На правах рукописи

Пудовкина Марина Александровна

СВОЙСТВА ПРОГРАММНО РЕАЛИЗУЕМЫХ ПОТОЧНЫХ ШИФРОВ (НА ПРИМЕРЕ RC4, GI, ВЕСТА)

Специальность: 05.13.19 - методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Автор: _____

Москва-2004

Работа выполнена в Московском государственном инженерно-физическом институте (государственном университете)

Научный руководитель: доктор физ.-мат. наук, профессор Борис Александрович Погорелов

Официальные оппоненты: доктор физ.-мат. наук, профессор Грушо Александр Александрович

кандидат физ.-мат. наук, доцент Велигура Александр Николаевич

Ведущая организация: Московский государственный институт электроники и математики Министерства образования и науки Российской Федерации

Защита состоится «29» апреля 2004 г. в 14 часов на заседании диссертационного совета ДМ 212.130.08 в МИФИ по адресу: 115409, Москва, Каширское шоссе, д.31.

С диссертацией можно ознакомиться в библиотеке МИФИ.
Автореферат разослан «29» июня 2004 г.

Просим принять участие в работе совета или прислать отзыв в одном экземпляре, заверенный печатью организации.

Ученый секретарь
диссертационного совет к.т. н. Горбатов В. С.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. В настоящее время в мире происходит активное развитие электронной коммерции. При разработке систем электронной коммерции фактор безопасности играет первостепенную роль. При этом типовой задачей является создание системы работы с партнерами, предоставляющей защищенный доступ к динамически обновляемой информации. В некоторых случаях система должна предоставлять возможность оформления заказа и резервирования товаров. Для этого она должна обеспечивать аутентификацию, невозможность отказа от авторства и конфиденциальность пользователя. Многие производители компьютеров, крупные компьютерные дистрибьюторы, а также многие из некомпьютерных международных корпораций уже используют системы с элементами обеспечения информационной безопасности в России. Одни системы разработаны специалистами самих фирм, другие используют стандартные решения.

Одним из эффективных способов создания подсистемы защиты в системах электронной коммерции на сегодняшний день является применение криптографических методов защиты информации (шифрование, идентификация, имитозащита). Основу большинства криптографических методов составляют симметричные алгоритмы шифрования, которые принято подразделять на блочные и поточные. Скорость работы поточных алгоритмов обычно значительно превышает скорость работы блочных. Значительная часть применяемых в настоящее время алгоритмов поточного шифрования продолжает оставаться неизвестной.

Поскольку Интернет технологии требуют высоких скоростей, то одной из актуальных задач в криптографии является разработка и реализация высокоскоростных программно реализуемых поточных алгоритмов шифрования, обеспечивающих высокую надежность защиты информации, с хорошими техническими и эксплуатационными свойствами. К настоящему

времени значительная часть предложенных в открытой литературе поточных шифров основана регистрах сдвига над полем GF(2).

В значительной степени это объясняется разработанной теорией и удобством при аппаратной реализации алгоритмов шифрования. Дня алгоритмов, основанных на регистрах сдвига над полем GF(2), достаточно хорошо разработаны методы синтеза и криптоанализа, в тоже время относительно мало изучены высокоскоростные программно реализуемые нерегистровые поточные алгоритмы и практически не развита теория их синтеза и методов криптоанализа, хотя при программной реализации они могут быть предпочтительнее.

Существенное повышение производительности микропроцессоров к 80-м годам вызвало в криптографии усиление интереса к программным методам реализации алгоритмов шифрования как возможной альтернативе аппаратным схемам. Одним из самых первых подобных алгоритмов шифрования, получившим широкое распространение в электронной коммерции, стал алгоритм поточного шифрования RC4 (также известный как алгоритм ARCFOUR). Он, например, используется во многих платежных системах. В России, кроме RC4, программно реализуемыми алгоритмами поточного шифрования, используемыми в электронной коммерции, являются Веста-2, Веста-2М. Поэтому основное внимание в диссертации уделяется как программно реализуемым неригистровым алгоритмам: RC4, предложенному в диссертации его обобщению (включающему IA, IBAA, ISAAC), Solitaire, так и регистровым: Веста-2, Веста-2М. Рассмотренные алгоритмы шифрования включают большинство наиболее распространенных среди программно реализуемых алгоритмов поточного шифрования. Единство исследований достигается общей математической моделью, изложенной на теоретико-автоматном языке, едиными математическими методами.

Целью диссертационной работы является разработка общих математических моделей, включающих ряд алгоритмов шифрования и методов их криптоанализа; исследование криптографических свойств (теоретико-

автоматных, теоретико-групповых, теоретико-вероятностных) программно реализуемых алгоритмов поточного шифрования RC4 и различных его обобщений GI (IA, IBAA, ISAAC), Solitaire, Веста.

Для достижения поставленной цели, используя единую теоретико-автоматную модель, были решены следующие задачи:

- Проведен обзор современных программно реализуемых поточных алгоритмов и методов их криптоанализа в рамках этой модели;
- Введено в рамках теоретико-автоматной модели семейство алгоритмов поточного шифрования GI, обобщающее ряд известных алгоритмов шифрования (IA, IBAA, ISAAC), предложенных в открытой литературе, и получена верхняя оценка числа знаков гаммы, необходимых для восстановления начального состояния по гамме алгоритма GI;
- Описан ряд криптографических свойств алгоритмов RC4, GI, Solitaire, Веста;
- Разработаны методы восстановления начального состояния по гамме алгоритмов семейств GI и Веста, основанные на их алгебраических свойствах.

Методы исследования: теоретическая криптография, теория автоматов, комбинаторный анализ, теория вероятностей, математическая статистика, теория групп, полугрупп.

Научная новизна работы заключается в следующем:

1. Описан ряд классов слабых состояний алгоритмов RC4, GI, Веста-2, и подсчитана их мощность;
2. Получены распределения первого, второго знака и биграмм в гамме RC4 при предположении о равновероятности выбора начальной подстановки из множества всех подстановок. Построены критерии различения гаммы RC4 от случайной равновероятной последовательности;
3. Подсчитано число ключей алгоритма RC4, приводящих к начальным подстановкам с произвольной фиксированной цикловой структурой;

4. Введена теоретико-автоматная модель семейства алгоритмов поточного шифрования GI, обобщающая ряд известных алгоритмов шифрования (IA, IBAA, ISAAC), предложенных в открытой литературе. Найдена верхняя оценка числа знаков, необходимых для восстановления начального состояния по гамме;
5. Разработаны методы восстановления по гамме начального состояния алгоритмов GI и Веста;
6. Описаны групповые свойства преобразований, связанных с алгоритмами Solitaire и Веста.

Результаты, выносимые на защиту.

- Разработка модели семейства алгоритмов шифрования GI и описание ее свойств;
- Определение распределения первого, второго знаков и биграмм в гамме RC4 при предположении о равновероятности выбора начальной подстановки из множества всех подстановок;
- Определение числа ключей алгоритма шифрования RC4, приводящих к начальным подстановкам с произвольной фиксированной цикловой структурой;
- Определение вида и числа слабых состояний алгоритмов шифрования RC4, Веста-2 и GI;
- Методы восстановления по гамме состояний автоматов, моделирующих алгоритмы шифрования GI и Веста, и оценка их трудоемкостей.

Практическую значимость представляют результаты работы, посвященные исследованию и описанию слабостей широко используемого в электронной коммерции шифра RC4, а также методы восстановления начального состояния используемых в электронной коммерции в России шифров Веста-2, Веста-2М. Результаты работы могут быть использованы для различения последовательностей, выработанных алгоритмом поточного шифрования RC4, от случайных равновероятных последовательностей и для восстановления состояний RC4, принадлежащих циклам длины $m(m-1)$.

Внедрение результатов исследований. Полученные свойства алгоритмов Веста-2, Веста-2М используются фирмой «ЛАН Крипто» при создании программно реализуемых шифров, предназначенных для электронной коммерции. Результаты исследований, связанные с анализом криптографических свойств поточных алгоритмов шифрования Solitaire, IA, IBAA, ISAAC внедрены в учебный процесс на факультете «Информационная безопасность» Московского государственного инженерно-физического института (государственного университета).

Публикации и апробация работы.

Результаты диссертации изложены в 40 публикациях и докладывались на конференциях и семинарах различного уровня, в том числе:

- на международной конференции EUROCRYPT'2001, EUROCRYPT'2002, EUROCRYPT2003 (rump sessions);
- в трудах международной конференции «First International IFIP TC-11 WG 11.4 Working Conference on NETWORK SECURITY», Leuven (Belgium) 2001;
- в трудах международной конференции «International Workshop on Computer Science and Information Technologies», 2001-2003 гг.;
- в трудах международной конференции «Discrete Mathematics and Theoretical Computer Science», France, 2003;
- на международной конференции «РусКрипто» в 1999-2004 гг.;
- в трудах XXIII конференция молодых ученых мехмата МГУ, Москва, 2001;
- в трудах XLIV юбилейной научной конференции МФТИ, 2001;
- в сборнике тезисов Российской научно-технической конференции «Методы и технические средства обеспечения безопасности информации», Санкт-Петербург, (2000-2003 гг.);
- на Всероссийской научно-практической конференции «Проблемы информационной безопасности в системе высшей школы», Москва, (2000-2003 гг.);

- в трудах седьмого международного научного семинара «Дискретная математика и ее приложения», МГУ, 2001;
- на общероссийской конференции с международным участием «Математика и безопасность информационных технологий» (МаБИТ-03), Москва;
- в трудах семинара «Информационная безопасность-Юг России», 2001;
- на международной конференции «Computer data analysis and modeling», Minsk, 2004.
- на III Сибирской научной школе-семинаре с международным участием «Компьютерная безопасность и криптография» - SIBECRYPT'04.
- написано 1 учебное пособие;
- на научных семинарах в ФАПСИ, ФСБ, ИКСИ, МИФИ, МГТУ.

Структура и объем работы.

Диссертация состоит из введения, пяти глав, списка литературы из 100 наименований. Работа изложена на 150 страницах с рисунками и таблицами.

СОДЕРЖАНИЕ РАБОТЫ

Во введении показана актуальность темы диссертации, определены цели и задачи проведенных исследований, отражены научная новизна и практическая значимость полученных результатов.

В имеющейся открытой литературе по поточным шифрам, как и шифрам вообще, в большинстве своем явно недостаточное использование в изложение доказательной базы, оно часто носит описательный характер. В связи с этим в первой главе, следуя работам отечественных математиков-криптографов, а также Рюппеля, в рамках единой математической модели на языке теории автоматов, выполнен обзор современных программно реализуемых алгоритмов поточного шифрования и методов их анализа.

Алгоритм поточного шифрования принято представлять в виде двух основных блоков. Первый блок, называемый управляющим блоком, предназначен для выработки последовательности, управляющей работой

второго шифрующего блока. Часто управляющая последовательность называется гаммой. Второй блок реализует в соответствии со знаком гаммы собственно функцию зашифрования текущего знака открытого текста. Ключом алгоритма поточного шифрования может являться, например, заполнение памятей узлов и блоков, составляющих управляющих блок, а в ряде случаев, и закон их функционирования.

Введем следующие обозначения. Пусть $Z_m=\{0,1,\ldots,m-1\}$-множество классов вычетов целых чисел по модулю m; $\overline{0,m-1}=0, 1,\ldots,m-1$; S_m – симметрическая группа подстановок множества $\{\overline{0,m-1}\}$; $s=<s[0], s[1],\ldots, s[m-1]>$– запись подстановки $s\in S_m$; $|x|$– модуль числа x, или мощность множества x; $\cong$– изоморфные алгебраические структуры; Р-конечный алфавит, элементы которого являются знаками открытого текста; Y— конечный алфавит, элементы которого являются знаками шифртекста; K — конечное множество всех ключей; Z - конечный алфавит, элементы которого являются знаками гаммы; $GF(q)$– конечное поле из q элементов; V -множество состояний автомата; э.о.— элементарная операция.

Для описания функционирования дискретных устройств, реализующих отдельные блоки шифратора, часто применяется язык теории автоматов.

Пусть $F: P\times K\times V\to V$, $f: P\times K\times V\to Y$–функции, где

$$F((x, k), v)=F_k(x, v)= F_{(x, k)}(v),$$

$$f((x, k), v)=f_k(x, v)= f_{(k,v)}(x).$$

Автомат $A=(P\times K, Y, V, F, f)$ называется *шифрующим автоматом*, если

1. автомат А регулярный, т.е. частичная функция $F_{(x, k)}: V\to V$ – биекция для любых пар $(x, k)\in P\times K$;
2. при фиксированном ключе $k\in K$ и состоянии $v\in V$ отображение $f_{(k,v)}: P\to Y$ является инъективным.

Алгоритм поточного шифрования моделируется шифрующим автоматом $A_п=(P\times K, V, Y, F, f)$, работа которого описывается в i-м такте, $i\geq 1$, уравнениями:

$$v_{i+1}=F_k(v_i, x_i);$$

$$y_i=f_k(v_i, x_i).$$

Условимся дальше под состоянием алгоритма шифрования понимать состояние автомата, моделирующего этот алгоритм.

Пусть $P=Z=Y$ и на Z задана групповая операция $\otimes$. Часто, например, в шифрах гаммирования, $f_k(v_i, x_i)$ зависит от x_i "линейно" относительно операции $\otimes$ на Z, т.е.

$$y_i=x_i\otimes f'_k(v_i).$$

Последовательность $\{z_i=f'_k(s_i): i\geq 1\}$ называется гаммой.

В последние 10 лет активно исследовались следующие программно реализуемые поточные алгоритмы: Pike, Scop, Dagger, Sober, Sober -t16, Bmgl, Sober-t32, RSC, Lili, Leviathan, RC4, Wake, Seal, Twoprime, ISAAC, IA, IBAA, Chameleon, Panama, Rabbit, Solitaire, Веста. Среди перечисленных алгоритмов только RC4, Wake, Веста используются практически. Алгоритмы шифрования RC4, IA, IBAA, ISAAC, Веста, Solitaire наиболее распространенные среди программно реализуемых поточных алгоритмов.

Поскольку RC4 один из наиболее применяемых в электронной коммерции алгоритмов, то основное внимание при обзоре алгоритмов поточного шифрования уделено результатам, полученным с 1993 по 2003 гг. по анализу алгоритма RC4 следующими авторами: Golic J., Shamir A., Knudsen L., Preneel B., Rijmen V., Fluhrer , McGrew D., Meier W., Verdoolaege S, Mantin I.. Отмечена связь их результатов с результатами диссертации.

Целью **второй главы** является исследование алгебраических и вероятностно-статистических свойств алгоритма поточного шифрования RC4. **Алгоритм RC4 зависит от параметра $m=2^n$, натуральное $n\geq 2$, (для практических приложений выбирается $m=256$). Алгоритм RC4 моделируется автономным автоматом $A=(Z_m\times Z_m\times S_m, Z_m, Z_m, F, f)$, где $F: Z_m\times Z_m\times S_m\to Z_m\times Z_m\times S_m$, $f: Z_m\times Z_m\times S_m\to Z_m$. Состоянием алгоритма в такте $t\geq 1$ является тройка $v_t=(i_t, j_t, s_t)\in Z_m\times Z_m\times S_m$, ключом длины L – слово $k=k_0,\ldots,k_{L-1}$, где $k_i\in Z_m$, $L\leq m$. Начальное состояние – $(i_0, j_0, s_0,)$, где $i_0=0$ и $j_0=0$.**

Функция ρ: $Z_m^* \to S_m$ формирования начального состояния (подстановки) из ключа $k \in Z_m^*$, т.е. $\rho(k) \in S_m$, задается следующим образом.

Пусть s_0' – тождественная подстановка, $i_0=j_0=0$. Для каждого t, $t=\overline{1,m}$, вычисляем:

1. $i_t=t-1$,
2. $j_t = j_{t-1} + s'_{t-1}[i_t] + k_{t-1 (mod\ L)} \pmod m$,
3. $s_t'[i_t]=s'_{t-1}[j_t]$, $s_t'[j_t]=s'_{t-1}[i_t]$, $s_t'[k]=s'_{t-1}[k]$ при $k \notin \{i_t, j_t\}$.

Приведем описание t–го (t=1,2….) такта работы алгоритма.

Функция переходов F

1. $i_t = i_{t-1}+1 \pmod m$;
2. $j_t = j_{t-1} + s_{t-1}[i_t] \pmod m$;
3. $s_t[i_t]= s_{t-1}[j_t]$, $s_t[j_t]= s_{t-1}[i_t]$, $s_t[k]=s_{t-1}[k]$ при $k \notin \{i_t, j_t\}$.

Функция выходов f

Выход: $z_t = s_t[(s_t[j_t]+ s_t[i_t])(\mathrm{mod}\ m)]$.

Шифрование t-го знака открытого текста $x_t=(x_{t,n-1},\ldots,x_{t,0}) \in Z_2^n$ имеет вид $c_t=x_t \oplus z_t$, где $\oplus$ – операция покоординатного сложения в поле F_2. Расшифрование t-го знака шифртекста определяется выражением $x_t=c_t \oplus z_t$.

Будем придерживаться следующих обозначений. Пусть $S(\alpha_1,\ldots,\alpha_m)$– множество всех подстановок из S_m с цикловой структурой $\{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\}$, где $1\cdot\alpha_1+2\cdot\alpha_2+\ldots+m\cdot\alpha_m=m$. Пусть подстановка s выбрана случайно равновероятно из S_m. Известно, что $|S(\alpha_1,\ldots,\alpha_m)|= \dfrac{m!}{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\cdot\alpha_1!\cdot\alpha_2!\ldots\cdot\alpha_m!}$.

Если распределение начальных подстановок $\rho(k)$, генерируемых алгоритмом RC4, равномерно на S_m, то

$$P\{\rho(k)\in S(\alpha_1,\ldots,\alpha_m)\}=\frac{1}{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\cdot\alpha_1!\cdot\alpha_2!\ldots\cdot\alpha_m!}$$

и среднее число ключей, приводящих к начальным подстановкам с фиксированной цикловой структурой, равно

$$N_m(\alpha_1,\ldots,\alpha_m)=\frac{m^m}{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\cdot\alpha_1!\,\alpha_2!\ldots\cdot\alpha_m!}.$$

В §2.1.1 описаны все самые короткие циклы длины $(m(m-1))$ графа Г функции переходов алгоритма RC4 и приведен метод восстановления по гамме состояний, порождающих эти циклы. Справедливы следующие утверждения.

Утверждение 1. *Все состояния* $(i_0, i_0+1, \langle s_0[0],\ldots,s_0[m-1]\rangle)$ *с* $s_0[i_0+1]=1$, $i_0\in Z_m$, *алгоритма RC4 принадлежат циклам длины* $(m-1)\cdot m$. *Число состояний с произвольной заданной парой координат* (i_0, i_0+1), *принадлежащих фиксированному циклу, равно* $m-1$.

Утверждение 2. *Пусть* $z_1, z_2, \ldots, z_{m(m-1)}$ – *гамма длины* $m(m-1)$. *Если* $v_0=(0, 1, \langle s_0[0], 1, s_0[2],\ldots,s_0[m-1]\rangle)$– *начальное состояние алгоритма RC4, то:*

1) $s_0[s_0[0]+k+1]=z_{k(m-1)}$, $k=\overline{1,m}$;
2) *существует единственное* $r\in\overline{1,m}$ *такое, что* $z_{r\cdot(m-1)}=1$; *при этом* $s_0[0]=m-r$, $s_0[k]=z_{(r+k-1)\cdot(m-1)}$, $k=\overline{1,m-1}$.

Из утверждения 2 следует способ восстановления начального состояния алгоритма RC4. Для этого среди знаков гаммы $z_{(m-1)}, z_{2(m-1)}, \ldots, z_{m(m-1)}$ находится элемент $z_{r\cdot(m-1)}=1$, затем полагают $s_0[0]=m-r$, $s_0[k]=z_{(r+k-1)\cdot(m-1)}$, $k=\overline{1,m-1}$.

В §2.1.2 построен изоморфизм (изоморфизм циклов понимается, как изоморфизм графов) между циклами в графе Г и показано, что число изоморфных циклов является делителем числа m.

Рассмотрим биективное преобразование $\varphi: (i, j, s)\to (i', j', s')=(i+1, j+1, \langle s[m-1], s[0],\ldots,s[m-2]\rangle)$. Группа $\langle\varphi\rangle$, порожденная преобразованием φ, является циклической порядка m.

Пусть $v^{\langle\varphi\rangle}$– орбита, содержащая состояние $v=(i, j, s)$. Будем говорить, что состояния $v=(i, j, s)$ и $v'=(i', j', s')$ алгоритма RC4 φ-связны, если они лежат на одной орбите группы $\langle\varphi\rangle$. Рассмотрим произвольное состояние v. На одной с ним орбите лежат состояния $\varphi(v),\ldots,\varphi^k(v),\ldots,\varphi^{m-1}(v)$, $\varphi^0(v)$, которые, очевидно,

все различны. Следовательно, $|v^{<\varphi>}|=m$, поэтому, число орбит равно $m!\cdot m$. Справедлива следующая теорема.

Теорема 3. *Если в графе* Γ *существует цикл, которому принадлежат ровно* k *состояний из множества* $v^{<\varphi>}$, *то* $k|m$ *и существует* $L=m/k$ *изоморфных циклов* $\Omega_1, \Omega_2, \Omega_3, \ldots, \Omega_L$. *Каждому из этих циклов принадлежит ровно* k *состояний из множества* $v^{<\varphi>}$.

Утверждение 4. *Пусть* $\Omega=\Omega_1\cup\ldots\cup\Omega_L$ – *множество состояний алгоритма RC4, лежащих на изоморфных циклах* $\Omega_1, \Omega_2, \ldots, \Omega_L$ *графа* Γ *и* $v\in\Omega$. *Тогда орбитой, содержащей состояние* v *в группе* $<\varphi, F>$, *является множество* Ω, *т.е.* $v^{<\varphi, F>}=\Omega$. *При ограничении действия группы* $<\varphi, F>$ *на множество* Ω *блоками импримитивности являются:* $\Omega_1, \Omega_2, \ldots, \Omega_L$.

Следствие 1. *Если* m–*простое число, то состояния из множества* $v^{<\varphi>}$ ($|v^{<\varphi>}|=m$) *либо принадлежат все одному циклу, либо имеется* m *изоморфных циклов, каждому из которых принадлежит только одно состояние из множества* $v^{<\varphi>}$.

Следствие 2. *Если* $m=p_1\cdot p_2$, *где* p_1, p_2– *простые числа, то состояния из множества* $v^{<\varphi>}$, *либо принадлежат все одному циклу, либо имеется* p_1 *изоморфных циклов, каждому из которых принадлежит* p_2 *состояний из множества* $v^{<\varphi>}$, *либо имеется* p_2 *изоморфных циклов, каждому из которых принадлежит* p_1 *состояний из множества* $v^{<\varphi>}$.

В §2.2 получены распределения частот k-грамм (k=1, 2) в гамме алгоритма RC4. Данный параграф является обобщением и усилением результатов, полученных A. Shamir, I. Mantin и S. Fluhrer, D. McGrew.

Теорема 5. *Пусть начальная подстановка* s_0 *выбирается случайно и равновероятно из* S_m, $i_1, j_0\in Z_m$ – *произвольные фиксированные числа. Тогда при* $m\to\infty$ *справедливы асимптотические равенства:*

I.

а) $P\{z_2=0\}=\frac{3}{m}+O(\frac{1}{m^2})$ *при* $i_2=j_0=0$,

b) $P\{z_2=k\}=\frac{1}{m}+O(\frac{1}{m^2})$ *при* $i_2=j_0=0,\ k\neq 0.$

II.

a) $P\{z_2=0\}=\frac{2}{m}+O(\frac{1}{m^2})$ *при* $j_0=0,\ i_2\neq 0,$

b) $P\{z_2=k\}=\frac{1}{m}+O(\frac{1}{m^2})$ *при* $j_0=0,\ i_2\neq 0,\ k\neq 0.$

В остальных случаях $P\{z_2=k\}=\frac{1}{m}+O(\frac{1}{m^2}),\ k=\overline{0,m-1}.$

Отметим, что первый частный случай ($P\{z_2=0\}=\frac{2}{m}$ при $j_0=i_0=0$, тогда $i_2=i_0+2=2$) теоремы 5 был получен A. Shamir, I. Mantin. Также найдено распределение биграмм (z_1, z_2) при предположении, что начальная подстановка s_0 выбирается случайно и равновероятно из S_m. Пусть имеется L последовательностей $\overline{z_1},\ldots,\overline{z_L}$, выработанных либо RC4, либо являющихся независимыми, случайными и равновероятными. Построены критерии отношения правдоподобия на основе обнаруженных неравновероятностях в распределениях первых знаков гаммы и биграмм, т.е. ищется неравновероятность в распределении знаков в L-граммах $(z_1^1,\ldots,z_1^L)$, $(z_2^1,\ldots,z_2^L)$ и $(z_1^1 z_2^1,\ldots,z_1^L z_2^L)$.

Теорема 6. *При различении последовательности, выработанной алгоритмом RC4, от случайной равновероятной последовательности:*

a) для критерия отношения правдоподобия с произвольно заданным уровнем значимости α *и мощностью* β*, основанного на неравновероятности первого знака* z_1 *гаммы (кроме случая* $i_1=2j_0$*) объем выборки равен* $n=O(m^3)$*;*

b) для критериев отношения правдоподобия с произвольно заданным уровнем значимости α *и мощностью* β*, основанных на неравновероятности второго знака* z_2 *гаммы, или на неравновероятности в распределение биграмм, или на*

неравновероятности первого знака z_1 гаммы при $i_1=2j_0$ объем выборки равен $n= O(m)$.

Из теоремы 6 получаем, что если основная гипотеза H_0: $p_0=\frac{1}{m}$ – последовательность случайная равновероятная, H_1: $p_1=\frac{2}{m}$ – последовательность выработана RC4, то для различения последовательности выработанной RC4 от случайной равновероятной последовательности при уровнях значимости $\alpha=0.05$ и $\beta=0.9$, объем выборки $n=12\,m-15$.

В §2.3 исследуется метод генерации начального состояния алгоритма RC4. А именно, найдено число всех ключей алгоритма длины m, приводящих к начальным подстановкам с произвольной фиксированной цикловой структурой. Поскольку каждому ключу $k\in Z_m^m$ алгоритма RC4 однозначно соответствует произведение транспозиций $(m-1,\ j_m)\ldots(1,j_2)(0,j_1)$, где $j_k\in\{\overline{0,m-1}\}$, $k=\overline{1,m}$, то задача оказалась эквивалентной проблеме порождения симметрической группы S_m системой транспозиций с ограничениями, которая решалась с применением методов комбинаторного анализа. Полученный результат показывает наличие большого числа эквивалентных ключей и приводится в теореме 7.

Теорема 7. *Пусть ключ к алгоритма RC4 выбирается случайно и равновероятно из* Z_m^m. *Тогда число • начальных подстановок с произвольной фиксированной цикловой структурой* $\{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\}$, $1\cdot\alpha_1+2\cdot\alpha_2+\ldots+m\cdot\alpha_m=m$, *равно*

$$\Omega(\alpha_1,\ldots,\alpha_m)= m!\cdot\prod_{r=1}^{m} r^{(r-1)\alpha_r}\cdot \sum_{\substack{(\vec{k}_1,\ldots,\vec{k}_m)\ 0\le k_{0j}\le\alpha_j,\\ 0<i\le j,\,k_{ij}\le\min(\alpha_j,\alpha_i),\\ \sum_{i=0}^{m}k_{ij}+k_{jj}=\alpha_j,\,k_{ij}=k_{ji}}} \prod_{i\le j}\frac{(C_{ij})^{k_{ij}}}{k_{ij}!\,(i+j)!^{k_{ij}}},$$

где суммирование проводится по первым $j+1$ *координатам* $(k_{0j}, k_{1j},\ldots, k_{jj})$ *вектора* $\vec{k}_j=(k_{0j}, k_{1j},\ldots, k_{mj})\in Z_m^{m+1}$, $j=\overline{1,m}$, *и*

$$C_{ij}=\sum_{n=2}^{i+j}\binom{i+j}{n}\left(\sum_{k=1}^{i}\binom{i+j-n}{i-k}\cdot k\cdot(n-k)\cdot i^{-k}\cdot j^{k-n}A_{n-1\,k-1}\right) \text{ при } 0<i\le j,$$

$$C_{0j}=1,$$

$$A_{n,k}=\sum_{t=0}^{k}(-1)^t\binom{n+1}{t}(k-t+1)^n,\ \ k=\overline{0,n-1}.$$

Укажем число ключей алгоритма RC4, приводящих к подстановкам с цикловыми структурами $\{1^0\ 2^0\ldots m^1\}$, $\{1^m\ 2^0\ldots m^0\}$, $\{1^{m-d}\ 2^0\ldots d^1\ldots m^0\}$. Так $\Omega(0,\ldots,0,1)=m^{m-1}$. Следовательно, вероятность того, что случайно равновероятно выбранный ключ из множества всех ключей алгоритма RC4 генерирует полноцикловую подстановку, равна $\frac{1}{m}$. Отметим, вероятность того, что при случайном равновероятном выборе подстановки из множества всех подстановок степени m она окажется полноцикловой, также равна $\frac{1}{m}$.

Число ключей, приводящих к тождественной подстановке, равно

$$\Omega(m,0,\ldots,0)=\sum_{k=0}^{m/2}\frac{m!}{k!\,(m-2\cdot k)!\,2^k}.$$

Отметим, что $\Omega(m,0,\ldots,0)$ также есть число решений уравнения $s^2=E$ в симметрической группе S_m, отличных от тождественной подстановки, т.е. число инволюций. Известно, что при $m\to\infty$

$$\Omega(m,0,\ldots,0)=\frac{1}{e^{1/4}\sqrt{2}}\left(\frac{m}{e}\right)^{m/2}\cdot e^{\sqrt{m}}\left(1+O\left(\frac{1}{\sqrt{m}}\right)\right).$$

Асимптотически число ключей, приводящих к начальным тождественным подстановкам алгоритма RC4 в $\sqrt{\pi}\,\frac{m^{(m+1)/2}}{e^{3/2m-\sqrt{m}+1/4}}$ раз больше ожидаемого значения $\frac{m^m}{m!}$, т.е. вероятность единичной подстановки много больше вероятности любой фиксированной подстановки.

Полученные результаты позволяют рекомендовать при применении метода полного перебора для восстановления начального состояния алгоритма

RC4 начинать опробование с тождественной подстановки, а затем опробовать подстановки с наибольшим числом неподвижных элементов.

В §2.4 введено понятие t-граммы подстановки и показано, что распределение t-граммы подстановки также является неравномерным.

Зафиксируем произвольно t попарно различных элементов $(r_1,\ldots,r_t)$ из Z_m. Будем называть t-граммой подстановки s образ $(r_1,\ldots, r_t)^s = (\alpha_1,\ldots, \alpha_t)$ при ее действии на множестве Z_m.

Пусть $p_d((r_1, \alpha_1),\ldots, (r_t, \alpha_t))=P\{s_d[r_1]=\alpha_1,\ldots, s_d[r_t]=\alpha_t\}$ и

$$\delta(A)=\begin{cases}1, \text{ если A истинно,}\\ 0, \text{ если A ложно.}\end{cases}$$

Теорема 8. *Пусть* $\{r_1,\ldots,r_t\}$, $\{\alpha_1,\ldots,\alpha_t\}$*–произвольные подмножества из* Z_m, $d\geq 1$, $i_d=d-1 \pmod m$. *Тогда выполняется рекуррентное соотношение:*

$$p_d((r_1, \alpha_1),\ldots, (r_t, \alpha_t))=0, \text{ если среди } \alpha_1,\ldots,\alpha_t \text{ встречаются равные,}$$

$$p_d((r_1, \alpha_1),\ldots, (r_t, \alpha_t))= p_{d-1}((r_1, \alpha_1),\ldots, (r_t, \alpha_t))\cdot$$

$$\cdot\Big(\sum_{c=1}^{t}\frac{\delta(i_d=r_c)}{m}+\delta(i_d\notin\{r_1,\ldots,r_t\})\Big(1-\frac{t}{m}\Big)\Big)+$$

$$+\delta(i_d\notin\{r_1,\ldots,r_t\})\frac{1}{m}\sum_{c=1}^{t}p_{d-1}((r_1,\alpha_1),\ldots,(i_d,\alpha_c),\ldots,(r_t,\alpha_t))+$$

$$+\sum_{c=1}^{t}\sum_{\gamma\in\{r_1,\ldots,r_t\}}\frac{\delta(i_d=r_c)}{m}p_{d-1}((r_1,\alpha_1),\ldots,(\gamma,\alpha_c),\ldots,(r_t,\alpha_t))+$$

$$+\sum_{c=1}^{t}\sum_{o\neq c}\frac{\delta(i_d=r_c)}{m}p_{d-1}((r_1,\alpha_1),\ldots,(r_c,\alpha_o),\ldots,(r_o,\alpha_c),\ldots,(r_t,\alpha_t)),$$

где $p_0((r_1, \alpha_1),\ldots, (r_t, \alpha_t))=1$ *при* $r_j=\alpha_j$, $p_0((r_1, \alpha_1),\ldots, (r_t, \alpha_t))=0$ *при* $r_j\neq\alpha_j$, $j=\overline{1,t}$.

С учетом найденных неравномарностей в распределении t-грамм подстановки в §2.5 получено распределение первого знака гаммы RC4 при предположении, что ключ выбирается случайно и равновероятно из Z_m^m.

После того, как в 1993 году стало известно описание алгоритма RC4, было предложено несколько поточных алгоритмов, являющихся

модификациями RC4 или основанными на общей идеи, которые также могут быть использованы в электронной коммерции.

В третьей главе введено семейство алгоритмов поточного шифрования GI, в результате чего, известные алгоритмы IA, IBAA, ISAAC представляются в общей теоретико-автоматной модели. Алгоритмы, принадлежащие семейству GI, определяются шестью функциями φ: $Z_m \times Z_{2^b} \times Z_{2^b}^m \to Z_{2^b}$, ρ: $Z_m \times Z_{2^b}^m \to Z_m$, σ: $Z_{2^b} \times Z_{2^b} \to Z_{2^b}$, δ: $Z_m \times Z_{2^b}^m \to Z_m$, χ: $Z_m \times Z_{2^b}^m \to Z_m$, ε: $N \to N$ (ε: $t \to t$ или ε: $t \to (t-1)$) и моделируются автономным автоматом $A_{GI}=(Z_m \times Z_{2^b} \times Z_{2^b} \times Z_{2^b}^m, Z_{2^b}, F_{GI}, f_{GI})$, где функции F_{GI}: $Z_m \times Z_{2^b} \times Z_{2^b} \times Z_{2^b}^m \to Z_m \times Z_{2^b} \times Z_{2^b} \times Z_{2^b}^m$, f_{GI}: $Z_m \times Z_{2^b} \times Z_{2^b} \times Z_{2^b}^m \to Z_{2^b}$ будут описаны ниже.

Семейство алгоритмов GI зависит от параметров $m=2^n$, $b \geq 2n$, n, $b \in \mathbf{N}$, которые для практических приложений выбираются равными $m=256$ (т.е. $n=8$) и $b=32$. Состоянием автомата A_{GI} в такте t ($t=0,1,\ldots$) является четверка $(i_t, a_t, q_t, s_t) \in Z_m \times Z_{2^b} \times Z_{2^b} \times Z_{2^b}^m$, где $s_t=\{s_t[0],\ldots,s_t[m-1]\}$–таблица из m n-мерных двоичных векторов. Начальным состоянием является четверка $(0, a_0, q_0, s_0)$, причем a_0, q_0 предполагаются известными параметрами GI.

Приведем описание t–го ($t=1,2\ldots.$) такта работы GI.

Функция переходов F_{GI}

$i_t = i_{t-1}+1 \pmod m$,

$a_t=\varphi(i_t, s_{t-1}, a_{t-1})$,

$s_t[i_t]=(s_{t-1}[\rho(i_t, s_{t-1})]+\sigma(a_t, q_{t-1})) \pmod{2^b}$, $s_t[k]=s_{t-1}[k]$ при $k=\overline{0,m-1}$ и $k \neq i_t$,

$q_t=(s_t[\delta(i_t, s_t)]+s_{\varepsilon(t)}[\chi(i_t, s_t)]) \pmod{2^b}$.

Функция выходов f_{GI}

$z_t=q_t$.

Шифрование t-го знака открытого текста $x_t=(x_{t,b-1},\ldots,x_{t,0}) \in Z_2^b$ имеет вид $c_t=x_t \oplus z_t$. Расшифрование t-го знака шифртекста определяется выражением $x_t=c_t \oplus z_t$.

В §3.1 и §3.2 показано, что задача восстановления по гамме начального состояния автомата A_{GI} сводится к решению системы уравнений и получена верхняя оценка числа знаков гаммы, необходимых для определения по гамме класса эквивалентных состояний.

Теорема 9. *При $m\to\infty$ оценка сверху числа знаков гаммы, при которой с вероятностью, стремящейся к единице, находится хотя бы одно состояние из класса эквивалентных состояний автомата A_{GI} не превышает* $\frac{m\ln m + x\cdot m}{2}$, *где $x\to\infty$ сколь угодно медленно.*

В §3.3 описаны ряд свойств GI. Так если в автомате A_{GI} для любого состояния $(i,a,q,s)\in Z_m\times Z_{2^b}\times Z_{2^b}\times Z_{2^b}^m$ выполняются равенства: $\sigma(a, q)=0(\mathrm{mod}\ 2)$, $\varphi(i, s, a)=0(\mathrm{mod}2)$, $s_0[i]=1\ (\mathrm{mod}\ 2)$, $i=\overline{0,m-1}$, $q_0=0(\mathrm{mod}\ 2)$, $a_0=0(\mathrm{mod}\ 2)$, то в любом такте $t\geq 1$ справедливы равенства $a_t=0(\mathrm{mod}\ 2)$, $z_t=q_t=0(\mathrm{mod}\ 2)$, и $s_t[i]=1(\mathrm{mod}\ 2)$, $i=\overline{0,m-1}$.

Доказаны следующие утверждения.

Утверждение 10. *Пусть существует такое натуральное число b_0, $n\leq b_0<b$, что для любого состояния $(i, a, q, s)\in Z_m\times Z_{2^b}\times Z_{2^b}\times Z_{2^b}^m$ автомата A_{GI} выполняются равенства*

$$\varphi(i, s, a) = \varphi(i, s(\mathrm{mod}\ 2^{b_0}), a(\mathrm{mod}\ 2^{b_0}))(\mathrm{mod}\ 2^{b_0}),$$

$$\sigma(a, q) =\sigma(a(\mathrm{mod}\ 2^{b_0}), q(\mathrm{mod}\ 2^{b_0}))(\mathrm{mod}\ 2^{b_0}),$$

$$\rho(i, s) =\rho(i, s(\mathrm{mod}\ 2^{b_0})),\ \delta(i, s) =\delta(i, s(\mathrm{mod}\ 2^{b_0})),$$

$$\chi(i, s) =\chi(i, s(\mathrm{mod}\ 2^{b_0})).$$

Будем считать b_0 наименьшим таким числом. Тогда для любых состояний (i_0, a_0, q_0, s_0) и (i_0', a_0', q_0', s_0') таких, что $i_0= i_0'\ (\mathrm{mod}\ m)$, $q_0= q_0'\ (\mathrm{mod}\ 2^{b_0})$, $s_0[i]=s_0'[i](\mathrm{mod}\ 2^{b_0})$, $i=\overline{0,m-1}$, в любом такте $t\geq 1$ выполняются равенства: $i_t=i_t'$, $a_t= a_t'\ (\mathrm{mod}\ 2^{b_0})$, $q_t= =q_t'(\mathrm{mod}2^{b_0})$, $s_t[i]= s_t'[i]\ (\mathrm{mod}\ 2^{b_0})$, $i=\overline{0,m-1}$.

Для формулировки следующего утверждения удобно обозначать GI=GI(b), где b один из параметров алгоритма GI.

Утверждение 11. *Пусть выполняются условия утверждения 10. Тогда при* $b>b_0\geq n$ *автомат* $A_{GI(b_0)}$ *есть гомоморфный образ автомата* $A_{GI(b)}$. *Гомоморфизм задается парой сюръективных отображений* (ψ, υ), *где*

$$\psi: Z_m\times Z_{2^b}\times Z_{2^b}\times Z_{2^b}^m \to Z_m\times Z_{2^{b_0}}\times Z_{2^{b_0}}\times Z_{2^{b_0}}^m,\ \upsilon: Z_{2^b}\to Z_{2^{b_0}},$$

и

$$\psi(i, a, q, s)= (i, a(\mathrm{mod}\ 2^{b_0}), q(\mathrm{mod}\ 2^{b_0}), s(\mathrm{mod}\ 2^{b_0})),\ \upsilon(z)=z\ (\mathrm{mod}\ 2^{b_0})$$

для любых $z\in Z_{2^b}$, $(i, a, q, s)\in Z_m\times Z_{2^b}\times Z_{2^b}\times Z_{2^b}^m$.

В §3.4 на основе построенного в утверждении 11 гомоморфизма предложен метод определения состояния GI по гамме. Получено, что трудоемкость метода равна $T_1=2^{b_0 m-1}m\ (2\ln m+2m+1)$ э.о.. Трудоемкость метода полного перебора равна $T_п=2^{b\,m-1}\cdot\frac{m\ln m+m^2}{2}$ э.о.. Для алгоритма IA трудоемкость определения начального состояния равна $2^{2n\,m-1}\cdot m\cdot(2\ln m+ +2m+1)$э.о.. В частности, для значений параметров b=32, n=8, m=256 имеем $T_1=7\cdot 10^{1237}$ э.о., $T_п=7{,}1\cdot 10^{2447}$ э.о.

Рассмотрены свойства функций φ, ρ, σ, δ, ε, χ, используемых в IA, и предложен метод восстановления начального состояния по гамме трудоемкостью $T_2=(\frac{m\ln m+3m^2}{2}+4m)\cdot 2^{n\,m-1}$ э.о.. В частности, для значений параметров b=32, n=8, m=256 имеем $T_2= 1{,}5\cdot 10^{621}$ э.о..

Описан класс слабых состояний алгоритма GI, трудоемкость восстановления которых существенно меньше, чем приведенная выше с использованием метода гомоморфизмов. Полученная трудоемкость восстановления слабых состояний равна $T_3=2^{b-b_0}(3\frac{m\ln m+m^2}{2}-2m)$. В случае

IA эту трудоемкость удалось существенно понизить и она не превышает $\frac{3}{2}m^2$ э.о.(при m=256 равна $98\cdot10^3$ э.о.).

В §3.5 и §3.6 рассмотрен метод восстановления начального состояния алгоритмов IBAA, ISAAC по гамме, основанный на методе частичной линеаризации и зависящий от φ. Получено, что трудоемкость восстановления начального состояния алгоритмов IBAA и ISAAC при $m\to\infty$ равна

$$T_4\sim 2^{(2n+\theta)m-1}\cdot m\cdot\left(\frac{\ln m+m}{2}+1\right)\cdot 2^{\ln 2m\left(1+\frac{\ln 2}{2}\right)(b-2n-\theta)} \text{ э.о.}$$

(θ=0 в алгоритме IBAA). Для значений параметров, используемых в алгоритмах IBAA и ISAAC уточняется оценка на T_4. Получено, что $T_{IBAA}\approx 5{,}0\cdot10^{1278}$ э.о., $T_{ISAAC}=5.9\cdot10^{1426}$ э.о. ($T_n\approx 7{,}1\cdot10^{2447}$ э.о.). В настоящее время методов с трудоемкостью меньшей, чем предложенная, неизвестно.

В четвертой главе рассмотрен алгоритм поточного шифрования "Solitaire"(" Пасьянс") предложенный Б. Шнайером в 1999 г.. Согласно замыслу автора он предназначен для применения в качестве ручного шифра ("paper-and-pencil cipher") и построен на основе перемешивания колоды игральных карт.

Пусть $n\geq 3$. Назовем элемент n–2 джокером A, а элемент n–1 джокером B. Будем записи джокеров в виде буквы или числа считать тождественными, т.е. n–2≡"A", n–1≡"B". Если в множестве Z_n выделены два элемента n–2, n–1, или один элемент n–2, или один элемент n–1, то будем, соответственно, использовать обозначения $Z^{AB}{}_n=\{\overline{0,n-3},\ A,\ B\}$, $Z^{A}{}_n=\{\overline{0,n-3},\ n-1,\ A\}$, $Z^{B}{}_n=\{\overline{0,n-2},\ B\}$. Поскольку элементы n–2, n–1 отождествляются с буквами A, B, то можно считать, что множества $Z^{AB}{}_n$, $Z^{A}{}_n$, $Z^{B}{}_n$ совпадают с множеством Z_n. Если конкретное расположение джокеров A, B в перестановке s не важно, то будем использовать запись $s=<s[0]\ s[1]\ldots s[n-1]>$ без указания положения джокеров в s.

Будем для удобства использовать одновременно две записи перестановки $<s[0]\ldots s[k_1-1]\ A\ s[k_1+1]\ldots\ s[k_2-1]\ B\ s[k_2+1]\ldots s[n-1]> \equiv <\delta[0]\ldots \delta[k_1-1]\ A\ \delta[k_1]\ldots\delta[k_2-2]\ B\ \delta[k_2-1]\ldots\delta[n-3]>_{AB}$, где $s[j]=\delta[j]$ при $j=\overline{0,k_1-1}$, $s[j+1]=\delta[j]$ при $j=\overline{k_1,k_2-2}$, $s[j+2]=\delta[j]$ при $j=\overline{k_2-1,n-3}$, и $<\delta[0]\ldots\delta[n-3]>\in S_{n-2}$.

Также будем понимать $<s[0]\ldots s[r-1]\ A\ s[r+1]\ldots s[n-1]>\equiv <\delta[0]\ldots\delta[r-1]\ A\ \delta[r]\ldots\delta[n-2]>_A$, где $s[j]=\delta[j]$ при $j=\overline{0,r-1}$, $s[j+1]=\delta[j]$ при $j=\overline{r+1,n-2}$ и $<\delta[0]\ldots\delta[n-2]>\in S_{n-1}(Z^A_n\setminus A)$. $<s[0]\ldots s[r-1]\ B\ s[r+1]\ldots s[n-1]>\equiv<\delta[0]\ldots\delta[r-1]\ B\ \delta[r]\ldots\delta[n-2]>_B$, где $s[j]=\delta[j]$ при $j=\overline{0,r-1}$, $s[j+1]=\delta[j]$ при $j=\overline{r+1,n-2}$ и $<\delta[0]\ldots\delta[n-2]>\in S_{n-1}$.

Алгоритм Solitaire моделируется автономным автоматом $A=(S_n, Z_m\cup\{\alpha\}, F, f)$, где $F: S_n \to S_n$, $f: S_n\to Z_m\cup\{\alpha\}$ и α– некоторый дополнительный символ. Алгоритм зависит от параметров $m, n\in\mathbf{N}$, которые предлагается брать равными $m=26, n=54$.

Перестановка $s_t\in S_n$ является состоянием алгоритма Solitaire в такте t ($t=0,1,\ldots$). Функция переходов F представима в виде композиции $F=F_1F_2F_3F_4$ четырех преобразований. В такте t выполняется равенство $s_{t+1}=(s_t)\,F_1\,F_2\,F_3\,F_4$.

Алгоритм ρ: $Z_m^*\times S_n\to S_n$ генерации начального состояния s_0 по ключевой фразе $k\in Z_m^*$ моделируется автоматом без выхода $A_\rho=(Z_m, S_n, P)$ с функцией переходов P: $Z_m\times S_n\to S_n$. Частичная функция переходов $P_r= F_1\,F_2\,F_3\,P_r$, $r\in Z_m$, где преобразование P_r: $S_n\to S_n$ есть

$$(<s[0]\ldots s[r]\ s[r+1]\ldots s[n-1]>)\,P_r = <s[r+1]\ldots s[n-1]\ s[0]\ldots s[r]>.$$

Начальным состоянием автомата A_ρ является тождественная перестановка $b_0=<0,1,\ldots, n-1>$.

Поскольку изучение преобразования F и полугруппы $<F>$ непосредственно является проблематичным, поэтому рассматривались свойства полугрупп $<F_1>$, $<F_2>$, $<F_1,F_2>$, $<F_1,F_2,F_3>$, $<F_1, F_2, F_3, F_4>$ и свойства групп, порожденных обратимыми модификациями преобразования F. Поскольку $<F>$

есть подполугруппа из $\langle F_1,F_2,F_3,F_4\rangle$, то ряд полученных результатов непосредственно переносится на $\langle F\rangle$.

В §4.2 рассмотрены свойства группы преобразований $\langle F_3, F_4\rangle$. Пусть $\{A_1, A_2\}=\{A, B\}$. Получено, что F_3 есть инволюция. В группе $\langle F_3\rangle$ длина орбит элементов вида $\langle A_1\ s[1]\dots s[n-2]\ A_2\rangle$ равна 1. Число орбит длины 1 равно $2(n-2)!$. Число орбит длины 2 равно $\frac{n!}{2}-(n-2)!$. В группе $\langle F_3, F_4\rangle$ длина орбит элементов вида $s=\langle A_1\ s[1]\dots s[n-2]\ A_2\rangle$ равна 1. Число таких элементов равно $2\cdot(n-2)!$.

В §4.3 описаны свойства полугруппы преобразований $\langle F_1, F_2, F_3, F_4\rangle$. Пусть $S(Z^A_n,\ r)=\{s\in S(Z^A_n) \mid s[r]=A\}$, $S(Z^B_n,\ r)=\{s\in S(Z^B_n) \mid s[r]=B\}$, $r=\overline{0,n-1}$, $S(Z^{AB}_n, r_1, r_2)=\{s\in S(Z^{AB}_n) \mid s[r_1]=A, s[r_2]=B\}$, $0\le r_1, r_2\le n-1$, $r_1\ne r_2$.

Множества $S(Z^A_n,\ r)$, $r=\overline{0,n-1}$, задают разбиение $S(Z^A_n)$; $S(Z^B_n,\ r)$, $r=\overline{0,n-1}$, задают разбиение $S(Z^B_n)$; $S(Z^{AB}_n, r_1, r_2)$, $0\le r_1,r_2\le n-1$, $r_1\ne r_2$, задают разбиение множества $S(Z^{AB}_n)$. Отметим, что $|S(Z^A_n,\ r)|=|S(Z^B_n,\ r)|=(n-1)!$, $|S(Z^{AB}_n, r_1, r_2)|=(n-2)!$, $0\le r, r_1, r_2\le n-1$, $r_1\ne r_2$.

Пусть $\sigma_A: S(Z^A_n)\to S(Z_{n-1})$, $\sigma_B: S(Z^B_n)\to S(Z_{n-1})$, где

$$\langle s[0]\dots s[k-1]\ A\ s[k]\dots s[n-2]\rangle_A\sigma_A=\langle s[0]\dots s[k-1]\ s[k]\dots s[n-2]\rangle,$$

$$\langle s[0]\dots s[k-1]\ B\ s[k]\dots s[n-2]\rangle_B\sigma_B=\langle s[0]\dots s[k-1]\ s[k]\dots s[n-2]\rangle,$$

где $k=\overline{0,n-1}$, $\langle s[0]\dots s[n-2]\rangle\in S_{n-1}$.

Пусть $\{A_1, A_2\}=\{A, B\}$, $\langle s[0]\dots s[n-3]\rangle\in S_{n-2}$ и преобразование $\sigma_{AB}: S(Z^{AB}_n)\to S(Z_{n-2})$ задается равенствами:

$$\langle s[0]\dots s[k_1-1]\ A_1\ s[k_1+1]\dots s[k_2-1]\ A_2\ s[k_2+1]\dots s[n-3]\rangle\sigma_{AB}=$$

$$=\langle s[0]\dots s[k_1-1]\ s[k_1+1]\dots\ s[k_2-1]\ s[k_2+1]\dots s[n-3]\rangle,$$

$$\langle s[0]\dots s[k-1]\ A_1\ A_2\ s[k+1]\dots s[n-3]\rangle\sigma_{AB}=\langle s[0]\dots s[k-1]\ s[k+1]\dots s[n-3]\rangle,$$

где $k_1\ne k_2$, $k_1=\overline{0,n-1}$, $k_2=\overline{0,n-1}$, $k=\overline{0,n-1}$.

Пусть $\mathrm{def}(G,\ X)=|X|-|X^G|$ – дефект полугруппы G преобразований, действующей на множестве X.

Получено, что $S(Z^A_n,0)\ F_1^{-1}=\varnothing$, $S(Z^B_n,\ 0)F_2^{-1}=\varnothing$, $\mathrm{def}(F_1,\ Z^{\Lambda}_n)=\ =(n-1)!$, $\mathrm{def}(F_2,\ Z^B_n)=(n-1)!$. $\langle F_1\rangle$, $\langle F_2\rangle$ являются циклическими полугруппами порядка n, индексом 1 и периодом $n-1$. Кроме того, ограничение действия полугрупп $\langle F_1\rangle$ на множество $\Omega_A=S(Z^A_n)\backslash S(Z^A_n,\ 0)$, а $\langle F_2\rangle$ на $\Omega_B=\!=S(Z^B_n)\backslash S(Z^B_n,\ 0)$ являются циклическими группами порядка $n-1$, $\langle F_1\rangle\cong\langle F_2\rangle$.

Пусть $\Omega_{AB}=S(Z^{AB}_n)\backslash(S(Z^A_n,0)\cup S(Z^B_n,0))$. Тогда $\Omega_{AB}<F_1,\ F_2>=\Omega_{AB}$, $\mathrm{def}(\langle F_1,\ F_2\rangle,\ Z^{AB}_n)=2(n-1)!$. Кроме того, $\mathrm{def}(F_1F_2,Z^{AB}_n)=\mathrm{def}(F_2F_1,Z^{AB}_n)=\ =2(n-1)!-(n-2)!$. Ограничение полугруппы преобразований $\langle F_1,\ F_2\rangle$ на множество Ω_{AB} является 1/2 транзитивной группой, $|(s)\langle F_1,\ F_2\rangle^{\Omega_{AB}}|=(n-1)(n-2)$. Число орбит группы $\langle F_1,\ F_2\rangle^{\Omega_{AB}}$ равно $(n-2)!$.

Получено, что (s) $F^1=\varnothing$, если состояниями алгоритма Solitaire являются перестановки следующих видов:

1) $s=\langle A, s[1]\ s[2]...s[n-2]\ B\rangle\in S(Z^{AB}_n)$;

2) $s=\langle s[0]\ s[1]\ s[2]...s[n-3]\ A\ B\rangle\in S(Z^{AB}_n)$;

3) $s=\langle s[0]...s[p-1]\ B\ s[p+1]...s[n-2]\ A\rangle\in S(Z^{AB}_n)$, где $p\in Z_{n-2}$;

4) $s=\langle s[0]...s[p-1]\ A\ s[\ p+1]...s[n-2]\ B\rangle\in S(Z^{AB}_n)$, где $p\in Z_{n-4}\cup\{n-2\}$.

Блоками импримитивности полугруппы $\langle F_1,\ F_2,\ F_3\rangle$ являются множества $S(Z^{AB}_n,\ r_1,\ r_2)$, $r_1,r_2=\overline{0,n-1}$, $r_1\neq r_2$. Число блоков импримитивности равно $n(n-1)$.

Пусть $s\in S(Z^{AB}_n)$, $\Delta_s=\{s'|\ s'\sim_{AB}s,\ s'\in S(Z^{AB}_n)\}$, $|\Delta_s|=n(n-1)$. Тогда полугруппа преобразования G, действующая на множестве блоков импримитивности, изоморфна полугруппе $\langle F_1,\ F_2\rangle^{\Delta_s}$. Также доказано, что полугруппа $\langle F_2\rangle$ делит полугруппу $\langle F_1,\ F_3\rangle$, а $\langle F_2\rangle$ – полугруппу $\langle F_1,\ F_3\rangle$. Кроме того, $\langle F_1,\ F_2,\ F_3\rangle\cong\langle F_1,\ F_3\rangle\cong\langle F_2,\ F_3\rangle$.

В §4.4 описаны свойства полугруппы преобразований автомата A_ρ. Для их формулировки введены преобразования $\langle\psi_A\rangle$, $\langle\vartheta_A\rangle$, являющиеся двумя биективными модификациями преобразования F_1, а $\langle\psi_B\rangle$, $\langle\vartheta_B\rangle$– преобразования F_2, причем $\langle\psi_A\rangle$, $\langle\vartheta_A\rangle$– циклические группы преобразований,

действующие на множестве $S(Z^A_n)$, а $\langle\psi_B\rangle$, $\langle\vartheta_B\rangle$– циклические группы преобразований, действующие на множестве $S(Z^B_n)$, η: $i \to i+1 \ (mod\ n)$.

Показано, что группа $\langle\psi_B\rangle$ изоморфно вложена в группу $\langle\eta, \psi_A, F_3\rangle$ и имеют место изоморфизмы групп $\langle\eta, \psi_A, \psi_B, F_3\rangle \cong \langle\theta_A, \theta_B, F_3\rangle \cong \langle\eta, \psi_A, F_3\rangle \cong \langle\theta_A, F_3\rangle$.

Пусть $\Omega=\{S(Z^{AB}_n, r_1, r_2) | \ r_1, r_2=\overline{0,n-1}, r_1 \neq r_2\}$. Тогда группа $\langle\eta, \psi_A, \psi_B, F_3\rangle$ является импримитивной на множестве $S(Z^{AB}_n)$, число блоков импримитивности равно $|\Omega|=n(n-1)$. Кроме того, группа G, действующая на множестве Ω блоков импримитивности группы $\langle\eta, \psi_A, \psi_B, F_3\rangle$, изоморфна группе $\langle\psi_A, \psi_B\rangle$. Известно, что импримитивность может являться слабостью алгоритма шифрования.

В пятой главе предложены методы восстановления начального состояния алгоритмов поточного шифрования Веста-2 и Веста-2М, разработанных в 1995–1998 гг. фирмой «ЛАН Крипто». Алгоритм Веста-2 используется в коммерческих продуктах этой фирмы, Веста-2М является стандартом газовой промышленности России. Также эти алгоритмы используются в электронной коммерции в России. Отличие алгоритма Веста-2М от алгоритма Веста-2 состоит в выборе функции обратной связи и расположением функций, осуществляющей перестановку координат двоичных 16-мерных векторов.

Алгоритмы семейства Веста моделируются автономным автоматом $A=(Z_p^{31} \times Z_{2^{16}} \times Z_{2^{16}} \times Z_{2^8}, Z_{2^8}, F, f)$, где $F: Z_p^{31} \times Z_{2^{16}} \times Z_{2^{16}} \times Z_{2^8} \to Z_p^{31} \times Z_{2^{16}} \times Z_{2^{16}} \times Z_{2^8}$, $f: Z_p^{31} \times Z_{2^{16}} \times Z_{2^{16}} \times Z_{2^8} \to Z_{2^8}$, p-простое число $2^{15}<p<2^{16}$. Состоянием алгоритмов семейства Веста в такте $t \geq 1$ является $((x_{t+30}, x_{t+29}, \ldots, x_t), w_t, v_t, u_t^{(1)}) \in Z_p^{31} \times Z_{2^{16}} \times Z_{2^{16}} \times Z_{2^8}$, где $(x_{t+30}, x_{t+29}, \ldots, x_t) \in Z_p^{31}$ есть 31-мерный вектор над полем $GF(p)$, являющийся состоянием линейного регистра сдвига в такте t с функцией обратной связи $g(x)=x^{31}+x^{10}-1$. Начальное состояние $((x_{30},$

$x_{29},\ldots,x_0)$, w_0, v_0, $u_0^{(1)}$) является ключом. Перестановка тс действует на множестве двоичных 16-мерных векторов.

Основная часть разработанных методов анализа фильтрующих генераторов с регистрами сдвига применима к анализу комбинирующих генераторов с линейными регистрами над полем *GF(2)*. Развитый подход позволяет частично линеаризовать некоторый класс фильтрующих генераторов над полем *GF(p)*. Его применение показано на примере алгоритмов Веста-2 и Веста-2М с тождественной перестановкой π. Поэтому в §5.3, §5.5 рассмотрены методы восстановления начального состояния $((x_{30}, x_{29},\ldots,x_0), w_0, v_0, u_0^{(1)}) \in Z_p^{31} \times Z_{2^{16}} \times Z_{2^{16}} \times Z_{2^8}$ алгоритмов Веста-2М и Веста-2 с тождественной перестановкой π, который сводится к восстановлению начального состояния $(x_{30}, x_{29},\ldots,x_0)$ линейного регистра по гамме $\bar{z}$.

Получено, что трудоемкость предложенных методов для алгоритмов Веста-2М, Веста-2 с тождественной перестановкой π одинакова и в среднем равна $T_м^{(1)}=2^{316}$ э.о.. Заметим, что трудоемкость метода полного перебора равна $T_{пер}=2^{544}$ э.о..

В §5.4, §5.6 решается задача определения начального состояния $((x_{30}, x_{29},\ldots,x_0), w_0, v_0, u_0^{(1)}) \in Z_p^{31} \times Z_{2^{16}} \times Z_{2^{16}} \times Z_{2^8}$ алгоритмов Веста-2М, Веста-2 с реальной перестановкой π. Она сводится к восстановлению начального состояния $(x_{30}, x_{29},\ldots,x_0)$ линейного регистра по гамме $\bar{z}$. Для алгоритмов Веста-2М, Веста-2 с реальной перестановкой π частичная линеаризация, как для тождественной перестановки π, не получается. Но в силу свойств алгоритмов, также удалось построить линейные соотношения. Получено, что трудоемкость предложенных методов для алгоритма Веста-2М в среднем равна $T_м^{(3)}=2^{343,4}$ э.о., для алгоритма **Веста-2** – $T_м^{(4)}=2^{345}$ э.о..

В §5.7 для алгоритма Веста-2 рассмотрен метод восстановлению начального состояния $(x_{30}, x_{29},\ldots,x_0)$ линейного регистра по гамме $\bar{z}$ для произвольной перестановки π. Метод основан на выведенных алгебраических

соотношениях. Получено, что трудоемкость предложенного метода в среднем равна $T_м^{(5)}=2^{381,2}$ э.о..

В §5.8 описан класс слабых состояний алгоритма Веста-2 и предложены методы их восстановления по гамме. Состояние вида $\left(x_{30},...,x_{20+k},\underbrace{0,...,0}_{k},x_{19},...,x_k,\underbrace{0,...,0}_{k}\right)$, где $1\le k\le 11$, линейного регистра сдвига назовем k-благоприятным. Пусть $\overline{z_t^k}=z_t\, z_{t+1}\ldots\, z_{t+k-1}$.

Утверждение 12. *Пусть состояние* $\overline{x_t}$ *линейного регистра сдвига является k-благоприятным,* $3<k\le 11$. *Тогда:*

а) состояние $\overline{x_{t+31}}$ *является (k–3)-благоприятным;*

б) $z_{t+1}=z_{t+2}=\ldots=z_{t+k-1}$.

Отметим, что число различных заполнений линейного регистра сдвига равно $p^{31}\approx 2^{16\ 31}$. Число 8-благоприятных состояний есть $p^{15,2}\approx\sqrt{p^{31}}$. Число 11-благоприятных состояний есть $p^{10,3}\approx\sqrt[3]{p^{31}}$.

Теорема 13. *Пусть состояние* $\overline{x}$ *линейного регистра сдвига является k-благоприятным. Тогда трудоемкость метода определения начального состояния алгоритма Веста-2 по гамме* $\overline{z}$:

а) при $k=11$ *не превышает* 2^{33} *элементарных операций;*

b) при $8\le k<11$ *не превышает* 2^{70} *элементарных операций.*

Для сравнения трудоемкость метода полного перебора равна $p^{31}\cdot 2^{39}\approx 2^{544}$. Таким образом, если в гамме $\overline{z}$ встречаются постоянные подпоследовательности $\overline{z_t^{k-1}}$, $\overline{z_{t+31}^{k-4}}$, то можно предположить, что состояние линейного регистра сдвига является k-благоприятным и определить состояние алгоритма Веста-2 с трудоемкостью существенно меньшей трудоемкости полного перебора.

В §5.9 описаны групповые свойства семейства Веста.

Утверждение 14. *Пусть D(g, p, d) – подстановка степени p^d, p– простое число, $d \in N$, реализуемая автономным регистром сдвига длины d с функцией обратной g над полем GF(p). Пусть $\mu_j, \nu_j \in S_m$, $j=\overline{0,p-1}$ и δ_{t_1,t_2}: $GF(p)^d \to S_m \times S_m$, где $(t_1, t_2) \in Z_d \times Z_d$, $t_1<t_2$, и $(a,b)^{\vec{x}\delta_{t_1,t_2}} = (a\mu_{x_{t_1}}, b\nu_{x_{t_2}})$. Пусть $M=<\mu_j \mid j=\overline{0,p-1}>$, $\Gamma=<\nu_j \mid j=\overline{0,p-1}>$. Тогда группа семейства Веста есть сплетение прямого произведения групп подстановок $M \times \Gamma$ и циклической группы $< D(g, p, n)>$, причем*

$$((a,b),\vec{x})^{(\delta_{t_1,t_2},D(g,p,d))} = ((a,b)^{\vec{x}\delta_{t_1,t_2}}, \vec{x}D(g,p,d)).$$

Основные публикации по теме диссертации:

1. Варфоломеев А. А., Жуков А.Е., Пудовкина М. А. Поточные криптосистемы. Основные свойства и методы анализа стойкости. Учебное пособие- ПАИМС - Москва, 2000.
2. Пудовкина М. А. Методы криптоанализа алгоритма поточного шифрования IA // Безопасность Информационных Технологий, 3- 2000- с. 40-46.
3. Пудовкина М.А. Криптоанализ алгоритма поточного шифрования ISAAC// В сб. научных трудов конференции Проблемы информационной безопасности в системе высшей школы - Москва-2001- с. 85-87.
4. Pudovkina M. Probabilistic relations for the jokers at the Solitaire keystream generator// In: Proceedings of first International IFIP TC-11 WG 11.4 Working Conference on NETWORK SECURITY - Leuven, Belgium- 2001-pp. 110—118.
5. Pudovkina M. Short cycles of the alleged RC4 keystream generator// In: Proceedings of 3nd International Workshop on Computer Science and Information Technologies- CSIT'2001- UFA- 2001-pp. 200-206.
6. Pudovkina M., Analysis of the IA keystream generator // In: Proceedings of 3nd International Workshop on Computer Science and Information Technologies- CSIT2001- UFA- 2001- pp. 207-215.

7. Pudovkina M., Varfolomeev A.A., A cycle structure of the Solitaire keystream generator// In: Proceedings of 3nd International Workshop on Computer Science and Information Technologies- CSIT'2001- UFA- 2001- 215-223.
8. Пудовкина М. А., О распределении первого выходного символа криптосхемы RC4 // В сб. научных трудов XLIV юбилейной научной конференции МФТИ. - Москва -Долгопрудный-2001-с. 201-202.
9. Пудовкина М. А., О распределении биграмм в криптосхеме RC4 // В сб. научных трудов конференции «Проблемы информационной безопасности в системе высшей школы»- Москва- 2002- с. 41-42.
10. Пудовкина М. А., Об одной системе образующих с ограничениями //В сб. научных трудов конференции «Проблемы информационной безопасности в системе высшей школы»- Москва- 2002- с. 43-44.
11. Пудовкина М. А., О группе преобразований криптосистемы Solitaire // В сб. тезисов конференции «Методы и технические средства обеспечения безопасности информации» - Санкт-Петербург- 2002- с. 66-67.
12. Пудовкина М. А., О слабых состояниях криптосистемы IA // В сб. научных трудов конференции «Проблемы информационной безопасности в системе высшей школы» - Москва- 2003- с. 37-38.
13. Pudovkina M., Statistical weaknesses in the alleged RC4 keystream generator // In: Proceedings of 4nd International Workshop on Computer Science and Information Technologies- CSIT'2002- Greece, Patras- 2002-pp. 301-307.
14. Пудовкина М. А., О слабых состояниях криптосистемы ВЕСТА-2 // В сб. тезисов конференции «Методы и технические средства обеспечения безопасности информации»- Санкт-Петербург— 2002- с. 68-69.
15. Pogorelov B., Pudovkina M, Properties of the transformation semigroup of the Solitaire stream cipher // Discrete Mathematics and Theoretical Computer Science-DTMCS'03 Proceedings, Springer-Verlag-2003-pp. 260-274.
16. Погорелов Б.А.„ Пудовкина М.А., О свойствах криптоалгоритма GI. //В трудах конференции Мабит'03-МГУ-2003-с.100-102.

17. Пудовкина М.А., О групповых свойствах криптоалгоритма Веста.//В трудах конференции Мабит'03-МГУ-2003-с.103-105.
18. Пудовкина М. А., Вероятностные свойства криптоалгоритма ВестаУ/В сб. тезисов конференции «Методы и технические средства обеспечения безопасности информации»- Санкт-Петербург- 2003-С.213-214.
19. Пудовкина М. А., О длине гаммы, необходимой для восстановления начального состояния криптоалгоритма GI.// В сб. тезисов конференции «Методы и технические средства обеспечения безопасности информации»- Санкт-Петербург- 2003- с.215-216.
20. Pudovkina M., On the transformation group generated by the key-schedule of the Solitaire stream cipher. // 5nd International Workshop on Computer Science and Information Technologies- CSIT'2003- UFA- 2003-pp. 304-309.
21. Pudovkina M., The number of initial states of the RC4 cipher with the same cycle structure. // 5nd International Workshop on Computer Science and Information Technologies-CSIT'2003- UFA- 2003- pp. 310-314.
22. Пудовкина М. А., О некоторых слабостях криптосистемы RC4// Защита информации, 2- 2002- с.50-56.
23. Пудовкина М. А., Свойства алгоритма поточного шифрования IA// Международный научный семинар «Дискретная математика и ее приложения»- Москва- МГУ-2001-с. 70-73.
24. Pudovkina M., The upper estimate of the unicity distance of the GI stream cipher // International conference computer data analysis and modeling - Minsk-2004- pp. 200-207.
25. Пудовкина М.А., О методе генерации начального состояния алгоритма поточного шифрования Solitaire.// В сб. научных трудов конференции «Проблемы информационной безопасности в системе высшей школы»- Москва- 2000- с. 10-11.

Отпечатано в копицентре
Москва, Ленинские горы, МГУ, 1 Гуманитарный корпус.
www.stprintru e-mail: zakaz@stprint.ru тел. 939-3338
Заказ № 22 тираж 60 экз. Подписано в печать 15.06.2004 г.