Научная криптография

(30-е - 60-е годы XX века)

Основные этапы развития криптографии

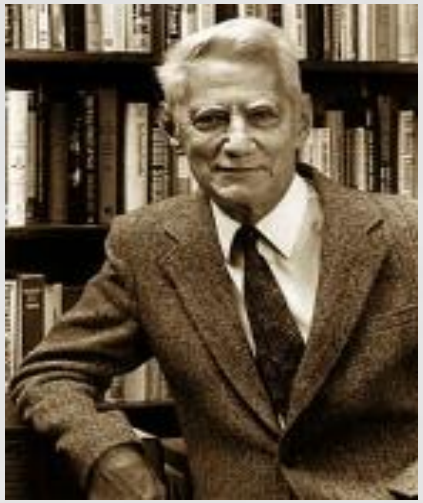
1. • Интуитивная (до начала XVI века)
2. • Формальная (конец XV века - начало XX века)
3. • Научная (30-е - 60-е годы XX века)
4. • Компьютерная (с 70-х годов XX века)

Вклад академика Котельникова Владимира Александровича

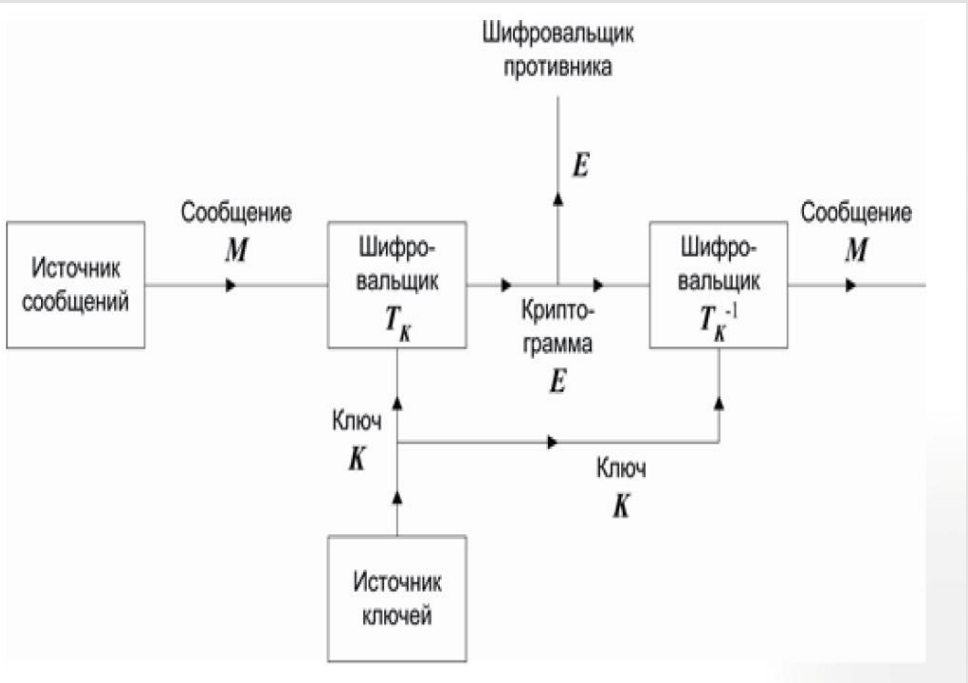
- Доказана теорема отчетов (1933) -непрерывный сигнал с ограниченным спектром можно точно восстановить по его дискретным отчетам, если они были взяты с частотой дискретизации, превышающей максимальную частоту сигнала минимум в два раза
- Сформулированы (1941) технические принципы построения стойкой (недешифруемой) системы засекречивания сигналов:
  - Система должна быть цифровой, а преобразование аналогового сигнала в цифровую форму должно основываться на теореме отсчётов.
  - Каждый знак сообщения должен засекречиваться выбираемым равновероятно знаком алфавита шифра.

Теория секретной связи Клода Шеннона 1949

- Изложена в работе «Теория связи в секретных системах» (Communication Theory of Secrecy Systems) опубликованной в 1949
- Сделано обобщение накопленный до него опыт разработки шифров
- Дано формальное определение абсолютно (теоретически) стойкого шифра и доказано, что такой шифр существует
- Разработаны принципы построения практически стойких шифров



Классическая схема секретной системы связи



- $M = \{M_i\}$-множество сообщений с априорными вероятностями $\{p(M_i)\}$
- $E = \{E_j\}$-множество криптограмм (шифротекстов) с априорными вероятностями $\{p(E_j)\}$
- $K = \{K_l\}$-множество ключей с априорными вероятностями $\{p(K_l)\}$
- T_K- отображение: $E = T_K(M); M = T_K^{-1}(E)$

- Секретная система - это семейство однозначно обратимых отображений T_K множества возможных сообщений M во множество криптограмм E, при этом каждое отображение из T_K имеет вероятность $p(K_l)$

Совершенная секретная система

- Секретная система называется совершенной, если знание криптограммы не дает противнику никакой дополнительной информации о сообщении:

$$p(M_i/E_j) = p(M_i),$$

$p(M_i/E_j)$- условная вероятность того, что именно сообщение M_i при зашифровании преобразовано в криптограмму E_j

$p(M_i)$ - априорная вероятность присутствия сообщения M_i

- Доказано, что совершенные секретные системы существуют. Например - блокнот одноразового использования (шифр Вернама):

$$E = M \oplus K$$

при условиях:

- равенство длины ключа и длины сообщения: $E, M, K \in \{0,1\}^N$
- полная случайность (равновероятность) ключа: $p(K_l)=\frac{1}{2^N}$
- однократность использования ключа

Матричное представление совершенно секретной системы

K

M

⊕	000	001	010	011	100	101	110	111
000	000	001	010	011	100	101	110	111
001	001	000	011	010	101	100	111	110
010	010	011	000	001	110	111	100	101
011	011	010	001	000	111	110	101	100
100	100	101	110	111	000	001	010	011
101	101	100	111	110	001	000	011	010
110	110	111	100	101	010	011	000	001
111	111	110	101	100	011	010	001	000

E

Латинский квадрат 2^n порядка (n=3)

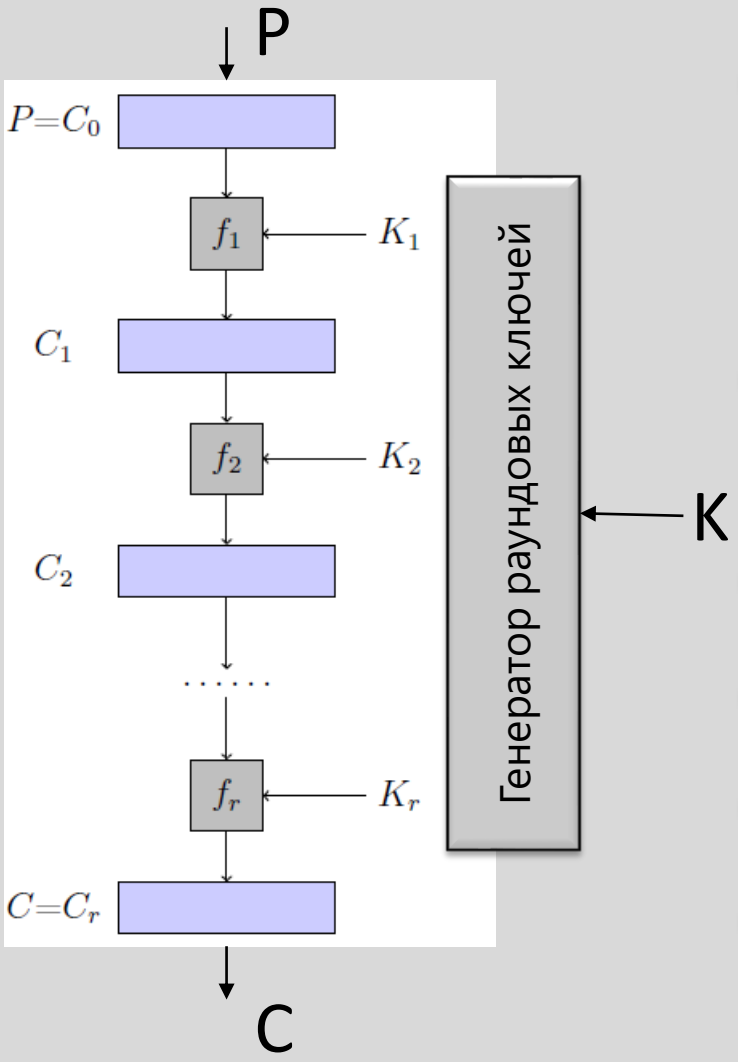
Понятие информационной энтропии

- **Информационная энтропия** - мера хаотичности информации или мера внутренней неупорядоченности информационной системы
- **Информация** – это снятая неопределенность
- **Избыточность** – это превышение количества информации, используемой для передачи или хранения сообщения, над его информационной энтропией
- Взаимосвязь энтропии и информации отражается в формуле ***H + I = const***

Принципы построения блочных шифров

- Предложено реализовывать блочные шифры путем многократного применения к блокам открытого текста некоторых базовых преобразований (модульная арифметика, замена, циклический сдвиг и др.), которые просто реализуемы, и при небольшом числе повторений обеспечивают сложные преобразования.
- Эти преобразования должны обеспечивать реализацию двух криптографических принципов – «рассеивание» и «перемешивание».
  - Цель «рассеивания» ***(Diffusion)*** состоит в том, чтобы распространять влияние одного символа открытого текста на как можно большее число знаков шифрованного текста. Это позволяет скрыть статистическую зависимость между символами открытого текста.
  - Цель «перемешивания» ***(Confusion)*** состоит в том, чтобы зависимость между открытым текстом, ключом и зашифрованным текстом сделать как можно более сложной, чтобы нельзя было обнаружить связь между открытым и зашифрованным текстом.

Схема блочного шифрования



- P- это блок открытого текста фиксированного размера (например, 64, 128, 256… бит)
- K – секретный ключ фиксированного размера (например, 64, 128, 256… бит)
- f - раундовое преобразование
- K_i - раундовый ключ
- C_i - промежуточный шифротекст
- C – это блок шифротекста

Примитивные операции

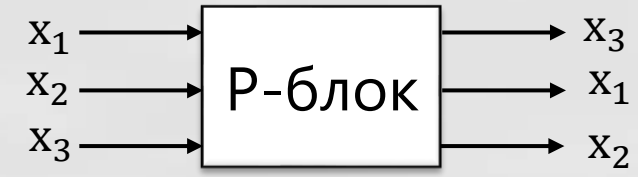
- Стандартное кодирование
- Исключающее ИЛИ (XOR)
- Сложение по модулю 2^n
- Циклический сдвиг
- Перестановка (Permutation)
- Подстановка (Substitution)

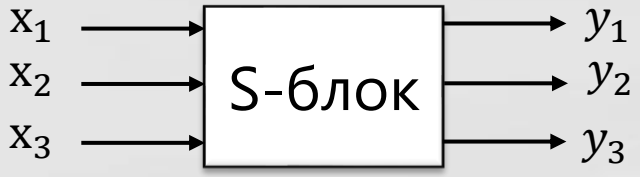
$(10101111100110110000001111010101) \rightarrow$ `af9b03d5`

$(1101) \oplus (1011) = (0110)$

$(1101) \boxplus (1011) = (1000)$

$(11010000) \ <<< \ 3 \ = \ (10000110)$



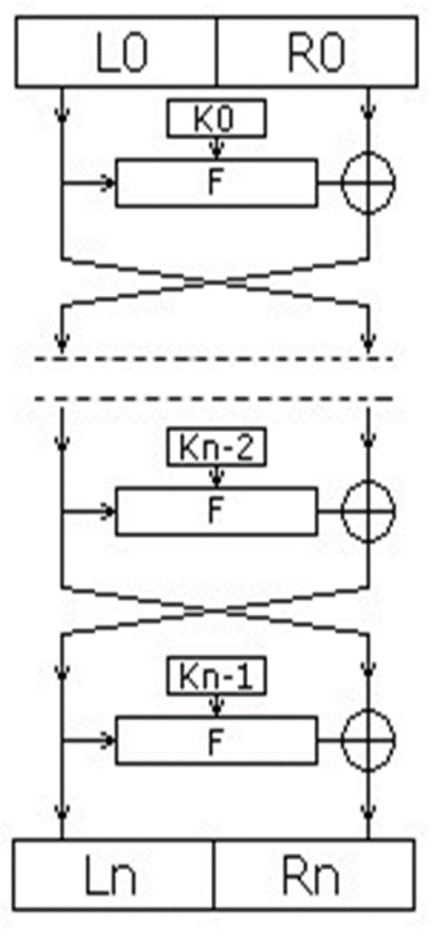


Принципы обеспечения криптостойкости шифров

- Криптографическая стойкость (или криптостойкость) — способность криптографического алгоритма противостоять криптоанализу.
- Обеспечение устойчивости шифра ко всем известным атакам - это до сих пор лучший из известных принципов построения симметричных шифров
- Сведение проблемы обеспечения секретности шифра к одной из известных вычислительно сложных задач (этот принцип используется при создании асимметричных шифров, но не используется для симметричных)

Шифр (сеть) Фейстеля 1971

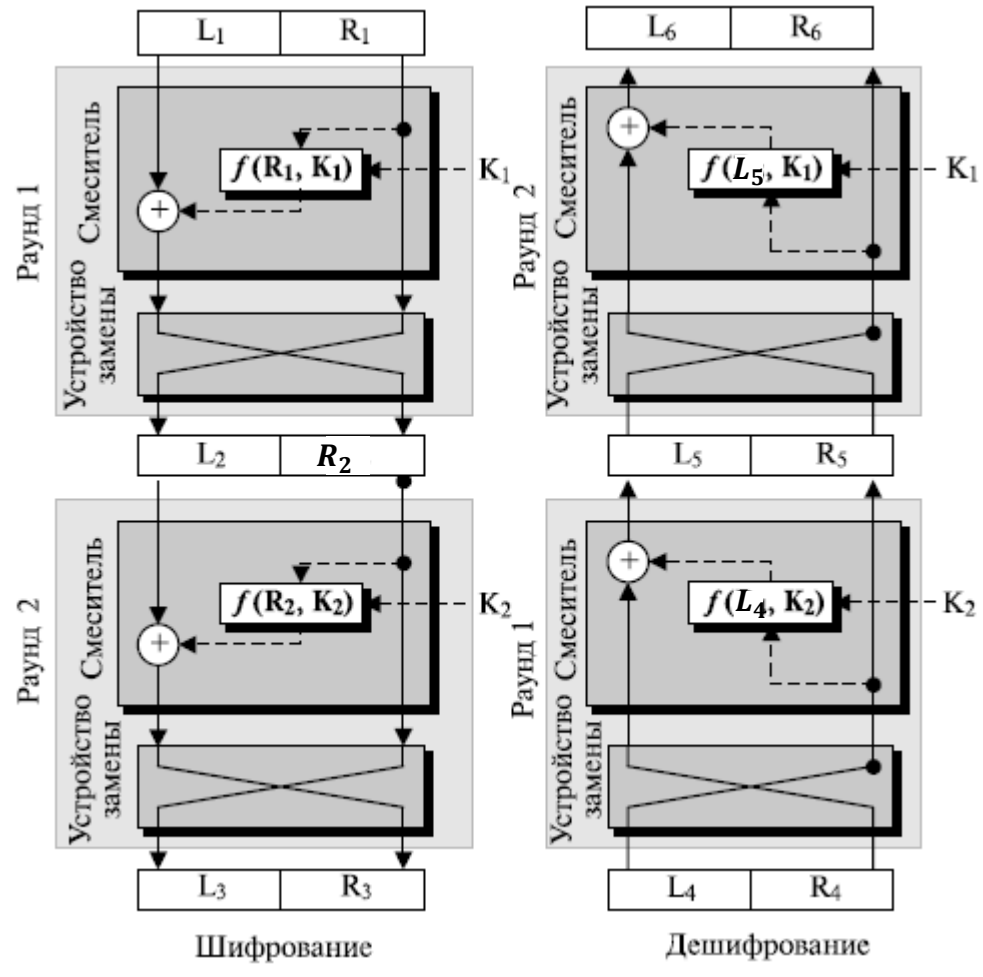
- Выбранный блок делится на два равных субблока — «левый» (L0) и «правый» (R0).
- «Левый субблок» L0 преобразуется функцией шифра F(L0,K0), после чего он складывается по модулю 2 (смешивается) с «правым субблоком» R0.
- Результат сложения заменяет «левый субблок» L1 (входное данное для следующего раунда), а «левый субблок» L0 заменяет «правый субблок» R1 (другое входное) данное для следующего раунда.
- Операция повторяется N-1 раз, при этом при переходе от одного раунда к другому меняются раундовые ключи (K0 на K1 и т. д.) по какому-либо математическому правилу, где N — количество раундов в заданном алгоритме
- Генерация раундовых ключей происходит на базе ключа шифрования и зависит от алгоритма шифрования



Свойства сети Фейстеля

- Увеличение количества раундов значительно увеличивает криптостойкость алгоритма
- Основной характеристикой алгоритма, построенного на основе сети Фейстеля, является функция F. Сеть Фейстеля является обратимой даже в том случае, если функция F не является таковой, так как для расшифрования не требуется вычислять функцию обратную F
- Необходимы начальное и конечного преобразования блока, для того, чтобы выполнить начальную рандомизацию входного текста. Подобные преобразования называются. забеливанием (***whitening***)

Обратимость сети Фейстеля



Пусть $L_3=L_4$ и $R_3=R_4$ (шифротекст передан без искажений), тогда:

$$R_5 = L_4 = L_3 = R_2$$

$$L_5 = R_4 \oplus f(L_4, K_2) = R_3 \oplus f(R_2, K_2) = L_2 \oplus f(R_2, K_2) \oplus f(R_2, K_2) = L_2$$

и окончательно имеем:

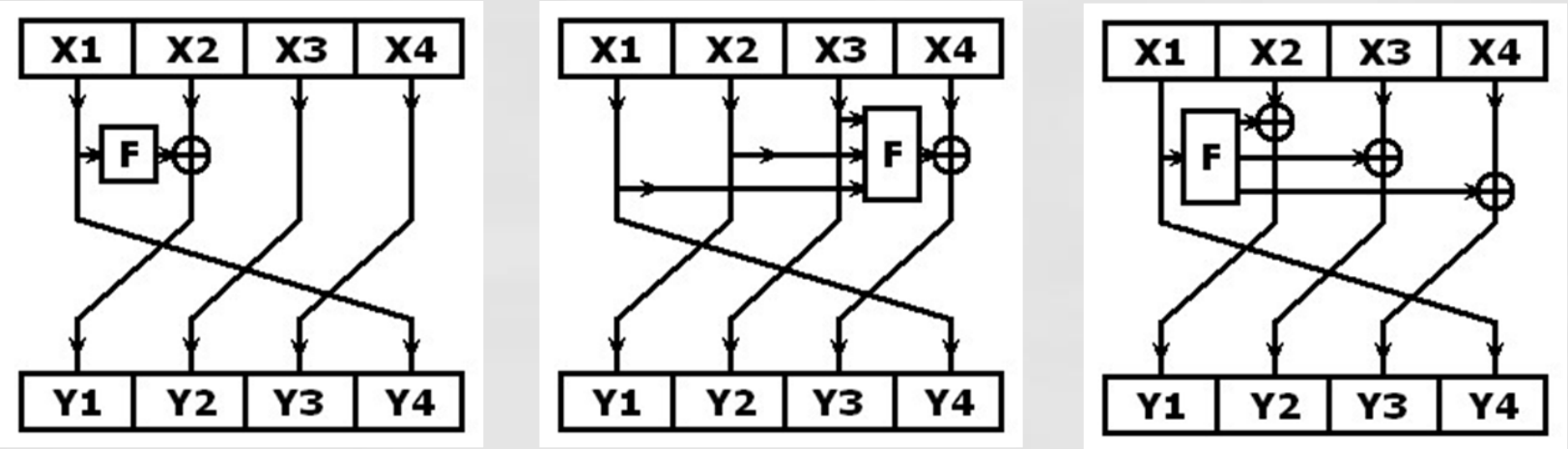
$$L_6 = R_5 \oplus f(L_5, K_1) = R_2 \oplus f(L_2, K_1) = L_1 \oplus f(R_1, K_1) \oplus f(R_1, K_1) = L_1$$

$$R_6 = L_5 = L_2 = R_1$$

т.е. открытый текст и расшифрованный шифротекст совпадают

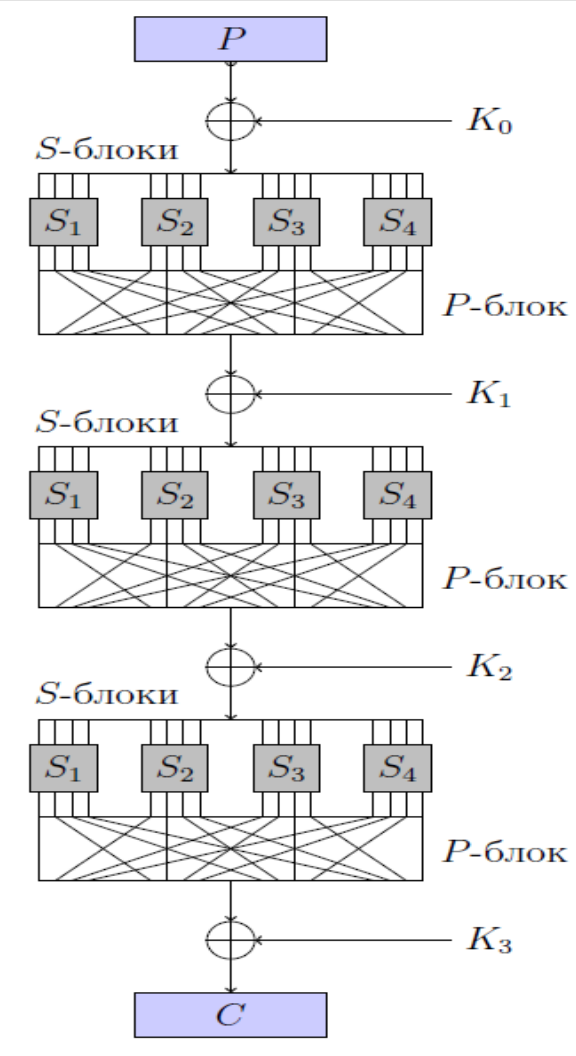
Модификации сети Фейстеля

В настоящее время все чаще используются различные разновидности сети Фейстеля для 128-битного блока с четырьмя ветвями.



Несимметричные сети Фейстеля с 4-мя ветвями

Перестановочно-подстановочная сеть (SP-сеть)



- S-блок (S-box, узел замены) — это в общем случае нелинейное отображение множества двоичных векторов длины n в множество двоичных векторов длины m. Числа n и m относительно малы (например, 4, 6, 8, 16, 32
- P-блок (P-box, permutation) — это перестановка на множестве элементов двоичного вектора длины n. Количество бит n, как правило, достаточно большое. Часто оно совпадает с длиной шифруемого блока (например, 64, 128 и т. п.)

Теория сложности вычислений

- Вычислительная сложность — понятие в информатике и теории алгоритмов, обозначающее функцию зависимости объёма работы, некоторого алгоритма от размера входных данных
- Объём работы обычно измеряется абстрактными понятиями времени и пространства, называемыми вычислительными ресурсами
- Одной из основных абстрактных моделей вычисления в теории сложности является Машина Тьюринга
- В качестве ресурсов обычно берутся время вычисления (количество рабочих тактов машины Тьюринга) или рабочая зона (количество использованных ячеек на ленте во время работы)

Теория сложности вычислений

- Классами сложности называются множества вычислительных задач, примерно одинаковых по сложности вычисления. Для каждого класса существует категория задач, которые являются «самыми сложными» и любая задача из класса сводится к такой задаче.
- Класс P *(Polynomial)* - множество задач, которые могут быть решены на детерминированной машине Тьюринга за полиномиальное время от длины входа
- Класс NP *(Non-deterministic polynomial)* - множество задачи распознавания , которые недетерминированная машина Тьюринга в состоянии решить за полиномиальное количество шагов от размера входных данных.
- Класс NP включает в себя класс P. Вопрос о равенстве этих двух классов считается одной из самых сложных открытых проблем в области теоретической информатики.

Классика

- Гэри М. , Джонсон Д. Вычислительные машины и труднорешаемые задачи .

Издательство Мир 1982 г.

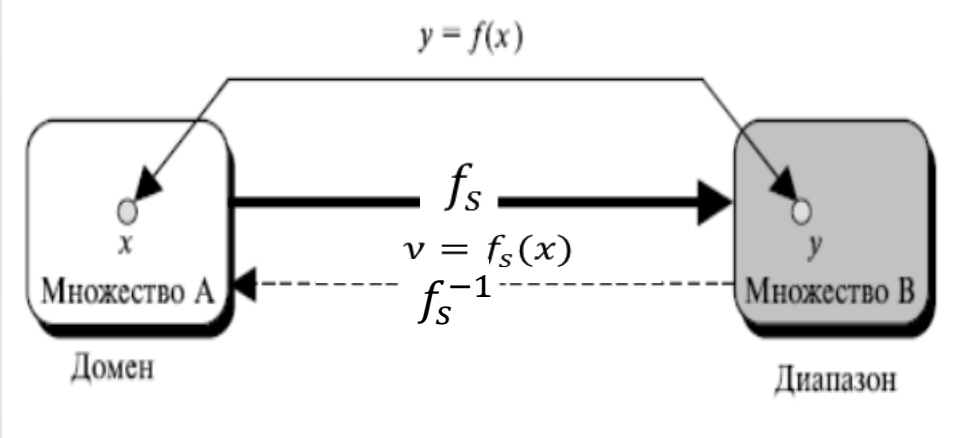
Монография американских ученых посвящена решению сложных (в том числе и NP-трудных) комбинаторных задач, возникающих в дискретной оптимизации, математическом программировании, алгебре, теории автоматов с примерами.

Асимметричная криптография

- Первой открытой публикацией в области асимметричной криптографии принято считать статью Уитфилда Диффи (Whitfield Diffie) и Мартина Хеллмана (Martin Heilman) «Новые направления в криптографии», опубликованную в 1976 г.
- Предложен алгоритм, позволяющий паре пользователей выработать числовой материал для общего секретный ключ, не обмениваясь секретными данными по небезопасному каналу связи (см. лабораторную работу № 7 из «Методическое пособие ….)
- В «новой криптографии» введено понятие одно-сторонней функции с секретом.

Односторонняя функция с секретом (люком)
(TOWF — Trapdoor One Way Function)

- Зная x, при любом s легко вычислить y=f_s(x)
- По известному значению y и s легко вычислить x=$f_s^{-1}(y)$
- Сложно вычислить x=$f_s^{-1}(y)$ по известному y, если секрет s не известен

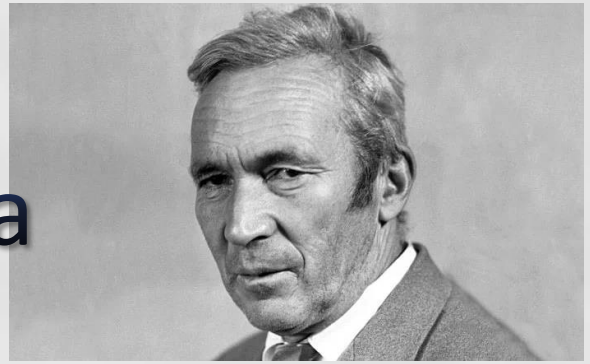


- Предположение о высокой сложности операции обратимости односторонних функций положено в основу криптографии в условиях зарождающейся компьютерной эпохи

Примеры используемых односторонних функций

- Протокол распределения ключей DH: $y = g^x mod\ n$ - вычислительно сложно найти *x*, зная *y,g,n* (задача дискретного логарифмирования)
- Протокол шифрования RSA: $y = x^e mod\ n$ – вычислительно сложно найти x, зная y,e,n и не зная разложения n=p*q на два простых сомножителя (задача факторизации)

Односторонняя функция и теория сложности академика Колмогорова Андрея Николаевича



- Ни для одной из ныне используемых функций не было определённо доказано, что они являются односторонними
- Доказано, что существование истинных односторонних функций зависит от решения задачи распознавания случайных строк чисел и строк чисел , содержащих некоторую информацию
- А. Н. Колмогоров определил сложность строки, как длину кратчайшей возможной программы, создающей на выходе эту строку. Чем больше код программы , тем случайнее набор цифр составляющих число

Проблема односторонности

- Доказано, что колмогоровская сложность не вычисляема - не существует программы, способной вычислить сложность любой возможной строки
- Однако, «если задача вычисления приблизительной, ограниченной по времени колмогоровской сложности является трудно решаемой, то истинные односторонние функции обязаны существовать» (Рафаэль Пасс, Яньи Лю, 2021)

Колмогоровская сложность

Один из показателей случайности основан на степени простоты описания строки чисел. Чем проще программа, способная вывести строку, тем меньше колмогоровская сложность строки.

СТРОКА	ПРОГРАММА	СЛОЖНОСТЬ
13579111315171 9212325272931 333537394143	**Вывод: первые 22 нечётных числа**	Меньше
99999999931415 9265358979323 846264338327	**Вывод: девять девяток, потом первые 30 символов пи**	
3238226618594 0767096401028 0409881234079	**Вывод:** 32382266185940767096 4010280409881234079	Больше

Вехи отечественной криптографии XX века

- 5 мая 1921 года была образована криптографическая служба при ВЧК. 5 мая в нашей стране ежегодно отмечается **День шифровальщика**
- 19 октября 1949 года было принято решение Центрального комитета ВКП(б) о создании ГУСС – координатора единой криптографической службы СССР. 19 октября в нашей стране ежегодно отмечается **День криптографа**
- Тогда же, в 1949 году, открылась Высшая школа криптографов с двухлетним обучением, обеспечивавшая получение второго высшего специального образования. В 1960 году ВШК была преобразована в технический факультет Высшей школы КГБ В 1992 году был создан **Институт криптографии, связи и информатики в составе Академии ФСБ России**

Ссылки по истории отечественной криптографии

- Соболева Татьяна. История шифровального дела в России – https://royallib.com/book/soboleva_tatyana/istoriya_shifrovalnogo_dela_v_rossii.html
- Синельников Андрей. Шифры советской разведки - http://www.hrono.ru/statii/2008/shifr0.html
- Шифровальные машины СССР 1931-1991 гг. (Обзор) - https://yarovan.ru/shifrovalnye-mashiny-sssr/