ТЕОРЕТИКО-ЧИСЛОВЫЕ МЕТОДЫ В КРИПТОГРАФИИ

Лекция 13

Псевдослучайные последовательности над конечным полем и их применение в криптографии

Объём по рабочей программе: 1 академический час. Основная часть посвящена линейным рекурсиям, периоду и криптографическому смыслу предсказуемости. Доказательства через расширение поля и автокорреляцию можно использовать для углублённого самостоятельного изучения.

Цель: установить связь между конечными полями, регистрами сдвига и периодическими последовательностями; научиться отличать максимальный период от криптографической стойкости.

Ожидаемый результат: студент записывает рекурсию по коэффициентам, определяет условия максимального периода, объясняет баланс и корреляцию двоичной последовательности и восстанавливает линейный закон по фрагменту её выхода.

Подготовка: конечные поля, многочлены, порядок элемента конечной группы, матрицы и решение линейных систем над полем.

1 Случайность и детерминированная генерация

Последовательность, полученная по фиксированному правилу из начального состояния, является детерминированной. Повторение правила и начальных данных воспроизводит тот же выход. При этом частоты символов и другие наблюдаемые свойства могут напоминать случайный источник. В теории последовательностей такие конструкции часто называют псевдослучайными.

В криптографии используется более сильное требование: **эффективный наблюдатель не должен отличать распределение выхода от равномерного распределения той же длины**. Это свойство распределения при случайном выборе начальных данных, а не отдельной фиксированной строки.

Пусть детерминированный генератор G за полиномиальное время преобразует равномерный начальный ключ из n бит в $\ell(n) > n$ бит. Для любого вероятностного полиномиального алгоритма D требуется пренебрежимо малая величина

$$\left|\Pr[D(G(U_n)) = 1] - \Pr[D(U_{\ell(n)}) = 1]\right|.$$

Здесь U_k — равномерная случайная строка длины k; вероятности учитывают также случайность самого алгоритма D. Пренебрежимость означает убывание быстрее любой обратной степени параметра n. Данная формулировка подробно рассматривается у Katz и Lindell, § 3.3, определение 3.14.

В этой лекции линейные последовательности служат примером того, что **большой период, хороший баланс и простая корреляция не обеспечивают непредсказуемость**. Их свойства полезны в математике и обработке сигналов, а ограничения важны для понимания криптографических генераторов.

2 Линейная рекурсия над конечным полем

Пусть $\mathbb{F}_q$ — конечное поле из q элементов, где $q = p^r$ для простого p. При $r = 1$ это поле вычетов по модулю p. При $r > 1$ поле строят как расширение, например с помощью неприводимого многочлена. Его нельзя подменять кольцом целых вычетов по модулю q: в частности, $\mathbb{F}_4$ и $\mathbb{Z}/4\mathbb{Z}$ — разные структуры.

Определение. Линейная рекурсия порядка m имеет вид

$$u_{t+m} = c_0 u_t + c_1 u_{t+1} + \cdots + c_{m-1} u_{t+m-1}, \qquad t \geq 0,$$

где все коэффициенты и значения принадлежат $\mathbb{F}_q$. Начальные значения $u_0, \ldots, u_{m-1}$ однозначно определяют всю последующую последовательность.

Состояние в момент t — вектор $S_t = (u_t, \ldots, u_{t+m-1})^{\mathsf{T}}$. Для определённости выходом считаем первый компонент u_t. При каждом шаге компоненты сдвигаются влево, а справа добавляется значение рекурсии. Это **регистр сдвига с линейной обратной связью**, или LFSR.

Переход записывается как $S_{t+1} = TS_t$, где

$$T = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ c_0 & c_1 & c_2 & \cdots & c_{m-1} \end{pmatrix}.$$

Для $\mathbb{F}_2$ сложение совпадает с XOR, а умножение коэффициента на бит означает включение или исключение соответствующего разряда из обратной связи. Везде далее коэффициенты перечисляются **от c_0 к c_{m-1}**, а состояние — **от текущего выходного бита к последнему разряду**.

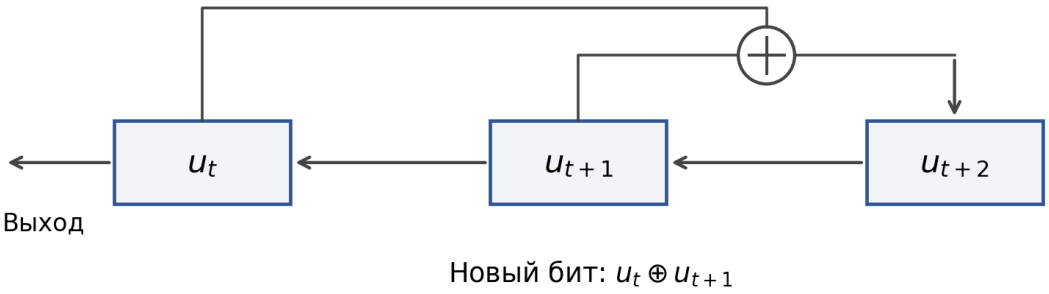


Рисунок 1 — Регистр для рекурсии $u_{t+3} = u_t + u_{t+1}$ над $\mathbb{F}_2$

Пример 1. При $c = (1,1,0)$ и начальном состоянии $(1,0,0)$ получаем цикл

$$100 \to 001 \to 010 \to 101 \to 011 \to 111 \to 110 \to 100.$$

Первые компоненты состояний образуют период 1001011. Период равен 7, потому что семь состояний различны, а следующий переход возвращает первое состояние.

3 Период и обратимость перехода

Состояний всего q^m. Поэтому любая траектория конечного детерминированного регистра когда-нибудь повторит состояние и далее будет двигаться по циклу. В общем случае перед циклом возможен **предпериод**.

Если $c_0 \neq 0$, переход обратим: из $u_{t+1}, \ldots, u_{t+m}$ восстанавливается

$$u_t = c_0^{-1}\left(u_{t+m} - \sum_{j=1}^{m-1} c_j\, u_{t+j}\right).$$

Это согласуется с $\det T = (-1)^{m-1}c_0$. Обратимое отображение конечного множества является перестановкой, поэтому предпериода нет: состояние с самого начала принадлежит циклу.

Нулевое состояние всегда неподвижно. При обратимом переходе ненулевое состояние не может попасть в ноль, следовательно, его период удовлетворяет

$$N \leq q^m - 1.$$

Период состояния совпадает с периодом выходной последовательности: состояние состоит из m последовательных выходных символов. Если выход имеет период N, то через N шагов повторяются все компоненты состояния; обратное следует из детерминированности переходов.

Пример 2. Для $u_{t+3} = u_t$ состояние 100 имеет цикл $100 \to 001 \to 010 \to 100$ и период 3, а состояние 111 неподвижно и имеет период 1. Ненулевой коэффициент c_0 гарантирует обратимость, но не максимальный период.

Если взять $u_{t+3} = u_{t+1}$ и начальное состояние 100, то следующий вектор равен 000 и остаётся нулевым. Здесь $c_0 = 0$, предпериод имеет длину 1, а период — длину 1. Поэтому предположение об обратимости нельзя опускать.

4 Характеристический и примитивный многочлены

Рекурсии сопоставляется характеристический многочлен

$$f(X) = X^m - c_{m-1}X^{m-1} - \cdots - c_1 X - c_0.$$

Он равен $\det(XI - T)$. В двоичном поле минус совпадает с плюсом. Например, $c = (1,1,0)$ соответствует $f(X) = X^3 + X + 1$. Порядок записи коэффициентов существенен: использование обратного порядка без изменения соглашения задаёт другую рекурсию.

Определение. Монический неприводимый многочлен f степени m с ненулевым свободным членом называется **примитивным**, если класс $\alpha = X \bmod f$ имеет порядок $q^m - 1$ в мультипликативной группе поля $\mathbb{F}_q[X]/(f)$.

Теорема. Если f примитивен, любое ненулевое начальное состояние соответствующего регистра имеет период $q^m - 1$. Если f неприводим и $f(0) \neq 0$, период каждого ненулевого состояния равен порядку α, который может быть меньше $q^m - 1$.

Доказательство. По теореме Гамильтона — Кэли $f(T) = 0$. Поэтому многочлен h действует на состояние как $h(T)$, причём действие зависит только от класса $h \bmod f$. Зафиксируем $S_0 \neq 0$. Отображение

$$\mathbb{F}_q[X]/(f) \to \mathbb{F}_q^m, \qquad h \bmod f \mapsto h(T)S_0$$

инъективно. Для ненулевого класса h существует обратный. Равенство $h(T)S_0 = 0$ тогда влечёт $S_0 = 0$ — противоречие. В обеих частях по q^m элементов, следовательно, отображение является биекцией.

При этой биекции $T^t S_0$ соответствует α^t. Возвращение к S_0 равносильно $\alpha^t = 1$, поэтому длина цикла точно равна порядку α. Для примитивного многочлена этот порядок равен $q^m - 1$.

Пример 3. Многочлен $X^3 + X + 1$ не имеет корней в $\mathbb{F}_2$: при подстановке 0 и 1 получается 1. Многочлен степени 3 без корней неприводим. Группа ненулевых элементов поля из 8 элементов имеет простой порядок 7. Класс α не равен 1, поэтому его порядок равен 7. Этим доказан максимальный период регистра из примера 1.

Неприводимость и примитивность различаются. Многочлен

$$g(X) = X^4 + X^3 + X^2 + X + 1$$

не имеет линейного делителя над $\mathbb{F}_2$. Единственный монический неприводимый квадратный многочлен — $X^2 + X + 1$; остаток g по нему равен $X + 1$, поэтому квадратного делителя также нет. Значит, g неприводим. Но $(X + 1)g(X) = X^5 + 1$, откуда в поле по модулю g выполняется $\alpha^5 = 1$. Поскольку $\alpha \neq 1$, его порядок равен 5, а не 15. Такой регистр имеет период 5 для каждого ненулевого состояния.

4.1 Проверка примитивности по степеням

Пусть f неприводим, $f(0) \neq 0$ и $M = q^m - 1$. Для каждого различного простого делителя r числа M проверим

$$X^{M/r} \not\equiv 1 \pmod{f}.$$

Эти условия эквивалентны примитивности. Порядок d класса X делит M. Если $d < M$, число $M/d > 1$ имеет простой делитель r, и тогда $d \mid M/r$; соответствующая проверка дала бы 1. Обратно, если хотя бы одна проверка даёт 1, порядок меньше M.

Пример 4. Для $f = X^4 + X + 1$ нет корней в $\mathbb{F}_2$, а остаток по $X^2 + X + 1$ равен 1, поэтому f неприводим. Так как $M = 15 = 3 \cdot 5$, достаточно проверить степени 5 и 3. Из $X^4 \equiv X + 1$ получаем $X^5 \equiv X^2 + X \neq 1$; также $X^3 \neq 1$. Следовательно, f примитивен и период равен 15.

5 Баланс и автокорреляция максимальной последовательности

Двоичная последовательность максимального периода $N = 2^m - 1$, полученная из ненулевого состояния примитивного регистра, называется **m-последовательностью**. За период её состояния пробегают все ненулевые двоичные векторы длины m.

Теорема о балансе. За один полный период число единиц и нулей равно

$$N_1 = 2^{m-1}, \qquad N_0 = 2^{m-1} - 1.$$

Доказательство. Среди всех 2^m векторов ровно половина имеет первый компонент 1 и половина — 0. Из траектории исключён только нулевой вектор, у которого первый компонент равен 0. Это уменьшает число нулей на один и не меняет число единиц.

Для изучения совпадений со сдвинутой последовательностью введём знаки $a_t = (-1)^{u_t} = 1 - 2u_t$ и **периодическую автокорреляцию**

$$C(k) = \sum_{t=0}^{N-1} a_t \, a_{t+k},$$

где индекс $t + k$ берётся по модулю N. Это ненормированная величина; нормированная автокорреляция равна $C(k)/N$.

Теорема. Для m-последовательности

$$C(0) = N, \qquad C(k) = -1 \quad \text{при } 1 \le k < N.$$

Доказательство. При нулевом сдвиге каждое произведение равно 1. Для ненулевого сдвига последовательность $v_t = u_t + u_{t+k}$ удовлетворяет той же линейной рекурсии. Её начальное состояние ненулевое: иначе исходные состояния S_0 и S_k совпали бы при $0 < k < N$. Поэтому v_t также является m-последовательностью с тем же балансом. Но $a_t a_{t+k} = (-1)^{v_t}$, откуда сумма равна числу нулей в v минус число единиц, то есть -1.

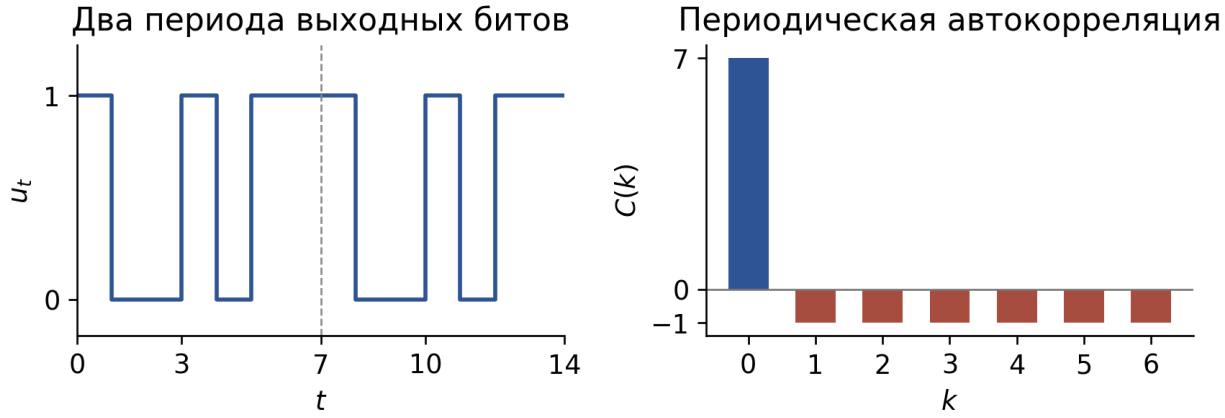


Рисунок 2 — Период и автокорреляция последовательности из примера 1

Эти формулы относятся к **целому периоду** и к **циклическим сдвигам**. Для произвольного короткого фрагмента или непериодической корреляции такие значения не гарантируются.

6 Линейная сложность и восстановление рекурсии

Линейная сложность последовательности — наименьшая длина линейного регистра, который порождает всю последовательность. Она может быть меньше числа разрядов конкретной реализации. Например, постоянная ненулевая последовательность получается регистром длины 1, даже если её поместить в более длинный регистр.

Для ненулевой последовательности с неприводимым характеристическим многочленом степени m линейная сложность равна m. Иначе существовал бы ненулевой многочлен h степени меньше m, задающий более короткую рекурсию. Тогда $h(T)S_0 = 0$. Но класс h обратим по модулю неприводимого f, что противоречит $S_0 \neq 0$.

Предположим, известны $2m$ последовательных выходных символов. Неизвестные коэффициенты входят в линейную систему

$$\begin{pmatrix} u_0 & u_1 & \cdots & u_{m-1} \\ u_1 & u_2 & \cdots & u_m \\ \vdots & \vdots & \ddots & \vdots \\ u_{m-1} & u_m & \cdots & u_{2m-2} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{m-1} \end{pmatrix} = \begin{pmatrix} u_m \\ u_{m+1} \\ \vdots \\ u_{2m-1} \end{pmatrix}.$$

Над $\mathbb{F}_2$ система решается исключением Гаусса, причём сложение строк выполняется операцией XOR. Если матрица имеет полный ранг, коэффициенты определяются однозначно.

Почему полный ранг обеспечен при минимальном порядке m. Если матрица вырождена, существует ненулевой вектор b её ядра. Тогда последовательность $w_t = \sum_{j=0}^{m-1} b_j\, u_{t+j}$ имеет первые m значений, равные нулю. Она удовлетворяет исходной рекурсии,

поэтому все дальнейшие значения также нулевые. Выбрав последний ненулевой коэффициент b_j, получаем рекурсию для u длины меньше m. Это противоречит минимальности.

Утверждение «любые $2m$ бит всегда однозначно восстанавливают любой регистр из m разрядов» неверно без условия о минимальности. Для нулевого выхода или регистра избыточной длины матрица может быть вырождена. Несовместная система означает, что выбранная длина не описывает наблюдаемый фрагмент. Несколько дополнительных битов позволяют проверить восстановленную модель.

Пример 5. Наблюдаются биты 100101. Для $m = 3$ система имеет вид

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Отсюда $c = (1,1,0)$. Начальное состояние — первые три бита, то есть 100. Следующие четыре бита после наблюдаемого фрагмента: 1100. Если фактический фрагмент имеет вид 1001011100, эти четыре бита подтверждают восстановленную рекурсию.

7 Применение в криптографии и границы линейной модели

В двоичном поточном шифровании гамма z_t складывается с открытым текстом P_t:

$$C_t = P_t \oplus z_t, \qquad P_t = C_t \oplus z_t.$$

Одинаковая операция шифрует и расшифровывает благодаря $z_t \oplus z_t = 0$. Если известен фрагмент открытого текста и соответствующий шифртекст, гамма этого фрагмента вычисляется как $z_t = P_t \oplus C_t$. Для одиночного линейного регистра её продолжение затем восстанавливается линейной системой из раздела 6.

Даже если коэффициенты регистра открыты, его выход можно эффективно отличать от равномерной строки. Наблюдатель проверяет, совпадает ли очередной бит с линейной комбинацией предыдущих m битов. Для выхода регистра проверка всегда успешна. Для равномерной строки последний бит независим от предыдущих и совпадает с заданной комбинацией с вероятностью $1/2$. Различие вероятностей равно $1/2$, поэтому такая схема не является криптографическим псевдослучайным генератором.

Повтор гаммы. Если одну и ту же гамму применить к двум сообщениям, то

$$C^{(1)} \oplus C^{(2)} = P^{(1)} \oplus P^{(2)}.$$

Гамма сокращается независимо от того, как она построена. Это отдельная проблема её повторного использования.

Линейные регистры используются как математические и аппаратные компоненты. Для криптографического применения необходима конструкция с обоснованной стойкостью; добавление произвольной нелинейной операции само по себе такого обоснования не даёт. Учебный регистр здесь нужен для исследования структуры последовательности и её восстановления.

Современная архитектура генерации случайных битов различает источник энтропии и детерминированный криптографический механизм. NIST SP 800-90C, опубликованный в сентябре 2025 года, описывает конструкции, объединяющие источники из SP 800-90B с механизмами из SP 800-90A. NIST SP 800-90C.

Упражнения для закрепления

1. Для $u_{t+3} = u_t + u_{t+1}$ над $\mathbb{F}_2$ и состояния 100 выпишите цикл, выход за период и числа нулей и единиц. Найдите всю периодическую автокорреляцию.
2. Докажите, что $X^4 + X^3 + X^2 + X + 1$ неприводим, но не примитивен над $\mathbb{F}_2$. Найдите период выхода регистра с состоянием 1000.
3. Сравните периоды состояний 100 и 111 для $u_{t+3} = u_t$. Затем найдите предпериод и период состояния 100 для $u_{t+3} = u_{t+1}$.
4. Восстановите рекурсию длины 3 по битам 100101. Предскажите следующие четыре бита. Объясните, почему для строки 111111 коэффициенты регистра длины 3 неоднозначны.
5. Зашифруйте строку 11001010 первыми восемью битами регистра из задачи 1 и проверьте расшифрование. Выведите равенство для XOR двух шифртекстов с общей гаммой.
6. Над $\mathbb{F}_3$ рассмотрите $u_{t+2} = u_t + u_{t+1}$ с состоянием $(1,0)$. Выпишите цикл состояний и выход за период; подсчитайте в нём символы 0, 1 и 2.

Контрольные вопросы

1. Чем криптографическое требование псевдослучайности отличается от проверки частот символов?
2. Почему при одинаковых начальных данных регистр выдаёт одинаковую последовательность?
3. Как определяется состояние и какой разряд принят за выход в этой лекции?
4. Какое условие обеспечивает обратимость перехода и отсутствие предпериода?
5. Почему максимальный период ненулевого состояния ограничен числом $q^m - 1$?
6. Как записывается характеристический многочлен по коэффициентам рекурсии?
7. Чем неприводимый многочлен отличается от примитивного?
8. Почему при проверке примитивности достаточно степеней с показателями M/r?
9. Как вывести баланс нулей и единиц из множества состояний?
10. Почему ненулевой циклический сдвиг m-последовательности даёт автокорреляцию -1?
11. При каком условии $2m$ наблюдаемых символов однозначно определяют коэффициенты?
12. Почему максимальный период одиночного линейного регистра не делает его стойким генератором гаммы?

Литература

1. Paar C., Pelzl J. **Understanding Cryptography.** Springer, 2010. § 2.3.1–2.3.2, с. 41–46: линейные регистры и восстановление по известному открытому тексту.
2. Trappe W., Washington L. C. **Introduction to Cryptography with Coding Theory.** 2nd ed., 2006. § 2.11, с. 43–50: последовательности LFSR и восстановление рекурсии.
3. Katz J., Lindell Y. **Introduction to Modern Cryptography.** 2nd ed., 2015. § 3.3, определение 3.14, с. 60–63: криптографическая псевдослучайность. Использовано предоставленное второе издание.
4. Barker E. et al. **Recommendation for Random Bit Generator Constructions.** NIST SP 800-90C, September 2025. Официальная публикация.