.16) получаем явное выражение частот $\{\bar{\nu}_j\}$ через частоты $\{\nu_j\}$:

$$\bar{\nu}_j(Y;\theta) = \sum_{i=1}^{n} \delta_{(j+\theta) \bmod N, y_i} = \nu_{(j+\theta) \bmod N}(Y;\theta).$$

Подставляя это выражение в (11.3.14), получаем (11.3.18). □

Заметим, что оценки (11.3.14) и (11.3.18) по методу максимального правдоподобия базируются на частотах символов $\{\bar{\nu}_j\}$ или $\{\nu_j\}$, поэтому данный метод часто называют *частотным методом криптоанализа.*

Для иллюстрации алгоритма криптоанализа (11.3.17) рассмотрим шифртекст длиной $n = 41$ (из п. 6.5):

PELCGBYBTLVFPELCGBTENCULNAQPELCGBNANYLFVF, (11.3.19)

полученный шифром Цезаря для неизвестного исходного английского текста X ($N = 26$) с неизвестным ключом $\theta^0 = 13 \in \{1, 2, \ldots, 25\}$. Напомним, что буквенный шифртекст (11.3.19) мы преобразуем в числовую последовательность $Y = (y_1, \ldots, y_n)$, где y_i — уменьшенный на единицу номер позиции i-й буквы шифртекста в английском алфавите:

$$Y = (15, 4, 11, 2, 6, 1, 24, 1, 19, 11, 21, 5, 15,$$
$$4, 11, 2, 6, 1, 19, 4, 13, 2, 20, 11, 12, 0, 16, 15,$$
$$4, 11, 2, 6, 1, 13, 0, 13, 24, 11, 5, 21, 5).$$

В табл. 11.2 приведены значения двоичных логарифмов вероятностей $\{\log p_1(j)\}$ всех 26 символов английского алфавита, полученные в результате статистического анализа объемных англоязычных текстов.

Таблица 11.2

Вероятностные характеристики алфавита

Буква	A	B	C	D	E	F	G	H	I
Символ $j \in V_N$	0	1	2	3	4	5	6	7	8
$\log p_1(j)$	–3.6	–6.0	–5.0	–4.7	–3.0	–5.4	–5.7	–4.2	–3.8

Окончание табл. 11.2

Буква	J	K	L	M	N	O	P	Q	R
Символ $j \in V_N$	9	10	11	12	13	14	15	16	17
$\log p_1(j)$	–9.3	–7.2	–4.6	–5.3	–3.8	–3.7	–5.6	–9.8	–4.0

Буква	S	T	U	V	W	X	Y	Z	
Символ $j \in V_N$	18	19	20	21	22	23	24	25	
$\log p_1(j)$	–3.9	–3.4	–5.2	–6.7	–5.7	–9.0	–5.9	–10.1	

В табл. 11.3 приведены вычисленные, согласно (11.3.16), частоты встречаемости символов в наблюдаемом шифртексте (11.3.19), удовлетворяющие условию нормировки.

Таблица 11.3

Частотные характеристики шифртекста

Буква	A	B	C	D	E	F	G	H	I	J	K	L	M
$j \in V_N$	0	1	2	3	4	5	6	7	8	9	10	11	12
$\nu_j(Y)$	2	4	4	0	4	3	3	0	0	0	0	6	0

Буква	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
$j \in V_N$	13	14	15	16	17	18	19	20	21	22	23	24	25
$\nu_j(Y)$	4	0	3	1	0	0	2	1	2	0	0	2	0

На рис. 11.1 изображен график логарифмической функции правдоподобия.

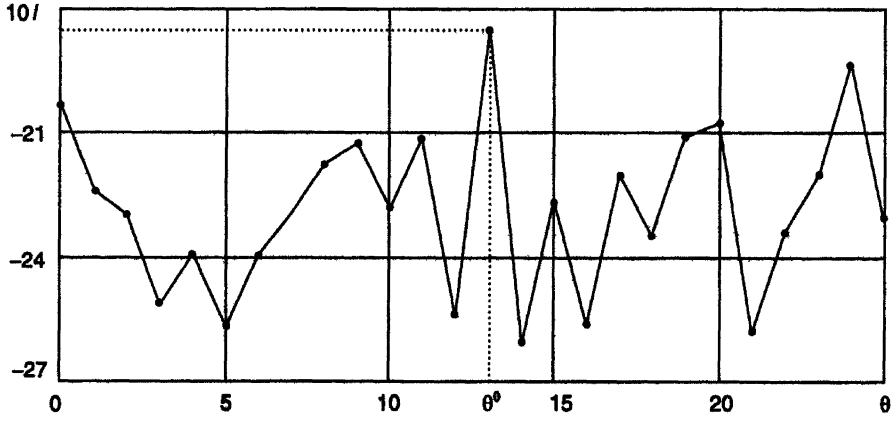


Рис. 11.1. График логарифмической функции правдоподобия

Из него видно, что оценка максимального правдоподобия (11.3.18) совпадает с истинным значением $\hat{\theta} = \theta^0 = 13$. При этом находится открытый текст $\hat{X} = f^{-1}(Y;\hat{\theta}) = X$:

CRYPTOLOGYISCRYPTOGRAPHYANDCRYPTOANALYSIS

Более сложные примеры криптоанализа приведены в [186, 46, 119].

Рассмотрим случай марковского источника дискретных сообщений X (см. главу 5). Наряду с (11.3.10) и (11.3.11) обозначим:

$$\bar{\nu}_{jk}(Y;\theta) = \sum_{i=1}^{n-1} \delta_{j,f_i^{-1}(Y;\theta)} \delta_{k,f_{i+1}^{-1}(Y;\theta)} \in \{0, 1, \ldots, n-1\}, \quad j,k \in \mathcal{A}_N \qquad (11.3.20)$$

есть частота встречаемости биграммы символов (j,k) в варианте $X(\theta) = f^{-1}(Y;\theta)$ расшифровывания шифртекста с использованием ключа $\theta \in \Theta$; при этом выполняется условие нормировки:

$$\sum_{j,k=0}^{N-1} \bar{\nu}_{jk}(Y;\theta) \equiv n-1.$$

Теорема 11.3.2. Если исходный текст $X = (x_1, \ldots, x_n)$ есть однородная цепь Маркова с пространством состояний V_N, вектором начального распределения вероятностей

$$p_1(u) = \mathbf{P}\{x_1 = u\}, \quad u \in V_N, \qquad (11.3.21)$$

и $(N \times N)$-матрицей вероятностей одношаговых переходов $P = (p_{jk})$:

$$p_{jk} = \mathbf{P}\{x_{i+1} = k \mid x_i = j\}, \quad j,k \in V_N; \quad i = 1, 2, \ldots, \qquad (11.3.22)$$

то оценка максимального правдоподобия для параметра $\theta^0 \in \Theta$ криптосистемы (11.3.1) по наблюдаемому шифртексту $Y = (y_1, \ldots, y_n) \in \mathcal{A}_N^n$ имеет вид

$$\hat{\theta} = \hat{\theta}(Y) = \arg\max_{\theta \in \Theta} \left(\sum_{j=0}^{N-1} \delta_{j,f_1^{-1}(Y;\theta)} \log p_1(j) + \sum_{j,k=0}^{N-1} \bar{\nu}_{jk}(Y;\theta) \log p_{jk} \right). \qquad (11.3.23)$$

Доказательство. В силу марковского свойства с помощью (11.3.21) и (11.3.22) найдем совместное n-мерное дискретное распределение вероятностей (11.3.2) исходного текста X:

$$p_n(a) = p_n(a_1, \ldots, a_n) = p_1(a_1) \prod_{i=1}^{n-1} p_{a_i, a_{i+1}}, \qquad a \in \mathcal{A}_N^n.$$

Поэтому в силу (11.3.8) и (11.3.10) логарифмическая функция правдоподобия примет вид

$$l(\theta) = \log p_1(f_1^{-1}(Y;\theta)) + \sum_{i=1}^{n-1} \log p_{f_i^{-1}(Y,\theta), f_{i+1}^{-1}(Y;\theta)}.$$

Используя обозначение (11.3.20), получаем эквивалентный вид

$$l(\theta) = \sum_{j=0}^{N-1} \delta_{j, f_1^{-1}(Y;\theta)} \log p_1(j) + \sum_{j,k=0}^{N-1} \bar{\nu}_{jk}(Y;\theta) \log p_{jk}. \qquad (11.3.24)$$

Подставляя выражение (11.3.24) логарифмической функции правдоподобия в (11.3.9), находим оценку $\hat{\theta}$ в виде (11.3.23). □

Проиллюстрируем применение теоремы 11.3.2 для простейшей криптосистемы Цезаря из п. 6.4. Обозначим:

$$\nu_{jk}(Y) = \sum_{i=1}^{n-1} \delta_{j,y_i} \delta_{k,y_{i+1}}, \quad j,k \in \mathcal{A}_N \qquad (11.3.25)$$

есть частота встречаемости биграммы символов (j,k) в наблюдаемом шифртексте Y; при этом

$$\sum_{j,k=0}^{N-1} \nu_{jk}(Y) \equiv n-1.$$

Следствие 11.3.2. Если используется криптосистема Цезаря (11.3.17) с неизвестным ключом $\theta^0 \in \Theta = \{1, 2, \ldots, N-1\}$, то алгоритм криптоанализа по методу максимального правдоподобия имеет вид

$$\hat{\theta} = \hat{\theta}(Y) = \arg\max_{\theta \in \Theta} \left(\sum_{j=0}^{N-1} \delta_{y_1,(j+\theta) \bmod N} \log p_1(j) + \right.$$

$$\left. + \sum_{j,k=0}^{N-1} \nu_{(j+\theta) \bmod N,(k+\theta) \bmod N}(Y) \log p_{jk} \right). \qquad (11.3.26)$$

Доказательство. Как и при доказательстве следствия 11.3.1, в силу (11.3.17) и (11.3.25) справедливы соотношения:

$$\delta_{j,(y_1-\theta) \bmod N} \equiv \delta_{y_1,(j+\theta) \bmod N},$$

$$\bar{\nu}_{jk}(Y;\theta) \equiv \nu_{(j+\theta) \bmod N,(k+\theta) \bmod N}.$$

Подставляя их в (11.3.23), получаем (11.3.26). □

Точность решения задачи криптоанализа, т.е. точность оценки $\hat{\theta}$ параметра θ^0, удобно характеризовать вероятностью ошибки:

$$r = r(n,\theta^0) = \mathbf{P}_{\theta^0}\{\hat{\theta}(Y) \neq \theta^0\}, \quad n \in \mathbb{N}, \quad \theta^0 \in \Theta.$$

Согласно (11.3.6), это выражение для вероятности ошибки можно записать в эквивалентном виде

$$r(n,\theta^0) = 1 - \sum_{b \in V_N^n} p_n(f^{-1}(b;\theta))\,\delta_{\hat{\theta}(b),\theta^0}.$$

Чем меньше величина r, тем точнее решается задача криптоанализа.

Условимся говорить, что алгоритм криптоанализа — *состоятельный*, если при увеличении длины шифртекста n вероятность ошибки r монотонно стремится к нулю, каково бы ни было истинное значение θ^0:

$$r(n,\theta^0) \to 0 \quad (n \to \infty), \quad \theta^0 \in \Theta.$$

Для практических приложений оказывается важным вычисление *минимальной длины шифртекста* n_ε, гарантирующей заданный уровень $\varepsilon \in (0,1)$ вероятности ошибки:

$$n_\varepsilon = \min\{n\colon \sup_{\theta^0 \in \Theta} r(n,\theta^0) \le \varepsilon\}.$$

Кратко остановимся еще на одном методе криптоанализа, использующем теорию статистических решений — *байесовском методе*. Он основан на двух модельных предположениях U_1 и U_2, задаваемых соотношениями (11.3.2) и (11.3.3). Характерным для байесовского подхода является предположение U_2 о том, что параметр $\theta^0 \in \Theta$ случаен, не зависит от X и имеет известное априорное распределение вероятностей $\pi(\theta)$, $\theta \in \Theta$.

Будем строить оценку $\hat{\theta}$ параметра θ^0 по наблюдаемому шифртексту $Y \in V_N^n$ с помощью некоторой статистики

$$\hat{\theta} = T(Y)\colon \mathcal{A}_N^n \to \Theta, \quad T(\cdot) \in \mathcal{T}, \tag{11.3.27}$$

где $\mathcal{T}$ — некоторое заданное семейство функций, осуществляющих отображение $\mathcal{A}_N^n \to \Theta$.

Как принято в байесовской статистике, определим *функцию потерь*:

$$w = w(\theta',\theta) \ge 0, \quad \theta',\theta \in \Theta,$$

где w — величина потерь, которые несет криптоаналитик в ситуации, когда истинное значение $\theta^0 = \theta'$, а он принял решение $\hat{\theta} = \theta$ (обычно полагается, что при правильном решении ($\theta = \theta'$) потери равны нулю: $w(\theta',\theta') \equiv 0$). Приведем три примера функций потерь:

$$w_1(\theta',\theta) = \begin{cases} 1, & \theta \ne \theta', \\ 0, & \theta = \theta'; \end{cases} \tag{11.3.28}$$

$$w_2(\theta',\theta) = \|\theta - \theta'\| = \sum_{k=1}^{L} (1 - \delta_{\theta_k,\theta'_k})$$

есть количество ошибочно определенных компонентов вектора θ^0;

$$w_3(\theta', \theta) = \begin{cases} 0, & \|\theta - \theta'\| \le \varepsilon, \\ 1, & \|\theta - \theta'\| > \varepsilon, \end{cases}$$

где $\varepsilon \ge 0$ — некоторое заданное пороговое значение. Функцию потерь (11.3.28) принято называть $(0-1)$-*функцией потерь*.

Определим *функционал риска* как математическое ожидание (среднее значение) функции потерь:

$$r = r(T(\cdot)) = \mathbf{E}\{w(\theta^0, T(Y))\} \ge 0, \quad T(\cdot) \in \mathcal{T}.$$

Используя условия U_1, U_2 и (11.3.5), (11.3.6), легко получить явный вид этого функционала:

$$r(T(\cdot)) = \sum_{\theta \in \Theta} \sum_{b \in V_N^n} w(\theta, T(b))\, \pi(\theta)\, p_n(f^{-1}(b; \theta)). \tag{11.3.29}$$

Чем меньше риск $r(T(\cdot))$, тем меньше средние потери при выполнении криптоанализа $\hat{\theta} = T(Y)$ с помощью статистики $T(\cdot)$.

Байесовской оценкой для θ^0 называется оценка $\hat{\theta}^* = T^*(Y)$, где $T^*(\cdot)$ — оптимальная функция, минимизирующая функционал риска:

$$r(T^*(\cdot)) = \inf_{T(\cdot) \in \mathcal{T}} r(T(\cdot)). \tag{11.3.30}$$

Теорема 11.3.3. Если используется $(0-1)$-функция потерь (11.3.28), а $\mathcal{T}$ есть множество всевозможных функций, задающих отображение $\mathcal{A}_N^n \to \Theta$, то байесовская оценка имеет вид

$$\hat{\theta}^* = T^*(Y) = \arg\max_{\theta \in \Theta} (\log \pi(\theta) + l(\theta)), \quad Y \in \mathcal{A}_N^n, \tag{11.3.31}$$

где $l(\theta) = \log p_n(f^{-1}(Y; \Theta))$ — логарифмическая функция правдоподобия, определяемая (11.3.8). При этом достигается минимум риска:

$$r_{\min} = r(T^*(\cdot)) = 1 - \sum_{b \in V_N^n} \max_{\theta \in \Theta} (\pi(\theta) p_n(f^{-1}(b; \theta))). \tag{11.3.32}$$

Доказательство. Подставляя (11.3.28) в (11.3.29), получаем явный вид экстремальной задачи (11.3.30):

$$r(T(\cdot)) = 1 - \sum_{b \in \mathcal{A}_N^n} (\pi(\theta) p_n(f^{-1}(d; \theta)))\Big|_{\theta = T(b)} \to \min_{T(\cdot)}, \tag{11.3.33}$$

которая эквивалентна задаче максимизации:

$$\sum_{b \in \mathcal{A}_N^n} (\pi(\theta) p_n(f^{-1}(d; \theta)))\Big|_{\theta = T(b)} \to \max_{T(\cdot)}.$$

Так как целевая функция этой задачи сепарабельна, а ограничения на $T(\cdot)$ носят лишь точечный характер, то

$$T^*(b) = \arg\max_{\theta\in\Theta}(\pi(\theta)\, p_n(f^{-1}(b;\theta))) =$$

$$= \arg\max_{\theta\in\Theta}(\log\pi(\theta) + \log p_n(f^{-1}(b;\theta))). \qquad (11.3.34)$$

Отсюда следует (11.3.31). Подставляя (11.3.34) в (11.3.33), получаем (11.3.32). □

Сравнивая байесовскую оценку (11.3.31) и оценку максимального правдоподобия (11.3.9), видим, что различие обусловлено наличием дополнительного слагаемого $\log\pi(\theta)$ в целевой функции. Это слагаемое учитывает априорную информацию о θ^0.

Следствие 11.3.3. Если априорное распределение вероятностей $\pi(\theta)$ параметра θ^0 является равномерным:

$$\pi(\theta) = \frac{1}{|\Theta|}, \quad \theta\in\Theta, \qquad (11.3.35)$$

то байесовская оценка (11.3.31) совпадает с оценкой (11.3.9) по методу максимального правдоподобия.

Доказательство. Подставляя (11.3.35) в (11.3.31), получаем

$$\hat{\theta}^* = T^*(Y) = \arg\max_{\theta\in\Theta} l(\theta) = \hat{\theta}, \quad Y\in\mathcal{A}_N^n. \qquad \square$$

В заключение рассмотрим вычислительную сложность алгоритмов криптоанализа, использующих теорию статистических решений.

Вычислительная сложность алгоритмов криптоанализа, основанных на методе максимального правдоподобия и байесовском методе, определяется сложностью алгоритма максимизации функции правдоподобия $l(\theta)$ в (11.3.9) и целевой функции $l_1(\theta) = \log\pi(\theta) + l(\theta)$ в (11.3.31). Задачи максимизации (11.3.9), (11.3.31) и их частные случаи (11.3.14), (11.3.18), (11.3.23), (11.3.26) являются задачами нелинейного целочисленного программирования. Если они решаются полным перебором значений $\theta\in\Theta$, то вычислительная сложность алгоритма криптоанализа легко находится:

$$W = c|\Theta|\, n(N^{s+1}+1), \qquad (11.3.36)$$

где c — некоторая константа, зависящая от конкретного способа реализации алгоритма криптоанализа; $s\in\{0,1\}$ — параметр: $s=0$ для модели независимых символов текста, $s=1$ для марковской модели. Для снижения вычислительной сложности (11.3.36) перспективным является применение специальных методов дискретной оптимизации, например, метода ветвей и границ.

11.4. Разностный криптоанализ

В 1990 г. Э. Байхэм и А. Шамир предложили метод разностного анализа алгоритма шифрования DES (см. описание в п. 10.1). Разработанный метод не позволил провести атаку на полный 16-тактовый DES — нестойкими к разностному анализу оказались лишь упрощенные версии криптоалгоритма с редуцированным количеством тактов. Однако проведенные криптологами исследования выявили интересные особенности в выборе функциональных узлов криптоалгоритма. Оказалось, что уже при небольших изменениях в таблицах подстановки DES появляется возможность проводить эффективные атаки. Действительно, после появления результатов Байхэма и Шамира компанией IBM — разработчиком DES — были опубликованы некоторые принципы построения криптоалгоритма, в которых закладывалась стойкость к методу разностного анализа — «мощной технике исследования, которая может быть применена ко многим криптосистемам» [177].

Разностный анализ является общим методом криптоанализа блочно-итерационных криптосистем. С использованием этого метода были построены эффективные атаки на криптоалгоритмы Lucifer, FEAL, LOKI, Khufu. Существуют модификации метода: *линейно-разностный анализ* (linear-differential analysis) [177], *анализ с использованием близких ключей* (related-key analysis) [52, 110], *разностный анализ с ошибками* (differential-fault analysis) [55].

Изложим основные принципы разностного анализа. Рассмотрим блочно-итерационную криптосистему, которая осуществляет взаимно-однозначное преобразование элементов множества A мощностью $N = |A|$. Преобразование открытого текста X в шифртекст Y такой системы является композицией r тактовых биективных преобразований $F_{\kappa^{(1)}}, \ldots, F_{\kappa^{(r)}}$:

$$Y = F_{\kappa^{(r)}} \cdots F_{\kappa^{(1)}}(X), \tag{11.4.1}$$

определяемых *тактовыми ключами* $\kappa^{(1)}, \ldots, \kappa^{(r)} \in K$, которые вырабатываются из исходного ключа θ^0.

Пусть на A задана групповая операция $\boxplus$, $\boxminus a$ — обратный к $a \in A$ элемент группы, 0 — нейтральный элемент. *Разность* ΔX элементов $X, X^* \in A$ определим как $X \boxplus (\boxminus X^*)$ или просто: $\Delta X = X \boxminus X^*$.

Пара (X, X^*) открытых текстов при последовательном выполнении преобразования (11.4.1) порождает последовательность разностей

$$\Delta Y(1) = Y(1) \boxminus Y^*(1), \ldots, \Delta Y(r) = Y(r) \boxminus Y^*(r),$$

где $Y(i)$, $Y^*(i)$ — результаты выполнения i тактов преобразования над X и X^* соответственно.

Последовательность $(\alpha, \beta(1), \ldots, \beta(s))$ возможных значений разностей ΔX, $\Delta Y(1), \ldots, \Delta Y(s)$ называется в разностном анализе *s-тактовой характеристикой*, а пара $(\alpha, \beta(s))$ возможных значений разностей ΔX, $\Delta Y(s)$ — *s-тактовым дифференциалом*, $1 \le s \le r$.

Заметим, что $\Delta Y(1) = \ldots = \Delta Y(r) = 0$, если $\Delta X = 0$. Тривиальные характеристики $(0, \ldots, 0)$ и дифференциалы $(0, 0)$ не используются в разностном анализе, и далее мы их не рассматриваем.

Предполагая, что X, X^* — случайные элементы A с равномерным распределением вероятностей, а случайные тактовые ключи $\kappa^{(1)}, \ldots, \kappa^{(r)}$ независимы и равномерно распределены на K, введем в рассмотрение вероятность характеристики

$$\mathbf{P}\{\Delta Y(1) = \beta(1), \ldots, \Delta Y(s) = \beta(s) \mid \Delta X = \alpha\}$$

и вероятность дифференциала

$$\mathbf{P}\{Y(s) = \beta(s) \mid \Delta X = \alpha\}.$$

Атака по методу разностного анализа на криптосистему (11.4.1) проводится при выбираемом открытом тексте, т.е. это атака типа A_3. Ее целью является восстановление неизвестного тактового ключа $\kappa^{(r)}$.

Предварительно исследователь находит $(r-1)$-тактовый дифференциал α, $\beta(r-1)$ с максимально возможной вероятностью $p > \dfrac{1}{N-1}$.

Затем в ходе T экспериментов исследователь генерирует случайный открытый текст X_t, вычисляет $X_t^* = X_t \boxplus \alpha$ и находит пару соответствующих шифр-текстов Y_t, Y_t^*, $t = 1, \ldots, T$.

В качестве оценки $\hat{\kappa}$ искомого ключа $\kappa^{(r)}$ выбирается элемент

$$\hat{\kappa} = \arg\max_{\kappa \in K} \sum_{t=1}^{T} \mathbf{I}\{F_\kappa^{-1}(Y_t) \boxminus F_\kappa^{-1}(Y_t^*) = \beta(r-1)\}. \qquad (11.4.2)$$

Здесь $\mathbf{I}\{\cdot\}$ — индикатор случайного события: $\mathbf{I}\{A\} = 1$, если событие A наступает, и 0 в противном случае.

Оценим вычислительную сложность метода, выраженную количеством шифрований $2T$, необходимых для успешного восстановления ключа $\kappa^{(r)}$.

Будем предполагать, что выполнены две гипотезы.

Гипотеза 1 (*стохастической эквивалентности* [120]). Почти для всех значений тактовых ключей $(\omega_1, \ldots, \omega_{r-1})$ имеет место

$$\mathbf{P}\{\Delta Y(r-1) = \beta(r-1) \mid \Delta X = \alpha\} \approx$$
$$\approx \mathbf{P}\{\Delta Y(r-1) = \beta(r-1) \mid \Delta X = \alpha, \kappa^{(1)} = \omega_1, \ldots, \kappa^{(r-1)} = \omega_{r-1}\}.$$

Гипотеза 2 (*различимости ключей*). Для пары (X, X^*) различных случайных элементов из A с равномерным распределением вероятностей и для двух различных ключей $\kappa_1, \kappa_2 \in K$ имеет место

$$\mathbf{P}\{F_{\kappa_2}^{-1} F_{\kappa_1}(X) \boxminus F_{\kappa_2}^{-1} F_{\kappa_1}(X^*) = \beta(r-1) \mid X \boxminus X^* = \beta(r-1)\} \approx$$
$$\approx \mathbf{P}\{F_{\kappa_2}^{-1} F_{\kappa_1}(X) \boxminus F_{\kappa_2}^{-1} F_{\kappa_1}(X^*) = \beta(r-1) \mid X \boxminus X^* \neq \beta(r-1)\}.$$

При введении этих гипотез мы «идеализируем» криптосистему и при этом в большинстве случаев не нарушаем точность получаемых выводов. Отметим, что при нарушении этих гипотез можно использовать метод *условного разностного анализа* [48]. Множество $\Theta = K^{r-1}$ всевозможных значений $(r-1)$-наборов тактовых ключей разбивается на классы $\Theta_1, \ldots, \Theta_L$. Для i-го класса находится свой высоковероятный дифференциал $(\alpha_i, \beta_i(r-1))$.

При выполнении гипотез для $\kappa = \kappa^{(r)}$ имеем

$$\mathbf{E}\left\{\sum_{t=1}^{T} \mathbf{I}\{F_\kappa^{-1}(Y_t) \boxminus F_\kappa^{-1}(Y_t^*) = \beta(r-1)\}\right\} = Tp,$$

а для $\kappa \neq \kappa^{(r)}$

$$\mathbf{E}\left\{\sum_{t=1}^{T} \mathbf{I}\{F_\kappa^{-1}(Y_t) \boxminus F_\kappa^{-1}(Y_t^*) = \beta(r-1)\}\right\} = \frac{T}{N-1}.$$

Для выполнения условия $\hat{\kappa} = \kappa^{(r)}$ необходимо выполнение неравенства

$$Tp \geq 1 + T/(N-1),$$

откуда получаем, что вычислительная сложность атаки по методу разностного анализа составляет не менее

$$W_{\min} = 2\left(p - \frac{1}{N-1}\right)^{-1}$$

шифрований вида (11.4.1).

Далее рассмотрим специальный класс криптосистем, которые представляют особый интерес для разностного анализа [120].

Пусть $a_1, \ldots, a_N$ — все различные элементы A, $a_1 = 0$. Семейству тактовых преобразований $Y = F_\kappa(X)$, $\kappa \in \mathrm{K}$, поставим в сооответствие $N \times N$ матрицу $P = (p_{ij})$, (i,j)-й элемент которой есть вероятность однотактовой характеристики (a_i, a_j):

$$p_{ij} = \mathbf{P}\{\Delta Y = a_j \mid \Delta X = a_i\}.$$

Легко показать, что матрица P является дважды стохастической, т.е. для всех i, $1 \leq i \leq N$,

$$\sum_{k=1}^{N} p_{ik} = \sum_{k=1}^{N} p_{ki}. \tag{11.4.3}$$

Блочно-итерационную криптосистему с тактовым преобразованием $Y = F_\kappa(X)$, $\kappa \in \mathrm{K}$ принято называть *марковской*, если для любых $\alpha, \beta, \gamma \in A$ выполняется равенство

$$\mathbf{P}\{\Delta Y = \beta \mid \Delta X = \alpha\} = \mathbf{P}\{\Delta Y = \beta \mid \Delta X = \alpha, X = \gamma\}$$

в предположении, что случайный ключ κ имеет равномерное на K распределение вероятностей.

Зададим для разностей пар открытых текстов распределение вероятностей $\pi = (\pi_1, \ldots, \pi_N)$: $\pi_i = \mathbf{P}\{\Delta X = a_i\}$ и будем предполагать, что тактовые ключи марковской криптосистемы (11.4.1) независимы и равномерно распределены на K. Покажем, что в этом случае последовательность разностей ΔX, $\Delta Y(1), \ldots, \Delta Y(r)$ является цепью Маркова с состояниями $a_1, \ldots, a_N$ и матрицей переходных вероятностей P.

Достаточно установить, что

$$\mathbf{P}\{\Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1), \Delta X = \alpha\} = \mathbf{P}\{\Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1)\}$$

для всех $\alpha, \beta(1), \beta(2) \in A$.

Действительно, по формулам полной вероятности и умножения вероятностей имеем цепочку равенств:

$$\begin{aligned}
&\mathbf{P}\{\Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1), \Delta X = \alpha\} = \\
&= \sum_{\gamma \in A} \mathbf{P}\{Y(1) = \gamma, \Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1), \Delta X = \alpha\} = \\
&= \sum_{\gamma \in A} \mathbf{P}\{Y(1) = \gamma \mid \Delta Y(1) = \beta(1), \Delta X = \alpha\} \times \\
&\times \mathbf{P}\{\Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1), Y(1) = \gamma, \Delta X = \alpha\} = \\
&= \sum_{\gamma \in A} \mathbf{P}\{Y(1) = \gamma \mid \Delta Y(1) = \beta(1), \Delta X = \alpha\} \mathbf{P}\{\Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1), Y(1) = \gamma\} = \\
&= \sum_{\gamma \in A} \mathbf{P}\{Y(1) = \gamma \mid \Delta Y(1) = \beta(1), \Delta X = \alpha\} \mathbf{P}\{\Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1)\} = \\
&= \mathbf{P}\{\Delta Y(2) = \beta(2) \mid \Delta Y(1) = \beta(1)\}.
\end{aligned}$$

Для марковской криптосистемы легко вычисляются вероятности характеристик и дифференциалов. Вероятность s-тактовой характеристики $a_{i_0}, a_{i_1}, \ldots, a_{i_s}$ есть $p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{s-1} i_s}$, а вероятность s-тактового дифференциала a_i, a_j равна (i, j)-му элементу матрицы P^s.

Далее рассмотрим класс DES-подобных криптосистем (см. главу 10). Положим $A = V_{2n}$, $N = 2^{2n}$, $\mathrm{K} = V_m$; $\oplus$ — операция покомпонентного сложения двоичных векторов по модулю 2; $a \,||\, b$ — конкатенация векторов a и b; $\mathbf{L}a$, $\mathbf{R}a$ — левый и правый n-блоки вектора $a \in V_2^{2n}$.

Зададим отображения $E: V_n \to V_m$ и $f: V_m \to V_n$. Действие тактового преобразования F_κ криптосистемы (11.4.1) определим следующим образом:

$$F_\kappa(X) = F_\kappa(\mathbf{L}X \,||\, \mathbf{R}X) = (\mathbf{R}X \,||\, \mathbf{L}X \oplus f(E(\mathbf{R}X) \oplus \kappa)). \qquad (11.4.4)$$

Покажем, что криптосистема вида (11.4.4) является марковской относительно операции $\oplus$, если E — линейное отображение, т.е. $E(x \oplus y) = E(x) \oplus E(y)$ для всех $x, y \in V_n$.

Для этого достаточно установить, что для всех $\alpha, \beta \in V_n$ и всех $\gamma \in V_m$:

$$\mathbf{P}\{f(z) \oplus f(z^*) = \beta \mid x \oplus x^* = \alpha, x = \gamma\} = \mathbf{P}\{f(z) \oplus f(z^*) = \beta \mid x \oplus x^* = \alpha\},$$

где $x, x^* \in V_n$, $z = E(x) \oplus \kappa$, $z^* = E(x^*) \oplus \kappa$, κ — случайный элемент V_m с равномерным распределением вероятностей.

Действительно,

$$\mathbf{P}\{f(z)\oplus f(z^*)=\beta\,|\,x\oplus x^*=\alpha, x=\gamma\}=$$

$$=\sum_{\lambda\in V_m}\mathbf{P}\{f(z)\oplus f(z^*)=\beta, z=\lambda\,|\,x\oplus x^*=\alpha, x=\gamma\}=$$

$$=\sum_{\lambda\in V_m}\mathbf{P}\{f(z)\oplus f(z^*)=\beta\,|\,x\oplus x^*=\alpha, x=\gamma, z=\lambda\}\,\mathbf{P}\{z=\lambda\,|\,x\oplus x^*=\alpha, x=\gamma\}=$$

$$=\sum_{\lambda\in V_m}\mathbf{P}\{f(z)\oplus f(z^*)=\beta\,|\,x\oplus x^*=\alpha, z=\lambda\}\,\mathbf{P}\{\kappa=E(\gamma)+\lambda\}=$$

$$=\frac{1}{|V_m|}\sum_{\lambda\in V_m}\mathbf{P}\{f(z)\oplus f(z^*)=\beta\,|\,x\oplus x^*=\alpha, z=\lambda\}=$$

$$=\mathbf{P}\{f(z)\oplus f(z^*)=\beta\,|\,x\oplus x^*=\alpha\},$$

что и требовалось доказать.

Пример. DES является 16-тактовой криптосистемой вида (11.4.1), (11.4.4). При этом $n=32$, $m=48$. Предложенная Байхэмом и Шамиром атака на DES основана на двухтактовой характеристике (α, β, α):

$$\alpha=19600000\ 00000000_{16},\quad \beta=00000000\ 19600000_{16}.$$

Вероятность такой характеристики

$$p=\frac{14}{64}\,\frac{8}{64}\,\frac{10}{64}\approx\frac{1}{234}.$$

Заметим, что начальная и конечная разности характеристики совпадают. Такие характеристики в разностном анализе принято называть *итеративными*.

Найденную двухтактовую итеративную характеристику можно использовать для построения 15-тактовой характеристики (α, β, α, β, α, β), вероятность которой оказывается равной $2^{-55.1}$.

11.5. Линейный криптоанализ

Этот метод криптоанализа разработан в 1993 г. известным японским криптологом Митсуру Матсуи [135, 136]. В первоначальном виде он был сформулирован применительно к криптоанализу криптосистемы DES (см. главу 10). Дадим вначале общее описание метода *линейного криптоанализа*, а затем — описание его применения к DES.

Пусть рассматривается произвольная симметричная криптосистема для защиты двоичных сообщений ($N=2$), общая математическая модель которой представлена в п. 11.2. Обозначим: $X=(x_1,\ldots,x_n)\in V_n$ — исходная двоичная последовательность из n символов; $\theta^0=(\theta_1^0,\ldots,\theta_L^0)\in\Theta\subseteq V_L$ — неизвестный

ключевой параметр, который предполагается двоичным вектором с L компонентами;

$$Y = (y_1, \ldots, y_n) = f(X; \theta^0) \in V_n \tag{11.5.1}$$

есть шифртекст, состоящий из n двоичных символов, где $f: V_n \Theta \to V_n$ — криптографическая функция, определяемая конкретным типом используемой криптосистемы.

Метод линейного криптоанализа применяется в криптоатаках типа A_3 (см. п. 11.1), когда криптоаналитик имеет возможность генерировать случайные исходные тексты $X^{(1)}, \ldots, X^{(M)}$ и наряду с ними наблюдать соответствующие им шифртексты $Y^{(1)} = f(X^{(1)}; \theta^0), \ldots, Y^{(M)} = f(X^{(M)}; \theta^0)$ для любого $M \in \mathbb{N}$.

При генерации исходного текста $X \in V_n$ будем предполагать, что случайные биты $x_1, \ldots, x_n \in V$ независимы в совокупности и равновероятны:

$$\mathbf{P}\{x_i = 0\} = \mathbf{P}\{x_i = 1\} = \frac{1}{2}, \quad i = \overline{1, n}. \tag{11.5.2}$$

Принято говорить, что выполняется *«приближенное линейное соотношение»* между X, Y, θ^0:

$$\lambda(X, Y) = \sum_{i=1}^{n} a_i x_i \oplus \sum_{i=1}^{n} b_i y_i = \sum_{k=1}^{L} c_k \theta_k^0, \tag{11.5.3}$$

где $a = (a_i) \in V_n$, $b = (b_i) \in V_n$, $c = (c_k) \in V_L$ — заданные неслучайные векторы весовых коэффициентов, если в условиях (11.5.1), (11.5.2) данное соотношение (11.5.3) выполняется с вероятностью $\frac{1}{2} + \varepsilon$:

$$\mathbf{P}\left\{\lambda(X, f(X; \theta^0)) = \sum_{i=1}^{L} c_i \theta_i^0\right\} = \frac{1}{2} + \varepsilon. \tag{11.5.4}$$

Отметим, что арифметические операции в (11.5.3) выполняются по модулю 2 (так что операция сложения — это обычная XOR-операция). Величина $|\varepsilon| \in \left[0, \frac{1}{2}\right]$ называется *эффективностью соотношения* (11.5.3), а коэффициенты a, b, c — его параметрами. Приближенное соотношение (11.5.3) представим в эквивалентном виде, записывая (11.5.1) покомпонентно:

$$\lambda(X, Y) = \sum_{i=1}^{n} (a_i x_i \oplus b_i f_i(X; \theta^0)) = \sum_{k=1}^{L} c_k \theta_k^0. \tag{11.5.5}$$

Приведем примеры приближенных линейных соотношений (11.5.3)—(11.5.5).

Пример 1. *Криптосистема Вернама (поточная криптосистема).* В этом случае $L = n$, а соотношение (11.5.1) имеет вид

$$y_i = f_i(X;\theta^0) = x_i \oplus \theta_i^0, \quad i = \overline{1,n}. \tag{11.5.6}$$

Тогда подстановка (11.5.6) в (11.5.5) дает

$$\lambda(X,Y) = \sum_{i=1}^{n}(a_i \oplus b_i)x_i \oplus \sum_{i=1}^{L} b_i\theta_i^0.$$

Учитывая равномерное распределение (11.5.2) и независимость битов $\{x_i\}$, обнаруживаем две ситуации:

а) если $b = a$, то $a \oplus b \equiv 0$ и $\lambda(X,Y) = \sum_{i=1}^{n} a_i\theta_i^0$ становится неслучайной величиной так, что (11.5.4) примет вид

$$\mathbf{P}\left\{\lambda(X, f(X;\theta^0)) = \sum_{i=1}^{n} a_i\theta_i^0\right\} = 1;$$

при этом эффективность линейного соотношения максимальна и равна $|\varepsilon| = \frac{1}{2}$;

б) если $b \neq a$, то $\sum_{i=1}^{n}(a_i \oplus b_i)\, x_i$ является случайным битом с тем же распределением (11.5.2) (см. свойство воспроизводимости РРСП, глава 7), поэтому (11.5.4) примет иной вид:

$$\mathbf{P}\left\{\lambda(X, f(X;\theta^0)) = \sum_{i=1}^{n} a_i\theta_i^0\right\} = \frac{1}{2};$$

при этом эффективность линейного соотношения минимальна и равна $|\varepsilon| = 0$; такое соотношение бесполезно для целей линейного криптоанализа.

Пример 2. *Нелинейное криптографическое преобразование* ($L = n = 3$, $i = \overline{1,3}$):

$$y_i = f_i(X;\theta^0) = x_i \oplus x_{(i+1)\bmod 3}x_{(i+2)\bmod 3} \oplus x_{(i+2)\bmod 3}x_{(i+3)\bmod 3} \oplus \theta_i^0.$$

Подставляя это выражение в (11.5.5), полагая $c = b = a$ и проводя тождественные преобразования, получаем левую часть (11.5.5) в виде

$$\lambda(X,Y) = \sum_{i=1}^{n} a_i\theta_i^0 \oplus \sum_{i=1}^{L} a_i x_{(i+2)\bmod 3}(x_{(i+1)\bmod 3} + x_i).$$

В частности, если вектор a является единичным ортом, т.е. имеет единицу только на некотором k-м месте ($k = 1; 2; 3$), то

$$\lambda(X,Y) = \theta_k^0 \oplus x_{(k+2)\bmod 3}(x_{(k+1)\bmod 3} + x_k).$$

Так как случайные биты в последнем выражении независимы и одинаково распределены, согласно (11.5.2), то

$$\mathbf{P}\{\lambda(X,Y)=\theta_k^0\}=\frac{3}{4}=\frac{1}{2}+\frac{1}{4},\quad \varepsilon=\frac{1}{4}.$$

Таким образом, найдено приближенное линейное соотношение вида (11.5.3) с характеристикой эффективности $|\varepsilon|=\frac{1}{4}$.

Теорема 11.5.1. Максимальное по коэффициентам a, b, c значение эффективности $|\varepsilon|=\frac{1}{2}$ для приближенного линейного соотношения (11.5.5) достигается тогда и только тогда, когда криптографическая функция (11.5.1) линейна по X и θ^0.

Доказательство. В силу (11.5.4) и (11.5.5) при $\varepsilon\pm\frac{1}{2}$ (для определенности $\varepsilon=\frac{1}{2}$) величина $\lambda(X,Y)$ становится неслучайной:

$$\sum_{i=1}^{n}(a_i x_i \oplus b_i f_i(X;\theta^0))\equiv\varphi(\theta^0), \tag{11.5.7}$$

где $\varphi(\theta^0)$ — некоторая функция аргумента θ^0 $(\Theta\to V)$. Это возможно тогда и только тогда, когда каждое слагаемое суммы не случайно:

$$a_i x_i \oplus b_i f_i(X;\theta^0)\equiv\varphi_i(\theta^0),\quad i=\overline{1,n};\quad \varphi(\theta^0)=\sum_{i=1}^{n}\varphi_i(\theta^0).$$

Это возможно тогда и только тогда, когда

$$f_i(X;\theta^0)\equiv d_i x_i\oplus\varphi_i(\theta^0),\quad i=\overline{1,n}.$$

Выбирая $\{a_i,b_i\}$ так, что

$$b_i d_i=a_i,\quad i=\overline{1,n},$$

приходим к (11.5.7). □

Теорема 11.5.1 и приведенные примеры показывают, что эффективность $|\varepsilon|$ характеризует степень линейности криптографической функции. Для метода линейного криптоанализа требуются приближенные линейные соотношения вида (11 5.3), коэффициенты a, b, c которого найдены из условия максимума эффективности:

$$|\varepsilon(a,b,c)|\to\max_{a,b,c}. \tag{11.5.8}$$

К сожалению, в настоящее время эффективные алгоритмы нахождения оптимальных коэффициентов a^*, b^*, c^* из условия (11 5.8) не известны.

Метод линейного криптоанализа заключается в выполнении последовательности шагов.

1. «Тщательно анализируется» криптографическая функция $f(\cdot)$ и находится L наилучших в смысле критерия (11.5.8) наборов весовых коэффициентов:

$$(a^{(1)}, b^{(1)}, c^{(1)}), \ldots, (a^{(L)}, b^{(L)}, c^{(L)}).$$

К этим наборам предъявляется еще одно требование: $(L \times L)$-матрица C, составленная из L строк $c^{(1)}, \ldots, c^{(L)}$,

$$C = \begin{pmatrix} c^{(1)} \\ \vdots \\ c^{(L)} \end{pmatrix},$$

должна быть невырожденной ($|C| \neq 0$). Иначе говоря, требуется, чтобы векторы $c^{(1)}, \ldots, c^{(L)} \in V_L$ были линейно независимы.

2. Генерируется M независимых исходных текстов $X^{(1)}, \ldots, X^{(M)}$, как было указано выше, и регистрируются соответствующие шифртексты $Y^{(1)}, \ldots, Y^{(M)}$.

3. Для каждой j-й пары $X^{(j)}, Y^{(j)}$, $j = \overline{1, M}$ вычисляется левая часть каждого из L соотношений вида (11.5.3):

$$\lambda^{(j,l)} = \lambda^{(l)}(X^{(j)}, Y^{(j)}) = \sum_{i=1}^{n} a_i^{(l)} x_i^{(j)} \oplus \sum_{i=1}^{n} b_i^{(l)} y_i^{(j)}, l = \overline{1, L}.$$

4. Вычисляется частота

$$\nu^{(l)} = \frac{1}{M} \sum_{j=1}^{M} \lambda^{(j,l)}$$

выпадения «1» среди M значений левой части l-го линейного соотношения ($l = \overline{1, L}$), и строится оценка максимального правдоподобия правой части l-го линейного соотношения:

$$d_l = \begin{cases} 1, & \text{если } \nu^{(l)} \geq 1/2, \\ 0, & \text{если } \nu^{(l)} < 1/2. \end{cases}$$

5. Строится и решается относительно $\theta \in V_L$ система L двоичных (булевых) линейных уравнений с использованием $\{d_l\}$ и правых частей линейных соотношений вида (11.5.3):

$$\sum_{k=1}^{L} c_k^{(l)} \theta_k = d_l, l = \overline{1, L}.$$

Единственное решение этой системы $\hat{\theta} \in V_L$ и используется в качестве оценки ключевого параметра θ^0.

Отметим, что количество M используемых в описываемой криптоатаке пар «исходный текст — шифртекст» должно быть тем больше, чем меньше величина $\min_{1\le l\le L}\left|\varepsilon^{(l)}\right|$ ($M \min_{1\le l\le L}\left|\varepsilon^{(l)}\right| \gg 1$).

В примере 1 для линейного криптоанализа достаточно использовать лишь одну пару «исходный текст — шифртекст» X, Y ($M = 1$). В качестве векторов $\{a^{(l)}, b^{(l)}, c^{(l)}\}$ удобно использовать орты:

$$c^{(l)} = b^{(l)} = a^{(l)} = (0, \ldots, 0, 1_l, 0, \ldots, 0), \quad l = \overline{1, n}.$$

При этом

$$\hat{\theta}_l = \lambda^{(l)}(X, Y) = x_l \oplus y_l, \quad l = \overline{1, L}. \tag{11.5.9}$$

Из (11.5.6) следует, что эта оценка (11.5.9) ключевого параметра безошибочна: $\hat{\theta} = \theta^0$.

Проиллюстрируем применение метода *линейного криптоанализа* при решении задачи криптоанализа DES [136]. Будем использовать обозначения, принятые в главе 7, а также следующие дополнительные обозначения Матсуи:

$P = X$ — 64-битовый исходный текст после начальной перестановки DES ($n = 64$);

$C = Y$ — 64-битовый шифртекст до конечной перестановки DES;

P_H, P_L — старшие и младшие 32 бита текста P соответственно;

C_H, C_L — старшие и младшие 32 бита шифртекста C соответственно;

$K = \theta^0$ — 56-битовый ключ после начальной перестановки его компонентов ($L = 56$);

K_r — 48-битовый ключ r-го такта, определяемый заданным в DES расписанием ключей ($r = \overline{1, 16}$);

$F_r(Z_r, K_r)$ — F-функция r-го такта;

$A[i]$ — i-й бит двоичного вектора A;

$A[i, j, \ldots, k] = A[i] \oplus A[j] \oplus \ldots \oplus A[k]$ — сумма указанных битов.

Здесь полагается, что в двоичной вектор-строке символов $(x_1, \ldots, x_n)$ правый (последний) бит x_n является младшим, а левый (первый) бит x_1 — старшим[1]. Так как начальные перестановки исходного текста P и ключа K, а также конечная перестановка шифртекста C в алгоритме DES фиксированы и известны, то для упрощения описания алгоритма линейного криптоанализа без потери общности эти перестановки можно опустить.

С помощью введенных обозначений Матсуи приближенное линейное соотношение (11.5.3) между P, C, K примет удобный вид

$$P[i_1, \ldots, i_u] \oplus C[j_1, \ldots, j_v] = K[k_1, \ldots, k_w], \tag{11.5.10}$$

где $1 \le i_1 < i_2 < \ldots < i_u \le 64, 1 \le j_1 < j_2 < \ldots < i_v \le 64, 1 \le k_1 < k_2 < \ldots < k_w \le 64$ — заданные упорядоченные наборы индексов «активных» битов, используемых

[1] В ряде книг по криптологии используется обратная этой рекомендуемой Матсуи интерпретация младших и старших битов. Это может привести к путанице, особенно при практической реализации систем.

в (11.5.10); $1 \le u \le 64$ — количество «активных» битов шифртекста, $1 \le w \le 56$ — количество «активных» битов ключа.

Проводя «тщательный анализ криптосистемы DES», Матсуи построил два приближенных линейных соотношения:

$$P_H[7,18,24] \oplus F_1(P_L,K_1)[7,18,24] \oplus C_H[15] \oplus C_L[7,18,24,29] \oplus F_{16}(C_L,K_{16})[15] =$$
$$= K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \oplus K_8[44] \oplus$$
$$\oplus K_9[22] \oplus K_{11}[22] \oplus K_{12}[44] \oplus K_{13}[22] \oplus K_{15}[22]; \qquad (11.5.11)$$

$$C_H[7,18,24] \oplus F_{16}(C_L,K_{16})[7,18,24] \oplus P_H[15] \oplus P_L[7,18,24,29] \oplus F_1(P_L,K_1)[15] =$$
$$= K_{14}[22] \oplus K_{13}[44] \oplus K_{12}[22] \oplus K_{10}[22] \oplus K_9[44] \oplus$$
$$\oplus K_8[22] \oplus K_6[22] \oplus K_5[44] \oplus K_4[22] \oplus K_2[22]. \qquad (11.5.12)$$

Для указанной выше модели случайного исходного текста P каждое соотношение выполняется с вероятностью $p = 1/2 - 1.19 \cdot 2^{-21}$ так, что эффективность соотношений (11.5.11) и (11.5.12) одинакова и равна $|\varepsilon| = 1.19 \cdot 2^{-21} \approx 0.57 \cdot 10^{-6}$.

Введем понятие *эффективного бита ключа*. Бит ключа (точнее его номер) называется эффективным относительно приближенного линейного соотношения, если он «влияет» на левую часть этого соотношения. Используя вид функций $F(\cdot)$, $F_{16}(\cdot)$ и расписание ключей DES (см. главу 7), находим 12 эффективных битов ключа для соотношения (11.5.11):

$$K_1[18], K_1[19], K_1[20], K_1[21], K_1[22], K_1[23], K_{16}[42],$$
$$K_{16}[43], K_{16}[44], K_{16}[45], K_{16}[46], K_{16}[47], \qquad (11.5.13)$$

а также 12 эффективных битов ключа для соотношения (11.5.12):

$$K_{16}[18], K_{16}[19], K_{16}[20], K_{16}[21], K_{16}[22], K_{16}[23],$$
$$K_1[42], K_1[43], K_1[44], K_1[45], K_1[46], K_1[47]. \qquad (11.5.14)$$

Для оценки этих 12 эффективных битов ключа (11.5.13) генерируется M независимых случайных исходных текстов и, используя шаги 2—4 описанного алгоритма линейного криптоанализа, перебором 2^{12} всевозможных вариантов битов (11.5.13) находится такой l^*-й вариант, для которого частота $\nu^{(l^*)}$ наиболее отличается от 1/2:

$$\nu^{(l^*)} = \max_{1 \le k \le 2^{12}} \left| \nu^{(l)} - 1/2 \right|.$$

При этом линейное соотношение (11.5.11) порождает (см. шаг 5 алгоритма) уравнение

$$K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \oplus K_8[44] \oplus K_9[22] \oplus$$
$$\oplus K_{11}[22] \oplus K_{12}[44] \oplus K_{13}[22] \oplus K_{15}[22] = d_{l^*},$$

позволяющее выразить 13-й бит ключа.

Аналогично проводится оценка 13 битов ключа с помощью приближенного линейного соотношения (11.5.12). Таким образом, удается оценить 26 битов ключа (см. расписание ключей DES в главе 10):

$$K[0], K[1], K[3], K[4], K[8], K[9], K[14], K[15], K[18], K[19], K[24],$$
$$K[25], K[31], K[32], K[38], K[39], K[41], K[42], K[44], K[45],$$
$$K[50], K[51], K[54], K[55]; K[5] \oplus K[13] \oplus K[17] \oplus K[20] \oplus K[46];$$
$$K[2] \oplus K[7] \oplus K[11] \oplus K[22] \oplus K[26] \oplus K[37] \oplus K[52].$$

Этим завершается первая стадия линейного криптоанализа DES.

На второй стадии оставшиеся 30 неизвестных битов 56-битового ключа K находятся с помощью полного перебора $2^{30} \approx 1.07 \cdot 10^9$ вариантов и их испытания на какой-нибудь фиксированной паре (P, C) «исходный текст — шифртекст». Для уменьшения вероятности ошибки линейного криптоанализа DES Матсуи предлагает на первой стадии отбирать несколько «квазиоптимальных» наборов 26 битов и для каждого из этих наборов на второй стадии осуществлять перебор всех 2^{30} вариантов остальных битов. По оценкам Матсуи вероятность ошибки $\mathbf{P}\{\hat{\theta} \neq \theta^0\} = 0.15$ достигается при количестве генерированных исходных текстов $M = 2^{43} \approx 9.16 \cdot 10^{12}$.

Экспериментальная атака с помощью линейного криптоанализа на полный 16-раундовый DES была реализована Матсуи в 1994 г. с помощью 12 параллельно работающих компьютеров HP9735/PA — RISC 99 MHz. Программа размером около 1000 строк была написана на языках C и Ассемблер. При $M = 2^{43}$ была получена безошибочная оценка $\hat{\theta} = \theta^0$ ключа. Вычисления заняли 50 суток, причем 40 из них было затрачено на генерацию M исходных текстов и получение соответствующих шифртекстов и только 10 суток — собственно на поиск ключа.

В заключение отметим, что линейный криптоанализ интенсивно развивается и создаются новые его модификации, в частности, разрабатывается разностно-линейный криптоанализ.

11.6. Задания

1. Разработать эффективный компьютерный алгоритм криптоанализа (A_2) «простой подстановки символов алфавита» методом «опробования». Оценить точность и быстродействие алгоритма.

2. Разработать эффективный компьютерный алгоритм криптоанализа (A_2) «перестановки символов с неизвестным периодом T» методом «опробования». Оценить точность и быстродействие алгоритма.

3. Разработать эффективный компьютерный алгоритм криптоанализа (A_2) шифра Виженера методом «опробования». Оценить точность и быстродействие алгоритма.

4. Разработать эффективный компьютерный алгоритм криптоанализа (A_2) шифра Вернама методом «опробования». Оценить точность и быстродействие алгоритма.

5. Обобщить теорему 11.3.2 на случай s-связной однородной цепи Маркова ($s = 0; 1; 2; \ldots$).

6. Для криптосистемы Цезаря (11.3.17) аналитически оценить вероятность ошибки $r(n, \theta^0) = \mathbf{P}_{\theta^0}\{\hat{\theta}(Y) \neq 0\}$ для простейшего алгоритма криптоанализа (11.3.18). Получить аналитическую оценку минимальной длины шифртекста n_ε.

7. Для шифртекста (11.3.19) применить алгоритм криптоанализа (11.3.26), основанный на методе максимального правдоподобия, используя марковскую модель источника сообщений (11.3.21), (11.3.22). Набор величин $\{\log p_1(j)\}$ приведен в табл. 11.2, а вероятности переходов (11.3.22) оценить самостоятельно, используя достаточно длинный англоязычный текст. Сравнить результаты с представленными на рис. 11.1.

8. Для шифртекста (11.3.19) применить байесовский алгоритм криптоанализа (11.3.31), выбирая различные априорные предположения о θ^0: а) $\pi(\theta)$ — равномерное на Θ распределение; б) распределение $\pi(\theta)$ сконцентрировано в окрестности истинного значения:

θ	0—9	10	11	12	13	14	15	16	17—25
$\pi(\theta)$	0	0.1	0.1	0.2	0.2	0.2	0.1	0.1	0

Как влияет априорная информация на точность оценки?

9. Рассмотреть возможность снижения вычислительной сложности (11.3.36) алгоритмов криптоанализа, использующих методы статистических решений, применяя специальные методы теории дискретного программирования при решении экстремальных задач (11.3.9), (11.3.1). Для иллюстрации рассмотреть алгоритм криптоанализа (11.3.18) и численный пример шифртекста (11.3.19).

10. Используя модель марковской криптосистемы из п. 11.4, показать, что если

$$\pi = \left(0, \frac{1}{N-1}, \ldots, \frac{1}{N-1}\right),$$

то распределение вероятностей $\Delta Y(s)$ является равномерным на $A \setminus \{0\}$, $s = 1, \ldots, r$.

11. В примере 2 (см. п. 11.5), перебирая вектор весовых коэффициентов $a \in V$, построить всевозможные приближенные линейные соотношения вида

(11.5.3) и оценить характеристику эффективности $|\varepsilon|$. Решить задачу поиска оптимальных линейных соотношений (11.5.8) для осуществления линейного криптоанализа. Обобщить результаты для произвольного n.

12. Построить оптимальные приближенные линейные соотношения для трехтактового DES ($r = 3$).

13. Применить метод линейного криптоанализа к трехтактовому DES ($r = 3$).

SXJKN8EYHR3I4HR987YDWFUI
HW9E8YR384H5O34596EUF9O8
WE76R83Y439846UHFE98YDSF
JH8374H4BR8276RUBNWDF87Y
6234BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7ERFH2
347FQ7WERF9UJ1349YUR7DFK
JHGJDFBVDGFJVBDSKFJHKJFV
KBDFKJVJKGFKJSBDFJBVKJDS
GFJKVBKSDJBFKVBDGFVAKDLK

Глава 12

КРИПТОСИСТЕМЫ С ОТКРЫТЫМ КЛЮЧОМ

SXJKN8EYHR3I4HR987YDWFUI
HW9E8YR384H5O34596EUF9O8
WE76R83Y439846UHFE98YDSF
JH8374H4BR8276RUBNWDF87Y
6234BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7ERFH2
347FQ7WERF9UJ1349YUR7DFK
JHGJDFBVDGFJVBDSKFJHKJFV
KBDFKJVJKGFKJSBDFJBVKJDS
GFJKVBKSDJBFKVBDGFVAKDLK

12.1. Описание RSA-криптосистемы

Идею, лежащую в основе криптосистем с открытым ключом, высказали в 1975 г. Диффи и Хеллмэн. Они ввели понятие односторонней функции с секретом. Более подробно об этом речь пойдет в следующей главе. Это дало принципиальную возможность разрабатывать криптосистемы с открытым ключом, в которых алгоритм шифрования является общедоступным, и поэтому нет необходимости в секретных каналах связи для предварительного обмена ключами.

Таким образом, эти криптосистемы принципиально отличаются от криптосистем, описанных в предыдущей главе. Их называют *асимметричными*, или *двухключевыми*, поскольку они имеют два ключа: несекретный — для шифрования и секретный — для дешифрования.

Метод шифрования RSA предложен в 1977 г. Ривестом, Шамиром и Адлеманом как реализация идеи Диффи и Хеллмэна.

Опишем процесс шифрования сообщений. Исходный текст должен быть переведен в числовую форму. Метод преобразования текста в числовую форму считается известным и необязательно держится в секрете. В результате текст представляется в виде одного большого числа. Затем полученное число разбивается на части так, чтобы каждая из них была числом в промежутке от 0 до N, где N будет выбрано ниже. Процесс шифрования одинаков для каждой части. Поэтому можно считать, что исходный текст представлен числом x таким, что $0 < x < N$.

Предположим, что некоторый пользователь (назовем его B) желает, чтобы ему передали секретное сообщение. Для этого он делает общедоступными два числа: N и e (открытый ключ), которые подчинены двум условиям:

1) $N = pq$, где p и q — большие простые числа, которые B держит в секрете. Числа p и q обычно выбираются порядком не ниже, чем 2^{256};

2) число e берется взаимно простым с $\varphi(N) = (p-1)(q-1)$.

Пользователь A, отправляющий сообщение x, шифрует его следующим образом: $E(x) \equiv x^e \pmod N$. Это и есть зашифрованный текст, который получает B.

Чтобы восстановить исходный текст, B поступает так:

1) находит число d такое, что $1 \le d \le N-1$ и $ed \equiv 1 (\mathrm{mod}\, \varphi(N))$. Это сравнение разрешимо единственным образом, поскольку $(e, \varphi(N)) = 1$. Здесь как раз

и проявляется особенность RSA. Для решения сравнения $ed \equiv 1 (\mathrm{mod}\, \varphi(N))$ пользователь B должен вычислить $\varphi(N)$, что для него не составит труда, так как $\varphi(N) = \varphi(pq) = \varphi(p)\varphi(q) = (p-1)(q-1)$. Любой другой пользователь, который знает только N, вынужден находить p и q, т.е. разлагать число N на простые множители, а эта задача при больших p и q имеет большую вычислительную сложность;

2) далее, имея в распоряжении число y, пользователь B вычисляет $D(y) \equiv y^d (\mathrm{mod}\, N)$, которая и есть представление x исходного текста. Действительно, применяя теорему Эйлера, получаем

$$y^d \equiv x^{ed} \equiv x^{\varphi(N)k+1} \equiv (x^{\varphi(N)})^k x \equiv x (\mathrm{mod}\, N).$$

Замечание 12.1.1. Из малой теоремы Ферма вытекает (см. главу 2), что $x^{p-1} \equiv 1 (\mathrm{mod}\, p)$ и $x^{q-1} \equiv 1 (\mathrm{mod}\, q)$. Тогда, если положить $l = \varphi(N) = (p-1)(q-1)$, то выполнено сравнение $x^l \equiv 1$ как по модулю p, так и по модулю q и, следовательно, по модулю $\varphi(N)$. Из этого вытекает, что любые числа e' и d' с условиями $e' \equiv e$ (mod l) и $d' \equiv d$ (mod l) также будут, соответственно, открытым и секретным ключами.

Корректность работы системы RSA. Пусть параметры ($N = pq$; e; d) задают RSA-криптосистему. Обозначим через $G = (\mathbb{Z}/(N))^*$ — мультипликативную группу кольца $\mathbb{Z}/(N)$. Она имеет порядок $\varphi(N) = (p-1)(q-1)$. Вероятность того, что случайно взятое число $0 < x < N$ не принадлежит G, равна $\frac{1}{p} + \frac{1}{q}$, что при больших p и q практически невозможно. С этой точки зрения можно считать, что все наши сообщения выбираются из G. Очевидно, что отображения E и D переводят элементы из G в элементы из G. Более того, они являются автоморфизмами G, т.е. взаимно-однозначными отображениями с условием $E(x_1x_2) = E(x_1)E(x_2)$ и $D(x_1x_2) = D(x_1)D(x_2)$. Кроме этого, E и D являются взаимно-обратными, т.е. $E(D(x)) = D(E(x)) = x$, что обеспечивается условием $ed \equiv 1 (\mathrm{mod}\, \varphi(N))$. Таким образом, секретный и несекретный ключи можно менять местами.

Обычно для обеспечения стойкости RSA предъявляют следующие требования к выбору параметров системы:

1) простые числа p и q должны быть большими;

2) разность $|p - q|$ должна быть большой;

3) числа $p \pm 1$, $q \pm 1$, $r - 1$, $s - 1$ должны содержать большой простой множитель;

4) число $(p-1, q-1)$ должно быть небольшим.

Также будет показано, что нельзя выбирать ключи e и d с короткой секретной экспонентой: необходимо придерживаться «нормального» случая, когда d — однопорядковая величина с $L = [p-1, q-1]$.

Если какое-либо из требований 1—4 не выполнено, то существует способ довольно быстро взломать эту криптосистему, что мы обсудим в следующих параграфах (см. задания к данной главе).

12.2. Возможные атаки на криптосистему RSA

Пусть параметры ($N = pq$; e; d) задают RSA-криптосистему. Для ее взлома, т.е. для прочтения секретного сообщения или подделки подписи, необходимо по известным входным N, e и зашифрованному или подписанному сообщению y найти такое $x \in G = (\mathbb{Z}/(N))^*$, что

$$y \equiv x^e \pmod N. \tag{12.2.1}$$

Например, можно пытаться решить сравнение (12.2.1) при конкретном y или y из некоторого класса $M \subset G$, чтобы затем, используя гомоморфность отображения $D(x)$, решить его на более широком множестве. Например, если $M = \{p_1, \ldots, p_k\}$ — множество малых простых, и каким-либо образом удалось решить сравнение (12.2.1), то это сравнение становится легко разрешимым для всех y, которые являются произведениями чисел из M. Более детально это означает следующее. Пусть имеем набор пар $XY = \{(x_1, y_1), \ldots, (x_k, y_k)\}$ с условием $x_i^e \equiv y_i \pmod N$. Пусть $1 < y < N$ и $(y, N) = 1$. Если каким-либо образом удастся представить y в виде

$$y \equiv y_1^{s_1} \cdots y_k^{s_k} \pmod N$$

с целыми s_j, то решением (12.2.1) будет

$$x \equiv x_1^{s_1} \cdots x_k^{s_k} \pmod N.$$

Однако данный подход не менее трудный, чем поиск алгоритма решения сравнения (12.2.1) при любом y.

Еще один вид атаки (он сводится к поиску ключа d' для дешифрования) — метод повторного шифрования. Он состоит в следующем. Пусть e — открытая экспонента, y — зашифрованное сообщение. Это значит, что для некоторого x верно (12.2.1). Строим последовательность y_i:

$$y_1 = y;$$

$$y_i \equiv y_{i-1}^e \pmod N, \quad i > 1.$$

Это значит, что

$$y_m \equiv y^{e^m} \pmod N.$$

Поскольку $(e, \varphi(N)) = 1$, то существует натуральное число m такое, что $e^m \equiv 1 \pmod{\varphi(N)}$. Но тогда

$$y^{e^m - 1} \equiv 1 \pmod N,$$

откуда вытекает

$$y^{e^m} \equiv y \pmod N.$$

Тогда y_{m-1} будет решением (12.2.1).

12.3. Стойкость RSA против метода повторного шифрования

Проанализируем более детально *метод повторного шифрования* и покажем необходимость соблюдения требований на выбор p и q для обеспечения стойкости.

Пусть $N = pq$, e, d — параметры системы RSA, y — зашифрованное сообщение, $y \in G = (\mathbb{Z}/(N))^*$, где G состоит из всех целых $x \in [1, N-1]$, взаимно простых с N.

Чтобы достичь успеха путем m-кратного шифрования, необходимо и достаточно, чтобы порядок $\operatorname{ord} y$ делил $e^m - 1$.

Теорема 12.3.1. Пусть $p-1 = rk$ и $q-1 = sl$, где r и s — различные простые, $(r,k)=1$, $(s,l)=1$. Тогда вероятность того, что случайно взятый элемент $y \in G$ имеет порядок t, делящийся на rs, равна $\left(1-\frac{1}{r}\right)\left(1-\frac{1}{s}\right)$.

Доказательство. Группа G есть прямое произведение двух своих циклических подгрупп G_1 и G_2, которые соответственно имеют порядки $p-1$ и $q-1$. Всякий $y \in G$ представляется однозначно в виде произведения $y = ab$, где $a \in G_1$, $b \in G_2$. Более того, его порядок равен $\operatorname{ord} y = [\operatorname{ord} a, \operatorname{ord} b]$. Тогда каждой паре (t_1, t_2), где $t_1 \mid k$, $t_2 \mid l$, соответствует $\varphi(rt_1)\varphi(st_2)$ элементов группы G порядка $rs(t_1,t_2)$. Различным таким парам соответствуют различные наборы элементов, и объединение таких наборов по всем парам (t_1,t_2) исчерпывает множество тех y, у которых порядок делится на rs. Ввиду того, что количество элементов в циклической группе, имеющих порядок w, равно $\varphi(w)$, получаем, что количество тех y, для которых $rs \mid \operatorname{ord} y$, равно

$$\sum_{t_1|k}\sum_{t_2|l}\varphi(rt_1)\varphi(st_2) = \sum_{t_1|k}\sum_{t_2|l}\varphi(r)\varphi(t_1)\varphi(s)\varphi(t_2) =$$

$$= (r-1)(s-1)\sum_{t_1|k}\varphi(t_1)\sum_{t_2|l}\varphi(t_2) = (r-1)(s-1)kl.$$

Всего в группе G имеется $rskl$ элементов. Тогда искомая вероятность равна

$$\frac{(r-1)(s-1)kl}{rskl} = \left(1-\frac{1}{r}\right)\left(1-\frac{1}{s}\right).$$

Из теоремы 12.3.1 вытекает, что требуется взять примерно $\min(r,s)$ элементов, чтобы порядок одного из них не делился хотя бы на одно из чисел r, s. Поэтому без ограничения общности можно считать, что $sr \mid \operatorname{ord} y$.

Аналогично доказывается, что вероятность того, что ни s, ни r не делят $\operatorname{ord} y$, равна $1 - 1/rs$.

Если предположить, что $sr \mid \operatorname{ord} y$, то успех в атаке методом повторного шифрования будет достигнут, только если $rs \mid (e^m - 1)$. Число $e < \varphi(N)$ выбирают

взаимно простым с $\varphi(N)$. Иначе говоря, e — элемент группы $H = \mathbb{Z}/\varphi(N)^*$. Если r_1 — большой простой делитель числа r, s_1 — большой простой делитель числа s, то аналогично показывается, что с вероятностью $\left(1-\frac{1}{r_1}\right)\left(1-\frac{1}{s_1}\right)$ число e имеет порядок в группе H, делящийся на $r_1 s_1$. В этом случае атака методом повторного шифрования будет успешной только при условии $r_1 s_1 \mid m$, что требует огромных вычислительных затрат.

Если два последних условия в требовании 3 не выполнены, то вероятность того, что случайный ключ e будет иметь небольшой порядок, возрастает. Это приводит к тому, что если в результате ряда атак с различными e при некотором $e = e'$ мы достигаем успеха, то с большой вероятностью можно ожидать, что $e_1^{m_1} \equiv 1 (\mathrm{mod}\, \varphi(N))$ с найденым относительно небольшим m_1. Тогда число $e_1^{m_1}$ можно использовать для факторизации N методом, описанным в следующем параграфе.

12.4. Поиск секретного ключа d и факторизация модуля N

Если известно разложение на простые множители числа N, то можно определить секретный ключ d с помощью полиномиального от длины N и e алгоритма. Оказывается, верно и обратное: существует эффективный вероятностный полиномиальный от длины d, e, N алгоритм разложения на множители числа N, если известно натуральное d с условием $ed \equiv 1 (\mathrm{mod}\, \varphi(N))$. Только в этом смысле можно говорить об эквивалентности задач взлома RSA путем определения секретного ключа d (или вообще любого числа d' такого, что $D(x) \equiv x^{d'} (\mathrm{mod}\, N)$ и путем факторизации. Это значит, что нет алгоритма, который бы определял секретный ключ d со сложностью, превосходящей сложность «самого лучшего» алгоритма факторизации в полиномиальное количество раз. Данный результат опирается на следующую теорему [118].

Теорема 12.4.1. Если тройка (N, e, d) образует RSA-криптосистему и известно натуральное d такое, что $ed \equiv 1 (\mathrm{mod}\, \varphi(N))$, то существует эффективный вероятностный алгоритм полиномиальной сложности для факторизации N.

Доказательство. Пусть известны параметры e и d, удовлетворяющие условию теоремы. Тогда $s = ed - 1$ делится на $\varphi(N)$. Следовательно, для любого $x \in (\mathbb{Z}/N)^* = G$ верно

$$x^s \equiv 1 (\mathrm{mod}\, N).$$

Запишем $s = 2^t u$, где u — нечетное, и рассмотрим множество $A = G \setminus B$, где B состоит из тех $x \in G$, для которых либо при некотором целом $j \in \{1, \ldots, t-1\}$ верно $x^{2^j u} \equiv -1 (\mathrm{mod}\, N)$, либо $x^u \equiv 1 (\mathrm{mod}\, N)$.

Для любого элемента $a \in A$ выберем число k наименьшим с условием $a^{2^k u} \equiv 1 \pmod N$. Поскольку $a \notin B$, то $k \geq 1$. Тогда положим $b = a^{2^{k-1}u} \pmod N$. Следовательно,

$$b^2 \equiv 1 \pmod N, \quad b \not\equiv \pm 1 \pmod N.$$

Поэтому $(b-1, N)$ — собственный делитель N. Тем самым достигается факторизация.

Далее запишем $p-1 = 2^{v_1} u_1$, $q-1 = 2^{v_2} u_2$, где u_1, u_2 — нечетные числа. Положим $v = \min(v_1, v_2)$ и $K = (u, u_1)(u, u_2)$. Используя теорию сравнений, можно получить оценки для количества решений сравнений $x^u \equiv 1 \pmod N$ и $x^{2^j u} \equiv -1 \pmod N$, $j \leq t-1$.

Сравнение $x^u \equiv 1 \pmod N$ равносильно системе

$$x^u \equiv 1 \pmod p, \quad x^u \equiv 1 \pmod q.$$

Проиндексируем первое сравнение. Получим $u \operatorname{ind} x \equiv 0 \pmod{2^{v_1} u_1}$. Следовательно, оно имеет (u, u_1) решений. Аналогично, второе сравнение имеет (u, u_2) решений. С учетом $(p, q) = 1$ и китайской теоремы об остатках сравнение $x^u \equiv 1 \pmod N$ имеет $K = (u, u_1)(u, u_2)$ решений. Этим же приемом можно найти и количество решений сравнения $x^{2^j u} \equiv -1 \pmod N$. Оно оказывается равным $4^j K$, $j < v$. Поэтому

$$|B| = (1 + 1 + 4 + 4^2 + \ldots + 4^{v-1})K = \left(1 + \frac{4^v - 1}{3}\right)K.$$

Путем элементарных вычислений находим, что

$$|B| = \left(1 + \frac{4^v - 1}{3}\right)K \leq \frac{\varphi(N)}{2} = \frac{1}{2}|G|.$$

Из этого вытекает, что вероятность того, что случайно взятый элемент $x \in G$ будет лежать в A, не менее $1/2$. Тогда за m попыток мы с вероятностью $\geq 1 - 1/2^m$ встретим элемент из A и найдем факторизацию N по алгоритму, вытекающему из доказательства. □

Замечание 12.4.1. Эквивалентность задач факторизации и поиска ключа d означает, что нельзя строить многопользовательскую RSA-криптосистему, чтобы разные пользователи имели свои ключи с одним и тем же модулем N. Потеря стойкости RSA-криптосистемы с параметрами N, e, d вследствие потери секретности ключа d влечет необходимость не только замены ключа d, но и модуля N.

12.5. Биты в RSA-криптосистеме

Изучим, как секретность отдельных битов сообщения x влияет на секретность сообщения в целом.

Пусть $y = E(x)$, где E — преобразование криптосистемы с параметрами $N = pq, e, d$. Введем функции:

$$p(y) = \begin{cases} 0, & x \text{ — четное}, \\ 1, & x \text{ — нечетное}; \end{cases}$$

$$h(y) = \begin{cases} 0, & 0 < x < N/2, \\ 1, & N/2 < x < N. \end{cases}$$

Теорема 12.5.1. Следующие утверждения эквивалентны:
1) существует эффективный алгоритм вычисления функции $p(y)$,
2) существует эффективный алгоритм вычисления функции $h(y)$,
3) существует эффективный алгоритм вычисления функции $D(y)$.

Доказательство. Пусть $0 < x < N$, тогда легко проверить утверждение

$$x < \frac{1}{2}N \Leftrightarrow 2x (\bmod N) \text{ — четно}. \tag{12.5.1}$$

Поэтому

$$h(y) = p(yE(2)).$$

Если в этом равенстве y заменить на $yE(2^{-1})$, то

$$p(y) = h(yE(2^{-1})).$$

Тем самым установлена эквивалентность первых двух утверждений теоремы.

Вычислим $x = D(y)$, если известен алгоритм для вычисления $h(y)$. Обобщая утверждение (12.5.1), можно показать, что

$$4x (\bmod N) \text{ — четно } \Leftrightarrow x \in \left[0, \frac{N}{4}\right) \cup \left[\frac{N}{2}, \frac{3N}{4}\right),$$

$$8x (\bmod N) \text{ — четно } \Leftrightarrow x \in \left[0, \frac{N}{8}\right) \cup \left[\frac{N}{4}, \frac{3N}{8}\right) \cup \left[\frac{N}{2}, \frac{5N}{8}\right) \cup \left[\frac{3N}{4}, \frac{7N}{8}\right)$$

и т.д. Располагая несекретным преобразованием E криптосистемы, можно легко вычислить $E(2)$, $E(4)$, $E(8)$, ... Тем самым находим $E(2x)$, $E(4x)$, $E(8x)$, ..., поскольку $E(x_1x_2) = E(x_1)E(x_2)$. Алгоритм h при этом будет давать на k-м шаге оценку для x:

$$\frac{i}{2^{k+1}}N < x < \frac{i+1}{2^{k+1}}N.$$

Этот интервал при достаточно большом K будет содержать лишь одно целое число. □

12.6. Система Рабина

Стойкость RSA-криптосистемы базируется на предполагаемой трудности задачи факторизации. Ответа на обратный вопрос на данный момент нет. Остается неясным, следует ли из наличия эффективного алгоритма для обращения RSA-функции существование эффективного алгоритма для факторизации модуля N.

Рабин предложил систему с открытым ключом, трудность которой доказуемо равносильна трудности проблемы факторизации [118].

В *системе Рабина* пользователи выбирают нечетные простые p и q. Они считаются секретными, а модуль $N = pq$ — несекретным. Далее каждый пользователь выбирает целое $b < N$. Функция шифрования задается формулой

$$E_{N,b}(x) \equiv x(x+b) \bmod N.$$

Дешифрование также несложно, если известна факторизация $N = pq$. Пусть m — зашифрованное сообщение. Используем известные алгоритмы решения сравнений $x(x+b) \equiv m \pmod p$ и $x(x+b) \equiv m \pmod q$. Обозначим их решения через r и s соответственно. Затем используем алгоритм Евклида, чтобы вычислить k и l такие, что $kp + lq = 1$. Тогда решением сравнения $x(x+b) \equiv m \pmod N$ будет $lqr + kps$. Заметим, что решение указанных сравнений легко сводится к решению сравнения вида $x^2 \equiv m \pmod p$. Оно легко разрешимо при $p \equiv 3 \pmod 4$. В этом случае

$$p = 4k+3, \quad (m^{k+1})^2 = m^{2k+2} = mm^{\frac{p-1}{2}} \equiv m \pmod p.$$

Поэтому при практическом использовании системы Рабина применяют простые числа вида $4k+3$. Заметим также, что сравнение $x^2 \equiv m \pmod N$ при $m \not\equiv 0 \pmod p$, $m \not\equiv 0 \pmod q$ имеет четыре решения.

Теорема 12.6.1. Пусть N — произведение двух нечетных простых. Следующие условия эквивалентны:

1) существует эффективный алгоритм решения сравнения $x^2 \equiv m \pmod N$;
2) существует эффективный алгоритм для факторизации N.

Доказательство. Случай $2 \Rightarrow 1$ фактически рассмотрен ранее. Осталось показать $1 \Rightarrow 2$. Выберем случайно целое a такое, что $(a, N) = 1$, и пусть $m \equiv a^2 \pmod N$. Пусть далее u — решение сравнения $x^2 \equiv m \pmod N$. Итак, если $u \notin \{a, N-a\}$, то $(N, u+a)$ — простой делитель N; а если $u \in \{a, N-a\}$, то выберем другое $u \notin \{a, N-a\}$ и повторим предыдущую процедуру. □

12.7. Рюкзачный метод шифрования

Выберем пару w, m натуральных взаимно простых чисел и будем считать их секретными. Число w назовем *множителем*, а m — *модулем*. Дополнительно выберем секретную последовательность $b = (b_1, \ldots, b_n)$ положительных целых с условием

$$b_i > \sum_{j=1}^{i-1} b_j, \quad \forall i \geq 1, \quad m > \sum_{j=1}^{n} b_j.$$

Такую последовательность будем называть *быстрорастущей*.

Последовательность $a = (a_1, \ldots, a_n)$, где

$$a_i \equiv wb_i \pmod{m}, \; i \geq 1,$$

считается несекретной.

Сообщение $x = (x_1, \ldots, x_n)$, являющееся набором нулей и единиц, шифруется по правилу

$$E(x) = \sum_{i=1}^{n} x_i a_i. \tag{12.7.1}$$

Для дешифрования полученного сообщения достаточно решить уравнение вида

$$S = \sum_{i=1}^{n} u_i a_i. \tag{12.7.2}$$

Уравнение (12.7.2) основано на *задаче о рюкзаке*, которая относится к классу NP-полных задач. Тем не менее следующая теорема указывает эффективный метод ее решения легальным пользователем для быстрорастущих последовательностей.

Теорема 12.7.1. Пусть $b_1, \ldots, b_n$ — быстрорастущая последовательность натуральных чисел и $S > 0$. Тогда уравнение $S = \sum_{i=1}^{n} x_i b_i$ имеет не более одного решения $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ с условием $S \leq \sum_{i=1}^{n} b_i$.

Доказательство. Любое решение уравнения $S = \sum_{i=1}^{n} x_i b_i$ должно удовлетворять условию

$$x_i = 1 \Leftrightarrow S \geq b_i + \sum_{j=i+1}^{n} x_j b_j, \quad i = 1, \ldots, n.$$

Необходимость условия очевидна, поскольку все b_i — положительны. Пусть $x_i = 0$. Тогда

$$S = \sum_{j=1}^{i-1} x_j b_j + \sum_{j=i+1}^{n} x_j b_j < b_i + \sum_{j=i+1}^{n} x_j b_j. \qquad \square$$

Остается показать, что функция дешифрования эффективно вычислима. При получении зашифрованного сообщения S пользователь вычисляет w^{-1} по модулю m и решает «задачу о рюкзаке»:

$$w^{-1}S = \sum_{i=1}^{n} x_i b_i.$$

Так как последовательность $b_1, \ldots, b_n$ быстрорастущая, это легко сделать с помощью теоремы 12.7.1.

12.8. Стойкость рюкзачного шифра

Пусть w, m, $b = (b_1, \ldots, b_n)$, $a = (a_1, \ldots, a_n)$ — параметры рюкзачной криптосистемы. Криптоаналитик знает последовательность $\{a_i\}$, но не знает w, m, b. Он может попытаться, и небезуспешно, найти пару чисел $\overline{w}$, $\overline{m}$ таких, что последовательность $\overline{a} = (\overline{a}_1, \ldots, \overline{a}_n)$, определяемая условием

$$\overline{a}_i \equiv a_i \overline{w} \pmod{m}, \tag{12.8.1}$$

является быстрорастущей и обладает свойством

$$\sum_{i=1}^{n} \overline{a}_i < \overline{m}. \tag{12.8.2}$$

Как видно из предыдущего параграфа, пару $\overline{w}$, $\overline{m}$ можно использовать для дешифрования с тем же успехом, что и секретную пару w, m.

Деля формулы (12.8.1) и (12.8.2) на $\overline{m}$, получаем

$$\frac{\overline{a}_i}{m} \equiv \left(a_i \frac{\overline{w}}{\overline{m}} \right) \bmod 1, \tag{12.8.3}$$

$$\sum_{i=1}^{n} a_i \overline{r}_i \pmod 1 < 1, \tag{12.8.4}$$

где $\overline{r} = \overline{w}/\overline{m}$. График функции $a_i \overline{r}_i \pmod 1$ представлен на рис. 12.1. Эту функцию принято называть *косозубой*.

Для вычисления пары $\overline{w}$, $\overline{m}$ сначала надо определить точку $\overline{r}_0$ на оси $\overline{r}$, чтобы выполнялось условие (12.8.4). Следовательно, существует и интервал $[r_1, r_2]$, для любой точки которого неравенство (12.8.4) выполнено.

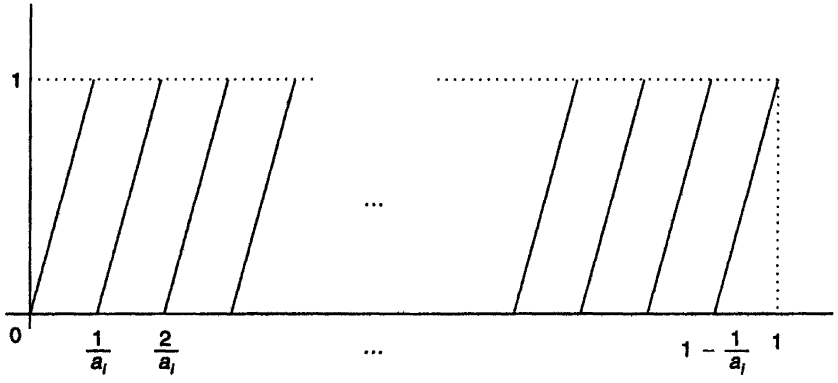


Рис. 12.1. График косозубой функции

Пусть p_i — p_i-й минимум i-й косозубой функции. Имеем две системы неравенств с целыми неизвестными $p_1, \ldots, p_n$:

$$\begin{array}{ll} 1 \le p_1 \le a_1 - 1, & -\varepsilon_2 \le p_1/a_1 - p_2/a_2 \le \varepsilon'_2, \\ 1 \le p_2 \le a_2 - 1, & -\varepsilon'_3 \le p_1/a_1 - p_3/a_3 \le \varepsilon'_3, \\ \vdots & \vdots \\ 1 \le p_n \le a_n - 1, & -\varepsilon_n \le p_1/a_1 - p_n/a_n \le \varepsilon'_n. \end{array} \tag{12.8.5}$$

Параметры ε_i, ε'_i нужно выбирать достаточно малыми, чтобы определить точку сгущения. Эту систему можно решить методом целочисленного линейного программирования.

Пусть p — одна из величин, определенная указанной процедурой, и $\bar{r}_1, \ldots, \bar{r}_k$ — точки разрыва всех косозубых функций такие, что

$$\bar{r}_1, \ldots, \bar{r}_k \in \left[\frac{p_1}{a_1}, \frac{p_1+1}{a_1}\right), \tag{12.8.6}$$

размещенные в убывающем порядке. Между двумя такими точками каждая косозубая функция задается отрезком прямой

$$\bar{r}a_i - q_i^t, \bar{r}_t \le \bar{r} < \bar{r}_{t+1}, \tag{12.8.7}$$

где q_i^t равно номеру минимума i-й косозубой кривой, лежащей в интервале $0, \bar{r}_t$. Таким образом, для каждого $1 \le t \le k$ условия (12.8.6) и (12.8.7) можно сформулировать в виде следующей системы линейных неравенств с неизвестным $\bar{r}$ таким, что $\bar{r}_t \le \bar{r} < \bar{r}_{t+1}$:

$$\sum_{i=1}^{n} (\bar{r}a_i - q_i^t) < 1, \tag{12.8.8}$$

$$(\bar{r}a_i - q_i^t) > \sum_{j=1}^{l-1}(\bar{r}a_j - g_j^t), \quad i = 1, \ldots, n. \tag{12.8.9}$$

Решение последней системы дает подинтервал $[\bar{r}_t, \bar{r}_{t+1})$. Любое $\bar{r} = \bar{w}/\bar{m}$, лежащее в этом подинтервале, дает искомую пару $\bar{w}$, $\bar{m}$.

12.9. Теорема М. Винера о малом секретном ключе

М. Винером было показано, что при небольшом секретном RSA-ключе d имеется возможность эффективно его вычислить. С учетом алгоритма, описанного в п. 9.4, это дает возможность факторизовать RSA-модуль. Отметим, что в RSA-криптографии простые числа p и q выбираются по возможности близкими, а это в силу теоремы Чебышёва автоматически приводит к выполнению условия $q < p < 2q$.

Теорема 12.9.1. Пусть задана RSA-криптосистема с алгоритмами $N = pq$, $ed \equiv 1 (\mathrm{mod}\, \varphi(N))$, $q < p < 2q$, $d < \frac{1}{3}N^{1/4}$. Тогда d эффективно вычислимо.

Доказательство. Оценим величину $N - \varphi(N)$.

$$\varphi(N) = (p-1)(q-1) = pq - p - q + 1 = N - p - q + 1.$$

Поэтому $N - \varphi(n) \le p + q - 1$. Так как $p + q < 3\sqrt{pq} = 3\sqrt{N}$, то $N - \varphi(N) < 3\sqrt{N}$. Положим $ed = 1 + k\varphi(N)$.

Далее оценим сверху разность

$$\left|\frac{e}{N} - \frac{k}{d}\right| = \left|\frac{ed - k\varphi(N) - kN + k\varphi(N)}{Nd}\right| =$$

$$= \left|\frac{1 - k(N - \varphi(N))}{Nd}\right| \le \left|\frac{3k\sqrt{N}}{N}\right| = \frac{3k}{d\sqrt{N}}.$$

Воспользуемся тем, что $k\varphi(N) = ed - 1$. Следовательно, $k\varphi(N) < ed$. Примем во внимание, что $e < \varphi(N)$. Окончательно:

$$\left|\frac{e}{N} - \frac{k}{d}\right| \le \frac{1}{d\sqrt[4]{N}} < \frac{1}{2d^2}.$$

Полученное неравенство означает, что величина $\frac{k}{d}$ является подходящей дробью для несекретной дроби $\frac{e}{N}$. Как показано в п. 2.18, всего подходящих дробей не более $2\log_2 N$ и все они эффективно вычислимы. □

Существует предположение, что все секретные экспоненты вплоть до $\sqrt{N}$ небезопасные.

12.10. Арифметика больших чисел

Задавшись натуральным *основанием* $b \geq 2$, неотрицательному целому числу u поставим в соответствие набор $(u_{k-1} \cdots u_1 u_0)_b$ целых чисел такой, что

$$u = u_{k-1}b^{k-1} + \ldots + u_1 b + u_0, \quad 0 \leq u_0, u_1, \ldots, u_{k-1} < b.$$

Если $u_{k-1} \neq 0$, то говорят, что u является *k-разрядным числом по основанию b* или *k-битовым числом*, если $b = 2$. Число u_0 называют *младшим разрядом u*, а число u_{k-1} — *старшим разрядом*. Разряды записи по основанию 16 принято обозначать цифровыми и буквенными символами (табл. 12.1).

Таблица 12.1

Разряды шестнадцатеричной записи

$0 = (0000)_2 = (0)_{16}$	$8 = (1000)_2 = (8)_{16}$
$1 = (0001)_2 = (1)_{16}$	$9 = (1001)_2 = (9)_{16}$
$2 = (0010)_2 = (2)_{16}$	$10 = (1010)_2 = (A)_{16}$
$3 = (0011)_2 = (3)_{16}$	$11 = (1011)_2 = (B)_{16}$
$4 = (0100)_2 = (4)_{16}$	$12 = (1100)_2 = (C)_{16}$
$5 = (0101)_2 = (5)_{16}$	$13 = (1101)_2 = (D)_{16}$
$6 = (0110)_2 = (6)_{16}$	$14 = (1110)_2 = (E)_{16}$
$7 = (0111)_2 = (7)_{16}$	$15 = (1111)_2 = (F)_{16}$

Приведем алгоритмы арифметических операций над числами с произвольным количеством разрядов (*большими числами*). При практической реализации таких алгоритмов основание b выбирается равным размеру машинного слова (как правило, это степень 2). Предполагается, что архитектура компьютера дает возможность выполнять элементарные арифметические операции над одно- и двухразрядными числами.

Сложение $u + v = (w_k \cdots w_1 w_0)_b$ чисел $u = (u_{k-1} \cdots u_1 u_0)_b$ и $v = (v_{k-1} \cdots v_1 v_0)_b$.

1. Установить $c \leftarrow 0$.
2. Для $i = 0, \ldots, k-1$ выполнить:
   а) $w_i \leftarrow (u_i + v_i + c) \bmod b$;
   б) положить $c \leftarrow 0$, если $u_i + v_i + c$, и $c \leftarrow 1$ в противном случае.
3. Установить $w_k \leftarrow c$.

Вычитание $u - v = (w_{k-1} \cdots w_1 w_0)_b$ числа $v = (v_{k-1} \cdots v_1 v_0)_b$ из числа $u = (u_{k-1} \cdots u_1 u_0)_b$, $u \geq v$.

1. Установить $c \leftarrow 0$.
2. Для $i = 0, \ldots, k-1$ выполнить:
   а) $w_i \leftarrow (u_i - v_i - c) \bmod b$;
   б) положить $c \leftarrow 0$, если $u_i - v_i - c \geq 0$, и $c \leftarrow 1$ в противном случае.

Если $u < v$, то по окончании выполнения алгоритма $c = 1$ и результат $w = u - v + b^k$.

Умножение $uv = (w_{k+l-1} \cdots w_1 w_0)_b$ числа $u = (u_{k-1} \cdots u_1 u_0)_b$ на число $v = (v_{l-1} \cdots v_1 v_0)_b$.

1. Для $i = 0, \ldots, k + l - 1$ установить $w_i \leftarrow 0$.
2. Для $i = 0, \ldots, l - 1$ выполнить:
   а) $c \leftarrow 0$;
   б) для $j = 0, \ldots, k - 1$ вычислить $(xy)_b \leftarrow w_{i+j} + u_j v_i + c$ и установить $w_{i+j} \leftarrow y$, $c \leftarrow x$;
   в) $w_{i+k} \leftarrow c$.

Деление числа $u = (u_{k+l-1} \cdots u_1 u_0)_b$ на число $v = (v_{k-1} \cdots v_1 v_0)_b$, $k \geq 2$, $v_{k-1} \neq 0$, т.е. нахождение частного $q = (q_l \cdots q_1 q_0)_b$ и остатка $r = (r_{k-1} \cdots r_1 r_0)_b$ таких, что $u = qv + r$ и $0 \leq r < v$.

1. Выбрать произвольное целое d, для которого $vd < b^k$ и $\left\lfloor \frac{b}{2} \right\rfloor \leq v_{k-1} d < b$. Установить:
   а) $u \leftarrow ud$, $u = (u_{k+l} \cdots u_1 u_0)_b$;
   б) $v \leftarrow vd$, $v = (v_{k-1} \cdots v_1 v_0)_b$.
2. Для $i = k + l, k + l - 1, \ldots, k$:
   а) вычислить пробное частное
   $$\hat{q} \leftarrow \min\left(\left\lfloor \frac{u_i b + u_{i-1}}{v_{k-1}} \right\rfloor, b - 1\right);$$
   б) пока $\hat{q}(v_{k-1} b + v_{k-2}) > u_i b^2 + u_{i-1} b + u_{i-2}$, выполнить $\hat{q} \leftarrow \hat{q} - 1$;
   в) $u \leftarrow u - \hat{q} v b^{i-k}$;
   г) если $u < 0$, то установить: $\hat{q} \leftarrow \hat{q} - 1$, $u \leftarrow u + v b^{i-k}$;
   д) $q_{i-k} \leftarrow \hat{q}$.
3. Установить $r \leftarrow u/d$.

Действия на шаге 1 алгоритма называются *нормализацией*. При нормализации цикл 2б выполняется не более двух раз. Корректирующее сложение на шаге 2г выполняется с вероятностью $\approx \frac{2}{b}$. Если основание b является степенью 2, то нормализация может состоять в сдвиге битов u и v влево на такое минимальное количество позиций, что старший бит v_{k-1} равен 1. В общем случае можно выбрать $d = \lfloor b/(v_{k-1} + 1) \rfloor$.

12.11. Модулярная арифметика

Располагая алгоритмами арифметики больших чисел, можно организовать вычисления в кольце $\mathbb{Z}_n$. Для приведения по модулю результата аддитивной операции достаточно прибавить или вычесть n, а при умножении следует найти остаток от деления результата на n. Кроме того, известны методы приведения по модулю, использующие вычислительно трудоемкое деление только на стадии предварительных вычислений. Опишем два таких метода.

Метод Барретта [45]. Пусть модуль n является k-разрядным числом по основанию b, $x = (x_{2k-1} \cdots x_1 x_0)_b$. Для вычисления $r = x \bmod n$ предварительно находят число $m = \left\lfloor b^{2k}/n \right\rfloor$. Если $x = qn + r$, то

$$q = \left\lfloor \frac{x}{b^{k-1}} \frac{b^{2k}}{n} \frac{1}{b^{k+1}} \right\rfloor \approx \hat{q} = \left\lfloor \left\lfloor \frac{x}{b^{k-1}} \right\rfloor \frac{m}{b^{k+1}} \right\rfloor.$$

Точнее, известно, что $q - 2 \le \hat{q} \le q$. Поэтому число $\hat{r} = x - \hat{q}n$ совпадает либо с r, либо с $r + n$, либо с $r + 2n$. Таким образом, для вычисления r требуется использовать два умножения $\left\lfloor x/b^{k-1} \right\rfloor m$, $\hat{q}n$ и не более трех вычитаний $x - \hat{q}n$, $\hat{r} - n$, $(\hat{r} - n) - n$ больших чисел. Деление на $b^{k\pm1}$ состоит в сдвиге разрядов делимого на $k \pm 1$ позиций вправо.

Метод Монтгомери [148]. Пусть R — натуральное число, взаимно простое с модулем n, $n' = -n^{-1} \bmod R$, $x \in \{0, 1, \ldots, nR - 1\}$. Вычисляется $xR^{-1} \bmod n$. Если $y = xn' \bmod R$, то:

1) $x + yn \equiv x(1 + nn') \equiv 0 (\bmod R)$ и $(x + yn)/R$ — целое число;
2) $(x + yn)/R \equiv xR^{-1} (\bmod n)$;
3) $x + yn < 2nR$.

Таким образом, $(x + yn)/R = xR^{-1} \bmod n$ или $(x + yn)/R = xR^{-1} \bmod n + n$. Если используется представление чисел по основанию b, а R является степенью b, то для вычисления $xR^{-1} \bmod n$ достаточно выполнить два умножения xn' и yn, одно сложение $x + yn$ и, возможно, одно вычитание $(x + yn)/R - n$. Деление на R состоит в сдвиге разрядов делимого вправо.

Полагаем далее, что $n = (n_{k-1} \cdots n_1 n_0)_b$, НОД $(n, d) = 1$. Определим $R = b^k$, $n' = -n^{-1} \bmod b$ и приведем алгоритмы умножения и возведения в степень, основанные на методе Монтгомери.

Умножение по Монтгомери — вычисление $w = uvR^{-1} \bmod n$ для $u = (u_{k-1} \cdots u_1 u_0)_b$, $v = (v_{k-1} \cdots v_1 v_0)_b$, $0 \le u, v < n$.

1. Установить $w \leftarrow 0$. Считаем, что $w = (w_k \cdots w_1 w_0)_b$.
2. Для $i = 0, 1, \ldots, k - 1$ выполнить:
   а) $y_i \leftarrow (w_0 + u_i v_0) n' \bmod b$;
   б) $w \leftarrow (w + u_i v + y_i n)/b$.
3. Если $w \ge n$, то $w \leftarrow w - n$.

При выполнении шага 2б $w_0 + u_i v_0 + y_i n \equiv 0 \bmod b$, т.е. число $w + u_i v + y_i n$ делится нацело на b. Если перед выполнением шага $w < 2n - 1$, то и после выполнения

$$w \le \frac{1}{b}(2n - 2 + (b-1)(n-1) + (b-1)n) = 2n - 1 - \frac{1}{b} < 2n - 1.$$

Таким образом, окончательный результат вычислений w всегда меньше n. Кроме того,

$$w \equiv v(u_0 + u_1 b + \ldots + u_{k-1} b^{k-1}) b^{-k} \equiv uvR^{-1} (\bmod n).$$

Возведение в степень по Монтгомери — определение числа $w = u^e \bmod n$ для $e = (e_{l-1} \cdots e_1 e_0)_2$, $e_{l-1} = 1$, $1 \le u < n$.

1. Установить $w \leftarrow R \bmod n$, $\tilde{u} \leftarrow$ **Mont** $(u, R^2 \bmod n)$.
2. Для $i = l-1, l-2, \ldots, 0$ выполнить:
   а) $w \leftarrow$ **Mont**(w, w);
   б) если $e_i = 1$, то $w \leftarrow$ **Mont**$(w, \tilde{u})$.
2. $w \leftarrow$ **Mont**$(w, 1)$.

Здесь **Mont**(u, v) есть результат $uvR^{-1} \bmod n$ умножения по Монтгомери чисел u и v.

12.12. Признаки простоты

Для криптографии с открытым ключом важно как построение тестов, позволяющих установить *простоту данного большого числа*, так и построение методов разложения (факторизации) большого числа на простые множители. Рассмотрим лишь простейшие методы.

Теорема 12.12.1 (Вильсон). Натуральное n тогда и только тогда является простым, когда $(n+1)! + 1 \equiv 0 \pmod n$.

Доказательство. По теореме Ферма сравнение $x^{p-1} - 1 \equiv 0 \pmod p$ имеет $p-1$ решений $1, \ldots, p-1$. Поэтому

$$x^{p-1} - 1 = (x-1)(x-2) \cdots (x-p+1) + pF(x).$$

Положим $x = 0$, тогда $(p+1)! \equiv -1 \pmod p$. Если же n — составное, то оно содержит простой множитель $q < n$, который является делителем $(n-1)!$, так что $(n-1)! + 1$ не делится на q, а значит, и на n. □

В качестве следствия отсюда получается следующая теорема.

Теорема 12.12.2. Числа n и $n+2$ тогда и только тогда являются простыми числами-близнецами, когда

$$4((n-1)! + 1) + n \equiv 0 (\bmod n(n+2)), \quad n > 1.$$

Доказательство. Если условие выполнено, то, поскольку $n \ne 2, 4$, должно быть $(n+1)! + 1 \equiv 0 \pmod n$. Остается применить теорему Вильсона, чтобы убедиться в простоте числа n. Далее $n \equiv -2 \pmod{n+2}$. Поэтому $(n+1)! \equiv$ $\equiv 2(n-1)! \pmod{n+2}$. Отсюда следует:

$$0 \equiv 4((n-1)! + 1) + n \equiv 2(n+1)! + 2 \equiv 2((n+1)! + n) \pmod{n+2}. \quad (12.12.1)$$

Поэтому простым является и число $n+2$.

Пусть, наоборот, n и $n+2$ — простые числа. Тогда $4((n+1)!+1)+n$ по теореме Вильсона делится на n, а вследствие (12.12.1) — и на $n+2$, так что оно делится и на $n(n+2)$. □

Еще один критерий простоты можно сформулировать в виде следующей теоремы.

Теорема 12.12.3. Если существует такое число a, взаимно простое с n, что $a^{n-1} \equiv 1 \pmod p$, но $a^{(n-1)/p} \not\equiv 1 \pmod n$ для любого простого делителя p числа $n-1$, то n — простое число.

Доказательство. Так как каждый нетривиальный делитель t числа $n-1$ входит в один из делителей $(n-1)/p$, то из предположения теоремы следует, что $a^t \not\equiv 1 \pmod n$, так как в противном случае существовало бы p, для которого было бы $a^{(n-1)/p} \equiv 1 \pmod n$. Если a принадлежит показателю d, то $a^d \equiv 1 \pmod n$ и d должно быть делителем $n-1$, что возможно лишь для $d = n-1$. Таким образом, $n-1$ является делителем $\varphi(n)$. Для составных n $\varphi(n) < n-1$. Поэтому n должно быть простым числом. Если, наоборот, $n = p$, то первообразный корень удовлетворяет условиям теоремы. □

12.13. Алгоритмы генерации простых чисел

В большинстве описанных выше криптографических схем используются простые числа. Причем стойкость применяемых алгоритмов существенным образом зависит от того, насколько велики эти простые числа. Общая схема для большинства методов получения простых чисел состоит из следующих двух шагов.

1. Выбор большого нечетного числа.

2. Тестирование этого числа на простоту. Если оно оказалось простым, то алгоритм заканчивает работу. В противном случае переходим к шагу 1.

Существует два подхода к реализации шага 2 — вероятностный и детерминированный. Результатом работы детерминированного теста является доказуемое высказывание о том, что тестируемое число простое (составное). Для вероятностного теста на простоту характерно то, что, как правило, со 100%-ной гарантией можно определить, является ли число составным и только с вероятностью, близкой к 1 (но не равной 1), можно определить случай простого числа.

В основе большинства тестов простоты лежит следующая известная малая теорема Ферма (см. главу 2).

Она дает необходимый признак простоты числа p. Отсюда вытекает алгоритм тестирования числа n на простоту. Очевидно, что если при некотором a таком, что $n \nmid a$ имеет место $a^{n-1} \not\equiv 1 \pmod n$, то число n составное.

Однако для доказательства простоты необходим какой-нибудь достаточный признак простоты числа. Оказалось, что теорема, обратная малой теореме Фер-

ма, не верна, а именно существуют составные числа n такие, что для любых $(a, n) = 1$ имеет место сравнение

$$a^{n-1} \equiv 1 (\bmod n).$$

Такие числа называются *числами Кармайкла* и имеют вид $p_1 \cdots p_k$, где все нечетные простые p_i различны, причем $(p_i - 1) | (n - 1)$. Наименьшее число Кармайкла равно $561 = 3 \cdot 7 \cdot 11$ (табл. 12.2). Однако такие числа встречаются редко и малую теорему Ферма можно применять для построения простых чисел.

Таблица 12.2

Числа Кармайкла

$561 = 3 \cdot 11 \cdot 17$
$1105 = 5 \cdot 13 \cdot 17$
$1729 = 7 \cdot 13 \cdot 31$
$2465 = 5 \cdot 17 \cdot 29$
$2821 = 7 \cdot 13 \cdot 31$
$6601 = 7 \cdot 32 \cdot 41$
$8911 = 7 \cdot 19 \cdot 67$
$41041 = 7 \cdot 11 \cdot 13 \cdot 41$
$825265 = 5 \cdot 7 \cdot 17 \cdot 19 \cdot 73$
$413631505 = 5 \cdot 7 \cdot 17 \cdot 73 \cdot 89 \cdot 107$

Пусть n — нечетное составное число и a — целое такое, что $(a, n) = 1$. Число n называется *псевдопростым по основанию* a, если

$$a^{n-1} \equiv 1 (\bmod n).$$

Оказывается, что если составное n является числом Кармайкла, то существует не менее $n/2$ оснований a, для которых n не псевдопростое. Таким образом, если заранее знать, что n не является числом Кармайкла, то тестировать n на простоту можно, выбрав случайное основание a и проверив, является ли n псевдопростым по основанию a. Если условие не выполнено, то, согласно малой теореме Ферма, n — составное. Допустим, что выдерживаются k проверок, тогда вероятность того, что при этом n является составным, равна $1/2^k$. Выбирая k достаточно большим, можно находить такие числа n, что вероятность того, что n — составное, сколь угодно мала. Очевидным недостатком этого теста является необходимость заранее знать, что n не является числом Кармайкла. От этого недостатка можно избавиться, используя теорему Эйлера (см. главу 2, теоремы 2.8.2, 2.14.1).

Теперь можно описать следующий вероятностный тест на простоту.

Тест Соловея — Штрассена

1. Выбрать k различных оснований a таких, что $1 < a < n$.

2. Для каждого a проверить условие (2.14.1), и если оно не выполнено хотя бы для одного a, то n — составное. Если же для всех a (2.14.1) выполнено, то n является составным с вероятностью, не превышающей $1/2^k$.

Дальнейшее усовершенствование данной методики приводит к следующему широко используемому вероятностному тесту простоты.

Нечетное составное число $n = 2^s t + 1$ (t — нечетное) называется *сильно псевдопростым по основанию* a, если выполняется одно из условий:

1) $a^t \equiv 1 \pmod n$;

2) существует $0 \le r < s$ такое, что $a^{2^r t} \equiv -1 \pmod n$.

Если число n является простым, то для него одно из этих условий всегда выполняется. Действительно, в этом случае $a^{n-1} \equiv 1 \pmod n$. Тогда, в силу простоты n, $a^{(n-1)/2} \pmod n$ равно 1 или –1. Если получено –1, то, очевидно, n — сильно псевдопростое. В случае $a^{(n-1)/2} \equiv 1 \pmod n$ рассматриваем число $a^{(n-1)/4} \pmod n$, которое в силу простоты n также равно 1 или –1. Проделав эту процедуру не более s раз, получим требуемое утверждение.

Теорема 12.13.1. Если n — нечетное составное число, то оно является сильно псевдопростым по основанию a для не более чем $n/4$ оснований a, $0 < a < n$.

Для доказательства теоремы потребуются следующие две леммы.

Лемма 12.13.1. Пусть $d = (k,m)$. Тогда существует в точности d элементов в группе $\{g, g^2, \ldots, g^m = 1\}$, удовлетворяющих уравнению $x^k = 1$.

Лемма 12.13.2. Пусть p — нечетное простое число, $p = 2^{s'}t' + 1$, где t' — нечетное. Тогда количество $x \in (\mathbb{Z}/p\mathbb{Z})^*$, удовлетворяющих $x^{2^r t} \equiv -1 \pmod p$ (t — нечетное), равно 0 в случае $r \ge s'$ и равно $d2^r$, если $r < s'$, здесь $d = (t,t')$.

Доказательство. Пусть g — порождающий элемент группы $(\mathbb{Z}/p\mathbb{Z})^*$ и $x = g^j$, где $0 \le j < p-1$. Так как $g^{(p-1)/2} \equiv -1 \pmod p$ и $p - 1 = 2^{s'}t'$, то сравнение в условии леммы эквивалентно сравнению

$$2^r tj \equiv 2^{s'-1}t' \pmod{2^{s'}t'}, \tag{12.13.1}$$

где j — неизвестное. Если $r > s' - 1$, то, очевидно, данное сравнение не имеет решений. Если $r \le s' - 1$, то, разделив модуль и обе части (12.13.1) на $2^r d$, получим

$$\frac{t}{d} j \equiv 2^{s'-r-1} \frac{t'}{d} \left(\bmod 2^{s'-r} \frac{t'}{d}\right). \tag{12.13.2}$$

Так как $(t/d, 2^{s'-r} t'/d) = 1$, то существует единственное решение j_0 сравнения (12.13.2). Тогда $j_0 k$, где $k = 1, 2, \ldots, 2^r d$, являются решениями сравнения из условия леммы. □

Доказательство теоремы 12.13.1. Рассмотрим три возможных случая.

Случай 1. Предположим, что n делится на квадрат некоторого простого p, т.е. $p^{\alpha} \mid n$, $\alpha \geq 2$. Покажем, что в этом случае n не может быть псевдопростым (тем более сильно псевдопростым) для более чем $(n-1)/4$ оснований a, $0 < a < n$. Для этого предположим $a^{n-1} \equiv 1 (\bmod n)$, откуда $a^{n-1} \equiv 1 (\bmod p^2)$. Так как $(\mathbb{Z}/p^2\mathbb{Z})^*$ — циклическая группа порядка $p(p-1)$, то существует целое g такое, что $(\mathbb{Z}/p^2\mathbb{Z})^* = \{g, g^2, \ldots, g^{p(p-1)}\}$. Согласно лемме 12.13.1, количество тех a, для которых $a^{n-1} \equiv 1 (\bmod p^2)$, равно $d = (p(p-1), n-1)$. Так как $p \mid n$, то $p \nmid n-1$ и $d \mid p-1$, откуда $d \leq p-1$. Это значит, что доля всех a, не делящихся на p^2 в интервале от 0 до n и удовлетворяющих сравнению $a^{n-1} \equiv (\bmod p^2)$, не превосходит величины

$$\frac{p-1}{p^2-1} \leq \frac{1}{p+1} \leq \frac{1}{4}.$$

Это доказывает первый случай теоремы.

Случай 2. Предположим, что n является произведением двух различных простых p и q, $n = pq$. Пусть $p = 2^{s_1} t_1 + 1$, $q = 2^{s_2} t_2 + 1$, t_1, t_2 — нечетные. Без ограничения общности можно предположить $s_1 \leq s_2$. Чтобы элемент a был основанием, по которому n было бы сильно псевдопростым, должно выполняться одно из следующих условий:

1) $a^t \equiv 1 (\bmod p)$ и $a^t \equiv 1 (\bmod q)$;

2) $a^{2^r t} \equiv -1 (\bmod p)$ и $a^{2^r t} \equiv -1 (\bmod q)$ для некоторого r, $0 \leq r \leq s$.

По лемме 12.13.1 количество тех a, для которых выполнено первое условие, равно произведению (t, t_1) (количество классов вычетов по модулю p) на (t, t_2) (количество классов вычетов по модулю q), что, очевидно, не превосходит $t_1 t_2$. По лемме 12.13.2 для каждого $r < \min(s_1, s_2) = s_1$ количество таких a, что $a^{2^r t} \equiv -1 (\bmod n)$, равно $2^r (t, t_1) 2^r (t, t_2) \leq 4^r t_1 t_2$. Из $n - 1 > \varphi(n) = 2^{s_1+s_2} t_1 t_2$ следует, что доля тех a, $0 < a < n$, для которых n является сильно псевдопростым, не превосходит

$$\frac{t_1 t_2 + t_1 t_2 + 4 t_1 t_2 + 4^2 t_1 t_2 + \ldots + 4^{s_1-1} t_1 t_2}{2^{s_1+s_2} t_1 t_2} = 2^{-s_1-s_2}\left(1 + \frac{4_1^s - 1}{4-1}\right).$$

Если $s_2 > s_1$, то указанная выше величина не превосходит $2^{-2s_1-1}(2/3 + 4^{s_1}/3) \leq$ $\leq 2^{-3}(2/3) + 1/6 = 1/4$. В случае $s_1 = s_2$ заметим, что одно из двух неравенств $(t, t_1) \leq t_1$, $(t, t_2) \leq t_2$ является строгим, как если бы $t_1 \mid t$ и $t_2 \mid t$, то из того, что $p \equiv 1 (\bmod t_1)$ и $2^s t = pq - 1$ получим $t_1 \mid q - 1 = 2^{s_2} t_2$, т.е. $t_1 \mid t_2$. Аналогично получаем $t_2 \mid t_1$, откуда $t_1 = t_2$ и $p = q$, что противоречит условию рассматриваемого случая. Следовательно, либо $(t, t_1) < t_1$, либо $(t, t_2) < t_2$, и поскольку мы имеем

дело с нечетными числами, то $(t,t_1)(t,t_2) \le t_1 t_2/3$. Таким образом, доля тех a, для которых n — сильно псевдопростое, оценивается сверху величиной

$$\frac{1}{3}2^{-2s_1}\left(\frac{2}{3}+\frac{4^{s_1}}{3}\right) \le \frac{1}{18}+\frac{1}{9}<\frac{1}{4}.$$

Тем самым второй случай доказан.

Случай 3. Наконец предположим, что n — произведение более чем двух различных простых чисел: $n = p_1 \cdots p_k$, $k \ge 3$. Пусть $p_j - 1 = 2^{s_j} t_j$, где t_j — нечетные числа, и пусть s_1 — минимальное из s_j. Поступая так же, как во втором случае, получим следующую верхнюю границу доли тех a, для которых n — сильно псевдопростое:

$$2^{-s_1-s_2-\ldots-s_k}\left(1+\frac{2^{ks_1}-1}{2^k-1}\right) \le 2^{-ks_1}\left(\frac{2^k-2}{2^k-1}+\frac{2^{ks_1}}{2^k-1}\right)=$$

$$=2^{-k}\frac{2^k-2}{2^k-1}+\frac{1}{2^k-1}=2^{1-k} \le \frac{1}{4}. \qquad \square$$

Тест Миллера — Рабина. Предположим, необходимо определить, простым или составным является нечетное число n. Запишем n в виде $n = 2^s + t$, где t — нечетное, и выберем случайное целое a в диапазоне $1 < a < n$. Вычислим $a^t \pmod n$. Если получаем 1 или –1, то заключаем, что n проходит тест для данного значения a, и переходим к выбору следующего a. В противном случае вычисляем $(a^t)^2 \pmod n$ и сравниваем с –1. Если это условие выполняется, то n проходит тест, и переходим к новому a. В противном случае возводим полученное выражение в квадрат и т.д. Если, дойдя до $a^{(n-1)/2}$, не получили –1, то n — составное. Если тест прошел для k различных значений a, то с вероятностью не более $1/4^k$ число n — составное.

Тест Миллера — Рабина применяется при генерации простых чисел для стандарта DSS NIST.

12.13.1. Генерация простых чисел в ГОСТ Р 34.10—94

При генерации простых чисел для ЭЦП в ГОСТ Р 34.10—94 применяется детерминированный тест простоты, основанный на следующем утверждении.

Теорема 12.13.2. Пусть q — нечетное простое и $p = qN + 1$, где N — четное. Если также $p < (2q+1)^2$ и выполняются условия:

1) $2^{qN} \equiv 1 \pmod p$;

2) $2^N \not\equiv 1 \pmod p$,

то p — простое.

Доказательство. Пусть $p = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$, d — порядок числа 2 по mod p. Тогда:

а) $d \mid p-1$ в силу условия 1;

б) $d \nmid \dfrac{p-1}{q}$ в силу условия 2;

в) $d \mid \varphi(p)$ по теореме Эйлера, где $\varphi(p)$ — функция Эйлера.

Из «а» и «б» следует, что $q \mid d$ и, значит, $q \mid \varphi(p)$. Известно, что

$$\varphi(p) = p_1^{\alpha_1 - 1} \cdots p_s^{\alpha_s - 1}(p_1 - 1) \cdots (p_s - 1).$$

Предположим, что q совпадает с одним из p_i. Тогда существует натуральное n такое, что $p = qn$, и, значит, в силу условия теоремы $qn = qN + 1$. Это противоречие, так как $q \nmid 1$.

Таким образом, q должен делить один из множителей $p_i - 1$, т.е. $p_i = qn + 1$ для некоторого n. Отсюда имеем $p = p_i m = (qn + 1)m = qN + 1$ и, следовательно, $m - 1 = q(N - nm)$. Итак,

$$p = (qn + 1)(qs + 1),$$

где s и n — четные числа, причем $n \geq 2$, $s \geq 0$.

Предположим, что p — составное. Тогда $s \geq 2$, откуда $p \geq (2q + 1)^2$. Из этого противоречия следует, что $s = 0$, значит, $p = p_i$ — простое. □

Алгоритм, приведенный в ГОСТ Р 34.10—94, позволяет строить простые числа, у которых длина двоичного разложения больше или равна 17. Идея алгоритма заключается в следующем. Пусть требуется построить простое число p длиной t битов ($t \geq 17$). Для этого построим убывающий набор натуральных чисел $t_0, \ldots, t_s$ таких, что $t_0 = t$, $t_s < 17$ и $t_{i+1} = [t_i/2]$, т.е. либо $t_{i-1} = 2t_i + 1$, либо $t_{i-1} = 2t_i$. В процессе работы алгоритма будем получать последовательно простые числа $p_s, p_{s-1}, \ldots, p_0$, причем длина каждого p_i в точности равна t_i битов. На первом шаге алгоритма посредством алгоритма пробного деления получаем простое p_s длиной $t_s < 17$.

Далее несколько раз выполняется следующий шаг итерации. Пусть имеется простое число p_i длиной t_i битов, тогда простое число p_{i-1} длиной t_{i-1} битов ищется в виде

$$p_{i-1} = p_i N + 1,$$

где N удовлетворяет следующим условиям:

1) N — четное;

2) N — такое, что длина числа $p_i N + 1$ в точности равна t_i битов.

Такое N получают с помощью датчика случайных чисел. После этого число $p_{i-1} = p_i N + 1$ тестируют на простоту с помощью следующего теста.

Проверяем два условия:

1) $2^{p_i N} \equiv 1 \pmod{p_{i-1}}$;

2) $2^N \not\equiv 1 \pmod{p_{i-1}}$.

Если оба они выполняются одновременно, то число p_{i-1} считается простым.

Если хотя бы одно из условий не выполнено, то число p_{i-1} считается составным, N увеличивается на 2 и тест повторяется для нового $p_{i-1} = p_i N + 1$.

Процедура повторяется до получения простого числа p_0 длиной t_0 битов.

12.14. Задача факторизации

Задача факторизации состоит в разложении заданного натурального числа n в произведение $p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, $\alpha_i \geq 1$, степеней различных простых чисел $p_1, .., p_k$. Очевидно, такое разложение можно получить, последовательно применяя алгоритм нахождения собственного делителя заданного натурального числа.

Рассмотрим метод факторизации Ферма. Идея состоит в попытке найти представление $n = x^2 - y^2$, которое сразу дает разложение $n = (x + y)(x - y)$. Найти такое разложение можно по следующему алгоритму. Поскольку больший из множителей n больше $\sqrt{n}$, начинаем с $x = \lceil \sqrt{n} \rceil$, где $\lceil a \rceil$ обозначает наименьшее целое, больше либо равное a. Полагаем $z = x^2 - n$.

1. Проверяем, является ли z квадратом натурального числа. Если ответ положительный, то $y = \sqrt{z}$, и мы получаем требуемое представление. В противном случае переходим к следующему пункту.

2. Находим следующие значения для z и x:

$$z + 2x + 1 = x^2 + 2x + 1 - n = (x + 1)^2 - n$$

и переходим к пункту 1.

Оценим количество циклов, необходимых для достижения результата. Если $n = ab$ — произведение двух простых (худший случай) и $a < b$, то алгоритм прекращает работу, когда $x = (a + b)/2$. Так как мы начинали с $x \approx \sqrt{n}$ и $b = n/a$, то число циклов

$$W \approx \frac{1}{2}\left(a + \frac{n}{a}\right) - \sqrt{n} = \frac{(\sqrt{n} - a)^2}{2a}.$$

Если предположить, что $a = \delta\sqrt{n}$, $0 < \delta < 1$, то получаем

$$W \approx \frac{(1 - \delta)^2}{2\delta}\sqrt{n}.$$

Еще одна оценка количества циклов возникает из того, что алгоритм прекращает работу, когда y достигает $(b - a)/2$. Начальное значение y находится в пределах от 0 до $2\sqrt{n} + 1$. Поэтому

$$\frac{b - a}{2} - 2\sqrt{n} - 1 < W < \frac{b - a}{2}.$$

Это означает, что для обеспечения стойкости RSA необходимо, чтобы разность $b - a$ была велика.

Будем говорить, что число m является *B-гладким*, если все простые делители m не превосходят $B \in \mathbb{N}$. Множество S_B, составленное из всех простых чисел, не больших B, назовем *базой множителей*.

Пусть число $p-1$ является B-гладким для некоторого (но не для каждого) простого делителя p нечетного числа n. *Метод* $p-1$, предложенный Поллардом [159], позволяет найти собственный делитель d числа n, затратив $O(B \ln n/\ln B)$ модулярных умножений. Пусть

$$Q = \prod_{q \in S_B} q^{\alpha_q},$$

где α_q — минимальное целое такое, что

$$q^{\alpha_q} \geq \sqrt{n} > \min_{p|n}(p-1).$$

Вычислим $2^Q \bmod n$ и определим $d = \text{НОД}\,(2^Q - 1, n)$. Имеем:

1) $p|n$;
2) $(p-1)|Q$;
3) по малой теореме Ферма $2^Q \equiv 2^{p-1} \equiv 1 \pmod p$ и $p|(2^Q - 1)$.

Таким образом, $p|d$ и число d является делителем n.

Метод эллиптических кривых [122] является обобщением метода $p-1$. Для факторизации n требуется в среднем $L_{1/2,1}(p)$ операций, где p — наименьший простой делитель n, а

$$L_{\varepsilon,c}(x) = O(\exp((c + 0(1))(\ln x)^{\varepsilon}(\ln \ln x)^{1-\varepsilon})), \quad 0 < \varepsilon < 1, \quad c > 0.$$

Квадратичные методы факторизации направлены на поиск решения $(x, y) \in \mathbb{Z}_n\mathbb{Z}_n$ сравнения

$$x^2 \equiv y^2 \pmod n.$$

Если $x \not\equiv \pm y \pmod n$, то n делит $(x-y)(x+y)$, но не делит ни $(x-y)$, ни $(x+y)$. Таким образом, $d = \text{НОД}\,(x-y, n)$ является собственным делителем n.

Для решения сравнения поступают следующим образом.

1. Выбирают базу множителей $S_B = \{p_1, \ldots, p_s\}$.

2. Находят T пар целых чисел (a_t, b_t) таких, что $a_t^2 \equiv b_t \pmod n$ и число $|b_t|$ является B-гладким, т.е.

$$b_t = (-1)^{\alpha_{t0}} p_1^{\alpha_{t1}} \cdots p_s^{\alpha_{ts}}.$$

3. Составляется $T \times (s+1)$ матрица $A = (\alpha_{tl})$. Если сумма некоторых строк A является вектором с четными координатами $2(\beta_0, \beta_1, \ldots, \beta_s)$ и $\mathcal{T}$ — множество номеров таких строк, то искомое сравнение найдено:

$$\left(\prod_{i \in \mathcal{T}} a_i\right)^2 \equiv \left(\prod_{i=1}^{s} p_i^{\beta_i}\right)^2 \pmod n.$$

4. Если найденное сравнение не позволяет факторизовать n, то находят новые пары (a_t, b_t) и (или) новое множество $\mathcal{T}$ и повторяют вычисления.

Основную вычислительную трудность при таком подходе составляет нахождение подходящих пар (a_t, b_t) на этапе 2. Метод *квадратичного решета* (Quadratic Sieve, QS [160]) направлен на нахождение пар вида $(x + m, Q(x))$, где $m = \lfloor \sqrt{n} \rfloor$, $Q(x) = (x + m)^2 - n$, $x = 0, \pm 1, \pm 2, \ldots$ Поиск подходящих пар проводится следующим образом:

1) определяется массив R, индексируемый числами $x = 0, \pm 1, \pm 2, \ldots, \pm M$, и заполняется значениями $R[x] \leftarrow \log|Q(x)|$;

2) для всякой степени $p^\alpha \le C$ элемента базы множителей p находятся решения сравнения $Q(x) \equiv 0 (\mathrm{mod}\, p^\alpha)$ относительно $x \in \{0, 1, \ldots, p^\alpha - 1\}$;

3) если x^* — одно из найденных решений, то $R[x] \leftarrow R[x] - \log p$ для всех x таких, что $x \equiv x^* (\mathrm{mod}\, p^\alpha)$;

4) если после обработки степеней всех чисел из базы множителей (*просеивания*) $R[x]$ равно 0, то число $|Q(x)|$ является B-гладким (доказать, почему) и $(x + m, Q(x))$ — искомая пара.

При практической реализации метода квадратичного решета значения логарифмов вычисляются приближенно. После просеивания для всех x таких, что $R[x] \approx 0$, проверяется B-гладкость числа $Q(x)$.

Доказано, что оптимальным в методе квадратичного решета является использование $s \approx L_{1/2,\,1/2}(n)$ элементов базы множителей. При этом сложность алгоритма факторизации составляет $L_{1/2,1}(n)$ операций в среднем. Для сравнения: наиболее эффективный на сегодняшний день *обобщенный метод решета числового поля* (General Number Field Sieve, GNFS) имеет сложность $L_{1/3,\,c}(n)$, где $c = \left(\frac{64}{9}\right)^{1/3} \approx 1.923$.

В табл. 12.3 приведены результаты факторизации чисел RSA-k.

Таблица 12.3

Рекорды факторизации

Число	Дата	Сложность, MIPS-лет[1]	Алгоритм
RSA-100	Апрель 1991	7	QS
RSA-110	Апрель 1992	75	QS
RSA-120	Июнь 1993	830	QS
RSA-129	Апрель 1994	5000	QS
RSA-130	Апрель 1996	500	GNFS
RSA-140	Февраль 1999	2000	GNFS
RSA-155	Август 1999	8000	GNFS

[1] MIPS — миллион инструкций в секунду.

Каждое число n типа RSA-k состоит из k десятичных разрядов и является произведением двух близких простых p и q, которые генерировались с использованием вероятностных тестов на простоту и уничтожались сразу после вычисления pq. Последнее достижение — факторизация числа RSA-155, потребовавшая:

— использования 160 рабочих станций SGI и Sun (175—400 МГц);
— 8 процессоров SGI Origin 2000 (250 МГц);
— 120 персональных компьютеров Pentium II (300—450 МГц);
— 4 кластеров Digital\Compaq (500 МГц);
— 9 недель для настройки параметров GNFS;
— 5.2 месяца для сбора пар (a_i, b_i);
— 3.2 Гб оперативной памяти, 224 часа вычислений на Cray C916 для обработки матрицы A размером 6699191 × 6711336 с 417132631 ненулевыми элементами.

12.15. Задания

1. Найти количество решений сравнения $x^m \equiv 1 \pmod{n}$.

2. Сколько решений имеет сравнение $x^m \equiv -1 \pmod{n}$ в случае его разрешимости?

3. Пусть E, D — взаимно обратные преобразования RSA-криптосистемы. Тогда выполняется $D(E(x)) = x$ для любого $x \in (Z/N)^*$. Показать, что это свойство справедливо при любом x.

4. Доказать, что всевозможные преобразования RSA-криптосистемы образуют группу относительно их композиции.

5. Зашифровано сообщение по правилу

$$y \equiv x^k \pmod{p},$$

где p — большое простое число, $1 \le x \le p-1$, k — целое число, $1 < k < p-1$. Показать, что если k выбрано взаимно простым с $p-1$, то алгоритм расшифровывания

$$d(y) = y^d \pmod{p}$$

является корректным с $d \equiv k^{-1} \pmod{(p-1)}$ и $d(y) = x$.

6. Что случится с криптосистемой в предыдущей задаче, если ошибочно взять целое число k, не взаимно простое с $p-1$?

7. Показать, что в схемах RSA и Рабина шифрование открытого текста длиной n битов в зашифрованный текст длиной N битов требует $o(n^3)$ операций.

8. Предположим, что пользователь RSA в качестве модуля N по ошибке выбрал большое простое число. Показать, что в этом случае расшифровать текст легко.

9. Заданы различные простые p, q и $n = pq$. Показать, что если при заданных взаимно простых d и n сравнение

$$x^2 \equiv d \pmod{n}$$

имеет одно решение, то оно имеет четыре решения.

10. Рассмотрим RSA систему с модулем N. Целое число $x, 1 \le x \le N-1$, назовем неподвижной точкой, если оно и в зашифрованном виде тоже x. Показать, что если x — неподвижная точка, то и $N - x$ также есть неподвижная точка.

11. Показать, что в схеме RSA с параметрами p, q, e, d имеется $r + s + rs$ неподвижных точек M, $1 \le M \le N-1$, где

$$r = (p-1, e-1), \quad s = (q-1, e-1).$$

Задача показывает, что в схемах шифрования с большим количеством неподвижных точек уже заложен недостаток, поэтому желательно выбирать p и q, для которых r и s малы.

12. (п. 12.10). Пусть $u = (u_{k-1} \cdots u_1 u_0)_b$, $v' = (v_{k-1} \cdots v_1 v_0)_b + 1$, причем $v_i = b - 1 - u_i$. Доказать, что

$$u + v' = 2^k.$$

13. (п. 12.10). Доказать корректность следующих *бинарных методов* возведения числа u в степень $e = (e_{l-1} \cdots e_1 e_0)_2$, $e_{l-1} \ne 0$.

Справа налево	Слева направо
1. Установить $w \leftarrow 1$, $v \leftarrow u$.	1. Установить $w \leftarrow 1$.
2. Для $i = 0, 1, \ldots, l-1$ выполнить:	2. Для $i = l-1, l-2, \ldots, 0$ выполнить:
а) если $e_i \ne 0$, то $w \leftarrow wv$;	а) $w \leftarrow ww$;
б) $v \leftarrow vv$.	б) если $e_i \ne 0$, то $w \leftarrow wu$.
3. Вернуть w.	3. Вернуть w.

14. (п. 12.10). Предложить представление больших чисел в машинной памяти. Разработать алгоритм деления на одноразрядное число. Реализовать алгоритмы сравнения, сложения, вычитания, умножения и деления больших чисел.

15. (п. 12.10). Разработать и реализовать алгоритмы обращения в $\mathbb{Z}_n^*$ для случая, когда n является степенью 2. Использовать, предварительно доказав, следующие утверждения:

1) если $n = 2^s$, a — нечетное, $b_1 = 1$ и

$$b_t = \begin{cases} b_{t-1}, & \text{если } (ab_{t-1} \bmod 2^t) < 2^{t-1}, \\ b_{t-1} + 2^{t-1} & \text{в противном случае,} \end{cases} \quad t = 2, 3, \ldots,$$

то $b_s = a^{-1} \bmod n$;

2) если $n = 2^{2^s}$, a — нечетное, $ж_1 = 1$ и

$$c_t = (2c_{t-1} - ac_{t-1}^2) \bmod n, \quad t = 2, 3, \ldots,$$

то $c_s = a^{-1} \bmod n$.

16. (п. 12.11). Для увеличения быстродействия метода Барретта число $\hat{r}$ при $b > 3$ определяют по следующему алгоритму:

1) $r_1 \leftarrow x \bmod b^{k+1}$, $r_2 \leftarrow \hat{q}n \bmod b^{k+1}$;
2) $\hat{r} \leftarrow r_1 - r_2$;
3) если $\hat{r} < 0$, то установить $\hat{r} \leftarrow \hat{r} + b^{k+1}$.

Доказать, что и в этом случае $\hat{r} = x - \hat{q}n$.

17. (п. 12.11). Используя задание 13, доказать корректность алгоритма возведения в степень по Монтгомери.

18. (п. 12.11). Реализовать алгоритмы умножения и возведения в степень по Монтгомери. Выбрать основание b как степень 2, а для вычисления n' использовать алгоритмы из задания 15.

19. (п. 12.11). Пусть $n = b^k - m$, где m — l-разрядное число по основанию b, $l < k$. Для натурального $x = q_0 b^k + r_0$, $0 \le r_0 < b^k$, по правилу

$$q_{i-1}m = q_i b^k + r_i, \quad 0 \le r_i < b^k,$$

определены числа q_i, r_i, $i = 1, 2, \ldots$ Доказать:

1) существует натуральное t, для которого $q_t = 0$;
2) $x + (q_0 + \ldots + q_{t-1})m = (q_0 + \ldots + q_{t-1})b^k + (r_0 + \ldots + r_t)$;
3) $x \equiv (r_0 + \ldots + r_t) \pmod n$.

Используя данные выводы, разработать алгоритм и написать программу вычисления $x \bmod n$.

20. (п. 12.11). Написать программу вычисления символа Якоби $\left(\frac{a}{n}\right)$, не требующую факторизации n. Использовать следующие свойства:

1) если $a \equiv b \pmod n$, то $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right)$;

2) $\left(\frac{1}{n}\right)=1$;

3) $\left(\frac{-1}{n}\right)=(-1)^{(n-1)/2}$;

4) $\left(\frac{ab}{n}\right)=\left(\frac{a}{n}\right)\left(\frac{b}{n}\right)$;

5) $\left(\frac{2}{n}\right)=(-1)^{(n^2-1)/8}$, иначе говоря, $\left(\frac{2}{n}\right)=1$, если $n\equiv\pm1 \pmod 8$, и $\left(\frac{2}{n}\right)=-1$, если $n\equiv\pm3 \pmod 8$;

6) если a и n взаимно простые нечетные числа, то

$$\left(\frac{a}{n}\right)=(-1)^{(a-1)(n-1)/4}\left(\frac{n}{a}\right).$$

21. (п. 12.12). Согласно теореме Эйлера, если n — нечетное простое и $(a,n)=1$, то

$$a^{(n-1)/2}\equiv\left(\frac{a}{n}\right) \pmod n. \qquad (12.15.1)$$

Доказать, что если n — нечетное составное число, то соотношение (12.15.1) не выполняется по крайней мере для половины всех a с условием $(a,n)=1$.

22. (п. 12.12). Доказать, что если (12.15.1) выполняется для a_1 и не выполняется для a_2, то (12.15.1) не выполняется для a_1a_2. Применить данное утверждение для доказательства того, что если (12.15.1) не выполняется хотя бы для одного a, то количество таких a не меньше количества тех a, для которых (12.15.1) выполнено.

23. (п. 12.12). Если n делится на квадрат простого числа, то показать, как найти целое a такое, что $a^{(n-1)/2}\not\equiv\pm1 \pmod n$.

24. (п. 12.12). Доказать, что если n — произведение различных простых чисел, p — одно из них, a обладает свойством $\left(\frac{a}{p}\right)=-1$ и $a\equiv1 \pmod{n/p}$, то (12.15.1) не выполняется для этого a. Показать, что такое a всегда существует.

25. (п. 12.12, числа Мерсенна). Числа вида $\mathcal{M}_n=2^n-1$, где $n=2,3,\ldots$, называют *числами Мерсенна*. Известно, что $\mathcal{M}_n$ является простым тогда и только тогда, когда:

1) n — простое;

2) $(n-2)$-й элемент последовательности

$$s_0=4,\quad s_t=(s_{t-1}^2-2) \bmod n,\quad t=1,2,\ldots,$$

равен 0.

Написать программу, доказывающую простоту первых 20 чисел Мерсенна из табл. 12.4.

Таблица 12.4

Последовательные простые числа Мерсенна $\mathcal{M}_n = 2^n - 1$

n	$\mathcal{M}_n$	n	n	n	n
2	3	61	2281	21701	859433
3	7	89	3217	23209	1257787
5	31	107	4253	44497	1398269
7	127	127	4423	86243	2976221
13	8191	521	9689	110503	3021377
17	131071	607	9941	132049	6972593(?)
19	524287	1279	11213	216091	13466917(?)
31	2147483647	2203	19937	756839	

26. (п. 12.14). Для метода $p-1$ проанализировать случаи, когда d не является собственным делителем n. Доказать:

    1) если $d = 1$, то $p-1$ не является B-гладким ни для одного простого делителя p числа n;

    2) если $d = n$, то $p-1$ является B-гладким для всякого простого делителя p числа n.

27. (п. 12.14). Доказать, что в методе квадратичного решета среди делителей $Q(x)$ нет таких нечетных простых p, что $\left(\frac{n}{p}\right) = -1$, и такие p следует исключить из базы множителей S_B.

28. (п. 12.14). Написать программу факторизации по методу $p-1$. Факторизовать число $n = 1846202297$, выбрав $B = 23$.

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV

Глава 13

ФУНКЦИИ ХЭШИРОВАНИЯ

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV

13.1. Определение и свойства

Функцией хэширования называется отображение $h: \mathcal{A}^* \to \mathcal{A}^n$ слов конечной длины в алфавите $\mathcal{A}$ в слова длиной n. Хэш-значение $Y = h(X)$ используется для контроля целостности X, например, в системах электронной цифровой подписи. Предполагается, что алгоритм вычисления хэш-значения является эффективным (имеет полиномиальную от n сложность) и общедоступным.

При построении функций хэширования требуют, чтобы вычислительно трудным являлось решение следующих криптоаналитических задач:

H1) по заданному $Y = h(X)$ определить X;

H2) для заданного X найти другое слово X' такое, что $h(X) = h(X')$;

H3) найти пару различных слов X, X' такую, что $h(X) = h(X')$.

Если не существует алгоритма полиномиальной от n сложности для решения H1, H2, H3, то h называется соответственно *односторонней*, *свободной от коллизий* и *строго свободной от коллизий*. Данные термины будем применять и к функциям $g: \mathcal{X} \to \mathcal{A}^n$, $\mathcal{X} \subseteq \mathcal{A}^*$.

Известно, что:

1) если g является строго свободной от коллизий, то g свободна от коллизий;

2) из того, что g является строго свободной от коллизий, не следует, что g является односторонней;

3) если g не является односторонней, то существует вероятностный полиномиальный алгоритм решения задачи H2.

Последнее утверждение вытекает из следующей теоремы.

Теорема 13.1.1. Пусть задано отображение $g: \mathcal{X} \to \mathcal{Y}$, где $\mathcal{X}$ и $\mathcal{Y}$ — конечные множества, причем $|\mathcal{X}| \geq 2|\mathcal{Y}|$. Если имеется эффективный алгоритм A вычисления X по известному $Y = f(X)$, то есть и эффективный вероятностный алгоритм, который находит коллизию для g с вероятностью не менее 1/2.

Доказательство. В качестве искомого алгоритма возьмем следующий. Выбираем произвольно $X \in \mathcal{X}$, находим $Y = g(X)$, $X' = \mathrm{A}(Y)$. Если $X \neq X'$, то имеем коллизию, так как $g(X) = g(X')$. Осталось показать, что вероятность успеха не меньше 1/2.

Обозначим через X полный прообраз $g(X)$, а через $\tilde{\mathcal{X}}$ — множество всех прообразов. Тогда вероятность успеха при выборе X равна $(|(X)| - 1)/|(X)|$. Вероят-

ность p успеха алгоритма в целом вычисляется усреднением по всем X. Таким образом,

$$p=\frac{1}{|\mathcal{X}|}\sum_{X\in\mathcal{X}}\frac{|(X)|-1}{|(X)|}=\frac{1}{|\mathcal{X}|}\sum_{X\in\tilde{\mathcal{X}}}(|(X)|-1)=\frac{1}{|\mathcal{X}|}\left(\sum_{X\in\tilde{\mathcal{X}}}|(X)|-\sum_{X\in\tilde{\mathcal{X}}}1\right)\geq$$

$$\geq\frac{|\mathcal{X}|-|\mathcal{Y}|}{|\mathcal{X}|}\geq\frac{|\mathcal{X}|-|\mathcal{X}|/2}{\mathcal{X}}=\frac{1}{2}. \qquad \square$$

13.2. Блочно-итерационные функции хэширования

Большинство используемых на практике функций хэширования являются *блочно-итерационными*. Значение $h(X)$ такой функции вычисляется следующим образом.

1. Слово X длиной L разбивается на блоки $X^{(1)},\ldots,X^{(d)}\in\mathcal{A}^m$. Если $l=L \bmod m\neq 0$, то предварительно к X дописывается $m-l$ заранее оговоренных символов.

2. Формируется слово $X^{(d+1)}\in\mathcal{A}^m$, содержащее представление числа L или числа l.

3. Выполняются итерации

$$Y^{(i)}=\varkappa(X^{(i)},Y^{(i-1)}),\quad i=1,\ldots,d+1,$$

использующие *шаговую функцию хэширования* $\varkappa$, действующую из $\mathcal{A}^m\times\mathcal{A}^n$ в $\mathcal{A}^n$, и фиксированное *стартовое хэш-значение* $Y^{(0)}$.

4. Устанавливается хэш-значение $h(X)=Y^{(d+1)}$.

13.3. Использование блочных криптосистем

Пусть блочно-итерационная функция хэширования h действует на множестве бинарных слов, т.е. $\mathcal{A}=\mathbb{F}_2$. Для построения шаговой функции хэширования будем использовать блочную криптосистему $F=\{F_\theta:\theta\in V_m\}$, где F_θ — биективное преобразование множества V_n, определяемое ключом θ. При $n=m$ определено семейство

$$\varkappa(X,Y)=F_{\alpha_1X\oplus\alpha_2Y}(\beta_1X\oplus\beta_2Y)\oplus\gamma_1X\oplus\gamma_2Y,\quad \alpha_i,\beta_i,\gamma_i\in\mathbb{F}_2,$$

шаговых функций хэширования.

Если, например, $\alpha_1=\gamma_1=0$, $\alpha_2=\beta_1=\beta_2=\gamma_2=1$, то

$$\varkappa(X,Y)=F_Y(X\oplus Y)\oplus Y$$

и получаемая хэш-функция h не является стойкой к коллизиям (доказать). Стойкие к известным методам криптоанализа конструкции шаговых функций хэширования приведены в табл. 13.1.

Таблица 13.1

Шаговые функции хэширования $\varkappa(X,Y)$, использующие блочную криптосистему F

$F_Y(X)\oplus X$	$F_X(Y)\oplus Y$	$F_{X\oplus Y}(X)\oplus X$
$F_Y(X\oplus Y)\oplus X\oplus Y$	$F_X(X\oplus Y)\oplus X\oplus Y$	$F_{X\oplus Y}(Y)\oplus Y$
$F_Y(X)\oplus X\oplus Y$	$F_X(Y)\oplus X\oplus Y$	$F_{X\oplus Y}(X)\oplus Y$
$F_Y(X\oplus Y)\oplus X$	$F_X(X\oplus Y)\oplus Y$	$F_{X\oplus Y}(Y)\oplus X$

13.3.1. Функция хэширования ГОСТ Р 34.11—94

Определяемая стандартом РФ ГОСТ Р 34.11—94 функция хэширования является блочно-итерационной с параметрами $\mathcal{A}=\mathbb{F}_2$, $n=m=256$. Хэш-функция построена с использованием блочной криптосистемы $F=\{F_\theta:\theta\in V_{256}\}$, $F_\theta\in\mathcal{S}(V_{64})$, стандарта ГОСТ 28147—89 (см. п. 10.3).

Стартовое хэш-значение $Y^{(0)}$ выбирается произвольным образом. Последний блок сообщения $X^{(d)}$ дополняется до 256 битов нулями. Дописываются два блока:

1) $X^{(d+1)}$ содержит контрольную сумму

$$\Sigma=X^{(1)}\boxplus X^{(2)}\boxplus\ldots\boxplus X^{(d)};$$

2) $X^{(d+2)}$ содержит длину L исходного сообщения.

Оба блока подвергаются преобразованию с использованием шаговой функции хэширования $\varkappa: V_{256}\times V_{256}\to V_{256}$. Искомое хэш-значение полагается равным $Y^{(d+2)}$.

Значения $Z=\varkappa(X,Y)$ вычисляются следующим образом.

1. Вычисляются ключи $\kappa=(\kappa_1,\kappa_2,\kappa_3,\kappa_4)$, где $\kappa_i=K_i(X,Y)$, K_i — аффинное отображение $V_{256}\times V_{256}\to V_{256}$.

2. Выполняется шифрование $S=F_\kappa(Y)$ по ГОСТ 28147—89, при котором первые 64 бита слова Y шифруются на ключе κ_1, вторые 64 бита — на ключе κ_2 и т.д.

3. Возвращается значение

$$Z=\psi^{74}(S)\oplus\psi^{62}(X)\oplus\psi^{61}(Y),$$

где ψ — линейная подстановка на множестве V_{256}.

13.4. Атака «дней рождения»

Атака *«дней рождения»* направлена на нахождение коллизии функции хэширования h (решение задачи Н3). Криптоаналитик образует T модификаций $X_1,\ldots,X_T$ *подлинного сообщения* $X\in\mathcal{A}^*$, которым соответствуют различные хэш-значения $Y_1=h(X_1),\ldots,Y_T=h(X_T)$. Модификации не изменяют сути X

и могут, например, состоять в добавлении или удалении незначащих символов. Затем криптоаналитик формирует T' вариантов $X'_1, \ldots, X'_{T'}$ *поддельного сообщения* $X' \in \mathcal{A}^*$, дающих различные хэш-значения $Y'_1, \ldots, Y'_{T'}$. Если множества $\mathcal{Y} = \{Y_1, \ldots, Y_T\}$ и $\mathcal{Y}' = \{Y'_1, \ldots, Y'_{T'}\}$ пересекаются, то коллизия найдена — хэш-значения некоторых модификаций подлинного и поддельного сообщений совпадают.

Если $\mathcal{Y}$ и $\mathcal{Y}'$ — случайные бесповторные выборки элементов $\mathcal{A}^n$, то коллизия будет найдена с вероятностью

$$1 - \frac{(N)_{T+T'}}{(N)_T (N)_{T'}} = 1 - \exp\left\{-\frac{TT'}{N}\right\}(1 + o(1)).$$

Здесь $N = |\mathcal{A}|^n$, $(N)_t = N!/(N-t)!$, а асимптотическая оценка в правой части справедлива при $N \to \infty$, T, $T' = o(N^{2/3})$. Например, после хэширования $\sqrt{N}$ подлинных и $\sqrt{N}$ поддельных сообщений коллизия появляется с вероятностью ≈ 0.63.

Название атаки связано с известным *парадоксом дней рождения*: при независимом равновероятном «размещении» 23 дней рождения по 365 дням года некоторые два дня рождения совпадут с вероятностью не менее 1/2.

Для хранения элементов $\mathcal{Y}$ и $\mathcal{Y}'$ требуется $O(\sqrt{N})$ битов памяти. Известны варианты проведения атаки «дней рождения», требующие хранения меньшего количества промежуточных хэш-значений. Приведем описание одного из них.

1. Строится разбиение $\mathcal{B}_1 \cup \mathcal{B}_2$ множества $\mathcal{A}^n$, $|\mathcal{B}_1| = |\mathcal{B}_2|$.

2. Задается отображение $\sigma\colon \mathcal{A}^n \to \mathcal{A}^n$,

$$\sigma(Y) = \begin{cases} h(\delta(X,Y)), & \text{если } Y \in \mathcal{B}_1, \\ h(\delta(X',Y)), & \text{если } Y \in \mathcal{B}_2, \end{cases}$$

где $\delta(Z,Y) \in \mathcal{A}^*$ — результат модификации слова $Z \in \mathcal{A}^*$, определяемой словом $Y \in \mathcal{A}^n$, причем $\delta(Z,Y) \neq \delta(Z,Y')$, если $Y \neq Y'$.

3. Выбирается произвольное слово $Y_0 \in \mathcal{A}^n$, и вычисляется последовательность $Y_t = \sigma(Y_{t-1})$, $t = 1, 2, \ldots$ Данная последовательность является периодической, и $Y_{t_0+r} = Y_{t_0}$ для предпериода t_0 и минимального периода r.

4. Если $t_0 \neq 0$ и слова Y_{t_0-1}, Y_{t_0+r-1} лежат в разных множествах $\mathcal{B}_1$, $\mathcal{B}_2$, то коллизия найдена. Действительно, если, например, $Y_{t_0-1} \in \mathcal{B}_1$, $Y_{t_0+r-1} \in \mathcal{B}_2$, то модификациям $\delta(X, Y_{t_0-1})$ и $\delta(X', Y_{t_0+r-1})$ истинного и поддельного сообщений соответствует одинаковое хэш-значение Y_{t_0}. Если коллизию найти не удалось, то следует вернуться к шагу 3.

Известно, что

$$\mathbf{E}\{t_0\} = \mathbf{E}\{r\} = \sqrt{\frac{\pi N}{8}}, \quad \mathbf{E}\{t_0 + r\} = \sqrt{\frac{\pi N}{2}}$$

для случайного σ с равномерным распределением на множестве всех отображений из $\mathcal{A}^n$ в $\mathcal{A}^n$.

Для поиска совпадения элементов последовательности $Y_0, Y_1, \ldots$ применяется *алгоритм Брента* [60]. На i-й итерации алгоритма, $i = 0, 1, \ldots$, проверяется совпадение сохраненного значения Y_t, $t = 2^i - 1$, с вычисляемыми значениями Y_τ, $\tau = 2^i, 2^i + 1, \ldots, 2^{i+1} - 1$. По первому найденному совпадению $Y_t = Y_\tau$ можно определить период $r = \tau - t$.

13.5. Криптосистемы аутентификации

При вычислении значения h может использоваться ключ θ. Совокупность $H = \{h_\theta : \theta \in \Theta\}$ ключезависимых функций хэширования называется *криптосистемой аутентификации*. Хэш-значение $Y = h_\theta(X)$ используется для контроля как целостности, так и происхождения X, поскольку Y могут вычислить лишь доверенные лица, знающие θ.

Система H подвержена атакам при известном, выбранном и выбираемом открытом тексте. Криптоаналитику требуется по парам «слово X_t — хэш-значение $Y_t = h_\theta(X_t)$», $t = 1, \ldots, T$, решить одну из следующих задач:

A1) определить ключ θ;

A2) не определяя θ, построить алгоритм вычисления хэш-значения $Y_{T+1} = h_\theta(X_{T+1})$ для любого заданного слова X_{T+1};

A3) построить алгоритм вычисления хэш-значения $Y_{T+1} = h_\theta(X_{T+1})$ для некоторого, отличного от $X_1, \ldots, X_T$, слова X_{T+1}.

В стандарте FIPS HMAC используются:

— блочно-итерационная функция хэширования $h: \{0,1\}^* \to V_n$ с длиной блока m;

— ключ $\theta \in V_m$;

— фиксированные векторы $\alpha, \beta \in V_m$.

Хэш-значение $Y = h_\theta(X)$ определяется по правилу

$$Y = h((\theta \oplus \alpha) \,||\, h((\theta \oplus \beta) \,||\, X)).$$

13.6. Функция хэширования СТБ 1176.1—99

Государственный стандарт Республики Беларусь СТБ 1176.1 «Информационная технология. Защита информации. Функция хэширования» введен в действие 30 сентября 1999 г. Определяемая стандартом функция хэширования h действует на последовательности байтов. Хэш-значение является двоичным словом длиной L из диапазона $142 \le L \le 256$. Используется 256-битовое стартовое хэш-значение, которое выбирается произвольным образом, единым для группы пользователей. Приведем алгоритм вычисления значений h.

Обозначения. Преобразования стандарта выполняются над числами из множества $Z(n) = \{0, 1, \ldots, 2^n - 1\}$. Отождествим число $a \in Z(n)$ с вектором из V_n,

содержащим двоичную запись a, и перенесем на $Z(n)$ действие операций $\oplus$, $\boxplus$, $\lll$. Разложение $a = \sum_{i=0}^{k-1} a_i (2^m)^i$, $a_i \in Z(m)$, числа a по основанию 2^m будем записывать в виде $a = a_{k-1} || \ldots || a_1 || a_0$.

Преобразования. При вычислении значений h используются следующие вспомогательные преобразования.

1. Преобразования $\rho_0, \rho_1, \rho_2, \rho_3 : Z(512) \to Z(512)$ действуют на число $X = x_{15} || \ldots || x_1 || x_0$, $x_i \in Z(32)$ по правилам:

$$\rho_0(X) = ((x_{15} \oplus x_{13} \oplus x_3 \oplus x_0) \boxplus C_0) || x_{15} || \ldots || x_1;$$

$$\rho_1(X) = ((x_{15} \oplus x_2 \oplus x_0) \boxplus C_1) || x_{15} || \ldots || x_1;$$

$$\rho_2(X) = ((x_9 \oplus x_4 \oplus x_0) \boxplus C_2) || x_{15} || \ldots || x_1;$$

$$\rho_3(X) = ((x_{13} \oplus x_8 \oplus x_0) \boxplus C_3) || x_{15} || \ldots || x_1,$$

где $C_0 = (2BDA732E)_{16}$, $C_1 = (3920FE85)_{16}$, $C_2 = (BC1641F9)_{16}$, $C_3 = (75FE243B)_{16}$.

2. В стандарте используются S-блоки $s: Z(8) \to Z(8)$, $j \mapsto s_j$, которые задаются числами $s_{255} || \ldots || s_1 || s_0 \in Z(2048)$. Преобразование ω числа $X \in Z(128)$ определяют восемь S-блоков, заданных числами $Y_1, \ldots, Y_8$. Образ $\omega(X; Y_1, \ldots, Y_8)$ является результатом 32-кратного действия подстановки Фейстеля:

$$X \mapsto X_0 || (X_1 \oplus f(X_0)), \quad X = X_1 || X_0, \quad X_i \in Z(64).$$

Используемая тактовая функция $f: Z(64) \to Z(64)$ действует по правилу

$$x_8 || \ldots || x_1 \mapsto (s_8(x_8) || \ldots || s_1(x_1)) \lll 3, \quad x_i \in Z(8),$$

где s_i — S-блок $Z(8) \to Z(8)$, определяемый числом Y_i.

3. Преобразование $\xi: Z(256) \to Z(256)$ ставит в соответствие числу $X = x_7 || \ldots || x_1 || x_0$, $x_i \in Z(32)$, число

$$\xi(X) = x_6 || \ldots || x_0 || (x_0 \oplus x_2 \oplus x_4 \oplus x_7).$$

4. Отображение $\varphi: Z(128) \times Z(256) \to Z(256)$ ставит в соответствие числам $X = x_3 || \ldots || x_0$ и $Y = y_7 || \ldots || y_0$, x_i и $y_i \in Z(32)$, число

$$\varphi(X,Y) = y_3 || \ldots || y_0 || ((y_3 \boxplus x_3) \oplus y_7) || \ldots || ((y_0 \boxplus x_0) \oplus y_4).$$

Входные данные. Подлежащее хэшированию сообщение M является последовательностью байтов $m_1, \ldots, m_l \in Z(8)$. Сообщение дополняется таким минимальным числом k нулевых байтов $m_{l+1}, \ldots, m_{l+k}$, что $n = (l + k)/32$ — целое число. Затем определяются числа

$$M_i = m_{32i} || \ldots || m_{32(i-1)+2} || m_{32(i-1)+1}, \quad i = 1, \ldots, n,$$

и число $M_{n+1} = l$.

Переменные. При вычислении хэш-значений используются следующие переменные.

1. Переменная $H \in Z(256)$, $H = H_1 \| H_0 = h_7 \| \ldots \| h_0$, где $H_i \in Z(128)$, $h_i \in Z(32)$. При выполнении алгоритма переменная H содержит текущее хэш-значение, а в начале вычислений — стартовое хэш-значение.

2. Переменные $T_0, T_1, \ldots, T_7 \in Z(2048)$. Значения переменных T_1, T_3, T_5, T_7 не изменяются при выполнении алгоритма и заданы в табл. 13.2, 13.3 числами $t_{ij} \in Z(32)$ так, что $T_i = t_{i,63} \| \ldots \| t_{i0}$. Значения переменных T_0, T_2, T_4, T_6 определяются числами $r_{ij} \in Z(512)$: $T_i = r_{i3} \| r_{i2} \| r_{i1} \| r_{i0}$.

3. Переменная $V \in Z(512)$, $V = v_{15} \| \ldots \| v_1 \| v_{10}$, $v_i \in Z(32)$. Начальные значения переменных v_i указаны в табл. 13.4.

4. Переменная $K \in Z(256)$, $K = K_1 \| K_0 = k_7 \| \ldots \| k_0$, где $K_i \in Z(128)$, $k_i \in Z(32)$.

5. Переменная $W \in Z(256)$, $W = W_1 \| W_0$, где $W_i \in Z(128)$.

Таблица 13.2

Значения t_{ij}, $i = 1, 3$ (шестнадцатеричная запись)

j	i		j	i		j	i	
	1	3		1	3		1	3
0	AA2AA82E	4DCDCF4F	22	BD3CEF6A	53D3C444	44	5B5C0459	188A8909
1	8A0A088E	69E9EB6B	23	CFBC1532	FF7F6EEE	45	CC5F5D16	29B9F63D
2	A222A026	65E5E767	24	EB3BE96D	1C9CDF5F	46	47945291	A406F721
3	82020086	41C1C343	25	9B396BCD	2AB8FB7B	47	35FEF813	853F8677
4	AE2CAC28	49C9CB4B	26	B130E336	E2700181	48	F272787C	07879515
5	8C0E0C88	6DEDEF6F	27	E11003B6	9E0D7CAD	49	907058FC	23BBA32B
6	A624A420	61E1E363	28	9E1C9C18	72F48357	50	147946B5	A5AF1A98
7	84061E9A	45C5C747	29	EACA4A6E	46D0A92D	51	DAD9C267	009A25BD
8	AB2BA92F	5DDDCC4C	30	4EDF4C99	8D5B0FDB	52	E765E511	31B19303
9	8B0B098F	79F9E868	31	5E7FB481	F3747E91	53	C0F0437D	0B19A7B5
10	A323A127	F17140C0	32	E868FA6C	54C6D456	54	F764C344	3C3E14BF
11	B32133A7	D55564E4	33	C8485ACE	20F0F222	55	C6FFF9C4	B33927B7
12	8D0F9F1B	51D1C242	34	E262E034	2CFC1EFE	56	F5775654	1F117624
13	AF2D3FBB	75F5EC6C	35	92424096	58D838BA	57	9545FD75	26948B35
14	07870583	FD7D4ECE	36	EE7EEC7A	50D28052	58	C1D8D376	991B9B97
15	17859D19	D95960EA	37	DD0D1FCB	30A0A232	59	C77441DE	9D34963B
16	BA3AB83E	5CDCDE5E	38	E666E460	3A825ADA	60	D25351D0	3337B0BE
17	98381ABE	6AF8FA78	39	971D4F93	08889092	61	C5D7D5F6	02101716
18	B231B025	66E6E062	40	FB697BED	04D6D705	62	50D6D157	12A69FB4
19	806312A5	48C88C0C	41	C949DB4D	73A1ACAE	63	D455DCF4	B61336B2
20	AD3DBF29	7AA8AA28	42	F36137F1	BC0AAB1D			
21	89B94B6F	8E0E4ACA	43	0173B771	8F2F842E			

Таблица 13.3

Значения t_{ij}, $i = 5, 7$ (шестнадцатеричная запись)

j	i		j	i		j	i	
	5	7		5	7		5	7
0	5557455D	B2B03212	22	420710D8	ACFE5623	44	C9CB9901	DDE62754
1	4715547C	921A13BF	23	DADCD5FB	9C263D24	45	CA86BF9B	850D4625
2	5644465C	BAB83A18	24	FDFFEDF7	A1217391	46	CE9C87A9	3972776E
3	1416177D	9A101BB7	25	EFBDDE36	839D502A	47	6D03ABC0	78E2416C
4	51534159	B6B43616	26	E8FCEAE6	FC2E7680	48	E7B5B76F	E06A6444
5	43115078	961E17BB	27	B8FEDD33	8ED2473B	49	A5A7A00A	C248D9EB
6	7072607A	BEBC1C3E	28	E9BAEBA8	954599F6	50	B6A29E2E	685FCB2B
7	F8FAF9D1	9E143F11	29	D9DBF212	A58A2C74	51	A4A6E38C	CAC95D29
8	5F4D4F05	B3B13303	30	4E0B2408	885A288F	52	B18EB389	E44FC870
9	1D1F5EF4	931902AE	31	B4E0C83C	F8FF092F	53	A1A3EC04	7BEDE975
10	585A4852	BDAF3715	32	D7C5C726	F2F02042	54	C2EE8B98	4D7F7DC0
11	4A185BF1	9F0706AA	33	9597D03A	D05851FD	55	9A880020	8C67DED7
12	494B1913	B5A73505	34	D4D6C41C	EAE87A4A	56	372527AD	E36365D5
13	1B091AF0	970E04B9	35	C694D379	DA405BF7	57	8A6B223D	D1C76FE1
14	7E6C6E74	ABA90131	36	C1C3917B	F4F5668B	58	02061E90	C379625C
15	F6E4F50D	9B0A3057	37	9381963E	861D890F	59	BC0C0E23	4ED6606D
16	7577657F	A2A02200	38	E2B0F369	7E2DD8FB	60	BE928FAC	CFDC6B4B
17	6735764C	900853AD	39	282A2C84	5E87F984	61	80AF3F83	69CDD4CE
18	64626668	A8FA380C	40	DFCDCFE1	F3F17143	62	39AE823B	C649614C
19	3430710F	98520BA3	41	CCD2B99D	8159D3EE	63	2D292F2B	C5CCC1C4
20	7332616A	A6A43455	42	AA9FB28D	EFE7E58D			
21	63313840	943C1F82	43	BBE58521	EC7CDFDB			

Таблица 13.4

Начальные значения v_i (шестнадцатеричная запись)

i	v_i	i	v_i	i	v_i	i	v_i
0	D1845AC6	4	F19A24CD	8	87D45A6F	12	2A8A1A76
1	AC3D25C6	5	B47D25C6	9	3D5721C6	13	DC6715C6
2	F467247D	6	D4522491	10	573714C8	14	B4F1257D
3	079294AB	7	0D817489	11	078274DB	15	0B1294AC

Алгоритм. Алгоритм вычисления хэш-значения $h(M)$ состоит из следующих шагов.

1. $d \leftarrow 1$.
2. $K \leftarrow M_d$.
3. $V \leftarrow (v_{15} \boxplus h_7) \,||\, \ldots \,||\, (v_8 \boxplus h_0) \,||\, (v_7 \boxplus k_7) \,||\, \ldots \,||\, (v_0 \boxplus k_0)$.

4. Для $i = 0, 2, 4, 6$ выполнить:
   а) $V \leftarrow \rho_0^{29}(V)$, $r_{i0} \leftarrow V$;
   б) $V \leftarrow \rho_1^{18}(V)$, $r_{i1} \leftarrow V$;
   в) $V \leftarrow \rho_2^{19}(V)$, $r_{i2} \leftarrow V$;
   г) $V \leftarrow \rho_3^{17}(V)$, $r_{i3} \leftarrow V$.
5. $W \leftarrow K \oplus H$.
6. $W_0 \leftarrow \omega(W_0; T_0, T_1, T_2, T_3, T_4, T_5, T_6, T_7)$.
7. $W_1 \leftarrow \omega(W_1; T_4, T_1, T_0, T_3, T_6, T_5, T_2, T_7)$.
8. $W \leftarrow \xi^{31}(W)$.
9. $W \leftarrow \varphi(H_0, \varphi(H_0, W))$.
10. $W \leftarrow \varphi(K_0, \varphi(K_0, W))$.
11. $W \leftarrow \varphi(H_1, \varphi(H_1, W))$.
12. $W \leftarrow \varphi(K_1, \varphi(K_1, W))$.
13. $H \leftarrow W$.
14. $d \leftarrow d + 1$.
15. Если $d < n + 2$, то вернуться к шагу 2. Иначе — перейти к следующему шагу.
16. Вернуть $H \bmod 2^L$.

13.7. Задания

1. (п. 13.2). Количество значений блочно-итерационной хэш-функции h не больше количества образов шаговой функции хэширования $\varkappa$. Доказать, что имеется

$$\sum_{i=0}^{N}(-1)^i\binom{N}{i}(N-i)^R$$

различных сюръективных отображений $\mathcal{A}^m \times \mathcal{A}^n \to \mathcal{A}^n$, где $N = |\mathcal{A}|^n$, $R = |\mathcal{A}|^{n+m}$.

2. (п. 13.2, конструкция Дамгарда [77]). Пусть функция $g\colon \mathcal{A}^{n+m+1} \to \mathcal{A}^n$ строго свободна от коллизий. Доказать, что если слово $X^{(d+1)}$ формируется в соответствии с п. 2б, а итерации выполняются по правилу

$$Y^{(1)} = g(\alpha^{n+1} \,||\, X^{(1)}), \quad Y^{(i)} = g(Y^{(i-1)} \,||\, \beta \,||\, X^{(i)}), \quad i = 2, \ldots, d+1,$$

где α и β — различные символы алфавита $\mathcal{A}$, то h также строго свободна от коллизий.

3. (п. 13.2). Пусть функция $h_1 : \mathcal{A}^{2n} \to \mathcal{A}^n$ является строго свободной от коллизий. Доказать, что строго свободной от коллизий будет всякая функция

$$h_i : \mathcal{A}^{2^i n} \to \mathcal{A}^n, \quad X_1 \,||\, X_2 \mapsto h_1(h_{i-1}(X_1) \,||\, h_{i-1}(X_2)).$$

Здесь $X_1, X_2 \in \mathcal{A}^{2^{i-1}n}$, $i = 2, 3, \ldots$

4. (п. 13.4, алгоритм Флойда [14]). *Алгоритм Флойда* поиска совпадения элементов периодической (с минимальным периодом r и предпериодом t_0) последовательности $Y_0, Y_1, \ldots$ состоит в проверке равенств $Y_t = Y_{2t}$ для $t = 1, 2, \ldots$Доказать, что номер первого совпадения $t = r(1 + \lfloor t_0/r \rfloor)$.

5. (п. 13.4). Если h — блочно-итерационная функция и шаговая функция хэширования $\varkappa$ является обратимой (допускает вычисление Y по известным $\varkappa(X,Y)$ и X), то для решения задачи Н2 применяется атака *«встретимся посередине»* [71] — модификация атаки «дней рождения». Дать описание возможного варианта такой атаки.

6. (п. 13.5). Пусть в криптосистеме аутентификации H используется блочно-итерационная хэш-функция, $\Theta = \mathcal{A}^n$, ключом является стартовое хэш-значение $Y^{(0)}$. Показать, что такая система не является стойкой к атаке при известном открытом тексте в постановке А3 задачи криптоанализа.

7. (п. 13.6). Написать программу хэширования по СТБ 1176.1—99. Установить нулевое стартовое хэш-значение, $L = 256$.

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFVKBDFKJ
VJKGFKJSBDFJBVKJDSGFJK

Глава 14

ЭЛЕКТРОННАЯ ЦИФРОВАЯ ПОДПИСЬ

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFVKBDFKJ
VJKGFKJSBDFJBVKJDSGFJK

Имеются следующие две тесно взаимосвязанные проблемы:

1) сохранение информации от незаконного использования;

2) подтверждение авторства (принадлежность информации конкретному лицу).

Первая проблема решается с помощью шифрования. Она рассмотрена в предыдущих главах.

Проблема подтверждения авторства тесно связана с первой и возникает при следующих обстоятельствах:

1) когда некоторый абонент A получает сообщение X предположительно от B, как подтвердить, что оно получено именно от него, а не от какого-либо третьего лица (не фальсифицировано противником);

2) когда A получает от B какое-то сообщение X, как подтвердить, что оно не было изменено кем-либо посторонним.

При обмене информацией в компьютерных сетях для подтверждения авторства были разработаны алгоритмы *электронной цифровой подписи* (ЭЦП). В основе большинства из них лежит идея шифрования с открытым ключом (см. главу 12).

14.1. Обобщенная модель ЭЦП

В практической деятельности важно не только защищать информацию от злоумышленника, но и иметь возможность проверить авторство данного сообщения, и что оно не было изменено посторонним лицом. Именно для решения этих проблем (аутентификации и целостности) разработан ряд алгоритмов ЭЦП. В основе большинства из них лежит идея использования односторонней функции с секретом.

Суть этой идеи состоит в использовании некоторой односторонней функции с секретом F_S для создания пары (x,y), где x — сообщение, y — решение уравнения $F_S(y) = x$.

Всякая информация, записанная в некотором алфавите, может быть представлена в виде двоичных слов, т.е. конечных последовательностей из нулей и единиц. Количество двоичных цифр в таком слове будем называть его *длиной*. Пусть X и Y — некоторые подмножества множества всех двоичных слов.

Односторонней функцией с секретом S называется функция $F_S : Y \to X$, зависящая от параметра S и обладающая следующими тремя свойствами:

1) при любом S существует полиномиальный алгоритм вычисления значений $F_S(y)$;

2) при неизвестном S не существует полиномиального алгоритма для решения уравнения $F_S(y) = x$ относительно y;

3) при известном S существует полиномиальный алгоритм для решения уравнения $F_S(y) = x$ относительно y.

До настоящего времени не известно ни одного примера односторонней функции с секретом, но для практических целей используют некоторые функции, которые могут оказаться односторонними. Для них второе свойство строго не доказано, но известно, что задача инвертирования эквивалентна некоторой трудно решаемой математической задаче.

Пусть A и B — некоторые пользователи, обменивающиеся информацией по открытому каналу связи. Пусть X — совокупность всевозможных сообщений, Y — некоторое множество «подписей». Пусть $F_k : Y \to X$ — функция, зависящая от параметра $k \in \mathcal{K}$, называемого *ключом*. Будем считать, что ключ k состоит из двух частей k_S и k_O, где k_S — секретная составляющая, известная только A, и k_O — открытая составляющая, известная «всем» (не держится в секрете). Пусть F_k является сюръекцией, т.е. для любого $x \in X$ существует прообраз $y = F_k^{-1}(x)$. Функцию F также считаем общеизвестной.

Предположим, что выполняются следующие свойства:

1) зная k_O, функцию $F_k(y)$ можно вычислить по алгоритму полиномиальной сложности;

2) зная k_S, функцию $F_k(y)$ можно инвертировать по алгоритму полиномиальной сложности;

3) зная k_O, но не зная k_S, функцию $F_k(y)$ сложно инвертировать, т.е. неизвестен или не существует полиномиальный алгоритм нахождения $F_k^{-1}(x)$.

Прообраз $y = F_k^{-1}(x)$ некоторого сообщения x называется *подписью этого сообщения*. Пара (x, y) называется *подписанным сообщением*.

В силу первого свойства всегда легко проверить, соответствует ли подпись сообщению, а в силу третьего — подделать подпись при достаточно большом ключе практически невозможно. Доказательство этого свойства позволило бы придать подписанным сообщениям юридическую силу.

Секретный и открытый ключи находятся во взаимно-однозначном соответствии, и (в силу третьего требования) нет полиномиального алгоритма вычисления секретной компоненты по открытой. В общем виде алгоритм ЭЦП выглядит так.

1. Для передаваемого сообщения x отправитель A находит

$$y = F_k^{-1}(x).$$

Знание секретного ключа k_S позволяет ему сделать это за приемлемое время.

2. Далее A передает B по какому-либо каналу связи пару (x, y), где x — сообщение, y — подпись.

3. Получив подписанное сообщение (x, y), B находит $x' = F_k(y)$. Знание открытого ключа k_O позволяет сделать это за приемлемое время.

4. Получатель B сверяет x и x'. Если они совпадают, то полученное сообщение считаем подлинным. В противном случае либо сообщение x изменено (фальшивое), либо подпись y неверная (поддельная).

Замечание 14.1.1. Предложенную модель можно дополнить предварительным шифрованием пересылаемого сообщения и итоговой расшифровкой.

Замечание 14.1.2. Роль функции F_k иногда играет некоторая схема шифрования с открытым ключом. В силу этого многие вопросы (стойкость, выбор ключей и др.) равносильны для схем ЭЦП и соответствующих криптосистем.

14.2. Схема ЭЦП Рабина

Стойкость схемы Рабина основана на трудности решения квадратичных сравнений по большому составному модулю. Как было показано в предыдущей главе, данная задача эквивалентна факторизации.

Вот одна из реализаций этой схемы.

Сначала выберем два простых числа p и q, оба сравнимые с 3 по модулю 4. Это простые секретные ключи, а их произведение $N = pq$ — открытый ключ.

Для шифрования сообщения m (m должно быть меньше, чем N) вычисляем

$$c = m^2 (\bmod N).$$

Расшифровать сообщения просто, но все же немного сложнее. Так как получатель знает p и q, он может решить два сравнения, используя китайскую теорему об остатках. Вычислим

$$m_1 \equiv c^{\frac{p+1}{4}} \bmod p, \qquad m_2 \equiv \left(p - c^{\frac{p+1}{4}}\right) \bmod p,$$

$$m_3 \equiv c^{\frac{q+1}{4}} \bmod q, \qquad m_4 \equiv \left(q - c^{\frac{q+1}{4}}\right) \bmod q.$$

Затем выберем целое $a \equiv q(q^{-1} \bmod p)$, целое $b = p(p^{-1} \bmod q)$ и найдем четыре возможных решения сравнений

$$M_1 \equiv (am_1 + bm_3) \bmod N,$$

$$M_2 \equiv (am_2 + bm_4) \bmod N,$$

$$M_3 \equiv (am_2 + bm_3) \bmod N,$$

$$M_4 \equiv (am_2 + bm_4) \bmod N.$$

Один из четырех результатов M_1, M_2, M_3 или M_4 равен m. Если сообщение является английским текстом, то правильное сообщение M_j выбрать легко. С другой стороны, если сообщение — случайная последовательность цифр, то нет возможности определить корректное сообщение M_j. Одно из решений этой проблемы — добавить к сообщению известный текст перед шифрованием.

Покажем, как схема Рабина может быть использована для аутентификации.

В качестве функции шифрования для сообщения m возьмем функцию $m^2 \bmod N$, где N — открытый ключ A. Если мы хотим быть уверенными, что это сообщение послано именно A, то случайно выберем большое целое число l и пошлем A сообщение $r = l^2 (\bmod N)$. A расшифровывает его, используя знание факторизации N на простые числа p и q, находит квадратный корень l_1 из числа r и возвращает его нам. Мы поверим, что связаны именно с A тогда и только тогда, когда вернувшееся сообщение l_1 удовлетворяет сравнению $l_1^2 \equiv r (\bmod N)$.

14.3. Схема Диффи — Лампорта

Для аутентификации пользователя можно применять произвольную симметрическую схему. Пусть отправитель сообщения A хочет подписать n-битовое бинарное сообщение

$$m = m_1 \cdots m_n \in V_2^n.$$

Он выбирает $2n$ ключей некоторой криптосистемы, которые хранятся в секрете. Обозначим их так: $a_1, \ldots, a_n$; $b_1, \ldots, b_n$.

Если E — алгоритм шифрования, то A генерирует $4n$ параметров проверки $\{(X_i, Y_i, U_i, V_i) : 1 \le i \le n\}$, где X_i и Y_i — величины, подаваемые на вход E и связанные соотношениями

$$U_i = E(X_i, a_i), \quad V_i = E(Y_i, b_i), \quad 1 \le i \le n.$$

Параметры проверки заранее посылаются получателю B и третьему доверенному лицу.

Чтобы подписать n-битовое сообщение $m = (m_1, \ldots, m_n)$, пользователь A применяет следующую процедуру: его подпись будет цепочкой

$$S = S_1 \cdots S_n,$$

где для каждого i

$$S_i = \begin{cases} a_i, & \text{если } m_i = 0, \\ b_i, & \text{если } m_i = 1. \end{cases}$$

Пользователь B проверяет подпись следующим образом: для каждого i $(1 \le i \le n)$ он использует бит m_i и ключ S_i, чтобы проверить:

$$\begin{cases} \text{если } m_i = 0, & \text{то } E(X_i, S_i) = U_i, \\ \text{если } m_i = 1, & \text{то } E(Y_i, S_i) = V_i. \end{cases} \tag{14.3.1}$$

Пользователь B принимает подписанное сообщение за истинное только в том случае, если при процедуре проверки (14.3.1) выполнено для каждого i.

Эта система проста в использовании, но у нее есть, по крайней мере, два очевидных недостатка. Во-первых, необходима предварительная передача параметров проверки. Во-вторых, что более важно, подпись сильно увеличивает длину сообщения. Если в криптосистеме используются ключи длиной, скажем, 64 бита, то длина подписанного сообщения увеличивается в 64 раза.

14.4. Вероятностная схема подписи Рабина

В 1978 г. Рабин предложил следующий способ подписи. Пусть E — функция шифрования некоторой криптосистемы, $(k_i, 1 \le i \le 2r)$ — последовательность случайно выбранных ключей, которые отправитель A держит в секрете. Пользователь B получает список параметров проверки (X_i, U_i) $(1 \le i \le 2r)$, где

$$E(X_i, k_i) = U_i \quad (1 \le i \le 2r).$$

Эти параметры хранятся в доступном для всех месте.

Предположим, что A хочет подписать сообщение m.

Его подписью будет цепочка

$$S = S_1 S_2 \cdots S_{2r},$$

где для каждого i $(1 \le i \le 2r)$ $S_i = E(m, k_i)$. B делает следующее. Сначала он выбирает случайным образом r ключей, которые A должен ему представить: $k_{i_1}, k_{i_2}, \ldots k_{i_r}$. При получении этих ключей от A он проверяет:

$$E(m, k_{i_1}) = S_{i_1}, \quad E(X_{i_1}, k_{i_1}) = U_{i_1}$$

и далее для всех индексов $i_2, i_3, \ldots, i_r$. Он считает действительной подпись от A только в том случае, когда выдерживаются все проверки. Безопасность получателя зависит от его уверенности в том, что только отправитель, знающий секретный ключ, может послать так подписанное сообщение. Что касается A, то предположим, что он отказывается от сообщения, которое по утверждению B он подписал. Тогда протокол для A следующий: он должен предъявить судье свои секретные ключи $k_1, k_2, \ldots k_{2r}$, и в присутствии обеих сторон делается $2r$ проверок:

$$E(m, k_i) = S_i, \quad E(X_i, k_i) = U_i.$$

Рассмотрим, что это означает в трех возможных случаях.

1. Корректными являются менее r проверок. Тогда B не должен был принимать подписанное сообщение.

2. Выдерживается точно r проверок, т.е. при формировании ключей B выбрал именно это подмножество из r ключей. Вероятность, что он угадает это подмножество, равна

$$p_r = \binom{2r}{r}^{-1},$$

для $r = 18$ $p_r \approx 10^{-10}$.

3. Выдерживается $r + 1$ и более проверок: отказ абонента A не принимается.

14.5. Стандарт ЭЦП DSS

В 1991 г. NIST (National Institute of Standards and Technology) предложил для обсуждения проект стандарта ЭЦП DSS (Digital Signature Standard), использующий алгоритм DSA (Digital Signature Algorithm). Стойкость данного алгоритма основана на сложности решения задачи дискретного логарифмирования в мультипликативной группе простого поля F_p (см. главу 3).

Формирование подписи. Формирование подписи происходит по следующему алгоритму.

1. Отправитель A сообщения M предоставляет широкому кругу абонентов (получателей его сообщений) доступ к следующим параметрам:

p — простое число, $2^{512} < p < 2^{1024}$, битовая длина p кратна 64;

q — простое число, $2^{159} < p < 2^{160}$, и делитель $p - 1$;

$g = h^{\frac{p-1}{q}} \pmod p$, где h — такое целое число, что $0 < h < p$ и $h^{\frac{p-1}{q}} \pmod p > 1$;

y — открытый ключ, сформированный по правилу $y = a^x \pmod p$. Здесь x — секретный ключ, известный только A, причем $0 < x < q$;

$H(M)$ — хэш-функция, которая по исходному сообщению M формирует целое число в диапазоне от 1 до q (см. п. 14.6).

2. Пользователь A генерирует случайное число k такое, что $0 < k < q$, держит его в секрете и уничтожает сразу после получения подписи.

3. A находит два числа r и s по следующему правилу:

$$r = (g^k \pmod p) \pmod q,$$

$$s = k^{-1}(xr + H(M)) \pmod q.$$

Подписью к сообщению M является пара (r, s).

Проверка подписи. Пользователь B получает от A сообщение M' и подпись (r', s') к нему. B должен убедиться, что M совпадает с M'. Для этого:

1) если хотя бы одно из условий $0 < s' < q$, $0 < r' < q$ не выполняется, то подпись считается недействительной;

2) B находит $v = (s')^{-1} \bmod q$;

3) B вычисляет

$$z_1 = H(M')v \pmod q, \quad z_2 = r'v \pmod q;$$

4) далее вычисляется

$$u = (g^{z_1} y^{z_2} (\bmod p)) (\bmod q);$$

5) B проверяет условие $r' = u$. Если оно выполняется, то подпись считается подлинной и сообщение не измененным, т.е. $M' = M$.

Корректность алгоритма DSA. Пусть $M = M'$, $s = s'$, $r = r'$. Покажем, что тогда $u = r$.

Итак, $v = s^{-1} (\bmod q)$, $z_1 = H(M) v (\bmod q)$, $z_2 = rv (\bmod q)$. Имеем

$$u = g^{z_1} y^{z_2} (\bmod p) (\bmod q) = g^{H(M)s^{-1}} g^{xrs^{-1}} (\bmod p) (\bmod q) =$$

$$= g^{k(xr+H(M))^{-1}(xr+H(M))} (\bmod p) (\bmod q) = g^k (\bmod p) (\bmod q) = r.$$

Таким образом, $u = r$, и корректность алгоритма доказана.

Для нахождения секретных параметров ЭЦП по открытым необходимо решить следующую систему сравнений:

$$\begin{cases} y \equiv a^x (\bmod p), \\ g^k + pn \equiv r' (\bmod p), \\ s' \equiv k^{-1}(xr + H(M')) (\bmod p), \end{cases}$$

где неизвестными являются x, n, k.

В работе [170] имеются некоторые замечания по стойкости алгоритма DSA.

1. В алгоритме формирования подписи есть недостаток: в редких случаях, когда $s = 0$, при проверке подписи будет сбой, так как в этом случае не существует s^{-1}. Эта ошибка легко устраняется при помощи дополнительной проверки, что и сделано в российском стандарте ЭЦП.

2. Алгоритм DSA медленный. В то время как скорость получения подписи сравнима со скоростью шифрования по схеме RSA, проверка подписи в большом количестве случаев примерно в 100 раз медленнее, чем RSA.

3. Тот факт, что один модуль p используется многими пользователями, ослабляет стойкость алгоритма, поскольку единственный «взлом» p нарушает безопасность сразу всех абонентов, пользующихся этим p. Под *взломом* понимается некое предвычисление, которое позволяет в дальнейшем легко решать проблему дискретного логарифмирования для данного p.

4. Величина 512 битов для p слишком мала. С учетом тенденции уменьшения стоимости вычислений стоимость взлома через несколько лет может сократиться до разумной величины, что для стандарта неприемлемо.

5. Существует целый класс простых чисел, для которых проблема дискретного логарифмирования решается легко. Причем построить такие числа также легко, однако затраты на проверку, является ли данное простое «слабым», превышают возможности среднего пользователя. Это значит, что тот, кто распределяет простые p, в принципе может знать секретные ключи своих клиентов.

6. Анализ алгоритма DSA показывает, что в данном случае проблема взлома подписи, вообще говоря, не сводится к проблеме дискретного логарифмирования, поскольку в алгоритме DSA g — не первообразный корень по модулю p, а лишь элемент порядка q, что намного меньше $p-1$, Таким образом, вполне возможно, что проблема взлома алгоритма ЭЦП легче общей проблемы дискретного логарифмирования.

14.6. Схема ЭЦП Эль-Гамаля

Пусть p — большое простое число, q — большой простой делитель $p-1$, $g \in (\mathbb{Z}/p\mathbb{Z})^*$ имеет порядок q. Пользователь A выбирает секретный ключ x $(1 < x < q)$ и находит $y = g^x (\mathrm{mod}\, p)$ — открытый ключ. Предположим, что он отправляет сообщение m $(1 \le m \le q)$. Подписью к этому сообщению будет пара (r, s), удовлетворяющая

$$g^m \equiv y^r r^s (\mathrm{mod}\, p) \tag{14.6.1}$$

с условиями $0 < r < p$ и $0 < s < q$. Она строится следующим образом. Выбираем случайное $k \in [1, q-1]$ и вычисляем

$$r = g^k \bmod p, \quad s = k^{-1}(m - xr) \bmod q.$$

Действительно, поскольку порядок g равен q, полученные r и s удовлетворяют (14.6.1):

$$y^r r^s \equiv g^{xr} g^{kk^{-1}(m-xr)} \equiv g^{xr+m-xr} \equiv g^m (\mathrm{mod}\, p).$$

Покажем, что при каждом k существует единственное $s < q$, удовлетворяющее (14.6.1). Из (14.6.1) имеем

$$g^m \equiv g^{xr} g^{ks} (\mathrm{mod}\, p) \Rightarrow g^{m-xr-ks} \equiv 1 (\mathrm{mod}\, p) \Rightarrow$$

$$\Rightarrow m - xr - ks \equiv 0 (\mathrm{mod}\, q) \Rightarrow m - xr \equiv ks (\mathrm{mod}\, q) \Rightarrow$$

$$\Rightarrow s \equiv k^{-1}(m - xr) (\mathrm{mod}\, q).$$

Если выбрано k, то для выработки подписи (без знания секретного ключа x) необходимо найти такое $s < q$, что верно (14.6.1) или

$$r^s \equiv b (\mathrm{mod}\, p), \tag{14.6.2}$$

где $b = g^{-r} \bmod p$. Так как числа k и q взаимно простые, то число $r = g^k$ имеет тот же порядок, что и g, т.е. q. Получается, что для взлома системы таким способом необходимо решить задачу дискретного логарифмирования в подгруппе G, образованной g и имеющей большой порядок q.

14.7. Арифметические свойства российского стандарта цифровой подписи

Механизм цифровой подписи. Задаются следующие параметры, используемые алгоритмом: p — простое число, $2^{509} < p < 2^{512}$ либо $2^{1020} < p < 2^{1024}$; q — простое число, $2^{254} < q < 2^{256}$ такое, что $q \mid p-1$; a — целое число в пределах $1 < a < p-1$ такое, что $a^q \bmod p = 1$; x — целое число в пределах $0 < x < q$; y — целое число, равное $a^x \bmod p$; M — целое число в пределах $0 < M < q$.

Число x называют *секретным ключом пользователя*, y — *открытым ключом пользователя*, M — *сообщением*. В соответствии с алгоритмом проверки подписи в ГОСТ Р 34.10—94 электронную подпись можно ввести следующим образом.

Пусть дано сообщение M. *Подписью* к M называется пара целых чисел (r, s) таких, что

$$1 < r < q, \quad 0 < s < q, \tag{14.7.1}$$

$$r = \left(a^{sM^{q-2} \bmod q} y^{-rM^{q-2} \bmod q} \bmod p \right) \bmod q. \tag{14.7.2}$$

Теорема 14.7.1. Пусть задано целое $M (0 < M < q)$. Тогда существует ровно q различных решений $(r_k, s_k)_{k=0}^{q-1}$ уравнения (14.7.2), причем

$$r_k = (a^k \bmod p) \bmod q, \tag{14.7.3}$$

$$s_k = (xr_k + kM) \bmod q. \tag{14.7.4}$$

Доказательство. Легко проверить путем подстановки, что все пары (r_k, s_k), заданные по формулам (14.7.3) и (14.7.4), удовлетворяют равенству (14.7.2). Все эти пары различны. Действительно, если $(r_i, s_i) = (r_j, s_j)$ при некоторых $i \neq j$, то, используя (14.7.4), получаем $xr_i + iM \equiv xr_j + jM (\bmod q)$, откуда следует: $(i-j)M \equiv 0 (\bmod q)$, что вместе с $(M, q) = 1$ дает $i \equiv j (\bmod q)$. Последнее сравнение возможно, только если $i = j$. Таким образом, чтобы доказать утверждение, достаточно показать, что сравнение (14.7.2) имеет не более q решений.

Будем рассматривать циклическую подгруппу группы $\mathbb{Z}_p^*$:

$$\langle a \rangle_p = \{a^k \bmod p : k = 0, \ldots, q-1\},$$

порожденную элементом a. Для $m \in \{0, \ldots, q-1\}$ определим функцию $\Psi(m)$ как число различных элементов группы $\langle a \rangle_p$, которые по модулю q дают остаток m. Очевидно, равенство $\sum_{m=0}^{q-1} \Psi(m) = q$. Пусть Φ — количество решений (14.7.2), $\Phi(r)$ — количество различных s $(0 \le s < q)$ таких, что пара (r, s) есть решение (14.7.2). Тогда $\Phi = \sum_{r=0}^{q-1} \Phi(r)$.

Зафиксируем r. Поскольку $y = a^x \bmod p$, то имеем равенство

$$a^{sM^{q-2} \bmod q} y^{-rM^{q-2} \bmod q} \bmod p = a^{(s-rx)M^{q-2} \bmod q} \bmod p.$$

Когда s пробегает значения от 0 до $q-1$, то $\nu = (s - rx)M^{q-2}$ также пробегает (в другом порядке) эти значения. Тогда $a^{\nu} \bmod p$ пробегает все элементы группы $\langle a \rangle_p$, когда s изменяется от 0 до $q-1$. Количество различных s таких, что

$$r = \left(a^{sM^{q-2} \bmod q} y^{-rM^{q-2} \bmod q} \bmod p \right) \bmod q$$

есть $\Phi(r)$, с одной стороны, и $\Psi(r)$ — с другой. Тогда

$$\Phi = \sum_{r=0}^{q-1} \Phi(r) = \sum_{r=0}^{q-1} \Psi(r) = q.$$

Итак, количество различных решений сравнения (14.7.2) равно q, и все эти решения описаны по формулам (14.7.3) и (14.7.4). □

Введем множества

$$R(M) = \{(r_k, s_k) : r_k = (a^k \bmod p) \bmod q,$$

$$s_k = (xr_k + kM) \bmod q, \ 0 \le k \le q-1\},$$

$$S(M) = \{(r,s) \in R(M) : (r-1)rs \ne 0\}.$$

Тогда множество $R(M)$ состоит из всех решений уравнения (14.7.2), а $S(M)$ — из всех подписей к M.

Чтобы сформировать подпись к сообщению M, необходимо решить сравнение (14.7.2) относительно (r,s). Такую пару можно вычислить по формулам (14.7.3) и (14.7.4). Но для этого необходимо знать значение секретного ключа x и параметра k. Таким образом, получаем следующие алгоритмы.

Алгоритм генерации подписи

1. Генерируем случайное k в интервале $[0, q-1]$ (его значение держится в секрете).
2. Находим r_k и s_k по формулам (14.7.3) и (14.7.4) (их вычисление осуществляется по алгоритмам полиномиальной сложности).
3. Значение числа k уничтожается.
4. Если $r_k s_k \ne 0$, то полученная пара (r_k, s_k) является подписью к M. В противном случае переходим к шагу 1.

Алгоритм проверки подписи

1. Если $r < 0$ или $r > q-1$, или $s < 0$, или $s > q-1$, то подпись недействительная, иначе переходим к следующему шагу.
2. Вычисляем $z_0 = M^{q-2} \bmod q$.

3. Вычисляем $z_1 = sz_0 \bmod q$.
4. Вычисляем $z_2 = -rz_0 \bmod q$.
5. Вычисляем $z_3 = a^{z_1} \bmod p$.
6. Вычисляем $z_4 = y^{z_2} \bmod p$.
7. Вычисляем $z_5 = z_3 z_4 \bmod p$.
8. Вычисляем $z_6 = z_5 \bmod q$.
9. Если $r \neq z_6$, то подпись недействительная, иначе подпись действительная.

Эквивалентное преобразование схемы ЭЦП. Рассмотрим уравнение

$$\bar{r} = \left(a^{\bar{s}M^{q-2} \bmod q} y^{-\bar{r}M^{q-2} \bmod q} \right) \bmod p \tag{14.7.5}$$

в целых $0 < \bar{r} < p$, $0 \le \bar{s} < q$.

Теорема 14.7.2. Пусть задано целое $M (0 < M < q)$. Тогда существует ровно q различных решений $(\bar{r}_k, \bar{s}_k)_{k=0}^{q-1}$ уравнения (14.7.5), причем

$$\bar{r}_k = (a^k \bmod p), \tag{14.7.6}$$

$$\bar{s}_k = (x\bar{r}_k + kM) \bmod q. \tag{14.7.7}$$

Доказательство. Из уравнения (14.7.5) очевидно, что $\bar{r}$ имеет вид (14.7.6) с некоторым целым k $(0 \le k < q)$, поскольку $y = a^x \bmod p$. Тогда имеем

$$a^k \equiv a^{\bar{s}M^{q-2} - \bar{r}xM^{q-2}} (\bmod p),$$

что выполнено тогда и только тогда, когда $k \equiv \bar{s}M^{q-2} - \bar{r}xM^{q-2} (\bmod q)$, откуда получаем (14.7.7). Итак, все решения уравнения (14.7.5) определяются по формулам (14.7.6) и (14.7.7). Все значения r_k различны и, следовательно, все решения различны. □

Определим $\bar{R}(M)$ — множество решений (14.7.5) и

$$\bar{S}(M) = \{(\bar{r}, \bar{s}) \in \bar{R}(M) : (r-1)rs \not\equiv 0 (\bmod q)\}.$$

Теорема 14.7.3. Пусть задано целое $M (0 < M < q)$. Тогда для всех k от 0 до $q-1$ справедливы следующие утверждения:

1) $r_k = \bar{r}_k \bmod q$;
2) $s_k = \bar{s}_k$;
3) $\bar{r}_k = \left(a^{s_k M^{q-2} \bmod q} y^{-r_k M^{q-2} \bmod q} \right) \bmod p$;
4) $(r_k, s_k) \in S(M) \Leftrightarrow (\bar{r}_k, \bar{s}_k) \in \bar{S}(M)$.

Доказательство. Первое равенство очевидно в силу определения r_k. Тогда имеем $s_k = (xr_k + kM) \bmod q = (x(\bar{r}_k \bmod q) + kM) \bmod q = (x\bar{r}_k + kM) \bmod q = \bar{s}_k$. Из того, что y имеет порядок q и $\bar{r}_k \equiv r_k (\bmod q)$, вытекает:

$$y^{-\bar{r}_k M^{q-2} \bmod q} = y^{-r_k M^{q-2} \bmod q}.$$

Подставляя вместо выражения из левой части выражение из правой части и s_k вместо $\bar{s}_k$ в (14.7.5), получаем третью формулу. Четвертое утверждение вытекает из первых двух. □

Теорема 14.7.4. Пусть задано целое M $(0 < M < q)$. Определим отображение $F_M : \bar{R}(M) \to R(M)$ по правилу $F_M(\bar{r}, \bar{s}) = (\bar{r} \bmod q, \bar{s})$. Тогда:

1) F_M устанавливает взаимно-однозначное соответствие между $\bar{R}(M)$ и $R(M)$;

2) обратное отображение $(\bar{r}, \bar{s}) = F_M^{-1}(r, s)$ имеет вид

$$\bar{s} = s, \quad \bar{r} = \left(a^{sM^{q-2} \bmod q} y^{-rM^{q-2} \bmod q} \right) \bmod p;$$

3) сужение F_M на $\bar{S}(M)$ устанавливает взаимно-однозначное соответствие между $\bar{S}(M)$ и $S(M)$.

Доказательство. Это утверждение вытекает из предыдущего. □

Вывод. В силу теоремы 14.7.3 задача нахождения решений уравнения (14.7.5) эквивалентна задаче нахождения решений уравненения (14.7.2) в том смысле, что по решению одной из этих задач (по алгоритмам полиномиальной сложности) легко строится решение другой.

Дискретное логарифмирование в схеме ЭЦП. Достаточным условием для формирования подписи к произвольным сообщениям является знание x, для нахождения которого достаточно решить задачу дискретного логарифмирования.

Выясним, к каким проблемам приводит задача построения подписи. В силу доказанной эквивалентности задач решения уравнений (14.7.2) и (14.7.5) будем исследовать задачу решения уравнения (14.7.5), которое можно переписать в виде

$$r^M y^r \equiv a^s (\bmod p). \tag{14.7.8}$$

Итак, чтобы сформировать подпись, необходимо решить уравнение (14.7.8). Множество решений этого уравнения есть однопараметрическое семейство пар. Если зафиксировать параметр r, то приходим к уравнению

$$a^s \equiv b (\bmod p),$$

где $b = r^M y^r \bmod p$, s — неизвестное. Его решение — задача дискретного логарифмирования.

Зададим s_0 и предположим, что уравнение (14.7.8) имеет решение (r,s) такое, что $s = s_0$. Тогда, чтобы найти r, необходимо решить уравнение $r^M y^r \equiv \equiv a^{s_0} \bmod p$ или $d^r \equiv br \pmod p$, где $b = a^{-s_0 M^{q-2} \bmod q} \bmod p$, $d = y^{-M^{q-2} \bmod q} \bmod p$. Алгоритмов для решения этой задачи, кроме почти полного перебора всех вариантов, на настоящее время не известно. Эту задачу можно интерпретировать иначе: необходимо найти такое r, что $rd_1^r = b_1$, где $d_1 = d^{-1} \bmod p$ и $b_1 = b^{-1} \bmod p$.

Совпадение ЭЦП для различных сообщений

Теорема 14.7.5. Пусть даны $M_1, M_2 \in \mathbb{Z}$ такие, что $0 < M_1, M_2 < q$, $M_1 \neq M_2$. Тогда

$$\overline{R}(M_1) \cap \overline{R}(M_2) = \{(1, x)\}, \tag{14.7.9}$$

$$\overline{S}(M_1) \cap \overline{S}(M_2) = \varnothing. \tag{14.7.10}$$

Доказательство. Пусть $(r,s) \in \overline{R}(M_1) \cap \overline{R}(M_2)$. Тогда $r = a^k \bmod p$ и $xr + kM_1 \equiv xr + kM_2 \pmod q$, откуда $k(M_1 - M_2) \equiv 0 \bmod q$, что возможно только при $k = 0$. Следовательно, $r = 1$ и $s = (x \cdot 1 + 0 \cdot M_1) \bmod q = x$. Соотношение (14.7.10) вытекает из (14.7.9). □

Замечание 14.7.1. Хотя $\overline{S}(M_1)$ и $\overline{S}(M_2)$ не пересекаются при различных M_1 и M_2, $S(M_1)$ и $S(M_2)$ могут пересекаться. Действительно, если $r_k = r_l$ при некоторых различных k и l, то пара сообщений $M_1 = (s - xr_k)k^{-1} \bmod q$ и $M_2 = (s - xr_l)l^{-1} \bmod q$, где $s \in \{1, \ldots, q-1\}$ — произвольно, будет иметь одинаковую подпись, причем $M_1 \neq M_2$. Это, очевидно, связано с тем, что r_k может принимать одинаковые значения при различных k.

Чтобы оценить вероятность совпадения подписей, для $m \in \{0, \ldots, q-1\}$ введем случайную величину η_m, равную количеству различных k $(0 \le k < q)$, для которых $m = (a^k \bmod p) \bmod q$. Также рассмотрим P_i, равные количеству различных m таких, что $\eta_m = i$. Введем вспомогательную функцию $\chi_i(j) = \delta_{ij}$, $(i,j) \in \{0, \ldots, q-1\}^2$. Легко заметить, что $\sum_{i=0}^{q-1} \chi_i(j) = 1$. Тогда величину η_m можно задать следующим образом:

$$\eta_m = \sum_{k=0}^{q-1} \chi_m(r_k),$$

где $r_k = (a^k \bmod p) \bmod q$. Найдем среднее значение η:

$$\sum_{m=0}^{q-1} \eta_m = \sum_{m=0}^{q-1} \sum_{k=0}^{q-1} \chi_m(r_k) = \sum_{k=0}^{q-1} \sum_{m=0}^{q-1} \chi_m(r_k) = \sum_{k=0}^{q-1} 1 = q, \tag{14.7.11}$$

откуда вытекает, что математическое ожидание $\mathbf{E}\{\eta\} = 1$.

Далее, P_i/q — вероятность события $\eta = i$, поэтому имеет место

$$\sum_{i=0}^{q-1} P_i = q. \tag{14.7.12}$$

Поскольку математическое ожидание η равно $\sum_{i=0}^{q-1} iP_i/q = 1$, то

$$\sum_{i=0}^{q-1} iP_i = q, \tag{14.7.13}$$

откуда $\sum_{i=0}^{q-1} P_i = \sum_{i=0}^{q-1} iP_i$, что эквивалентно $P_0 = \sum_{i=2}^{q-1}(i-1)P_i$. Пусть w — количество различных значений случайной величины η. В этом случае η принимает значение i тогда и только тогда, когда $P_i \neq 0$. Пусть $i_1, \ldots, i_w$ — все различные значения η, причем $i_j < i_{j+1}$. Тогда, используя (14.7.13), получаем

$$q = \sum_{j=1}^{w} i_j P_{i_j} \geq \sum_{j=1}^{w} i_j \geq \sum_{i=0}^{w-1} i = \frac{w(w-1)}{2} > \frac{(w-1)^2}{2}.$$

Следовательно,

$$w \leq [\sqrt{2q}] + 1. \tag{14.7.14}$$

Пример. Пусть при $p = 22643563$, $q = 1109$ имеем следующее распределение η_m: $P_0 = 425$, $P_1 = 382$, $P_2 = 202$, $P_3 = 80$, $P_4 = 17$, $P_5 = 3$ и $P_i = 0$ при $i \geq 5$, или:

η	0	1	2	3	4	5	> 5
P	$\frac{425}{1109}$	$\frac{382}{1109}$	$\frac{202}{1109}$	$\frac{80}{1109}$	$\frac{17}{1109}$	$\frac{3}{1109}$	0

Наиболее вероятно естественное предположение о том, что η принимает значения $0, 1, \ldots, w-1$. Но в этом случае получаем, что $P_i = 0$ при $i \geq w$. Значит, среди значений r_k одно и то же не может встретиться более w раз. Скорее всего, вероятности P_i должны быстро убывать (как показывает пример). В этом случае на w можно получить гораздо более сильные оценки. Это будет свидетельствовать о том, что одни и те же значения r_k встречаются очень редко и r_k распределена в некотором смысле равномерно.

Замечание 14.7.2. В качестве подписи к некоторому сообщению M можно было бы принять пару $(\bar{r}, \bar{s}) \in \bar{S}(M)$, чтобы избежать опасности совпадения подписей на разных сообщениях или опасности искажения сообщения с сохранением подписи. Но опыт показывает, что подобные события маловероятны и практически не влияют на стойкость. Кроме того, приведение по двойному модулю обладает неоспоримым преимуществом: построенная таким образом подпись на 30 % короче ее расширенного аналога $(\bar{r}, \bar{s})$ и тем самым существенно экономит объем пересылаемых информационных сообщений. Чтобы дать

строгое математическое обоснование, необходимо изучить распределение значений последовательности $(a^j \bmod q) \bmod p$. В следующем параграфе проводится предварительный анализ этого распределения.

14.8. Эквивалентность задач фальсификации подписи в DSS и схеме Эль-Гамаля

Теорема 14.8.1. Существование эффективного алгоритма подписи по схеме DSS эквивалентно существованию эффективного алгоритма подписи по схеме Эль-Гамаля.

Доказательство. Считаем, что параметры p, q и a в обеих схемах одинаковы. Пусть существует эффективный алгоритм подписи по схеме DSS и задан открытый ключ y_2 для схемы Эль-Гамаля. Положим $y_1 = y_2^{-1} \bmod p$ — открытый ключ для схемы DSS. По гипотетическому алгоритму находим подпись $(\tilde{r}, \tilde{s})$ заданного сообщения M. Подпись для схемы Эль-Гамаля построим по формулам

$$\begin{cases} s = \tilde{s}, \\ r = a^{M\tilde{s}^{q-2} \bmod q} y_1^{\tilde{r}\tilde{s}^{q-2} \bmod q} \bmod p. \end{cases}$$

Действительно, полученная пара удовлетворяет условиям определения.

Обратно, пусть существует эффективный алгоритм подписи по схеме Эль-Гамаля и задан открытый ключ y_1 для схемы DSS. Положим $y_2 = y_1^{-1} \bmod p$ — открытый ключ для схемы Эль-Гамаля. По гипотетическому алгоритму находим подпись $(\tilde{r}, \tilde{s})$ заданного сообщения M. Подпись для схемы DSS построим по формулам

$$\begin{cases} s = \tilde{s}, \\ r = \tilde{r} \bmod q. \end{cases}$$

Полученная пара удовлетворяет определению. □

Замечание 14.8.1. Теорема 14.8.1 аналогична третьему пункту теоремы 14.7.3, который устанавливает эквивалентность схемы цифровой подписи в стандарте РФ, и аналогична схеме Эль-Гамаля.

В силу эквивалентности DSS и Эль-Гамаля достаточно изучить проблему взлома последней. Существует аналогия со стандартом РФ (см. п. 14.7): чтобы найти корректную подпись (r, s) при заданном $r = r_0$, необходимо решить сравнение вида $r^s \equiv b (\bmod p)$, где $b = a^M y^{-r} \bmod p$ и s — неизвестное. Его решение есть задача дискретного логарифмирования. В свою очередь, чтобы корректную подпись (r, s) при заданном $s = s_0 \neq 0$, необходимо решить сравнение вида

$$rd^r \equiv b (\bmod p),$$

где $b = a^{Ms^{q-2} \bmod q} \bmod p$, $d = y^{s^{q-2} \bmod q} \bmod p$.

Замечание 14.8.2. Одной из возможных задач при взломе какой-либо схемы цифровой подписи может быть задача нахождения какой-нибудь корректной пары (M,S), где M — сообщение, S — подпись. С этой точки зрения все описанные схемы (российский стандарт, DSS, алгоритм Эль-Гамаля) эквивалентны. Действительно, в случае российского стандарта необходимо и достаточно решить уравнение (14.7.8), а в случае DSS и Эль-Гамаля — уравнение (14.6.1). Но очевидно, что эти уравнения переходят друг в друга путем перестановки переменных m и s.

Замечание 14.8.3. При оценке стойкости любой криптосистемы встает вопрос о влиянии количества известных корректных пар (M,S), где M — сообщение, S — подпись, на стойкость криптографического алгоритма. Алгоритмы, построенные на операции возведения в степень в кольце $\mathbb{Z}/N$ (типа RSA), позволяют найти любое количество таких пар. Для этого достаточно знать одну пару и преобразовывать ее, используя мультипликативную структуру операции возведения в степень, или по заданному S вычислять M. Такой прием в используемых схемах, основанных на операции экспонирования, сталкивается с непреодолимой трудностью — необходимо решать задачу дискретного логарифмирования.

14.9. Электронная цифровая подпись СТБ 1176.2—99

Введенный в 1999 г. стандарт Республики Беларусь СТБ 1176.2 «Информационная технология. Защита информации. Процедуры выработки и проверки электронной цифровой подписи» базируется на *схеме ЭЦП Шнорра* [179], которая, в свою очередь, является модификацией схемы Эль-Гамаля.

Параметры. При выработке и проверке подписи используются l-битовое простое число p и r-битовое простое число q, $q \mid (p-1)$. Допустимые значения параметров l и r указаны в табл. 14.1. Применяется определяемая СТБ 1176.1 функция хэширования h, параметр L которой устанавливается равным $(r-1)$ (см. п. 13.6).

Таблица 14.1

Допустимые значения параметров r и l

Уровень стойкости	r	l	Уровень стойкости	r	l
1	143	638	6	208	1534
2	154	766	7	222	1790
3	175	1022	8	235	2046
4	182	1118	9	249	2334
5	195	1310	10	257	2462

Часть преобразований стандарта выполняется в группе G, определяемой множеством $B_p = \{1, 2, \ldots, p-1\}$ и операцией $\circ$: $u \circ v = uvR^{-1} \bmod p$, где $R = 2^{l+2}$. Использование операции $\circ$ вместо обычного умножения по модулю p упрощает применение алгоритма Монтгомери (см. п. 12.11). Далее, $u^{(m)}$ — m-я степень числа $u \in B_p$ как элемента G.

В стандарте приведены алгоритмы генерации чисел p, q и элемента $a \in B_p$, имеющего порядок q в группе G. Числа p, q, a являются долговременными параметрами СТБ 1176.2—99, едиными для группы пользователей.

Входные данные. Всякое сообщение M, подпись к которому вырабатывается или проверяется, задается последовательностью байтов. Если t — n-разрядное число по основанию $2^8 = 256$, т.е.

$$t = \sum_{i=0}^{n-1} t_i (256)^i, \quad 0 \le t_i < 256, \quad t_{n-1} \ne 0,$$

то $t \,||\, M$ — сообщение, полученное вставкой байтов $t_0, t_1, \ldots, t_{n-1}$ в начало M.

Выработка подписи. При выработке подписи S к сообщению M используется личный ключ x, $0 < x < q$. Выполняются следующие шаги.

1. Выработать случайное секретное число k, $1 < k < q$.
2. $t \leftarrow a^{(k)}$.
3. $U \leftarrow h(t \,||\, M)$. Если $U = 0$, то вернуться к шагу 1.
4. $V \leftarrow (k - xU) \bmod q$. Если $V = 0$, то вернуться к шагу 1.
5. $S \leftarrow U2^r + V$.

Проверка подписи. При проверке подписи S к сообщению M используется открытый ключ $y = a^{(x)}$. Алгоритм проверки подписи состоит из следующих шагов.

1. $V \leftarrow S \bmod 2^r$.
2. $U \leftarrow (S - V)/2^r$.
3. Проверить условия $0 < U < 2^{r-1}$, $0 < V < q$. Если хотя бы одно из условий нарушается, то подпись признается недействительной и выполнение алгоритма завершается.
4. $t \leftarrow a^{(V)} \circ y^{(U)}$.
5. $W \leftarrow h(t \,||\, M)$.
6. Если $W \ne U$, то подпись признается недействительной.

   Если $W = U$, то принимаются решения о том, что:

   а) подпись S была создана с помощью личного ключа x, связанного с открытым ключом y;

   б) подпись S и сообщение M не были изменены с момента их создания.

14.10. Задача дискретного логарифмирования

Пусть G — конечная мультипликативная группа. *Задача дискретного логарифмирования* состоит в решении при заданных $a, b \in G$ уравнения

$$a^x = b$$

относительно целого x, $0 \le x < n = \operatorname{ord} a$. Решение x называется *дискретным логарифмом*, или *индексом*, b по основанию a и обозначается $\operatorname{ind}_a b$.

Пусть $m = \lceil \sqrt{n} \rceil$. *Метод больших-малых шагов* состоит в нахождении совпадения элемента последовательности

$$1, a, a^2, \ldots, a^{m-1}$$

с элементом последовательности

$$b, ba^{-m}, ba^{-2m}, \ldots, ba^{-(m-1)m}.$$

Для вычисления последовательностей требуется $O(m + \log_2 m)$ умножений на элементы G. Если найдено совпадение $a^j = ba^{-im}$, то $a^{im+j} = b$ и $\operatorname{ind}_a b = (im + j) \bmod n$.

Пусть $n = qr$, q — простое, $r > 1$. *Метод Нечаева — Полига — Хеллмана* [24, 157] решения уравнения $a^x = b$ состоит в следующем.

1. По методу больших-малых шагов находится решение x_1, $0 \le x_1 < q = \operatorname{ord} a^r$, уравнения

$$(a^r)^{x_1} = b^r.$$

2. Находится решение x_2, $0 \le x_2 < r = \operatorname{ord} a^q$, уравнения

$$(a^q)^{x_2} = ba^{-x_1}.$$

3. Определяется решение $x = x_2 q + x_1$.

Если r — простое, то уравнение на шаге 2 решается по методу больших-малых шагов. При составном r уравнение снова заменяется двумя уравнениями и т.д. Если известно разложение $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ порядка элемента a на простые множители, то для вычисления $\operatorname{ind}_a b$ требуется

$$O\left(\log n + \alpha_i \sum_{i=1}^{k} \sqrt{p_i}\right)$$

операций умножения на элементы G.

Для нахождения дискретного логарифма в поле $\mathbb{F}_q$, q — простое число, по *индекс-методу* [40] поступают следующим образом.

1. Выбирают базу множителей $S_B = \{p_1, \ldots, p_s\}$ (см. п. 12.14).

2. Находят T пар натуральных чисел (b_t, m_t) таких, что $1 \le m_t \le q-2$, $b_t = a^{m_t} \bmod q$ является B-гладким, т.е.

$$b_t = p_1^{\alpha_{t1}} \cdots p_s^{\alpha_{ts}}.$$

3. Составляется система из T уравнений

$$\alpha_{t1} x_1 + \ldots + \alpha_{ts} x_s \equiv m_t (\mathrm{mod}(q-1)), \quad t = 1, \ldots, T,$$

относительно неизвестных $x_i = \mathrm{ind}_a p_i$, $i = 1, \ldots, s$. На шаге 2 находят такое количество пар T, что решение полученной системы определено однозначно.

4. Находят пару натуральных чисел (m', b') таких, что $1 \le m' \le q-2$, $b' = ba^{m'} \bmod q$ является B-гладким, т.е.

$$b' = p_1^{\beta_1} \cdots p_s^{\beta_s}.$$

5. Искомый логарифм

$$x = (\beta_1 x_1 + \ldots + \beta_s x_s - m') \bmod (q-1).$$

Для нахождения дискретного логарифма по индекс-методу требуется затратить $L_{1/2,c}(q)$ операций, где c — положительная константа. Наиболее эффективная на сегодняшний день модификация метода — метод решета числового поля — позволяет вычислить дискретный логарифм за время $L_{1/3,1.923}(q)$.

14.11. Задания

1. В системе аутентификации, основанной на схеме RSA, пользователь A выбрал открытый ключ $e = 7$ и $n = 77$. Если он получил от B число 23, то что A должен ответить, чтобы аутентифицировать себя?

2. В системе аутентификации, основанной на схеме Рабина, A выбирает в качестве открытых ключей $B = 2$, $N = 200$. C посылает A число $R = 168$. Как должен ответить A, чтобы убедить C, что именно ему попало сообщение?

3. Используя схему подписи Даффи — Лампорта, основанную на шифровании

$$e(x, k) = x^k (\mathrm{mod}\, N),$$

A выбирает четыре ключа: $a_1 = 2$, $a_2 = 5$, $b_1 = 7$, $b_2 = 3$.
   а) Если $N = 13$, то какие утвержденные параметры A посылает B?
   б) Как бы A подписал сообщение $M = 10$?

4. В той же схеме B получает сообщение $m_1 m_2 73$. Если оно подлинное, то что представляют собой m_1 и m_2?

5. В схеме подписи, основанной на RSA, пользователи A и B имеют открытые ключи $e_A = 3$, $n_A = 15$; $e_B = 7$, $n_B = 77$ соответственно. A хочет послать сооб-

щение $M = 4$ как подпись к некоторому тексту. Какое целое число он посылает?

6. Доказать, что в схеме подписи, основанной на RSA, при подписи S_1 сообщения m_1 и подписи S_2 сообщения m_2, сообщение m_1m_2 можно подписать S_1S_2.

7. В схеме подписи Рабина (2) A имеет открытый ключ N_A и желает подписать сообщения $m_1 = 9$ и $m_2 = 29$. Какими будут соответствующие подписи S_1 и S_2?

8. Если S — подпись к сообщению M, то как следует подписать сообщение $4M$ в схеме подписи Рабина?

9. В схеме подписи RSA пользователь A имеет открытый ключ $e = 11$, $n = 899$. Как он подпишет сообщение 876?

10. В вероятностной схеме подписи Рабина с функцией шифрования

$$e(x,k) = x^k \pmod{13}$$

A посылает B четыре утвержденных параметра: (2, 10), (11, 7), (7, 8), (6, 3). Как следует подписать сообщение $m = 8$?

11. В предыдущей задаче почему возникли бы трудности, если бы A отправил утвержденный параметр (3,1) вместо (2,10)?

12. Предположим, что мы знаем подписи S_1 и S_2 сообщений m_1 и m_2. Используя схему Рабина

$$S_i = E(m_i) \equiv m_i(m_i + B) \pmod{N}, \quad i = 1, 2,$$

показать, что для некоторых значений открытого ключа B легко подписать сообщение m_1m_2.

13. (п. 14.9). Найти единицу группы G. Разработать алгоритм поиска элементов $g \in G$, имеющих порядок q.

14. (п. 14.9). Сравнить вычислительную сложность систем электронной цифровой подписи DSS, ГОСТ Р 34.10—94, СТБ 1176.2—99. Точнее, сколько требуется выполнить умножений (возведений в степень) по модулю p (или q) для выработки (проверки) подписи?

15. (п. 14.10). Доказать, что всякое число из множества $\{0, 1, \ldots, n-1\}$ можно представить в виде $im + j$, $0 \le i, j < m$, и совпадение $a^j = ba^{-im}$ в методе больших-малых шагов всегда будет найдено.

16. (п. 14.10). Написать программу вычисления дискретного логарифма по методу больших-малых шагов. Вычислить $\mathrm{ind}_7 5$ в группе $\mathbb{Z}^*_{250837}$.

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFVKBDFKJ
VJKGFKJSBDFJBVKJDSGFJK

Глава 15

ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ В КРИПТОГРАФИИ

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFVKBDFKJ
VJKGFKJSBDFJBVKJDSGFJK

15.1. Цифровая подпись на эллиптических кривых

Криптосистемы на эллиптических кривых предложены в 1985 г. В. Миллером и Н. Коблицем. Основные преимущества, которые позволили говорить о криптографии на эллиптических кривых с практической точки зрения, — это, во-первых, большие возможности выбора группы, в которой производятся вычисления, и, во-вторых, отсутствие субэкспоненциального алгоритма дискретного логарифмирования в группе точек на эллиптической кривой (за исключением некоторых частных случаев). Большинство криптографических алгоритмов, стойкость которых базируется на сложности дискретного логарифмирования в конечном поле, достаточно легко переносятся на случай эллиптических кривых. В данном параграфе рассмотрим американский алгоритм цифровой подписи на эллиптических кривых (ECDSA), принятый в качестве международного стандарта. В упомянутом стандарте используются эллиптические кривые над полем характеристики 2. Однако криптографически стойких кривых над такими полями сравнительно мало. Поэтому ограничимся рассмотрением случая эллиптических кривых, заданных над простым полем большей характеристики.

15.1.1. Начальные параметры алгоритма цифровой подписи

Основными начальными параметрами в рассматриваемом алгоритме являются эллиптическая кривая E, определенная над конечным полем $\mathbb{F}_p$ характеристики p, и базовая точка $G \in E(\mathbb{F}_p)$, имеющая большой простой порядок в группе точек на данной эллиптической кривой. Эти параметры могут использоваться как совокупностью лиц, так и одним пользователем.

Эллиптическая кривая задается уравнением

$$y^2 = x^3 + ax + b.$$

Таким образом, задание кривой состоит в выборе двух элементов a, b из поля $\mathbb{F}_p$, которые определяют это уравнение. Различные пары параметров (a,b) могут определять изоморфные эллиптические кривые. При выборе эллиптической кривой можно сначала задать j-инвариант этой кривой, а затем по нему построить коэффициенты a и b.

Точка G на эллиптической кривой определяется парой элементов x_G, y_G из $\mathbb{F}_p$: $G = (x_G, y_G)$. Эта точка выбирается случайно. Один из способов выбора — зафиксировать случайное x и затем найти y как корень второй степени из $x^3 + ax + b$ в поле $\mathbb{F}_p$, если он существует. Существование корня проверяется путем вычисления символа Лежандра $\left(\frac{x^3 + ax + b}{p}\right)$ (см. п. 2.14).

Для получения криптографически стойкой системы цифровой подписи должны выполняться следующие требования:

1) порядок точки G должен быть равен простому числу $n > \max\{2^{160}, 4\sqrt{p}\}$;

2) $\# E(\mathbb{F}_p) \neq p + 1$, т.е. кривая не должна быть *суперсингулярной*;

3) $p^k \not\equiv 1 \pmod{n}$ для всех $k \in \{1, \ldots, C\}$, где C настолько велико, что вычислить дискретный логарифм в $\mathbb{F}_{p^C}$ за приемлемое время невозможно (обычно берут $C = 20$);

4) $\# E(\mathbb{F}_p) \neq p$, т.е. кривая не должна быть аномальной.

Заметим, что условие 3 подразумевает условие 2. Возможный способ защититься от известных атак и от возможных атак для специальных классов кривых, которые могут быть обнаружены в будущем, — выбирать кривую E случайным образом так, чтобы выполнялись указанные условия.

Выбор эллиптической кривой подразумевает решение ряда трудоемких вспомогательных задач. Прежде всего — это подсчет количества точек на эллиптической кривой (об этом речь пойдет в конце главы). После того как порядок N кривой определен, требуется найти большой простой делитель n порядка кривой. Такой делитель может, в принципе, не существовать, и тогда потребуется повторять процедуру выбора кривой до тех пор, пока не выполнятся все требуемые условия. Поиск числа n может потребовать как разложения на множители числа N, так и доказательства простоты полученного множителя n.

Точку G можно выбрать следующим образом. Найдем случайную точку $G' \in E(\mathbb{F}_p)$ и вычислим $G = [N/n]G'$ [см. формулу (15.2.1)]. Будем повторять эту операцию до тех пор, пока точка G не станет отличной от точки $\mathcal{O}$.

Описанные параметры могут быть общими для совокупности пользователей. Для генерации и проверки подписи требуются еще и индивидуальные параметры пользователя — так называемые секретный и открытый ключи.

Ключ подписи (секретный ключ) — это случайное число d в интервале $0 < d < n$.

Ключ проверки подписи (открытый ключ) — это точка на эллиптической кривой $Q = [d]G$.

Алгоритм цифровой подписи также использует хэш-функцию, которая обозначается h.

15.1.2. Генерация и проверка цифровой подписи

Алгоритм 15.1.1 (генерация подписи). Входные данные: сообщение m, исходные параметры и ключ подписи. Выходные данные: подпись (r, s).

1) Выбрать случайное число k в интервале $1 \le k \le n-1$.
2) Вычислить $(x_1, y_1) := [k]G$.
3) Вычислить $r := x_1 \bmod n$.
4) Если $r = 0$, то вернуться к шагу 1.
5) Вычислить $z := k^{-1} \bmod n$.
6) Вычислить $e := h(m)$.
7) Вычислить $s := z(e + dr) \bmod n$.
8) Если $s = 0$, то вернуться к шагу 1.
9) Вывести пару (r, s) — подпись к m.

Алгоритм 15.1.2 (проверка подписи). Входные данные: сообщение m, исходные параметры, ключ проверки подписи и подпись к m. Выходные данные: утверждение, что подпись действительная или фальшивая.

1) Если условия $1 \le r, s \le n-1$ нарушаются, то вывести «подпись фальшивая» и завершить работу алгоритма.
2) Вычислить $e := h(m)$.
3) Вычислить $w := s^{-1} \bmod n$.
4) Вычислить $u_1 := ew \bmod n$.
5) Вычислить $u_2 := rw \bmod n$.
6) Вычислить $X := [u_1]G + [u_2]Q = (x_1, y_1)$.
7) Если $r = x_1 \bmod n$, то вывести «подпись действительная», иначе — «подпись фальшивая», и завершить работу алгоритма.

Корректность алгоритма генерации подписи. Докажем, что любая подпись, сгенерированная по алгоритму 15.1.1, будет «действительной» согласно алгоритму проверки подписи 15.1.2.

Прежде всего заметим, что параметры r и s, получаемые в алгоритме 15.1.1, не превосходят $n-1$, как остатки при делении на n целых чисел. С другой стороны, выполняется проверка того, что $r, s \ne 0$ на шагах 4 и 8 алгоритма 15.1.1. Следовательно, условия шага 1 алгоритма проверки подписи будут выполнены всякий раз, когда r, s получены по алгоритму генерации подписи.

Далее, согласно шагам 5 и 7 алгоритма генерации подписи, имеем $ks \equiv e + dr \pmod n$. Поскольку $w = s^{-1} \bmod n$ (шаг 3 алгоритма проверки подписи), то $k \equiv we + wrd \pmod n$. Так как точка G имеет порядок n, то

$$[k]G = [we + wrd]G = [we]G + [wr][d]G =$$

$$= [we]G + [wr]Q = [u_1]G + [u_2]Q = X.$$

Таким образом, точка X, получаемая на шаге 6 алгоритма проверки подписи, совпадет с точкой $[k]G$, сгенерированной при получении подписи по алгоритму генерации. Первая координата X будет равна x_1, и ее остаток $\bmod n$ будет равен r (согласно шагу 3 алгоритма генерации подписи). Корректность доказана. □

15.2. Особенности скалярного умножения на эллиптических кривых

Под скалярным умножением точки P некоторой эллиптической кривой на целое число $m > 0$ понимается вычисление точки

$$[m]P = \underbrace{P + \ldots + P}_{m}. \qquad (15.2.1)$$

При m полагают $[m]P = -([-m]P)$, а при $m = 0$ — $[0]P = \mathcal{O}$.

Эта операция — аналог возведения в степень в мультипликативных группах. Умение эффективно выполнять скалярное умножение весьма важно, так как именно эта операция наиболее трудоемкая во многих криптографических алгоритмах, в том числе и в алгоритме цифровой подписи, приведенном выше.

В группе точек на эллиптической кривой скалярное умножение можно осуществить *бинарным методом*. Рассмотрим разложение m в двоичной системе счисления:

$$m = a_0 + a_1 2 + a_2 2^2 + \ldots + a_{k-1} 2^{k-1}, \quad a_i \in \{0, 1\},$$

где $k = [\log_2 m] + 1$. Тогда для нахождения значения $[m]P$ (P — точка на эллиптической кривой) достаточно вычислить $[2^i]P$ для $i = 0, \ldots, k-1$ и просуммировать лишь те $[2^i]P$, для которых $a_i = 1$. Всего нам потребуется $k-1$ раз удвоить точку, начиная с P. Затем сделаем не более $k-1$ сложений (это число достигается, когда $m = 2^k - 1$).

Однако специфика эллиптических кривых позволяет уменьшить количество сложений в бинарном методе. Идея метода состоит в том, чтобы блок из последовательных единиц в двоичном представлении m заменять на разность степеней двойки. Например, вместо вычисления $[2^i]P + [2^{i+1}]P + \ldots + [2^{i+j}]P$ можно вычислить $[2^{i+j+1}]P - [2^i]P$, сэкономив $(j-1)$ сложений, поскольку операция нахождения обратной точки, вследствие специфики эллиптических кривых, практически не требует времени. Еще одно улучшение касается случая, когда два блока из последовательных единиц разделены нулем:

$$m = \underbrace{11\ldots1}_{a}\,0\,\underbrace{11\ldots1}_{b}\ldots$$

Тогда, легко видеть, достаточно двух вычитаний:

$$[m]P = [2^{a+b+1}]P - [1]P - [2^b]P.$$

15.3. Вычисление порядка эллиптической кривой

Здесь речь пойдет о проблеме подсчета точек на эллиптической кривой E над конечным полем $\mathbb{F}_q$. Безопасность криптосистем, построенных на эллиптических кривых, основана на трудности вычисления дискретного логарифма

в группе точек на эллиптических кривых. Сложность логарифмирования в группе точек на эллиптической кривой оценивается квадратным корнем из наибольшего простого делителя порядка этой группы [56]. Поэтому в вопросах безопасности криптосистем на эллиптических кривых важно знать порядок группы и даже его разложение на множители. Некоторые вопросы подсчета количества точек на эллиптической кривой были рассмотрены в главе 5.

15.3.1. Метод больших-малых шагов для вычисления порядка

Рассмотрим следующий вопрос: как определить, является ли данное число m порядком эллиптической кривой. Далее рассмотрим метод, позволяющий найти «числа-кандидаты» на то, чтобы быть порядком. Поэтому важно уметь выбрать среди них то единственное, которое и является порядком группы точек на эллиптической кривой.

Пусть $\overline{\mathbb{F}}_q$ — алгебраическое замыкание $\mathbb{F}_q$. Напомним, что для поля K, $\mathbb{F}_q \subset K \subset \overline{\mathbb{F}}_q$, $E(K)$ обозначает множество K-рациональных точек.

Если число $m = \# E(\mathbb{F}_q)$, то должны выполняться следующие условия:

$$\begin{cases} q+1-2\sqrt{q} \le m \le q+1+2\sqrt{q}, \\ P \in E(\mathbb{F}_q) \Rightarrow [m]P = \mathcal{O}. \end{cases} \tag{15.3.1}$$

Таким образом, чтобы отсеять лишних кандидатов из списка, в котором наверняка есть правильный ответ, можно проверять условия (15.3.1), перебирая разные точки P. Единственное оставшееся число в этом списке составит искомый порядок.

Рассмотрим алгоритм, позволяющий найти сравнительно небольшое количество кандидатов на порядок группы. По теореме Хассе

$$\# E(\mathbb{F}_q) = q+1-t, |t| \le 2\sqrt{q}. \tag{15.3.2}$$

Таким образом, чтобы установить порядок эллиптической кривой, требуется найти число t, для которого имеется $2[2\sqrt{q}]+1$ возможностей.

Идея алгоритма состоит в нахождении для точки $P \in E(\mathbb{F}_q)$, выбранной случайным образом, всех чисел $m \in (q+1-2\sqrt{q}),\ q+1+2\sqrt{q}$ таких, что $[m]P = \mathcal{O}$. Если такое число в заданном интервале единственное, то по теореме Хассе $m = \# E(\mathbb{F}_q)$. Если таких чисел несколько, то можно уменьшить количество кандидатов на целый порядок из полученного списка методом, описанным выше.

Итак, точка $P \in E(\mathbb{F}_q)$ выбирается случайно. Положим $s = \lceil \sqrt[4]{q} \rceil$, где $\lceil x \rceil = \min\{a \in \mathbb{Z} : \ge x\}$. Сначала составляется и сортируется таблица из $2s+1$ точек:

$$\mathcal{O}, \pm P, \ldots, \pm [s]P. \tag{15.3.3}$$

Далее определим точки

$$Q = [2s+1]P, \quad R = [q+1]P. \tag{15.3.4}$$

Пусть $m = \# E(\mathbb{F}_q) = q + 1 - t$. Тогда

$$\mathcal{O} = [m]P = [q + 1 - t]P = R - [t]P. \tag{15.3.5}$$

Легко видеть, что существует единственное целое число j такое, что $|j| \le s$ и $t \equiv j \pmod{2s+1}$. Тогда число $i = (j - t)/(2s + 1)$ также целое, причем

$$|i| \le \frac{s + 2\sqrt{q}}{2s+1} < \frac{s + 2s^2}{2s+1} = s.$$

Таким образом, найдутся целые i, j такие, что $|i|$, $|j| \le s$ и $-t = (2s + 1)i - j$. Подставляя это выражение в (15.3.5), получаем

$$\mathcal{O} = R + [i(2s + 1)]P - [j]P = R + [i]Q - [j]P. \tag{15.3.6}$$

Следовательно, вычисляя точки вида $R \pm [i]Q$, где $i = \overline{0, s}$, и проверяя, есть ли такая точка в массиве (15.3.3), найдем все пары i, j такие, что выполнено (15.3.6). В частности, одной из этих пар будет соответствовать искомый порядок m.

15.3.2. Метод Шуфа

Данный метод основан на использовании свойств так называемого *эндоморфизма Фробениуса*:

$$\varphi_E : E(\overline{\mathbb{F}}_q) \to E(\overline{\mathbb{F}}_q),$$
$$(x, y) \mapsto (x^q, y^q),$$
$$\mathcal{O} \mapsto \mathcal{O}.$$

Согласно теореме Хассе [56], эндоморфизм Фробениуса удовлетворяет уравнению $f(\varphi_E) = 0$ в кольце эндоморфизмов $\mathrm{End}(E)$, где $f(x) = x^2 - tx + q \in \mathbb{Z}[x]$, $t = q + 1 - \# E(\mathbb{F}_q)$, $t \le 2\sqrt{q}$. Число t называется *следом Фробениуса*. Таким образом, для любой точки $P = (x, y) \in E$

$$\varphi_E^2(P) - [t]\varphi_E(P) + [q]P = \mathcal{O}. \tag{15.3.7}$$

Идея метода Шуфа сводится к определению остатков числа t по простым модулям l, $l \le l_{\max}$, где $l_{\max}$ — наименьшее простое такое, что

$$\prod_{2 \le l \le l_{\max}} l > 4\sqrt{q}.$$

Когда все такие остатки найдены, восстановить число t по китайской теореме об остатках не представляет особого труда.

При $l = 2$ легко определить $t \bmod 2$. Если характеристика поля нечетная, $t \equiv \# E(\mathbb{F}_q) \pmod 2$, то $\# E(\mathbb{F}_q) \equiv 1 \pmod 2 \Leftrightarrow x^3 + ax + b$ неразложим над $\mathbb{F}_q$, что эквивалентно $(x^3 + ax + b, x^q - x) = 1$. Если $\mathrm{char}(\mathbb{F}_q) = 2$ и кривая не является вырожденной, то $t \equiv 1 \pmod 2$.

Рассмотрим случай $l > 2$.

Множество $E[l] \subset E$, состоящее из точек эллиптической кривой, порядок которых делит l, т.е. из точек $P \in E(\overline{\mathbb{F}}_q)$ таких, что $[l]P = \mathcal{O}$, называется *l-кручением кривой E*.

Известно, что

$$E[l] \cong (\mathbb{Z}/l\,\mathbb{Z}) \times (\mathbb{Z}/l\,\mathbb{Z}).$$

Также легко проверяется, что $E[l]$ — подгруппа E.

Пусть $q_l = q \bmod l$. Для нахождения $t_l = t \bmod l$ выбирается точка $P = (x,y) \in E[l] \setminus \{\mathcal{O}\}$ и ищется $\tau \in 0, 1, \ldots, l-1$ такое, что

$$\varphi_E^2(P) + [q_l]P = [\tau]\varphi_E(P). \tag{15.3.8}$$

Такое τ единственное, так как l — простое, $P \neq \mathcal{O}$. Кроме того, в силу выбора точки P (15.3.8) выполнено при $\tau = t_l$. Поэтому решение (15.3.8) существует и единственно. Найти его можно даже любым переборным методом (например, методом больших-малых шагов, аналогичным описанному выше).

15.4. Задания

1. Дать верхнюю оценку количества производимых операций в модификации бинарного метода скалярного умножения, предварительно написав алгоритм скалярного умножения по этому методу (см. п. 15.2).

2. Существенно ли требование $r, s \neq 0$ в алгоритме подписи (см. п. 15.1) и почему?

3. Пусть P — точка эллиптической кривой $E/\mathbb{F}_q$ и порядок этой точки $> 4\sqrt{q}$. Сколько различных целых чисел m удовлетворяет условиям (15.3.1)?

4. Пусть задана эллиптическая кривая $E/\mathbb{F}_q$ и $E(\mathbb{F}_q)$ — циклическая группа. Оценить вероятность того, что для случайно взятой точки $P \in E$ найдется по крайней мере два различных целых числа m, удовлетворяющих условиям (15.3.1).

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7DFJ
BHVIEURKDHFVIUDKFJBVKD
FVKJDBFKVJHB0954TOIRF7
ERFH2347FQ7WERF9UJ1349
YUR7DFKJHGJDFBVDGFJVBD
SKFJHKJFVKBDFKJVJKGFKJ
SBDFJBVKJDSGFJKVBKSDJB
FKVBDGFVAKDLKWEF90UQ3I

Глава 16

ПРОТОКОЛЫ УПРАВЛЕНИЯ КРИПТОГРАФИЧЕСКИМИ КЛЮЧАМИ

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7DFJ
BHVIEURKDHFVIUDKFJBVKD
FVKJDBFKVJHB0954TOIRF7
ERFH2347FQ7WERF9UJ1349
YUR7DFKJHGJDFBVDGFJVBD
SKFJHKJFVKBDFKJVJKGFKJ
SBDFJBVKJDSGFJKVBKSDJB
FKVBDGFVAKDLKWEF90UQ3I

Эффективность криптографической защиты информации в компьютерных сетях определяется стойкостью используемых алгоритмов криптографических преобразований и надежностью протоколов управления ключами.

Протоколы управления криптографическими ключами включают в себя протоколы генерации, распределения, хранения, смены и уничтожения криптографических ключей. В данной главе основное внимание уделяется протоколам генерации и распределения ключей на основе симметричных и асимметричных криптосистем. Протоколы распределения ключей должны обеспечивать взаимную аутентификацию сторон, целостность сообщений и защиту от повторных запросов.

16.1. Протоколы генерации ключей

Используемые в компьютерной сети ключи принято разделять на три группы: 1) *главные ключи km* сети, серверов, компьютеров, терминалов, которые используются для шифрования и генерации ключей; 2) *ключи шифрования ключей файлов* k_F; 3) *сеансовые ключи* $k_{A,B}$, которые используются для шифрования данных, передаваемых в сеансе связи между пользователями сети *A* и *B*.

Для генерации главных ключей *km*, или *мастер-ключей*, используются физические генераторы случайных последовательностей (см. главу 6). Мафтик [18] отмечает, что главные ключи распределяются и устанавливаются специальным образом.

Для генерации ключей шифрования ключей и сеансовых ключей наряду с физическими датчиками используются псевдослучайные методы с секретными параметрами.

Рассмотрим протоколы генерации сеансовых ключей на основе криптосистемы DES. В качестве исходных данных для генерации ключей используются случайные числа *RN*, вектор «дата — время» *DT* и номер генерируемого ключа *i*, представленные в виде 64-разрядных двоичных векторов.

Матиас и Мейер [138] предложили следующий *протокол генерации сеансовых ключей*:

$$k_{A,B} = D(km, RN),$$

где $D(km, RN)$ — расшифрованное на мастер-ключе km случайное число RN. В данном протоколе для генерации RN рекомендуется использовать физический датчик.

В стандарте ИСО 8731 [19] для генерации сеансовых ключей предлагается использовать следующий протокол.

Шаг 1. Вычисление промежуточного значения:

$$J = E(km, DT),$$

т. е. зашифрованного на мастер-ключе km текста DT.

Шаг 2. Генерация сеансового ключа:

$$k_{A,B} = E(km, J \oplus V),$$

где V — 64-битовый вектор инициализации, сохраняемый в тайне.

Шаг 3. Генерация нового значения вектора инициализации:

$$V = E(km, k_{A,B} \oplus J).$$

Согласно ИСО 8731 значение вектора инициализации V должно храниться в секрете.

Матиас и Мейер [138] предложили также *протокол генерации ключей шифрования ключей файлов* псевдослучайным методом.

Шаг 1. Генерация вспомогательного ключа:

$$k_V = D(km_1, RN),$$

где km_1 — дополнительный главный ключ.

Шаг 2. Генерация ключа шифрования ключей файлов:

$$k_F = E(k_V, D(km, E(k_V, D(km, (DT + i))))).$$

16.2. Протоколы взаимной аутентификации

Для обеспечения подлинности сеансов связи при управлении ключами используются *протоколы взаимной аутентификации*, основанные на симметричных и асимметричных криптосистемах.

В данном параграфе приводятся протоколы взаимной аутентификации без участия третьей доверенной стороны и с ее участием.

Для взаимной аутентификации участников сеанса связи без использования третьей доверенной стороны можно использовать следующую модификацию *протокола рукопожатия*.

Шаг 1. Пользователь A посылает пользователю B число x_A:

$$A \to B : x_A.$$

Шаг 2. Пользователь B вычисляет значение $y_A = f(x_A)$ и отправляет y_A и x_B пользователю A:

$$B \to A : y_A, x_B.$$

Шаг 3. Пользователь A вычисляет $f(x_A)$ и сравнивает с полученным значением y_A. В случае равенства пользователь B является истинным, легальным. Затем пользователь A вычисляет значение $y_B = f(x_B)$ и отправляет пользователю B:

$$A \to B : y_B.$$

Шаг 4. Пользователь B вычисляет значение $f(x_B)$ и сравнивает с полученным значением y_B. В случае равенства пользователь A является подлинным.

Функция $f(\cdot)$ должна быть достаточно сложной, чтобы криптоаналитик не смог по значениям x и $f(x)$ восстановить функцию $f(\cdot)$. В качестве $f(x)$ можно использовать функцию вида $f(x) = E(K_{AB}, x)$, где E — симметричный криптоалгоритм, K_{AB} — ключ, известный только пользователям A и B.

Для взаимной аутентификации пользователей с участием третьей доверенной стороны Нидхэм и Шредер [150] предложили протоколы взаимной аутентификации на основе симметричных и асимметричных криптосистем. В качестве третьей доверенной стороны в данном параграфе используется сервер аутентификации AS.

Протокол взаимной аутентификации на основе симметричных криптосистем имеет следующий вид.

Шаг 1. Пользователь A посылает AS сообщение, содержащее идентификаторы пользователей A, B и свое сообщение I_A:

$$A \to AS : A, B, I_A.$$

Шаг 2. AS формирует для пользователей A и B сообщения, содержащие сеансовый ключ $k_{A,B}$, зашифровывает их на ключах k_A и k_B соответственно и отправляет пользователю A:

$$AS \to A : E(k_A, \{I_A, B, k_{A,B}, Y_B\}),$$

где $Y_B = E(k_B, \{k_{A,B}, A\})$.

Шаг 3. Пользователь A расшифровывает с помощью ключа k_A сообщение, полученное от AS, и отправляет пользователю B сообщение AS для B:

$$A \to B : E(k_B, \{k_{A,B}, A\}).$$

Шаг 4. Пользователь B с помощью ключа k_B расшифровывает полученное от пользователя A сообщение и направляет ему сообщение I_B, зашифрованное сеансовым ключом $k_{A,B}$:

$$B \to A : E(k_{A,B}, \{I_B\}).$$

Шаг 5. Пользователь A с помощью сеансового ключа $k_{A,B}$ расшифровывает сообщение, полученное от B, и отправляет ему следующее сообщение:

$$A \to B : E(k_{A,B}, \{I_B - 1\}).$$

Протокол взаимной аутентификации Нидхэма — Шредера на основе асимметричных криптосистем включает следующие шаги.

Шаг 1. Пользователь A посылает AS сообщение, содержащее идентификаторы пользователей A и B:

$$A \to AS : A, B.$$

Шаг 2. AS формирует для пользователя A сообщение, содержащее открытый ключ pk_B и идентификатор пользователя B, зашифровывает его с помощью секретного ключа сервера аутентификации sk_{AS} и отправляет пользователю A:

$$AS \to A : E(sk_{AS}, \{pk_B, B\}).$$

Шаг 3. Пользователь A с помощью открытого ключа сервера аутентификации pk_{AS} расшифровывает сообщение, полученное от AS, и отправляет пользователю B сообщение

$$A \to B : E(pk_B, \{I_A, A\}).$$

Шаг 4. Пользователь B с помощью секретного ключа sk_B расшифровывает полученное от A сообщение и направляет серверу аутентификации сообщение

$$B \to AS : A, B.$$

Шаг 5. AS формирует для пользователя B сообщение, содержащее открытый ключ pk_A и идентификатор пользователя A, зашифровывает его с помощью секретного ключа сервера аутентификации sk_{AS} и отправляет пользователю B:

$$AS \to B : E(sk_{AS}, \{pk_A, A\}).$$

Шаг 6. Пользователь B с помощью открытого ключа сервера аутентификации pk_{AS} расшифровывает сообщение, полученное от AS, и отправляет пользователю A сообщение

$$B \to A : E(pk_A, \{I_A, I_B\}).$$

Шаг 7. Пользователь A расшифровывает с помощью секретного ключа sk_A расшифровывает полученное от B сообщение и отправляет ему сообщение, зашифрованное с помощью открытого ключа pk_B:

$$A \to B : (pk_B, \{I_B\}).$$

16.3. Протоколы прямого обмена ключами

После взаимной проверки подлинности, например с помощью протоколов взаимной аутентификации (см. п. 16.2), пользователи A и B для получения сеансового ключа могут применять следующую модификацию *протокола Диффи — Хеллмана.*

Шаг 1. Пользователь A выбирает число x_A, которое держит в секрете, вычисляет $y_A = a^{x_A} \bmod p$ и отправляет y_A пользователю B:

$$A \to B : y_A.$$

Шаг 2. Пользователь B выбирает число x_B, которое держит в секрете, вычисляет $y_B = a^{x_B} \bmod p$ и отправляет y_B пользователю A:

$$B \to A : y_B.$$

Шаг 3. Пользователь A вычисляет $k_A = (y_B)^{x_A} \bmod p$ и получает $k_{A,B} = F(k_A)$.

Шаг 4. Пользователь B вычисляет $k_B = (y_A)^{x_B} \bmod p$ и получает $k_{A,B} = F(k_B)$.

В качестве отображения F можно использовать, например, композицию отображений, реализующих функцию хэширования и операцию усечения (см. главу 10).

Приведем *протоколы взаимного обмена сеансовым ключом*, основанные на симметричных и асимметричных криптосистемах, в которых совмещены процедуры взаимной аутентификации и распределения сеансового ключа. Предположим, что пользователи A и B имеют общий секретный ключ K_{AB}. Тогда для обмена сеансовым ключом $k_{A,B}$ можно использовать следующий протокол.

Шаг 1. Пользователь B посылает A сообщение, содержащее идентификатор пользователя B и запрос $I1$:

$$B \to A : B, I1.$$

Шаг 2. Пользователь A отправляет B свой идентификатор и зашифрованное на ключе K_{AB} сообщение, включающее ответ $f(I1)$, запрос $I2$, идентификатор A и сеансовый ключ $k_{A,B}$:

$$A \to B : A, E(K_{AB}, \{f(I1), I2, A, k_{A,B}\}).$$

Шаг 3. Пользователь B с помощью ключа K_{AB} расшифровывает полученное сообщение и отправляет пользователю A свой идентификатор и зашифрованное на ключе $k_{A,B}$ сообщение, содержащее ответ $f(I2)$ и идентификатор B:

$$B \to A : B, E(k_{A,B}, \{f(I2), B\}).$$

Приведем *протокол взаимного обмена сеансовым ключом на основе асимметричных криптосистем*. Предположим, что пользователи A и B обменялись своими открытыми ключами pk_A и pk_B. Тогда протокол взаимного обмена сеансовым ключом имеет следующий вид.

Шаг 1. Пользователь B посылает пользователю A свой идентификатор и зашифрованное на ключе pk_A сообщение, содержащее запрос $I1$ и идентификатор B:

$$B \to A : B, E(pk_A, \{I1, B\}).$$

Шаг 2. Пользователь A с помощью секретного ключа sk_A расшифровывает зашифрованную часть полученного сообщения и отправляет пользователю B

зашифрованное на ключе pk_B сообщение, содержащее ответ $f(I1)$, запрос $I2$, идентификатор пользователя A и сеансовый ключ $k_{A,B}$:

$$A \to B : E(pk_B, \{f(I1), I2, A, k_{A,B}\}).$$

Шаг 3. Пользователь B с помощью секретного ключа sk_B расшифровывает полученное сообщение и отправляет A идентификатор пользователя B и зашифрованное на ключе $k_{A,B}$ сообщение, содержащее ответ $f(I2)$ и идентификатор B:

$$B \to A : B, E(k_{A,B}, \{f_{I2}, B\}).$$

16.4. Протоколы распределения сеансовых ключей с использованием центра распределения ключей

В качестве *протокола распределения сеансовых ключей* с использованием центра распределения ключей рассмотрим *протокол Цербер* [151] (от греч. Kerberos — трехголовый злой пес с хвостом и гривой из змей в древнегреческой мифологии, охранявший вход в подземное царство). Схема протокола Цербер приведена на рис. 16.1.

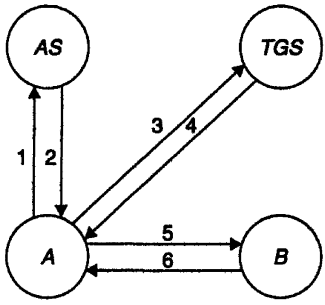


Рис. 16.1. Схема протокола Цербер

Протокол Цербер разработан в Массачусетском технологическом институте на основе протокола Нидхэма — Шредера (см. п. 16.2). В протоколе Цербер реализованы взаимная аутентификация и механизм отметки времени, гарантирующие подлинность сеанса распределения сеансового ключа для пользователей A и B. В качестве центра распределения ключей в протоколе Цербер выступает сервер TGS.

Протокол Цербер включает следующие шаги.

Шаг 1. Пользователь A посылает на сервер аутентификации AS запрос на билет $T_{A,TGS}$ пользователя A для сервера TGS, включающий идентификатор пользователя A, идентификатор TGS и отметку времени t:

$$A \to AS : A, TGS, t.$$

Шаг 2. Сервер аутентификации AS отправляет пользователю A пакет ответа. Последний включает сообщение для A, зашифрованное на ключе k_A и содержащее сеансовый ключ $k_{A,TGS}$ для связи пользователя A и сервера TGS, идентификатор сервера TGS, сетевой адрес S_{TGS}, отметку времени t, и зашифрованный на ключе k_{TGS} билет, включающий идентификаторы пользователя A и сервера TGS, сетевой адрес S_A пользователя A, отметку времени t, время действия билета tv и сеансовый ключ $k_{A,TGS}$:

$$AS \to A: E(k_A, \{k_{A,TGS}, TGS, S_{TGS}, t\}), E(k_{TGS}, \{T_{A,TGS}\}),$$

где $T_{A,TGS} = \{A, TGS, S_A, t, tv, k_{A,TGS}\}$.

Шаг 3. Пользователь A отправляет на сервер TGS пакет запроса, включающий зашифрованный на ключе $k_{A,TGS}$ аутентификатор пользователя A, содержащий идентификатор пользователя A, сетевой адрес S_A, отметку времени t, и билет $T_{A,TGS}$ для связи пользователя A с сервером TGS, зашифрованный на ключе k_{TGS}, а также идентификатор пользователя B и отметку времени t:

$$A \to TGS: E(k_{A,TGS}, \{A, S_A, t\}), E(k_{TGS}, \{T_{A,TGS}\}), B, t.$$

Шаг 4. Сервер TGS, используя сеансовый ключ $k_{A,TGS}$ и секретный ключ k_B пользователя B, готовит пакет ответа и отправляет пользователю A. Пакет ответа включает зашифрованное на ключе $k_{A,TGS}$ сообщение для пользователя A, содержащее сеансовый ключ $k_{A,B}$ для связи пользователей A и B, идентификатор пользователя B, сетевой адрес S_B, отметку времени t, а также сообщение для пользователя B, содержащее зашифрованный на ключе k_B билет $T_{A,B}$, включающий идентификаторы пользователей A и B, сетевой адрес S_A, отметку времени t, время действия билета tv и сеансовый ключ $k_{A,B}$:

$$TGS \to A: E(k_{A,TGS}, \{k_{A,B}, B, S_B, t\}), E(k_B, \{T_{A,B}\}),$$

где $T_{A,B} = \{A, B, S_A, t, tv, k_{A,B}\}$.

Шаг 5. Пользователь A, получив пакет ответа, расшифровывает с помощью ключа $k_{A,TGS}$ сообщение для пользователя A. Затем A проверяет подлинность сообщения, используя отметку времени t и идентификатор B. Пользователь A создает аутентификатор, содержащий идентификатор пользователя A, сетевой адрес S_A, отметку времени t, и зашифровывает его с помощью сеансового ключа $k_{A,B}$. A отправляет пользователю B аутентификатор и зашифрованный на ключе k_B билет $T_{A,B}$:

$$A \to B: E(k_{A,B}, \{A, S_A, t\}), E(k_B\{T_{A,B}\}).$$

Шаг 6. Пользователь B расшифровывает билет $T_{A,B}$ и проверяет подлинность сообщения по отметке времени и идентификаторам пользователей A и B. Применяя сеансовый ключ связи $k_{A,B}$, пользователь B расшифровывает аутентификатор пользователя A и убеждается в подлинности пользователя A. Пользователь B создает аутентификатор, содержащий идентификатор пользователя

B, сетевой адрес S_B, отметку времени t, зашифровывает его с помощью сеансового ключа $k_{A,B}$ и отправляет пользователю A:

$$B \to A: E(k_{A,B}, \{B, S_B, t\}).$$

Шаг 7. Пользователь A, применяя сеансовый ключ $k_{A,B}$, расшифровывает аутентификатор, полученный от пользователя B, и убеждается в подлинности пользователя B.

В результате выполнения шагов протокола Цербер пользователи A и B получают сеансовый ключ $k_{A,B}$ для обмена сообщениями.

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFVKBDFKJ
VJKGFKJSBDFJBVKJDSGFJK

Глава 17

НОВЫЕ НАПРАВЛЕНИЯ В КРИПТОЛОГИИ

SXJKN8EYHR3I4HR987YDWF
UIHW9E8YR384H5O34596EU
F9O8WE76R83Y439846UHFE
98YDSFJH8374H4BR8276RU
BNWDF87Y6234BN9WE872H3
WRTJOJRTGM905JRO9G7095
4TOIRF7ERFH2347FQ7WERF
9UJ1349YUR7DFKJHGJDFBV
DGFJVBDSKFJHKJFVKBDFKJ
VJKGFKJSBDFJBVKJDSGFJK

17.1. Возможности квантовой криптографии

Криптология является многогранным направлением прикладной математики, интенсивно использующим такие фундаментальные понятия современной математической науки, как информация, случайность, алгоритм, сложность. Так как неопределенность и стохастичность окружающего мира имеют квантовую природу, то криптология связана с квантовой механикой и с ее основным постулатом — принципом неопределенности Гейзенберга.

Впервые идея шифрования с использованием квантового канала сформулирована в 1970 г. С. Уиснером и развита в 80-е гг. Ч. Беннетом, Ж. Брассаром и С. Брейдбардом [50].

Для пересылки последовательности двоичных битов в квантовом канале связи используются специальным образом поляризованные фотоны. Фотон — элементарная квантовая система, которая характеризуется определенным направлением поляризации $r = (r_1, r_2)' = (\cos\alpha, \sin\alpha)'$ (рис. 17.1).

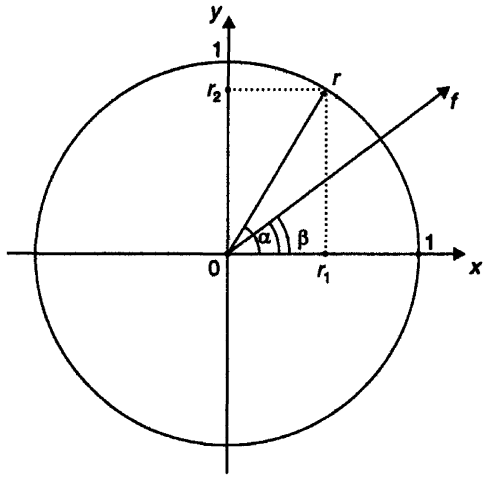


Рис. 17.1. Поляризация фотона

Здесь r — единичный вектор, указывающий поляризацию фотона и направленный под углом α; $r_1 = \cos\alpha$, $r_2 = \sin\alpha$ — проекции; f — направление поляризации фильтра-анализатора. Из законов квантовой механики следует, что при попадании в такой фильтр-анализатор с параметром β фотон ведет себя «дихотомическим и совершенно непредсказуемым образом» [6], проходя через фильтр без изменения с вероятностью $p = p(\alpha - \beta) = \cos^2(\alpha - \beta)$ и поглощаясь в нем с дополнительной вероятностью $q = q(\alpha - \beta) = 1 - p = \sin^2(\alpha - \beta)$. Это квантово-механическое явление проявляется и на макроуровне для интенсивности пучка поляризованного света при прохождении через фильтр в виде известного физического закона Малюса. Отметим, что если разность угла поляризации и угла ориентации фильтра-анализатора принимает значение

$$\gamma = \alpha - \beta \in \left\{\frac{\pi}{4}, \frac{3\pi}{4}\right\},$$

то $p(\gamma) = q(\gamma) = 1/2$ и фотон проходит фильтр или поглощается им с одной и той же вероятностью 1/2. Этот факт и используется при выборе осей поляризации для квантового шифрования.

Принято говорить, что имеет место *прямоугольная поляризация* фотона, если его вектор поляризации

$$r = r^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{или} \quad r = r^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Прямоугольная поляризация обозначается символом «+». Если же

$$r = r^{(3)} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \text{или} \quad r = r^{(2)} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix},$$

то говорят о *диагональной поляризации* и обозначают ее символом «×».

Проиллюстрируем применение квантовых сигналов в криптографии на примере передачи секретного ключа по открытому квантовому каналу. Рассмотрим классическую схему криптографической передачи информации, в которой участвуют три персонажа: 1) Алиса, желающая передать Бобу секретный ключ, который в дальнейшем будет использоваться, например, для обмена информацией с помощью некоторой симметричной криптосистемы; 2) Боб, который нуждается в секретном ключе; 3) Ева — криптоаналитик, которая хотела бы скрытно завладеть копией передаваемого секретного ключа.

Алиса генерирует чисто случайную секретную двоичную ключевую последовательность длиной N:

$$K = (K_1, \ldots, K_N) \in V_N$$

и чисто случайную секретную индексную последовательность той же длины:

$$J = (J_1, \ldots, J_N) \in V_N.$$

Затем Алиса генерирует в квантовый канал связи последовательность N фотонов: i-му фотону F_i, несущему сигнал K_i, дается поляризация «+» (если $J_i = 0$) и поляризация «×» (если $J_i = 1$).

Боб, имея два приемника фотонов (с поляризациями «+» и «×»), принимает поток фотонов. Если бы он знал индексную последовательность J, то, переключая приемники, добился бы того, что $\gamma_i \equiv 0$, $p_i \equiv 1$, $q_i \equiv 0$, т.е. что последовательность фотонов была бы воспринята безошибочно, и, следовательно, получен ключ K. Так как Боб не знает J, он использует свою, наудачу выбранную индексную последовательность $J' \in V_N$. Очевидно, что в среднем совпадет $N/2$ символов у J и J'. Следовательно, лишь половина (в среднем) фотонов будет зарегистрирована безошибочно, а остальные фотоны не несут никакой информации о K:

$$K' = (K'_1, \ldots, K'_N) \in V_N,$$

$$K'_{i_l} = K_{i_l}, \quad l = 1, 2, \ldots, n, \quad \mathbf{E}\{n\} = \frac{n}{2}.$$

Заметим, что если Ева извлечет некоторые фотоны из последовательности для измерения своими приемниками, то Боб сразу же обнаружит потерю фотонов.

Следующие шаги протокола выполняются в обычном канале связи. Прежде всего, через этот канал Алиса и Боб определяют посредством открытого обмена сообщениями, какие фотоны зарегистрированы и какие из них соответствуют истинной поляризации $J_{i'} = J_i$. Если Ева не нарушала квантовой передачи (т.е. получено столько же фотонов, сколько отправлено), то биты $K_{i_1}, \ldots, K_{i_n}$, значения которых не обсуждались по открытому каналу, становятся общим секретным ключом для Алисы и Боба:

$$k = (k_1, \ldots, k_n) = (K_{i_1}, \ldots, K_{i_n}).$$

Из-за возможных действий Евы (в том числе связанных со «вставкой» фотонов) Алиса и Боб должны убедиться, что их получившиеся битовые строки идентичны. Простое решение состоит в том, чтобы Алиса и Боб открыто сравнили некоторые m из n битов ($m \ll n$), относительно которых, как они думают, необходимо прийти к соглашению. Позиции этих «открыто сверяемых» битов должны быть выбраны после того, как квантовая передача завершена, чтобы лишить Еву информации о том, какие фотоны она может измерять без опаски. Оставшиеся $n - m$ битов могут использоваться в качестве секретного ключа для последующей связи по открытому каналу с помощью одного из симметричных криптоалгоритмов. На рис. 17.2 аналогично [6] приведен протокол выработки секретного ключа с помощью квантового канала связи при $N = 15$, $n = 6$, $m = 2$, состоящий из следующих шагов:

1) передача по квантовому каналу:
   а) случайная битовая строка K, посылаемая Алисой;
   б) последовательность поляризаций, задаваемая J;
   в) посланная Алисой последовательность фотонов;
   г) последовательность поляризаций J', использованная Бобом;
   д) битовая строка, зарегистрированная Бобом;
2) обсуждение по открытому каналу:
   а) Боб сообщает поляризацию зарегистрированных фотонов;
   б) Алиса отмечает, какие поляризации были угаданы правильно;

в) последовательность $n = 6$ секретных битов, которую можно использовать в качестве секретного ключа (если не было нарушения квантовой передачи);

г) Боб указывает номера $m = 2$ «открыто сверяемых» битов ключа;

д) Алиса подтверждает эти биты;

3) результат: оставшиеся $n - m = 4$ секретных битов.

Шаг	Информация														
1	0	1	1	0	1	1	0	0	1	0	1	1	0	0	1
2	×	+	×	+	+	+	+	+	×	×	+	×	×	×	+
3	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}
4	+	×	×	+	+	×	×	+	×	+	×	×	×	×	+
5	1		1		1	0	0	0		1	1	1		0	1
6	+		×		+	×	×	+		+	×	×		×	+
7			√		√			√				√		√	√
8			1		1			0				1		0	1
9					1									0	
10					√									√	
11			1					0				1			1

Рис. 17.2. Протокол выработки секретного ключа с помощью квантового канала связи

17.2. Математическое разделение секрета

Рассмотрим разделение секрета между несколькими законными пользователями.

Будем говорить, что t участников A_i, $i = 1, \ldots, t$ (законных пользователей) *k-хранят* секрет c, $1 \le k \le t$, если выполняются следующие три условия:

1) каждый A_i знает некоторую информацию (частичный секрет) a_i, неизвестную любому другому участнику;

2) секрет c может быть легко вычислен на основе любых k частичных секретов $a_{i_1}, \ldots, a_{i_k}$;

3) знание любых $k - 1$ частичных секретов a_i не дает такой возможности.

Множество $\{a_1, \ldots, a_t\}$, удовлетворяющее этим условиям, называется *(k, t)-пороговой схемой*.

Практическое применение таких схем очевидно. Сначала дадим подход к построению пороговых схем на основе теории сравнений (см. главу 2). Пусть k фиксировано, $1 < k \le t$. Обозначим через M_1 наименьшее произведение k различных модулей. Разместив модули в порядке возрастания, получим $M_1 = m_1 \cdots m_k$.

Обозначим через M_2 наибольшее произведение $k-1$ по возможности модулей. Модули следует выбирать так, чтобы разность $M_1 - M_2$ была велика. Далее требуется, чтобы $M_2 < c < M_1$. В качестве секретов участников a_i возьмем наименьшие неотрицательные вычеты секрета c по модулям m_i так, что

$$a_i \equiv c \pmod{m_i}, \quad i = 1, \ldots, t.$$

Теорема 17.2.1. Множество $a_1, \ldots, a_t$ есть (k, t)-пороговая схема для секрета c.

Доказательство. Участники с номерами 1, 2, ..., k могут найти секрет, решив систему сравнений

$$x \equiv a_i \pmod{m_i}, \quad i = 1, \ldots, k.$$

По китайской теореме об остатках эта система имеет единственное решение в промежутке

$$[0, M_1), \quad M_1 = m_1 m_2 \cdots m_k.$$

Это и есть секрет c, поскольку $(M_2, M_1) \subset [0, M_1)$. Другие k или более участников также найдут секрет, поскольку произведение их модулей будет больше M_1. Если количество участников меньше k, то произведение модулей M удовлетворяет условию $M < M_2 < c$. Следовательно, они в состоянии найти лишь некоторое число c' такое, что

$$c' < c, \quad c' \equiv c \pmod{M}.$$

Чисел с таким свойством в промежутке (M_2, M_1) находится примерно $(M_2 - M_1)M^{-1}$. □

Рекомендуем самостоятельно проверить корректность (3, 5)-пороговой схемы, где

$$m_1 = 97, \ m_2 = 98, \ m_3 = 99, \ m_4 = 101, \ m_5 = 103,$$

$$a_1 = 73, \ a_2 = 15, \ a_3 = 61, \ a_4 = 61, \ a_5 = 49,$$

и найти секрет c.

Еще один способ разделения секрета, предложенный Шамиром [9], состоит в следующем. Берут некоторый многочлен

$$f(x) = a_{k-1}x^{k-1} + a_{k-2}x^{k-2} + \ldots + a_1 x + a_0$$

над большим конечным полем. Секретным считается свободный член a_0. Секреты t участников определяются как значения этого многочлена в некоторых точках, т.е. участники знают $c_1 = f(x_1)$, $c_2 = f(x_2)$, ..., $c_t = f(x_t)$.

Очевидно, любые k участников, воспользовавшись интерполяционной формулой Лагранжа, могут найти многочлен $f(x)$, а значит, и элемент a_0. Нетрудно видеть, что любые $k-1$ участников никакой информацией об a_0 обладать не могут.

Рассмотрим еще один простой пример пороговой (t, t)-схемы. Будем отождествлять секреты с элементами кольца вычетов Z_n. Для разделения секрета $c \in Z_n$ секреты участников $A_1, A_2, \ldots, A_{t-1}$ определим случайным образом:

$a_1, a_2, \ldots, a_{t-1} \in Z_n$, а последний секрет определим по формуле $a_t = c - (a_1 + a_2 + \ldots + a_{t-1})$ с тем, чтобы $a_1 + a_2 + \ldots + a_t = c$.

Этот пример играет важную роль в разделении секрета в произвольных структурах доступа. Под структурой доступа понимают некоторую совокупность подмножеств множества всех участников. Эти подмножества называются разрешенными. Задача состоит в нахождении такого разделения секрета среди всех участников, чтобы его могли вычислить по общему согласию лишь участники разрешенных подмножеств.

Рассмотрим пример. Пусть $t = 4$, а разрешенными являются следующие подмножества участников:

$$\{A_1, A_2, A_4\}, \{A_1, A_3, A_4\}, \{A_2, A_3\}.$$

Пусть $c \subset Z_n$. Разделим этот секрет так:

A_1 знает a_1, b_1;

A_2 знает a_2, d_1;

A_3 знает $b_2, c - d_1$;

A_4 знает $c - (a_1 + a_2)$, $c - (b_1 + b_2)$.

Очевидно, что участники любого разрешенного подмножества без труда найдут секрет c. По существу, этот пример показывает, как построить разделение секрета в произвольной схеме доступа, включая пороговые схемы. При разделении секрета стараются, чтобы эти схемы были совершенными, т.е. такими, что участники неразрешенного множества не могут получить о секрете никакой информации, кроме априорной. Читателю рекомендуется определить, какие рассмотренные выше схемы являются совершенными. Не всегда схема разделения секрета, будучи совершенной, является экономной. Имеется в виду схема, у которой секреты участников имеют большие размеры относительно основного секрета. Изучение этого вопроса требует привлечения теории матроидов [9].

17.3. Стеганография и ее применение

Если криптологию трактовать в широком смысле, то стенография (steganography) — направление криптологии, имеющее целью сокрытие (hiding) сообщения в потоке передаваемой информации. Интересно проследить историю возникновения терминов «криптография» (от греч. kryptos — скрытый) и «стеганография» (от греч. steganos — прикрытый). Термин «криптография», обозначающий «секретное письмо», впервые использован в 1641 г. Дж. Уилкинсом — одним из основателей Лондонского королевского общества. При этом имелось в виду «открытое секретное письмо» (overt secret writing): открытое в том смысле, что очевидна зашифрованность письма. Термин «стеганография» введен в 1665 г. К. Шоттом. При этом имелось в виду «скрытое секретное письмо» (covert secret writing): скрытое в том смысле, что тайной является сам факт наличия секретного сообщения в потоке передаваемой информации.

В докомпьютерной стеганографии Ф.Л. Бауэр [46] выделяет два направления: лингвистическую и техническую стеганографию. Первая использует, нап-

ример, неопределенности написания букв, промежутков между словами для сокрытия сообщения. Классическим примером технической стеганографии является написание скрытого сообщения «молоком на белой бумаге».

Современная стеганография применяется для тайной передачи сообщений внутри «безобидных» данных, для скрытой маркировки данных с помощью «водяных знаков» (watermarking) [172]. Отметим, что «водяные знаки» используются для защиты электронных произведений (книг, музыки, видео) от пиратского копирования.

Качество стеганографии характеризуется тремя факторами: 1) обнаруживаемость (detectability); 2) робастность (robustness); 3) информационная емкость (bit rate). Обнаруживаемость — главная проблема тайной передачи информации, и поэтому передаваемая информация обычно предварительно шифруется. Робастность (устойчивость) ко всем ожидаемым типам обработки переданной информации (например, к фильтрации, усечению, масштабированию) является важнейшим требованием при изготовлении «водяных знаков». Информационная емкость — максимальное количество битов скрываемого сообщения, при котором передаваемый сигнал еще не получает заметных искажений. В качестве сигналов-носителей тайных сообщений в современной стаганографии обычно используются аудиосигналы (музыка, речь), а также изображения и видео.

Типичная схема стеганографической предачи сообщения представлена на рис. 17.3.

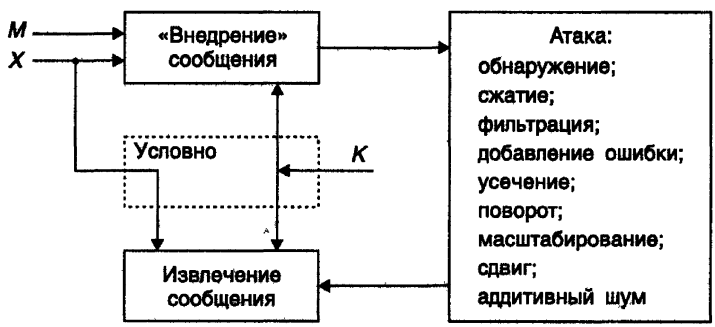


Рис. 17.3. Схема стеганографической передачи сообщения

Здесь X — исходный «безобидный» сигнал-носитель, который иногда называют контейнером; M — исходное сообщение, которое надо «спрятать» в сигнале X; K — ключ для шифрования сообщения. Сначала осуществляют шифрование:

$$m = E_k(M).$$

Затем зашифрованное сообщение m «незаметно» внедряется в сигнал X неким специальным способом (эти способы будут рассмотрены ниже):

$$X_m = f(X, m).$$

После этого «загруженный секретом» сигнал X_m пересылается по каналу связи, в котором этот сигнал может быть подвергнут различным атакам (см. рис. 18.3). В результате получатель принимает искаженный сигнал $\hat{X}_m$ и пытается выделить «спрятанное» сообщение M или хотя бы обнаружить его наличие. Использование ключа K является условным.

Остановимся на способах задания функции $f(\cdot)$, т.е. на способах «внедрения» сообщения m в сигнал-контейнер X. Выделяются два класса этих способов: пространственные (временные) и частотные (спектральные). Примером пространственного способа «внедрения» является способ, в котором m используется для изменения пикселей, номера которых определяет ключ K [49]. Другой способ состоит в модификации всех пикселей на «незаметную» величину [104]. Например, в способе LSB изменениям подвергаются младшие биты (Least Significant Bits).

Оказывается, что пространственные способы не позволяют достичь требуемого уровня «неприметности» искажений сигнала-контейнера X, поэтому наиболее широко используются частотные способы. При этом X вначале подвергается одному из спектральных преобразований (быстрое преобразование Фурье, дискретное косинус-преобразование, преобразование Уолша):

$$Y = g(X).$$

Затем сообщение m «внедряется» в массив коэффициентов преобразования Y одним из указанных выше пространственных способов:

$$Y_m = f(Y, m),$$

после чего осуществляет обратное спектральное преобразование:

$$X_m = g^{-1}(Y_m).$$

17.4. Активный криптоанализ

Криптографические алгоритмы реализуются в *устройствах* — интеллектуальных карточках, аппаратных комплексах защиты информации, программных криптографических модулях. Такие устройства выполняют криптографические преобразования с использованием секретных параметров, хранящихся в защищенной памяти. Устройства вполне могут быть доступны злоумышленнику для экспериментов и манипуляций, не санкционированных разработчиками.

Методы *инженерного криптоанализа* направлены на определение архитектуры устройства, содержимого защищенной памяти, реализованных в устройстве секретных алгоритмов. При этом исследователи пользуются дорогостоящими физическими приборами и инструментами: специальными микроскопами и микрозондами, устройствами лазерной резки, источниками ионных и электронных пучков. Методами инженерной криптографии был восстановлен, например, алгоритм поточного шифрования A5, реализованный в SIM-картах GSM [96].

Методы *активного криптоанализа* менее дорогостоящие и занимают промежуточное место между математическими и инженерными методами. Криптоаналитик

— *наблюдает* за выполнением алгоритма в аппаратном или программном устройстве: замеряет временные показатели и уровни напряжения, фиксирует флаги переноса и количество обращений к кэш-памяти, запоминает содержимое регистров в момент переключения между задачами процессора и др.;

— *кратковременно воздействует* на выполнение алгоритма: облучает устройство для изменения содержимого регистров, меняет тактовую частоту с целью внесения ошибок в порядок выполнения инструкций процессором и т.д.;

— *разрушающе воздействует* на выполнение алгоритма: обнуляет ультрафиолетовыми лучами ячейки постоянной памяти, отрезает проводники, ведущие к ячейкам рабочих регистров, и т.д.

Собранная информация и полученные при внесенных сбоях результаты выполнения алгоритма позволяют упростить решение математических задач криптоанализа.

Далее считаем, что в устройстве реализован следующий алгоритм возведения чисел $u \in \mathbb{Z}_n$, $\lceil \log_2 n \rceil = l$, в степень $d = (d_{l-1} \dots d_0)_2$.

1. Установить $w \leftarrow 1$, $v \leftarrow u$.
2. Для $i = 0, 1, \dots, l-1$ выполнить:
   а) если $d_i = 1$, то $w \leftarrow (wv) \bmod n$;
   б) $v \leftarrow v^2 \bmod n$.
3. Вернуть w.

Предположим, что устройство используется для подписи сообщений $x \in \mathbb{Z}_n$ по RSA. При этом n — составной модуль RSA, d — секретная экспонента, которая хранится в защищенной памяти устройства.

В классической постановке задачи криптоанализа злоумышленнику известны:

1) открытая экспонента e;
2) подписываемые сообщения x;
3) подписи $y = x^d \bmod n$.

Требуется определить секретный ключ d. Рассмотрим некоторые активные методы решения данной задачи.

Использование замеров времени. Пусть криптоаналитик располагает возможностью измерять время $R(x, d)$ вычисления подписи $x^d \bmod n$, а также оценивать время $R(x, \hat{d})$ для произвольной выбираемой экспоненты $\hat{d}$.

Для случайного $x \in \mathbb{Z}_n$ можно предположить, что время вычисления $R(x, e)$ является суммой независимых случайных величин

$$r_i = d_i \xi_i + \zeta_i, \quad i = 0, \dots, l-1,$$

где ξ_i, ζ_i — время выполнения шагов 2а и 2б алгоритма соответственно.

Пусть криптоаналитик подбирает оценку $\hat{d}_0$ бита d_0. Обозначим

$$\Delta(x, \hat{d}_0) = R(x, d) - R(x, \hat{d}_0).$$

Если $\hat{d}_0 = d_0$, то

$$\Delta(x, \hat{d}_0) = \sum_{i=1}^{l-1} r_i, \quad \mathbf{D}\{\Delta(x, \hat{d}_0)\} = \sum_{i=1}^{l-1} \mathbf{D}\{r_i\}.$$

Если же $\hat{d}_0 \neq d_0$, то

$$\Delta(x, \hat{d}_0) = \sum_{i=1}^{l-1} r_i \pm \xi_0, \quad \mathbf{D}\{\Delta(x, \hat{d}_0)\} = \sum_{i=1}^{l-1} \mathbf{D}\{r_i\} + \mathbf{D}\{\xi_0\}.$$

Таким образом, при изменении бита $\hat{d}_0$ с истинного значения на ложное дисперсия случайной величины $\Delta(x, \hat{d}_0)$ увеличивается на $\mathbf{D}\{\xi_0\}$.

Используя данные соображения, криптоаналитик проводит атаку по определению бита d_0 следующим образом.

1. Перехватываются сообщения x_t, $t = 1, \ldots, T$, и фиксируется время $R(x_t, d)$ их возведения в степень.

2. Для $b = 0, 1$ формируются выборки $\Delta(x_1, b), \ldots, \Delta(x_T, b)$ и определяются их выборочные дисперсии s_b^2.

3. Если $s_0^2 > s_1^2$, то искомая оценка $\hat{d}_0$. Если же $s_0^2 < s_1^2$, то $\hat{d}_0 = 1$.

Определив d_0, далее таким же образом можно определить бит d_1, затем $d_2, \ldots, d_{l-1}$.

Использование временных сбоев. Пусть при вычислении подписи $x^d \bmod n$ в устройстве происходит единичный сбой — на i-й итерации алгоритма в регистре $w = (w_{l-1} \ldots w_0)_2$ изменяется значение бита w_j. Будем считать, что i и j — случайные величины с равномерным распределением на множестве $\{0, 1, \ldots, l-1\}$.

Обозначим

$$d[i] = (d_{i-1} \ldots d_0)_2, \quad D[i] = (d_{l-1} \ldots d_i 0 \ldots 0)_2$$

(считаем, что $d[0] = D[l] = 0$). Тогда истинная подпись

$$y = x^{d[i] + D[i]} \bmod n,$$

а подпись, полученная при сбое,

$$y' = (x^{d[i]} \pm 2^j)\, x^{D[i]} \bmod n = (y \pm 2^j x^{D[i]}) \bmod n.$$

Криптоанализ проводится следующим образом.

1. Фиксируются сообщения x_t, $t = 1, \ldots, T$ и соответствующие им сбойные подписи y'_t.

2. Для $\hat{i} = l-1, l-2, \ldots$

   а) выбираются все возможные номера $\hat{j}$ бита ошибки;

   б) выбираются все возможные значения $\hat{D}[\hat{i}]$;

   в) для всех $t = 1, \ldots, T$ проверяется совпадение

$$\left(y'_t \mp 2^{\hat{j}} x_t^{\hat{D}[\hat{i}]}\right)^e \bmod n = x_t.$$

3. Если совпадение найдено, то с большой вероятностью $\hat{D}[\hat{i}] = D[\hat{i}]$ и определены $1 - \hat{i}$ битов секретной экспоненты d.

Пусть $i_1, \ldots, i_T$ — неизвестные номера сбойных итераций, упорядоченные по неубыванию. Тогда

$$\mathbf{P}\{i_T < l - m\} = 1 - \left(1 - \frac{m}{l}\right)^T, \quad 0 \le m < l,$$

и на шаге 2 с такой вероятностью потребуется перебрать не более 2^m значений $\hat{D}[\hat{i}]$. Ясно, что атаку можно продолжить и восстановить следующую порцию битов d.

Использование необратимых сбоев. Пусть секретная экспонента хранится в защищенной памяти EEPROM (Electrically Erasable Programmable Read-Only Memory) как бинарное слово θ_0. Такая память является асимметричной — при воздействии на нее ультрафиолетовым излучением заряд (бит ключа) скорее обнулится, чем примет значение 1.

Криптоаналитик действует следующим образом.

1. Задает произвольное сообщение $x \in \mathbb{Z}_n^*$ и получает значение подписи $y = x^d \bmod n$.

2. Для $t = 1, 2, \ldots$ криптоаналитик облучает устройство, формируя таким образом ключ θ_t, отличающийся от θ_{t-1} на небольшое количество битов. Затем для сообщения x снова вырабатывается подпись y_t, но уже на ключе θ_t. Вычисления проводятся до тех пор, пока не встретится подпись $y_T = y_{T+1} = \ldots = 1$.

3. Веса Хэмминга (количества ненулевых битов) ключей удовлетворяют цепочке равенств

$$\mathrm{w}(\theta_0) \ge \mathrm{w}(\theta_1) \ge \ldots \ge \mathrm{w}(\theta_{T-1}) > \mathrm{w}(\theta_T) = 0.$$

Для $t = T - 1, \ldots, 0$ ключ θ_t ищется в малой окрестности ключа θ_{t+1}. Для каждого кандидата $\hat{\theta}_t$ проверяется совпадение соответствующей подписи $\hat{y}_t$ со значением y_t.

SXJKN8EYHR3I4HR987YDWFUIHW9E
8YR384H5O34596EUF9O8WE76R83Y
439846UHFE98YDSFJH8374H4BR82
76RUBNWDF87Y6234BN9WE872H3WR
TJOJRTGM905JRO9G70954TOIRF7E
RFH2347FQ7WERF9UJ1349YUR7DFK
JHGJDFBVDGFJVBDSKFJHKJFVKBDF
KJVJKGFKJSBDFJBVKJDSGFJKVBKS

Приложение А

ТАБЛИЦЫ

SXJKN8EYHR3I4HR987YDWFUIHW9E
8YR384H5O34596EUF9O8WE76R83Y
439846UHFE98YDSFJH8374H4BR82
76RUBNWDF87Y6234BN9WE872H3WR
TJOJRTGM905JRO9G70954TOIRF7E
RFH2347FQ7WERF9UJ1349YUR7DFK
JHGJDFBVDGFJVBDSKFJHKJFVKBDF
KJVJKGFKJSBDFJBVKJDSGFJKVBKS

Таблица А.1

Большие числа и астрономические величины

Астрономическая величина	Число
Возраст Земли, лет	2^{30}
Возраст Вселенной, лет	2^{34}
Число атомов на Земле	2^{170}
Число атомов на Солнце	2^{190}
Число атомов во Вселенной	2^{265}
Объем Вселенной, см	2^{280}

Таблица А.2

Большие числа

n	2^n	n	2^n
10	1024	120	$1.32923 \cdot 10^{36}$
20	$1.04858 \cdot 10^{6}$	140	$1.3938 \cdot 10^{42}$
30	$1.07374 \cdot 10^{9}$	160	$1.4615 \cdot 10^{48}$
40	$1.09951 \cdot 10^{12}$	180	$1.5325 \cdot 10^{54}$
50	$1.1259 \cdot 10^{15}$	200	$1.60694 \cdot 10^{60}$
60	$1.15292 \cdot 10^{18}$	220	$1.685 \cdot 10^{66}$
70	$1.18059 \cdot 10^{21}$	240	$1.76685 \cdot 10^{72}$
80	$1.20893 \cdot 10^{24}$	260	$1.85267 \cdot 10^{78}$
90	$1.23794 \cdot 10^{27}$	280	$1.94267 \cdot 10^{84}$
100	$1.26765 \cdot 10^{30}$	300	$2.03704 \cdot 10^{90}$

Таблица А.3

Точные значения больших чисел

n	2^n	n	2^n
1	2	33	8589934592
2	4	34	17179869184
3	8	35	34359738368
4	16	36	68719476736
5	32	37	137438953472
6	64	38	274877906944
7	128	39	549755813888
8	256	40	1099511627776
9	512	41	2199023255552
10	1024	42	4398046511104
11	2048	43	8796093022208
12	4096	44	17592186044416
13	8192	45	35184372088832
14	16384	46	70368744177664
15	32768	47	140737488355328
16	65536	48	281474976710656
17	131072	49	562949953421312
18	262144	50	1125899906842624
19	524288	51	2251799813685248
20	1048576	52	4503599627370496
21	2097152	53	9007199254740992
22	4194304	54	18014398509481984
23	8388608	55	36028797018963968
24	16777216	56	72057594037927936
25	33554432	57	144115188075855872
26	67108864	58	288230376151711744
27	134217728	59	576460752303423488
28	268435456	60	1152921504606846976
29	536870912	61	2305843009213693952
30	1073741824	62	4611686018427387904
31	2147483648	63	9223372036854775808
32	4294967296	64	18446744073709551616

SXJKN8EYHR3I4HR987YDWFUI
HW9E8YR384H5O34596EUF908
WE76R83Y439846UHFE98YDSF
JH8374H4BR8276RUBNWDF87Y
6234BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7ERFH2
347FQ7WERF9UJ1349YUR7DFK
JHGJDFBVDGFJVBDSKFJHKJFV
KBDFKJVJKGFKJSBDFJBVKJDS
GFJKVBKSDJBFKVBDGFVAKDLK

Приложение Б

АРХИВ ДИСКРЕТНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

SXJKN8EYHR3I4HR987YDWFUI
HW9E8YR384H5O34596EUF908
WE76R83Y439846UHFE98YDSF
JH8374H4BR8276RUBNWDF87Y
6234BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7ERFH2
347FQ7WERF9UJ1349YUR7DFK
JHGJDFBVDGFJVBDSKFJHKJFV
KBDFKJVJKGFKJSBDFJBVKJDS
GFJKVBKSDJBFKVBDGFVAKDLK

Архив содержит 18 файлов дискретных последовательностей (модельных и реальных), которые используются при выполнении заданий данного практикума. Этот архив можно скачать со страницы www.wnk.biz/crypto.

1. **File00** — «эталонная чисто случайная» двоичная последовательность Марсальи $x_t \in \{0, 1\}$ длиной 65536 элементов.
2. **File01** — двоичная псевдослучайная последовательность, порожденная С-генератором rand, длиной 65536 элементов.
3. **File02.1** — двоичная стационарная цепь Маркова длиной 65536 элементов с матрицей вероятностей одношаговых переходов $P(0.2)$, где

$$P(\varepsilon) = \begin{pmatrix} 0.5+\varepsilon & 0.5-\varepsilon \\ 0.5-\varepsilon & 0.5+\varepsilon \end{pmatrix}, \quad -0.5 \le \varepsilon \le 0.5.$$

4. **File02.2** — двоичная стационарная цепь Маркова длиной 65536 элементов с матрицей переходов $P(0.1)$.
5. **File02.3** — двоичная стационарная цепь Маркова длиной 65536 элементов с матрицей переходов $P(0.05)$.
6. **File02.4** — двоичная стационарная цепь Маркова длиной 65536 элементов с матрицей переходов $P(0.01)$.
7. **File02.5** — двоичная стационарная цепь Маркова длиной 65536 элементов с матрицей переходов $P(0.001)$.
8. **File02.6** — двоичная стационарная цепь Маркова длиной 65536 элементов с матрицей переходов $P(0.0001)$.
9. **File03.1** — двоичная линейная рекуррента порядка $s = 7$: $x_t = x_{t-3} \oplus x_{t-7}$, $t = 1, 2, \ldots$, длиной 65536 элементов.
10. **File03.2** — двоичная линейная рекуррента порядка $s = 15$: $x_t = x_{t-7} \oplus x_{t-15}$, $t = 1, 2, \ldots$, длиной 65536 элементов.
11. **File03.3** — двоичная линейная рекуррента порядка $s = 31$: $x_t = x_{t-13} \oplus x_{t-31}$, $t = 1, 2, \ldots$, длиной 65536 элементов.
12. **File04.1** — бинарная авторегрессия BAR(7) длиной 65536 элементов порядка $s = 7$: $x_t = x_{t-3} \oplus x_{t-7} \oplus \xi_t$, $t = 1, 2, \ldots$, где $\xi_t \in \{0, 1\}$ — последовательность независимых одинаково распределенных случайных величин Бернулли: $\mathbf{P}\{\xi_t = 0\} = 0.6$, $\mathbf{P}\{\xi_t = 1\} = 0.4$.
13. **File04.2** — бинарная авторегрессия BAR(15) длиной 65536 элементов порядка $s = 15$: $x_t = x_{t-7} \oplus x_{t-15} \oplus \xi_t$, $t = 1, 2, \ldots$, где $\xi_t \in \{0, 1\}$ — последовательность независимых одинаково распределенных случайных величин Бернулли: $\mathbf{P}\{\xi_t = 0\} = 0.6$, $\mathbf{P}\{\xi_t = 1\} = 0.4$.

14. **File04.3** — бинарная авторегрессия BAR(31) длиной 65536 элементов порядка $s = 31$: $x_t = x_{t-13} \oplus x_{t-31} \oplus \xi_t$, $t = 1, 2, \ldots$, где $\xi_t \in \{0, 1\}$ — последовательность независимых одинаково распределенных случайных величин Бернулли: $\mathbf{P}\{\xi_t = 0\} = 0.6$, $\mathbf{P}\{\xi_t = 1\} = 0.4$.
15. **File05** — двоичная последовательность длиной 65536 элементов, представляющая собой фрагмент оцифрованного музыкального файла.
16. **File06** — двоичная последовательность длиной 65536 элементов, представляющая собой фрагмент оцифрованного изображения.
17. **File07** — «эталонная чисто случайная» N-значная последовательность $x_t \in \{0, 1, \ldots, N-1\}$ длиной 65536 элементов при $N = 8$.
18. **File08** — стационарная цепь Маркова с $N = 8$ состояниями $V = 0, 1, \ldots, 7$ длиной 65536 элементов с матрицей вероятностей одношаговых переходов

$$P = (p_{ij}), \quad p_{ii} = \frac{1}{2}, \quad p_{ij} = \frac{1}{14} \quad (i \neq j), \quad i, j = 1, \ldots, 8.$$

SXJKN8EYHR3I4HR987YD
WFUIHW9E8YR384H5O345
96EUF9O8WE76R83Y4398
46UHFE98YDSFJH8374H4
BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7E
RFH2347FQ7WERF9UJ134
9YUR7DFKJHGJDFBVDGFJ
VBDSKFJHKJFVKBDFKJVJ

Приложение В

КОМПЬЮТЕРНЫЙ ПРАКТИКУМ «КРИПТОЛАБОРАТОРИЯ»

SXJKN8EYHR3I4HR987YD
WFUIHW9E8YR384H5O345
96EUF9O8WE76R83Y4398
46UHFE98YDSFJH8374H4
BR8276RUBNWDF87Y6234
BN9WE872H3WRTJOJRTGM
905JRO9G70954TOIRF7E
RFH2347FQ7WERF9UJ134
9YUR7DFKJHGJDFBVDGFJ
VBDSKFJHKJFVKBDFKJVJ

Общие сведения

Компьютерный практикум «КриптоЛаборатория» (КЛ) предназначен для изучения и исследования в рамках учебного процесса широкого класса криптографических систем.

Для установки и работы с КЛ требуется:

— компьютер под управлением операционной системы Windows 95, Windows NT 4.0 и выше;

— 4 Мб свободного места на жестком диске;

— 8 Мб оперативной памяти.

КЛ состоит из трех взаимосвязанных компонентов: рабочего плана (проекта) криптосистемы, инструментов для работы с ней и протокола сеанса работы.

Дополнительно с КЛ поставляется мультимедийный файл справки *КриптоАльбом*, в котором описываются некоторые блочные криптосистемы, а также простейшие докомпьютерные шифры: Цезаря, Вернама, модулярный, гомофонический и т.д.

Рабочий план «КриптоЛаборатории»

Рабочий план (*проект*) содержит информацию о типе исследуемой криптосистемы, режиме ее работы, виде криптопреобразования и о параметрах преобразования.

В КЛ организована поддержка шести *режимов* работы:

1) простой замены;
2) гаммирования;
3) цепной обработки;
4) гаммирования с обратной связью;
5) выработки значения функции хэширования;
6) выработки (проверки) электронной цифровой подписи.

Каждая из реализованных в КЛ криптосистем может поддерживать один или несколько из перечисленных режимов и выполнять один из *видов* преобразования (за исключением режима выработки значения функции хэширования):

— шифрование или выработка электронной цифровой подписи;

— расшифровывание или проверка электронной цифровой подписи.

Параметры криптосистемы в КЛ делятся на *долговременные* (ключевая информация длительного использования) и *кратковременные* (ключевая информация разового, сеансового использования).

Перечислим проекты КЛ.

Блочная криптосистема DES. Предусмотрена возможность изменения количества тактов d криптосистемы DES. Если заданное пользователем количество $d < 16$, то используются первые d тактовых ключей, рассчитанных по расписанию DES.

Режимы — простой замены, гаммирования, цепной обработки, гаммирования с обратной связью.

Долговременный параметр — количество тактов d, $1 \le d \le 16$.

Кратковременные параметры — ключ $\theta \in V_{56}$ и синхропосылка $Y_0 \in V_{64}$ (не используется в режиме простой замены).

Для задания θ пользователь вводит 64 бита, каждый 8-й бит игнорируется. Контроль четности байтов, предусмотренный в DES, не осуществляется. В режимах простой замены и цепной обработки преобразованию подвергаются полные 8-байтовые блоки данных. Последний неполный блок не обрабатывается.

Универсальная блочная система (УБС). УБС — d-тактовая, $1 \le d \le 16$, блочная криптосистема Фейстеля, действующая на множестве V_{64}. Ключ $\theta \in V_{32d}$ является конкатенацией $\kappa_1 \| \kappa_2 \| \ldots \| \kappa_d$ тактовых ключей $\kappa_1, \kappa_2, \ldots, \kappa_d \in K$.

Действие тактовой функции $f_\kappa : V \to V_{32}$ задается правилом

$$f_\kappa(X) = S(X \otimes \kappa) \lll r,$$

где $\otimes$ — бинарная операция на V_{32} ($\oplus$ или $\boxplus$), $r \in \{0, \ldots, 31\}$;

$$S(x_1 \| x_2 \| \ldots \| x_8) = S_1(x_1) \| S_2(x_2) \| \ldots \| S_8(x_8),$$

$$x_i \in V_4, \quad S_i \in \mathcal{S}(V_4).$$

Например, при $\otimes = \boxplus$ и $r = 11$ получаем тактовую функцию ГОСТ 28147—89.

Режимы — простой замены, гаммирования, цепной обработки, гаммирования с обратной связью.

Долговременные параметры — количество тактов d, операция $\otimes$, S-блоки $S_1, \ldots, S_8$, величина циклического сдвига r.

Кратковременные параметры — тактовые ключи $\kappa_1, \ldots, \kappa_d$ и 64-битовая синхропосылка (не используется в режиме простой замены).

В режимах простой замены и цепной обработки преобразованию подвергаются только полные 8-байтовые блоки данных.

Поточная криптосистема, построенная на генераторе линейной рекуррентной последовательности (ЛРП). В качестве гаммы поточной криптосистемы используется ЛРП с заданным характеристическим многочленом $f(\lambda) \in \mathbb{F}_2[\lambda]$, $\deg f(\lambda) < 32$.

Режим — гаммирования.

Долговременный параметр — характеристический многочлен $f(\lambda)$.

Кратковременный параметр — n-вектор стартового состояния ЛРП.

Многочлен $f(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$ задается двоичным словом $1a_{n-1} \cdots a_1a_0$, записанным в шестнадцатеричной системе счисления. Например, многочлену $\lambda^3 + \lambda + 1$ соответствует слово $(1011)_2 = (B)_{16}$.

Криптосистема с открытым ключом RSA. Используются модули n криптосистемы RSA из диапазона $2^{95} < n < 2^{4096}$. Реализована процедура генерации модулей.

Режим — простой замены.

Долговременные параметры — модуль n, открытая экспонента e, секретная экспонента d.

Внешняя криптосистема. Проект позволяет подключать программные реализации криптосистем, удовлетворяющие требованиям КЛ и организованные в виде библиотек динамической компоновки (DLL) операционной системы Windows. Требования к реализациям задает специальный программный интерфейс, описание которого поставляется вместе с КЛ.

Для подключения реализации пользователю следует указать имя динамической библиотеки и трех функций из этой библиотеки, реализующих: а) инициализацию контекста шифрования; б) криптопреобразование с использованием контекста; в) завершение работы с контекстом шифрования. *Контекст шифрования* — это структура данных, содержащая информацию о параметрах криптосистемы, режиме работы и типе преобразования.

Инструменты «КриптоЛаборатории»

Инструменты предназначены для решения заданий КЛ. К инструментам относятся утилиты и КриптоСтудия.

Утилиты предоставляют пользователю возможность выполнения некоторых сопутствующих действий при работе с КЛ. *Утилита проверки на простоту* числа n использует тест Ферма на простоту с основаниями 2, 3. Предварительно проверяется делимость n на простые числа, не превосходящие 1009. *Утилита генерации простых чисел* позволяет получать простые числа p из диапазона $2^{31} < p < 2^{2056}$. *Утилита псевдослучайной генерации* предназначена для получения файлов псевдослучайных чисел.

Инструмент *КриптоСтудия* является основным средством решения заданий КЛ и предназначен для проведения вычислительных экспериментов с криптосистемой. В рамках этих экспериментов *криптоаналитик* может:

— изменять параметры θ выбранного криптопреобразования F_θ, доступ к которым разрешен *криптографом*;

— задавать входные данные X;

— выполнять криптопреобразование $Y = F_\theta(X)$;

— наблюдать и проводить анализ выходных данных Y.

Подготовка и обработка входных (выходных) данных осуществляется на *панелях* Источник (Назначение) КриптоСтудии. Панели можно переставлять, изменяя тем самым направление криптопреобразования. На каждой панели реализованы следующие функции.

1. Генерация данных с заданными статистическими свойствами. Включены алгоритмы:

— генерации псевдослучайной последовательности битов, тетрад битов, байтов;

— генерации однородной марковской цепи байтов (переходные вероятности цепи соответствуют русскому или английскому языку и, кроме того, могут быть оценены по заданному пользователем файлу данных).

2. Проверка статистических свойств набора данных. Реализованы:

— тест многомерной дискретной равномерности по непересекающимся отрезкам;

— тест многомерной дискретной равномерности по пересекающимся отрезкам;

— покер-тест;
— тест серий.
3. Редактирование данных. Поддерживаются операции:
— создания, чтения, сохранения, изменения данных;
— вставки (удаления) заданного количества байтов в набор данных;
— изменения способа представления данных.

Задания «КриптоЛаборатории»

С проектом КЛ может быть связано некоторое задание. Сформулировать его в ходе создания проекта может каждый пользователь, выступающий при этом в роли криптографа.

КЛ снабжен средствами доступа к кратковременным и долговременным параметрам проекта — криптограф в ходе создания проекта имеет возможность «защитить», т.е. сделать недоступными для других пользователей, те или иные параметры. Проект вместе с заданием сохраняется в дисковом файле. Информация о созданном проекте отражается в протоколе сеанса работы с КЛ. Другие пользователи КЛ, выступающие в роли криптоаналитиков, имеют доступ к открытой информации проекта, а также к самому криптопреобразованию. Это означает, что криптоаналитик с помощью средств КЛ может выполнять криптопреобразование над произвольными данными, т.е. проводить атаку при выбираемом открытом тексте.

Задания КЛ бывают двух типов — исследовательские и практические. В ходе выполнения заданий первого типа криптоаналитику предлагается исследовать криптосистемы с заданными в проекте параметрами: установить какие-либо свойства, выявить закономерности, выполнить статистические тесты и т.д. При выполнении практического задания криптоаналитик с помощью средств КЛ должен восстановить некоторые параметры криптосистемы, т.е. провести частичный или полный криптоанализ. Ответ криптоаналитика проверяется в КЛ.

Приведем некоторые задания, поставляемые с КЛ.

1. При помощи средств КЛ исследовать влияние изменения одного бита в открытом тексте на результат криптопреобразования УБС. Рассмотреть режим простой замены, гаммирования и гаммирования с обратной связью.

2. Восстановить ключ двухтактовой УБС в режиме простой замены.

3. При помощи средств КЛ установить свойство дополнения для DES.

4. Исследовать влияние выбора количества тактов DES на статистические свойства зашифрованного текста. Открытый текст генерируется как однородная цепь Маркова с переходными вероятностями, соответствующими английскому языку. Исследовать режим простой замены с количеством тактов 1, 2.

5. Выполнить все тесты дискретной равномерности, включенные в КЛ. В качестве исходных данных использовать зашифрованный по RSA файл rsa.txt, поставляемый с КЛ. Для какого теста p-значение минимально?

6. Зашифровать в режимах простой замены, гаммирования, гаммирования с обратной связью и цепной обработки 8096 нулевых байтов. Для полученных последовательностей выполнить покер-тест с параметрами: длина отрезка 8, уровень значимости 0.05. Для каких режимов гипотеза о дискретной равномерности не выполняется?

SXJKN8EYHR3I4HR987YDWFUIHW9E
8YR384H5O34596EUF9O8WE76R83Y
439846UHFE98YDSFJH8374H4BR82
76RUBNWDF87Y6234BN9WE872H3WR
TJOJRTGM905JRO9G70954TOIRF7E

ЛИТЕРАТУРА

SXJKN8EYHR3I4HR987YDWFUIHW9E
8YR384H5O34596EUF9O8WE76R83Y
439846UHFE98YDSFJH8374H4BR82
76RUBNWDF87Y6234BN9WE872H3WR
TJOJRTGM905JRO9G70954TOIRF7E

1. *Айвазян С.А., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика. М.: Статистика, 1983.
2. *Андерсон Т.В.* Статистический анализ временных рядов. М.: Мир, 1984.
3. *Биркгоф Г., Барти Т.* Современная прикладная алгебра. М.: Мир, 1976.
4. *Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики. М.: Наука, 1983.
5. *Боровков А.А.* Математическая статистика. М.: Наука, 1988.
6. *Брассар Ж.* Современная криптология. М.: Полимед, 1999.
7. *Брикелл Э.Ф., Одлижко Э.М.* Криптоанализ: Обзор новейших результатов // ТИИЭР, 1988. Т. 76. № 5. С. 75—93.
8. *Бэйкер А.* Введение в теорию чисел. Мн.: Вышэйш. шк., 1995.
9. Введение в криптографию / Под ред. В.В. Ященко. М.: МЦНМО, 2001.
10. *Верн Ж.* Жангада. Мн.: Университетское, 1992.
11. *Герасименко В.А., Малюк А.А.* Основы защиты информации: Учеб. пособие. М.: МИФИ, 1997.
12. *Жельников В.* Криптография от папируса до компьютера. М.: ABF, 1996.
13. *Кейперс Л., Нидеррайтер Г.* Равномерное распределение последовательностей. М.: Наука, 1985.
14. *Кнут Д.* Искусство программирования: В 3 т. М.: Мир, 1992.
15. *Колмогоров А.Н., Успенский В.А.* Алгоритмы и случайность // Теория вероятностей и ее применения. 1987. Т. 32. Вып. 3. С. 425—455.
16. *Левитан Ю.Л., Соболь И.М.* О датчике псевдослучайных чисел для персональных компьютеров // Мат. моделирование. 1990. Т. 2. № 8. С. 119—126.
17. *Лидл Р., Нидеррайтер Г.* Конечные поля: В 2 т. М.: Мир, 1988.
18. *Мафтик С.* Механизмы защиты в сетях ЭВМ. М.: Мир, 1993.
19. Международные банковские стандарты. ИСО 8731.
20. *Мельников В.В.* Защита информации в компьютерных системах. М.: Финансы и статистика, 1997.
21. *Молдовян Н.А.* Проблематика и методы криптологии. СПб.: Изд-во СПбГУ, 1998.
22. *Молдовян Н.А.* Скоростные и блочные шифры. СПб.: Изд-во СПбГУ, 1998.
23. *Мур Дж.* Несостоятельность протоколов криптосистем // ТИИЭР, 1988. Т. 76. № 5. С. 94—104.
24. *Нечаев В.И.* Элементы криптографии. Основы теории защиты информации. М.: Высш. шк., 1999.
25. *Ноден П., Китте К.* Алгебраическая алгоритмика. М.: Мир, 1999.
26. *Орлов В.А., Филиппов Л.И.* Теория информации в упражнениях и задачах. М.: Высш. шк., 1976.

27. *Прудников А.П., Брычков Ю.А., Маричев О.И.* Интегралы и ряды. М.: Наука, 1981.
28. *Сачков В.Н.* Введение в комбинаторные методы дискретной математики. М.: Наука, 1982.
29. Справочник по теории вероятностей и математической статистике / Под ред. В.С. Королюка. М.: Наука, 1992.
30. *Стратонович Р.Л.* Теория информации. М.: Наука, 1975.
31. *Сяо Д., Керр Д., Мэдник С.* Защита ЭВМ. М.: Мир, 1982.
32. *Тарасенко Ф.П.* Введение в курс теории информации. Томск: ТГУ, 1973.
33. *Феллер В.* Введение в теорию вероятностей и ее применения: В 2 т. М.: Наука, 1984.
34. *Харин Ю.С., Агиевич С.В.* Компьютерный практикум по математическим методам защиты информации. Мн.: БГУ, 2001.
35. *Харин Ю.С., Берник В.И., Матвеев Г.А.* Математические основы криптологии. Мн.: БГУ, 1999.
36. *Харин Ю.С.* и др. Основы имитационного и статистического моделирования. Мн.: ДизайнПРО, 1997.
37. *Харин Ю.С., Мартиневский А.В.* Обнаружение n-мерной равномерности в двоичных последовательностях с использованием экстремальных статистик // Pattern Recognition and Information Processing. 1999. V. 2. P. 358—362.
38. *Харин Ю.С., Степанова М.Д.* Практикум на ЭВМ по математической статистике. Мн.: Университетское, 1987.
39. *Хинчин А.Я.* Цепные дроби. М.: Физматгиз, 1964.
40. *Adleman L.* A subexponential algorithm for the discrete logarithm problem with applications to cryptography // IEEE 18th Ann. Symp. on Found. of Comp. Sc. 1979. P. 55—60.
41. *Adleman L., Huang M.* Recognition primes in random polynomial time // Proc. 19th. Ann. ACM 1987. P. 462—469.
42. *Agnew G., Mullin R., Onyschuk T., Vaustone S.* An implementation for a fast Public-Key cryptosystem // J. of Cryptology, 1991. P.63—79.
43. *Agnew G., Mullin R., Vaustone S.* An interactive data exchange protocol based on discrete exponentiation // Adv. in Cryptology: Proc. of Eurocrypt '88, LNCS 330. 1988. P. 159—166.
44. *Arvillias A., Maritsas A.* Partiting the period of m-sequences and applications to pseudorandom number generation // J. ACM, V. 25. 1978. P. 675—686.
45. *Barrett P.* Implementing the Riviest Shamir and Adleman public key encryption algorithm on a standard digital signal processor // Adv. in Cryptology: Proc. of Crypto '86, LNCS 263. 1987. P. 311—323.
46. *Bauer F.* Decrypted Secrets: methods and maxims of cryptology. N.Y.: Springer, 1997.
47. *Beker H., Piper F.* Cipher Systems. London: Northwod Books, 1982.
48. *Ben-Aroya I., Biham E.* Differential cryptanalysis of Lucifer // J. of Cryptology. 1996. V. 9. № 4. P. 21—34.
49. *Bender W., Gruhl D., Morimoto N., Lu A.* Techniques for data hiding // IBM Syst. J., 1996. V. 35. № 3—4. P. 313—336.
50. *Bennet C.H., Brassard G., Breidbart S., Wiesner S.* Quantum cryptography, or unfrgeable subway tokens // Proc. of Crypto '82. N.Y.: Plenum Press, 1982. P. 267—275.

51. *Beth T., Piper F.C.* The stop-and-go generator // Adv. in Cryptology: Proc. of Eurocrypt '84, 1985. LNCS 209. P. 88—92.
52. *Biham E.* New Types of Cryptanalytic Attack Using Related Keys // J. of Cryptology. 1994. V. 7. № 4. P. 229—246.
53. *Biham E., Shamir A.* Differential cryptanalysis of DES-like cryptosystems // J. of Cryptology. 1991. V. 4. P. 3—72.
54. *Biham E., Shamir A.* Differential Cryptanalysis of the Full 16-round DES // Adv. in Cryptology: Proc. of Crypto '92. N.Y.: Springer, 1993. P. 487—496.
55. *Biham E., Shamir A.* Differential Fault Analysis of Secret Key Cryptosystems // Adv. in Cryptology: Proc. of Crypto '97, LNCS 1294. 1997. P. 513—525.
56. *Blake I., Seroussi G., Smart N.* Elliptic Curves in Cryptography. Cambridge Univ. Press, 1999.
57. *Bocma W.* Primality testing using elliptic curves // Report 85—12. Amsterdam, 1985.
58. *Boer B., Bosselaers A.* Collisions for the Compresion Function of MD5// Adv. in Cryptology. N.Y.: Springer, 1994. P. 293—304.
59. *Brassard G.* Modern Cryptology // LNCS 325. Berlin: Springer, 1988.
60. *Brent R.P.* An improved Monte Carlo factorization algorithm // BIT. 1980. V. 20. P. 176—184.
61. *Bright H., Enison R.* Quasi-random number sequences from a long perod TLP generator with remarks on applications to cryptography // Comp. Surveys. 1979. V. 11. № 2. P. 357—370.
62. *Brown M., Solomon H.* On combining pseudorandom number generators // Technical Report. Dept. of Statistics, Stanford University. № 233, 1976.
63. *Brüer J.O.* On pseudo random sequences as crypto generators // Proc. of the International Zurich Seminar on Dig. Com. Switzerland, 1984.
64. *Campbell K.W., Wiener M.J.* DES is not a group // Adv. in Cryptology: Proc. of Crypto '92, LNCS 740. 1992. P. 512—520.
65. *Chambers W.G., Gollmann D.* Lock-in effect in cascades of clock-controlled shift-registers // Adv. in Cryptology: Proc. of Eurocrypt '88, LNCS 330. 1988. P. 331—343.
66. *Chassaing P.* An optimal random number generator on $\mathbb{Z}_p$ // Statistics & Probability Letters. 1989. V. 7. P. 307—309.
67. *Chin Y., Wang P., Hwang J.* Cryptanalysis on Schwartz graphical encryption method // Cryptologia. 1993. V. 17. № 3. P. 301—304.
68. *Chudnovsky P., Chudnovsky G.* Sequeus of numbers and new primality and factorization tests // Adv. in Appl. Math. 1986. V. 7. P. 187—237.
69. *Cillogly J.* Ciphertext — only cryptanalysis of ENIGMA // Cryptologia. 1995. V. 19. № 4. P. 405—413.
70. *Cohen H.* A Course in Computational Algebraic Number Theory. GTM 138, Springer, 1993.
71. *Coppersmith D.* Another birthday attack // Adv. in Cryptology: Proc. of Crypto '85, LNCS 218. 1986. P. 14—17.
72. *Coppersmith D., Krawchuk Y., Mansour Y.* The shrinking generator // Adv. in Cryptology: Proc. of Crypto '93, LNCS 773. 1994. P. 22—39.
73. *Coppersmith P.* Fast evaluation of logarithms on fields of character two // IEEE Inf. Theory. 1984. V. 30. P. 587—594.
74. *Couvreur C., Quisquater J.* Fast decipherment algorithm for RSA public-key cryptosystem // Electron. Lett. 1982. V. 18. № 21. P. 905—907.

75. Cryptography. Proc. of the Workshop on Cryptography // LNCS. V. 149. N.Y.: Springer, 1983.
76. Cryptology: Yesterday, Today, and Tomorrow / Ed. Deavours C. Artech House, Norwood, 1987.
77. *Damgard I.B.* A design principle for hash functions // Adv. in Cryptology: Proc. of Crypto '89, LNCS 435. 1990. P. 416—427.
78. Data Encryption Standard. FIPS Publication 46. — Springfield: Nat. Institute of Standards and Technology, 1977.
79. *Deavours C., Kruh L.* Machine Cryptography and Modern Cryptanalysis. Dedham, MA: Artech House, 1985.
80. *Denning D.* Cryptography and Data Security. Reading, MA: Addison-Wesley, 1982.
81. *Denning D., Sacco G.* Timestamps in Key Distribution Protocols // Comm. ACM. 1981. V. 24. № 4. P. 533—536.
82. *Diffie W., Hellman M.* New directions in cryptography // IEEE Trans. Inf. Theory. 1976. V. 22. № 11. P. 644—654.
83. *Dillon J.F.* A survey of bent functions // The NSA Technical J. 1972. P. 191—215.
84. *Dixon B., Lenstra A.K.* Factoring Integers Using SIMD Sieves // Adv. in Cryptology: Proc. of Eurocrypt '93, LNCS 765. 1993. P. 28—39.
85. *Dixon B., Lenstra A.K.* Massively parallel eliptic curve factoring // Adv. in Cryptology: Proc. of Eurocrypt '92, LNCS 658. 1992. P. 183—193.
86. *Dixon J.* Assymptotically fast factorization of integers // Math. Comp. 1981. V. 36. № 3. P. 255—260.
87. *Ehrsam W., Matyas S., Meyer C., Tuchman W.* A cryptographic key management scheme for implementing the Data Encryption Standard // IBM Syst. J. 1978. V. 17. № 2. P. 106—124.
88. *Eichenauer J., Lehn J., Topuzoglu A.* A nonlinear congruential pseudorandom number generator with power of two modulus // Mathematics of Comp. 1988. V. 51. P. 757—759.
89. *ElGamal T.* A public key cryptosystem and a signature scheme based on discrete logarithms // IEEE Trans. on Inf. Theory. 1985. V. 31. P. 469—472.
90. *Elias P.* The efficient construction of an unbiased random sequence // The Math. Statistics. 1972. V. 43. P. 865—870.
91. *Feistel H., Notz W.A., Smith J.L.* Some cryptographic techniques for machine-to-machine data communications // Proc. of the IEEE. 1975. V. 63. P. 1545—1554.
92. *Forre R.* The strict avalanche criterion: spectral properties of Boolean functions and an extended definition // Adv. in Cryptology: Crypto '88, LNCS 403. 1990. P. 450—468.
93. *Fushimi M., Tezuka B.* The k-distribution of generalized feedback shift register pseudorandom numbers // Communications ACM. 1983. V. 26. № 5. P. 516—523.
94. *Games R.A., Chan A.H.* A fast algorithm for determining the complexity of a binary sequence with period 2^n // IEEE Trans. on Inf. Theory. 1983. V. 29. P. 144—146.
95. *Geffe P.* How to protect data with ciphers that are really hard to break // Electronics. 1973. V. 46. P. 99—101.
96. *Golic J.D.* Cryptanalysis of alleged A5 stream cipher // Adv. in Cryptology: Proc. of Eurocrypt '97, LNCS 1233. 1997. P. 239—255.
97. *Golomb S.W.* Shift Register Sequences. San-Francisco: Holden-Day, 1967.
98. *Günther C.G.* Alternating steps generators controlled by de Bruijn sequences // Adv. in Cryptology: Proc. of Eurocrypt '87, LNCS 304. 1988. P. 5—14.

99. *Haas A.* The multiple prime random number generator // ACM Trans. on Math. Software. 1987. V. 13. № 4.
100. Handbook of Theoretical Computer Science / Ed. J. van Leeuwen. N.Y.: Elsevier, 1990.
101. *Hartung F., Girod B.* Watermarking of compressed and uncompressed video // Signal Processing, 1998. V. 66. № 3. P. 283—301.
102. *Hellman M., Reyneri J.* Fast computation of discrete logarithms over $GF(q)$ // Adv. in cryptography. N.Y.: Plenum Publ. Comp., 1983. P. 3—13.
103. *Henk van Tilborg C.* An introduction to cryptology. Eindhoven University of Technology, 1988.
104. *Jakobsen T.* A fast method of cryptanalysis of substitution ciphers // Cryptologia. 1995. V. 19. № 3. P. 265—274.
105. *Jakobsen T., Knudsen L.R.* The interpolation attack on block cipher // Fast Software Encryption — 4th International Workshop, LNCS 1267. Berlin: Springer, 1997. P. 28—40.
106. *Jennings S.M.* Multiplexed sequences // Cryptography: Proc. of the Workshop on Cryptography, Burg Feurerstein, LNCS 149. Berlin: Springer, 1983. P. 189—206.
107. *Kahn D.* The Codebreakers. N.Y.: Macmillan, 1967.
108. *Kaliski B.S., Rivest R.L., Sherman A.T.* Is the Data Encryption Standard a Group? (Results of Cycling Experiments on DES) // J. of Cryptology. 1988. V. 1. P. 3—36.
109. *Kelsey J., Schneier B., Ferguson N.* Yarrow-160: Notes on the design and analysis of the Yarrow cryprographic pseudorandom number generator // 6th Ann. Workshop on Selected Areas in Cryptography, Springer, 1999.
110. *Kelsey J., Schneier B., Wagner D.* Related-Key Cryptanalysis of 3-Way, Biham-DES, CAST, DES-X, New-DES, RC2, and TEA // ICICS '97 Proc. N.Y.: Springer, 1997. P. 233—246.
111. *Kendall D., Babington-Smith B.* Randomness and random sampling numbers // J. Royal Statist. Soc. 1938. V. 101. № 2. P. 146—166.
112. *Kiny J.* Algorithm for the complete automated cryptanalysis of periodic polyalphabetic substitution ciphers // Cryptologia. 1994. V. 18. № 4. P. 332—355.
113. *Knudsen L.R.* Truncated and higher order differentials // Fast Software Encryption — Second International Workshop, LNCS 1008. 1995. P. 196—211.
114. *Kocher P.C.* Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems // Adv. in Cryptology: Proc. of Crypto '96, LNCS 1109. 1994. P. 104—113.
115. *Koblitz N.* A Course in Number Theory and Cryptography. Springer, 1994.
116. *Koblitz N.* Algebraic Aspects of Cryptography. Algorithms and Computation in Math., 3, Springer, 1998.
117. *Konheim A.* Cryptography. N.Y.: Wiley, 1981.
118. *Kranakis E.* Primality and Cryptography. N.Y.: Wiley, 1983.
119. *Kullback S.* Statistical Methods in Cryptanalysis. Laguna Hills, CA: Aegean Park Press, 1976.
120. *Lai X., Massey J., Murphy S.* Markov cipher and differential analysis // Adv. in Cryptology: Proc. of Eurocrypt '92, LNCS 547. 1991. P. 17—38.
121. *Lenstra A., Lenstra H., Manasse M., Pollard J.* The factorization of the ninth fermat number // Math. of Comp., 1993. V. 61. № 203. P. 319—349.
122. *Lenstra H.W.Jr.* Factoring integers with elliptic curves // Ann. of Math., 1987. V. 126. P. 649—673.

123. *Lewis T., Payne W.* Generalized feedback shift register pseudrandom number algorithms // J. ACM. 1973. V. 20. № 5. P. 456—468.
124. *Lint J.* Introduction to Coding Theory. N.Y.: Springer, 1982.
125. *MacLaren M., Marsaglia G.* Uniform random number generators // J. ACM 12. 1965. № 1. P. 83—89.
126. *Marsaglia G.* A current view of random number generators // Proc. Comp. Sci. and Statistics: 16th Symposium on the Interface. N.Y.: Elsevier, 1985.
127. *Marsaglia G., Bray T.A.* One line random number generators and their use in combinatorics// Comm. ACM. 1968. V. 11. P.757—759.
128. *Marsaglia G.* Marsaglia random number CD-ROM, 1996.
129. *Marsaglia G.* Random number fall mainly in the plans // Proc. Nat. Acad. Sci. 1968. V. 61. № 61. P. 25—28.
130. *Marsaglia G.* The mathematics of random number generators // Proc. of Symp. on Appl. Math. 1992. V. 46. P. 73—89.
131. *Marsaglia G.* The structure of linear congruential sequences. N.Y.: Acad. Press, 1972.
132. *Marsaglia G., Tsay L.* Matrices and the structure of random number sequences // Linear Algebra and Its Appl., 1985. V. 67. № 1. P. 147—156.
133. *Marsaglia G., Zaman A.* A new class of random number generators// Ann. of Appl. Prob. 1991. V. 1. № 3. P. 462—480.
134. *Marshall A., Olkin J.* Inequalities: Theory of Majorization and its Applications. N.Y.: Acad. Press, 1979.
135. *Matsui M.* Linear Cryptanalysis Method for DES Cipher // Adv. in Cryptology: Proc. of Eurocrypt '93, LNCS 765. 1994. P. 386—397.
136. *Matsui M.* The first experimental cryptanalysis of the Data Encryption Standard // Adv. in Cryptology: Proc. of Crypto '94, LNCS 839. 1994. P. 1—11.
137. *Matthews R.* The use of genetic algorithm in cryptanalysis // Cryptologia. 1993. V. 17. № 2. P. 187—201.
138. *Matyas S., Meyer C.* Generation, distribution and installation of cryptographic keys // IBM Syst. J. 1978. V. 17. № 2. P. 126—137.
139. *Maurer U.* A universal statistical test for random bit generators // Adv. in Cryptology: Proc. of Crypto '90, LNCS 537. 1991. P. 409—420.
140. *McEliece R.* The Theory of Information and Coding. CUP, 1977.
141. *Meier W., Staffelbach O.* Fast correlation attacks on stream ciphers // Adv. in Cryptology: Proc. of Eurocrypt '88, LNCS 330. 1988. P. 301—314.
142. *Menezes A.* Elliptic curve public key cryptosystems. Kluwer Acad. Publishers, 1993.
143. *Menezes A., Oorschot P., Vanstone S.* Handbook of Applied Cryptology. N.Y.: CRC Press, 1997.
144. *Merkly R.* A Fast Software One-Way Hash Function // J. of Cryptology. 1990. V. 3. № 1. P. 43—58.
145. *Meyer C., Matyas St.* Cryptography. N.Y.: Wiley, 1982.
146. *Miller G.L.* Riemann's hypothesis and tests for primality // J. of Comp. and Syst. Sci., 1976. V. 13. P. 300—317.
147. *Miyagashi S., Ohta K., Iwata M.* Confirmation that some hash functions are not collision free // Adv. in Cryptology. N.Y.: Springer, 1991. P. 326—343.
148. *Montgomery P.* Modular multiplication without trial division // Math. of Comp., 1985. V. 44. P. 519—521.
149. *Mullin R., Onyszchuk T., Vaustone S., Wilson R.* Optimal Normal Bases in $GF(p^n)$ // Discr. Appl. Math. 1988—1989. V. 22. P. 149—161.

150. *Needham P., Schroeder M.* Using encryption for authentification in large networks of computers // Com. of the ACM. 1978. V. 21. № 12. P. 993—999.
151. *Neuman B., Ts'o T.* Kerberos: An Authentification Service for Computer Networks// IEEE Com. Mag., 1994. № 9. P. 33—38.
152. *Nyberg K.* Differentially uniform mappings for cryptography // Adv. in Cryptology: Proc. of Eurocrypt '93, LNCS 765. 1993. P. 55—64.
153. *Nyberg K.* Perfect nonlinear *S*-boxes // Adv. in Cryptology: Proc. of Eurocrypt '91, LNCS 547. 1991. P. 378—386.
154. *O'Connor L.* On the distribution of the characteristics in the bijective mappings // Adv. in Cryptology: Proc. of Eurocrypt '93, LNCS 765. 1993. P. 360—370.
155. *Odlyzko A.* Discrete logarithms in finite fields and there cryptographic significance // Adv. in Cryptology: Proc. of Eurocrypt '84, LNCS 209. 1985. P. 224—314.
156. *Pincus S., Singer B.H.* Randomness and degrees of irregularity // Proc. Nat. Acad. Sci., 1993. P. 2083—2088.
157. *Pohlig S.C., Hellman M.E.* An improved algorithm for computing logarithms over $GF(p)$ and its cryptographic significance // IEEE Trans. Inf. Theory, 1985. V. 24. P. 106—110.
158. *Pollard J.* A Monte Carlo Method for Factorization Nord // Tid. for Inf. 1975. V. 15. № 2. P. 331—334.
159. *Pollard J.M.* Theorems on factorization and primality testing // Proc. of the Cambridge Phil. Soc., 1974. V. 76. P. 521—528.
160. *Pomerance C.* The quadratic sieve factoring algorithm // Adv. in Cryptology: Proc. of Eurocrypt '84, LNCS 209. 1985. P. 169—182.
161. *Pomerance C.* The role of smooth numbers in number theoretic algorithms // Proc. Int. Cong. Math. Zurich, Switzerland, 1994.
162. *Preneel B., Van Leekwijck W., Van Linden L., Govaerts R., Vandewalle J.* Propagation characteristics of Boolean functions // Adv. in Cryptology: Proc. of Eurocrypt '90, LNCS 473. 1991. P. 161—173.
163. *Rabin M.* "Digital Signatures" Foundation of Secyre Communication. N.Y.: Acad. Press, 1978. P. 155—168.
164. *Rabin M.O.* Digitalized signatures and public-key functions as intractable as factorization // MIT/LCS/TR-212, MIT Laboratory for Comp. Sci., 1979.
165. *Rabin M.O.* Probabilistic algorithm for testing primality // J. of Number Theory. 1990. V. 12. P. 128—138.
166. *Ramesh R., Athitham G., Thiruvengaden K.* An automated approach to solve simple substitution ciphers // Cryptologia. 1993. V. 17. № 2. P. 202—218.
167. *Reeds J.* "Cracking" a random number generator // Cryptologia. 1977. V. 1. № 1. P. 20—26.
168. *Riesel H.* Prime Numbers and Computer Methods for Factorization // Progr. in Math. 57, Birkhauser, 1985.
169. *Rivest L.R., Shamir A., Adleman L.* A method for abtaining digital signatures and public-key cryptosystems // Comm. ACM. 1978. V. 21. P. 120—126.
170. *Rivest R.L.* On NIST's Proposed Digital Signature Standard, 1992.
171. *Rivest R.* The MD5 Message Digest Algorithm, RFC 1321. Apr., 1992.
172. *Robie D.L., Merserean R.M.* Video error correction using steganography // Eur. J. Appl. Sign. Processing, 2002, V. 2002. № 2. P. 164—173.
173. *Rueppel R.A.* Analysis and Design of Stream Ciphers. Berlin: Springer, 1986.

174. *Rueppel R.A., Staffelbach O.J.* Products of linear recurring sequences with maximum complexity // IEEE Trans. Inf. Theory, 1987. V. 33. P. 124—131.
175. *Rukhin A.L.* Approximate Entropy for Testing Randomness // www.math.umbc.edu/ftp/tr98-07.html.
176. *Salomaa A.* Public-Key Cryptography. Berlin: Springer, 1993.
177. *Schneier B.* Applied Cryptography: Protocols, Algorithms, and Source Code in C. N.Y.: John Wiley & Sons, 1996.
178. *Schnorr C., Lenstra H.* A monte carlo factozing algorithms with linear storage // Math.Comp. 1984. V. 43. P. 289—311.
179. *Schnorr C.P.* Efficient signature generation by smart cards // J. of Cryptology. 1991. V. 4. P. 161—174.
180. *Sherman L.* Factoring Large Integers // Math. Comp. 1974. V. 28. P. 637—646.
181. *Siegenthaler T.* Correlation-immunity of nonlinear combining functions for cryptographic applications // IEEE Trans. Inf. Theory. 1984. V. 30. P. 776—780.
182. *Siegenthaler T.* Decrypting a class of stream ciphers using ciphertext only // IEEE Trans. Comp., 1985. V. C-34. P. 81—85.
183. *Silverman J.* Adv. Topics in the Arithmetic of Elliptic Curves. GTM 151. N.Y.: Springer, 1994.
184. *Silverman J.H.* The Arithmetic of Elliptic Curves. N.Y.: Springer, 1986.
185. *Simmons G.J.* Subliminal communication is easy using the DSA // Adv. in Cryptology: Proc. of Eurocrypt '93, LNCS 765. 1994. P. 218—232.
186. *Sinkov A.* Elementary Cryptanalysis. Mathematical Approach. N.Y.: Random House, 1956.
187. *Solovay R., Strassen V.* A fast Monte-Carlo test for primality // SIAM J. on Computing. 1977. V. 6. P. 84—85.
188. *Stinson D.R.* Cryptography: Theory and Practice. Boca Raton, Florida: CRC Press, 1995.
189. *Webster A.F., Tavares S.E.* On the design of *S*-boxes // Adv. in Cryptology: Proc. of Crypto '85, LNCS 218. 1986. P. 523—534.
190. *Whittlesey J.* On the multidimensional uniformity of pseudorandom number generators // Com. ACM. 1969. V. 12. P. 247—256.
191. *Wichmann B., Hill I.* An efficient and portable pseudo-random number generator // Appl. Statistics. 1982. V. 31. № 2. P. 188—190.
192. *Wiener M.* Cryptanalysis of short RSA secret exponents // IEEE Trans. Inf. Theory, 1990. V. 36. № 3. P. 553—558.
193. *Willems F.* Universal data compression and repetition times // IEEE Trans. Inf. Theory, 1989. V. 35. P. 54—58.
194. *Youssef A.M., Tavares S.E.* Resistance of balanced *S*-boxes to linear and differential cryptanalysis // Inf. Processing Letters, 1995. V. 56. P. 249—252.
195. *Zeng K., Huang M.* An improved linear syndrome algorithm in cryptanalysis with applications // Adv. in Cryptology: Proc. of Crypto '90, LNC S537. 1990. P. 34—47.
196. *Ziv J., Lempel A.* On the complexity of finite sequences // IEEE Trans. Inf. Theory, 1976. V. 22. P. 75—81.

SXJKN8EYHR3I4HR987YDWFUIHW9
E8YR384H5O34596EUF9O8WE76R8
3Y439846UHFE98YDSFJH8374H4B
R8276RUBNWDF87Y6234BN9WE872
H3WRTJOJRTGM905JRO9G70954TO
IRF7ERFH2347FQ7WERF9UJ1349Y

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

SXJKN8EYHR3I4HR987YDWFUIHW9
E8YR384H5O34596EUF9O8WE76R8
3Y439846UHFE98YDSFJH8374H4B
R8276RUBNWDF87Y6234BN9WE872
H3WRTJOJRTGM905JRO9G70954TO
IRF7ERFH2347FQ7WERF9UJ1349Y

P

P-значение 157

S

S-блок 232, 239, 242
s-тактовый дифференциал 261

А

Автоморфизм 54, 61
Алгоритм
  Берлекэмпа — Месси 219
  Брента 309
  деления с остатком 25
  Евклида 26, 41
  статистического тестирования 156
  Флойда 314
Алфавит
  мощность 89
  символ 89
Арифметика
  больших чисел 287
  модулярная 288

Б

База множителей 298
Батарея тестов
  Кнута 158
  Дж Марсальи 162
Блок шифрования 229

В

Вес Хэмминга 168, 361
Водяной знак 357
Возведение в степень по Монтгомери 290
Вычет
  по модулю 30
  квадратичный 36

Г

Гармограмма 164
Генератор РРСП 155
  ANSI X9.17 201
  BBS (Blum — Blum — Shub)-алгоритм 205
  FIPS-186 202
  RSA-алгоритм 204
  Yarrow-160 203
  Беса — Пайпера 224
  Геффе 221
  Голлманна 222
  квадратичный конгруэнтный 196
  классификация алгоритмов генерации 194
  конгруэнтный со случайными параметрами 210
  конгруэнтный, использующий умножение с переносом 197
  линейный конгруэнтный 193
  мультипликативный конгруэнтный 193
  на основе односторонних функций 200
  нелинейный конгруэнтный 196
  основанный на проблемах теории чисел 204
  переключающий шаг-стоп 223
  подключей 239
  пороговый 222
  программный 155, 156, 193
  прореживающий 223
  смешанный конгруэнтный 193
  табличный 155
  Фибоначчи 200
  физический 155
  Эйхенауэра — Лена с обращением 197
Гомоморфизм 54, 61
Гомоморфный образ 54
Группа 49
  абелева 49
  действие на множестве 57
  подгруппа 50
  подстановок 55
  полная линейная 226
  порядок 50
  разложение 51
  симметрическая *n*-й степени 55
  транзитивная 58
  фактор-группа 54
  циклическая 50, 52, 61

Д

Делитель нуля 58
Дискретное логарифмирование 332
  в схеме ЭЦП 326
  индекс-метод 332
  метод больших-малых шагов 332
  метод Нечаева — Полига — Хеллмана 332
  метод решета числового поля 333
Дискретное преобразование Фурье 164
Дискретный логарифм 35, 36, 332
Дробь
  подходящая 42
  цепная 41

З

Задача о рюкзаке 283
Защита информации 17

И

Идеал
  главный 59, 61
  двусторонний 59
  максимальный 63
  простой 63
ИДС, см. *Источник дискретных сообщений*
Изоморфизм 54, 61
Инволюция 234
Индекс см. *Дискретный логарифм*
ИНС, см. *Источник непрерывных сообщений*
Источник дискретных сообщений 89
  асимптотические свойства 125
  без памяти 90
  стационарный 90
  энтропия 91
Источник непрерывных сообщений 89

К

Квадратичный
  вычет 36
  закон взаимности 37, 38
  невычет 36
Квантовый канал связи 351
Класс вычетов 30, 49
  по модулю идеала 60

Ключ 9, 18
главный 343
мультипликативное обращение 236
протокол управления 343
размер 229
сеансовый 343
шифрования ключей 343
Ключевое расписание см. *Расписание ключей*
Количество информации по Шеннону 135, 139
Кольцо 58
евклидово 62
коммутативное 58
с единицей 58
центр 60
Комбинирование генераторов
LFSR-генераторов 207
метод Макларена — Марсальи 207
полиномиальное 207
с помощью псевдослучайного прореживания 209
Кривая
аномальная 336
суперсингулярная 336
эллиптическая 79, 335
Криптоанализ 9, 17, 245
DES 270
активный 358, 359
байесовский метод 258
инженерный 358
корреляционный 223
линейный 246, 265
метод максимального правдоподобия 251
метод «опробования» 246, 247
принцип Живерже 245
принцип Керкхоффа 245
принцип Керкхоффа — Шеннона 245
разностный 246, 261
с использованием теории статистических решений 246, 250
Криптоатака 246
«встретимся посередине» 314
«дней рождения» 307
на RSA 277
Криптография 9, 17, 356
квантовая 351
Криптология 9, 17
Криптопреобразование 229
Фейстеля 230
Криптосистема 17
DES 229, 235, 368
DSA 320
ECDSA 335
Rijndael 240, 241
IDEA 229, 235
Rijndael 229, 240, 241
RSA 275, 282, 359, 368
аутентификация 309
взвешенная сумма 145
ГОСТ 28147—89 229, 239, 307
криптостойкость 145
марковская 263
поточная 215
произведение 145
Рабина 282
с открытым ключом 275
симметричная 229
совершенно криптостойкая 146
Фейстеля 368
шеннoновские модели 140
Криптостойкость 145
Критерий
простоты 291
Эйлера 37

Л

Лемма Гаусса 38
Линейная рекуррентная последовательность 70, 197, 216
Линейная сложность 208, 219
профиль 220
Линейный регистр сдвига с обратной связью 198
ЛРП, см. *Линейная рекуррентная последовательность*

М

Многочлен
аннулирующий 65
возвратный 217
минимальный 65, 217
порядок 69,70
характеристический 71, 216
Модель сообщений 89
Модуль 29
Мономорфизм 61

Н

Наибольший общий делитель 25
Наилучшее приближение 44
Наименьшее общее кратное 27

О

Область целостности 58
Однородная цепь Маркова 104
Определитель Ганкеля 218
Орбита 57
ОЦМ, см. *Однородная цепь Маркова*

П

Первообразный корень 34
Перестановка символов 142
Подгруппа 50
знакопеременная группа 56
индекс 51
нормальная 54
порядок 52
стабилизатор 57
циклическая 52
Подкольцо 59
Подписанное сообщение 316
Подстановка
биграммная 145
нечетная 56
символов 142
четная 56
Поле 59
Галуа 59, 61, 68
конечное 68
конечное расширение 65
надполе 65
подполе 65
примитивный элемент 69
разложения 67
расширение 65
степень расширения 65
Полная система вычетов 30
Поляризация фотона 352
Пороговая схема 354, 355
Порядок
многочлена 69, 70
элемента 52
Последовательность 70
m-последовательность 217
де Брейна 216
линейная рекуррентная 70, 197, 216, 368
линейная сложность 173, 208, 219
максимального периода 71, 217
псевдослучайная 156
равномерно распределенная случайная 153
рекуррентная 216
случайная по Колмогорову 182
случайная по Мартин-Лёфу 182
стохастическая 182, 183
хаотическая 182
Приведение по модулю 288
метод Барретта 289
метод Монтгомери 289
Приведенная система вычетов 37
Признак простоты 290
Простое число
алгоритм генерации 291
генерация в ГОСТ Р 34.10—94 295
Протокол
взаимной аутентификации 344
генерации ключей 343
Диффи — Хеллмана 346
Нидхэма — Шредера 346
прямого обмена ключами 346
распределения ключей 348
рукопожатия 344
Цербер 348

Р

Равномерно распределенная случайная последовательность 153
Разделение секрета 354
Расписание ключей 236, 240
Расстояние
единственности 149, 150
Колмогорова 167

Расширение
базис 67
конечное 65
примитивный элемент 66
простое алгебраическое 66
степень 65
Расшифровывание 236
Регистр сдвига с обратной связью 217
РРСП, см. *Равномерно распределенная случайная последовательность*

С

Символ
Лежандра 36, 39
Якоби 39
Сравнение первой степени 32
Стандарт ЭЦП DSS 320
Стеганография 356
Схема
вероятностная подписи Рабина 319
Диффи — Лампорта 318
Фейстеля 230, 241
ЭЦП Рабина 317
ЭЦП Шнорра 330
ЭЦП Эль-Гамаля 322

Т

Таблица Кэли 50, 61
Теорема
Вильсона 290
Евклида 28
китайская об остатках 33
Кэли 57
Лагранжа 52
Левина — Шнорра 182
М. Винера 286
Мартин-Лёфа 182
Маурера 169
о высоковероятном подмножестве 126
о гомоморфизме 55
Стратоновича 131
Ферма 31
Хассе 85
Чебышева 286
Эйлера 32, 296
Тест
n-серий 158
интервалов 159
классический покер-тест 159
Лемпеля — Зива 172
Миллера — Рабина 295
на основе приращений энтропии 170
обобщенный покер-тест 159
основанный на линейной сложности 173
основанный на рангах двоичных матриц 163
пересекающихся *n*-грамм 162
перестановок 161
случайного блуждания 167
«собирателя купонов» 160
Соловея — Штрассена 292
спектральный 164
статистики дельта-произведения 178
статистики скалярного произведения 175
статистический Маурера 169
Точка
F-рациональная 79
аффинная 79
на бесконечности 79
невырожденная (гладкая) 79
Транспозиция 55

У

Умножение по Монтгомери 289
Уравнение Вейерштрасса 79

Ф

Факторизация 280, 297
квадратичный метод 298
метод квадратичного решета 299
метод эллиптических кривых 298
модуля 279
обобщенный метод решета числового поля 299
рекорды 299
Фактор-кольцо 60, 241, 243
Функция
односторонняя с секретом 316
распределения 166
Эйлера 30
Функция хэширования 305
блочно-итерационная 306
ГОСТ Р 34 11—94 307
односторонняя 305
свободная от коллизий 305
СТБ 1176 1—99 309
строго свободная от коллизий 305
шаговая 306

Ц

Цикл 55
Цифровая подпись
генерация 336, 337
на эллиптических кривых 335
проверка 336, 337

Ч

Число
большое 287, 363
взаимно простое 25, 27
каноническое разложение 28
Кармайкла 292
Мерсенна 303
неполное частное 42
простое 28
целое гауссовское 62

Ш

Шифр
Бофора 144
Вернама 22, 144
Виженера 22, 143
модулярный 19
перестановочный 22
Плейфера 20
подстановки 19
поточный, см. *Вернама*
Цезаря 19, 143
Шифрование 230
биграммное 20
гомофоническое 20
многоалфавитное подстановочное 21
повторное 278
полиграммное 20
поточное 235
рюкзачный метод 283

Э

Электронная цифровая подпись 315
ГОСТ Р 34.10—94 323
СТБ 1176 2—99 330
Элемент
единичный 49
обратный 49
порядок 52
Эллиптическая кривая 79, 335
j-инвариант 80
дзета-функция 85
дискриминант 80
изоморфная 80
над конечными полями 84
невырожденная (гладкая) 80
порядок 338
скалярное умножение 338
скручивание 87
сложение точек 82
Эндоморфизм 61
Фробениуса 340
Энтропийная устойчивость случайной последовательности 130
Энтропия 91, 220
безусловная 135
выборочная 170
дифференциальная 109
иерархическая аддитивность 94, 110
случайного сообщения 108
удельная 99
условная 94, 136
условная дифференциальная 110
Хартли 92
Шеннона 92
Эпиморфизм 54, 61
канонический 67
ЭЦП, см. *Электронная цифровая подпись*

Учебное издание

Харин Юрий Семенович
Берник Василий Иванович
Матвеев Геннадий Васильевич
Агиевич Сергей Валерьевич

Математические и компьютерные основы криптологии

Учебное пособие

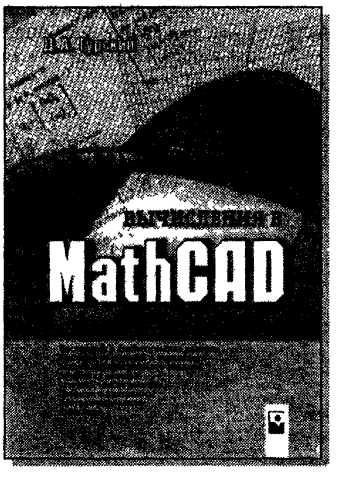
Редактор *М.Э. Моисеева*
Корректор *Л.К. Мисуно*
Художник обложки *С.В. Ковалевский*
Компьютерная верстка *С.И. Лученок*

Подписано в печать с готовых диапозитивов 07.07.2003.
Формат 70×100 $^1/_{16}$. Бумага газетная. Гарнитура Ньютон.
Печать офсетная. Усл. печ. л. 31,2. Уч.-изд. л. 24,01.
Тираж 2010 экз. Заказ № 1386.

Общество с ограниченной ответственностью «Новое знание».
ЛВ № 310 от 01.07.2003. Минск, ул. Академическая, д. 28, к. 112.
Почтовый адрес: 220050, Минск, а/я 79.
Телефон/факс: (10-375-17) 211-50-38, 284-03-23.
Москва, ул. Маросейка, 10/1.
Телефон (095) 921-67-21.
E-mail: nk@wnk.biz
http://wnk.biz

Республиканское унитарное предприятие
«Издательство «Белорусский Дом печати».
220013, Минск, пр. Ф.Скорины, 79.

ИЗДАТЕЛЬСТВО "НОВОЕ ЗНАНИЕ"



Вычисления в MathCAD

Д. Гурский

Практическое руководство
800 с.

Описана эффективная работа в MathCAD — самом популярном программном продукте, ориентированном на математические вычисления. Читатель научится выбирать оптимальный метод для конкретной вычислительной задачи. Рассмотрен широкий круг проблем, которые можно решить с помощью MathCAD: численные и символьные вычисления, создание программ, операции с матрицами, дифференцирование и интегрирование и т.д Приведена оценка точности различных вычислительных методов.

Книга предназначена для студентов математических и естественно-научных специальностей, преподавателей и научных работников, а также для всех, кто интересуется математическими пакетами.

ISBN 985-475-013-2

Наши координаты:
в Москве: (095) 921-67-21, e-mail: ru@wnk.biz
в Минске: (10-375-17) 211-50-38, 284-03-23, e-mail: nk@wnk.biz

http://wnk.biz