МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
ХАРЬКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ РАДИОЭЛЕКТРОНИКИ

РАДИОТЕХНИКА

Всеукраинский межведомственный научно-технический сборник

Тематический выпуск
"ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ"

Основан в 1965 г.

ВЫПУСК 119

Харківський державний технічний
університет радіоелектроніки

2001

УДК 621.3

Сборник включен в список специальных изданий ВАК Украины по физико-математическим и техническим наукам.

Рассматриваются проблемные вопросы теории и практики защиты информации в различных информационных технологиях, системах и сетях. Обсуждаются противоречия, возникшие в информационных технологиях, анализируется их сущность, определяются методы их разрешения. Излагаются концептуальные взгляды и точки зрения на комплексные системы защиты информации, методы и механизмы реализации услуг.

Представлен ряд результатов, посвященных криптографическим преобразованиям в группах точек эллиптической кривой, методам аутентификации, криптоанализу симметричных шифров, построению и исследованию свойств случайных и псевдослучайных чисел последовательностей, а также результаты системно-технического характера и протоколы.

Ответственность за содержание статей несет автор.

Рекомендовано Ученым советом Харьковского техничного университета радиоэлектроники.
Протокол № 30 від 30.03.2001

Адрес редакционной коллегии: Украина, 61166 Харьков-166, просп. Ленина, 14,
Харьковский государственный технический университет радиоэлектроники (ХТУРЭ), тел. 40-93-97

ЗБІРНИК НАУКОВИХ ПРАЦЬ
РАДІОТЕХНІКА
Тематический выпуск
"ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ"
Випуск 119

СБОРНИК НАУЧНЫХ ТРУДОВ
РАДИОТЕХНИКА
Тематичний випуск
"ІНФОРМАЦІЙНА БЕЗПЕКА"
Выпуск 119

Редактор *Денисова Л.Н.* Компьютерная верстка *Карпинской Е.Д.*

Підп. до друку 10.05.2001. Формат 60x90/8.
Папір офсетний. Друк офсетний. Умов. друк. арк. 28,9. Обл. вид. арк. 26,4.
Тираж 300 прим. Зам. № 14/2001. Ціна договірна.
Харківський державний технічний університет радіоелектроніки (ХТУРЕ)
Україна, 61166 Харків, просп. Леніна, 14.
Оригінал-макет збірника підготовлено и надруковано у видавництві "Медицина і..." (ПФ "Крокус")
Україна, 61183, Харків, вул. Дружби народів, 277, к. 93

Сидят слева направо: А.З. Степченко, проф. В.А. Краснобаев, проф. В.И. Долгов,
зав. каф. БИТ проф. И.Д. Горбенко, проф. Н.С. Лесная, ректор ХТУРЭ проф. М.Ф. Бондаренко,
проректор ХТУРЭ проф. Н.И. Слипченко, директор института РТЭ ХТУРЭ проф. В.М. Шокало, проф. Е.Г. Прошкин, Г.В. Груша.
Стоят слева направо: А.В. Свинарев, Я.Ю. Стасева, А.А. Казьмин, А.В. Потий, А.А. Гайович, Ю.И. Горбенко, А.А. Торба,
В.Н. Вервейко, С.В. Полчанинов, А.А. Кузнецов, А.Г. Пащенко, Г.З. Халимов, В.В. Лесняк, М.А. Федотов, С.И. Бабкин,
В.М. Карташов, М.А. Омаров, А.А. Поляков, С.И. Збитнев, Т.А. Шатовская, Р.В. Олейников, А.И. Лучанинов, С.А. Головашич,
М.А. Волк, Т.В. Цепурит, А.А. Замула, Т.А. Гриненко, Д.А. Чекалин, М.А. Кривошлык, С.Г. Елаков, А.А. Ткач, Д.Ю. Максименко

Уважаемые читатели!

Специализированный выпуск настоящего научно-технического сборника посвящен ряду направлений теории и практики информационной безопасности. Публикуемые в сборнике статьи, по мнению специалистов ХТУРЭ и АТ «Институт информационных технологий», г. Харьков отражают ряд остро стоящих проблем и направлены на поиск путей, разработки методов, методик, алгоритмов и средств их разрешения. Дело в том, что конец XX и начало XXI века характерны возникновением в ряде случаев неуправляемых процессов информатизации сообщества, которые сегодня называют компьютеризацией. Примером тому сегодня являются системы Internet, Intranet, Extranet, с появлением которых по существу создано единое мировое информационное пространство. К сожалению, но по существу диалектически закономерно, в этом информационном пространстве непрерывно ведутся информационные войны, создается и совершенствуется информационное оружие. Сегодня даже специалисты не в состоянии оценить их опасность, уровень потерь, так как они могут привести при определенных условиях даже к гибели земной цивилизации. Разрешение этих противоречий может быть достигнуто за счет обеспечения в различных информационных технологиях необходимого уровня ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ (ИБ).

Существует ряд определений ИБ. На наш взгляд приемлемым является следующее. ИБ – есть защищенность информации и поддерживающей информационной структуры от случайных или преднамеренных действий естественного или искусственного характера, в результате которых наносится ущерб владельцам и/или пользователям информации, а также поддерживающий информационной структуре. Необходимый уровень ИБ достигается за счет создания и применения Систем защиты информации в виде реализуемого комплекса мер, мероприятий и средств.

При этом для предоставления пользователям услуг конфиденциальности, целостности, доступности и наблюдаемости необходимо использовать криптографические средства защиты информации. На сегодня проблемным и необходимым с точки зрения практики является ряд научных и практических направлений. Основными их них являются:

- разработка методологических основ теории и практики создания защищенных информационных технологий;
- разработка перспективных и анализ существующих алгоритмов криптографических преобразований (цифровая подпись, блочное шифрование, направленное шифрование), принятие соответствующих стандартов;
- разработка теории и создание комплексов криптоанализа существующих и перспективных криптоалгоритмов;
- анализ стойкости и сложности криптографических преобразований, выполняемых в группах точек эллиптических кривых;
- анализ и синтез состоятельных криптографических протоколов аутентификации, управления ключами, разделения секрета и др.;
- разработка алгоритмов и средств формирования случайных и псевдослучайных последовательностей, освоение существующих и разработка перспективных методик исследования их свойств;
- освоение и развитие теории и практики криптоанализа существующих симметричных криптоалгоритмов.

В настоящем сборнике и представлен ряд результатов исследований и разработок специалистов ХТУРЭ и АТ «Институт информационных технологий» в указанных направлениях. Издание сборника такой направленности позволит довести до специалистов и интересующихся проблемами защиты информации ряд новых результатов и достижений, развернуть дискуссии.

Конечно же, статьи отражают, прежде всего, мнения и взгляды авторов по обсуждаемым проблемам, в то же время статьи к опубликованию кафедрой «Безопасность информационных технологий» и ХТУРЭ. В целом считаем, что публикация статей в тематическом сборнике позволит ускорить процессы освоения и использования известных методов и средств криптографической защиты информации, обоснования требований к разработке перспективных. С учетом этого и производится отбор статей.

С уважением и благодарностью к специалистам и читателям, которые интересуются проблемами информационной безопасности.

Ректор ХТУРЭ, профессор М.Ф. Бондаренко

Заведующий кафедрой БИТ, профессор И.Д. Горбенко

УДК 681.3.06:519.248.681

М. Ф. БОНДАРЕНКО, д-р техн. наук, С. П. ЧЕРНЫХ, И. Д. ГОРБЕНКО, д-р техн. наук, А. А. ЗАМУЛА, канд. техн. наук, А. А. ТКАЧ

МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ КОНЦЕПЦИИ И ПОЛИТИКИ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Введение

В постоянно расширяющейся области использования средств вычислительной техники и передачи данных появляются все новые и новые проблемы сохранения конфиденциальности, целостности, наблюдаемости и доступности информации, ограждения её от посягательств злоумышленников. Наиболее надежную защиту информации в автоматизированных системах и сетях (далее – Системах) различных классов можно обеспечить только с помощью системного подхода. Он предполагает, что решение задачи должно достигаться за счет использования совокупности организационных и организационно-технических мер и мероприятий, а также криптографических систем и средств. Системный подход ориентирован на создание комплексной системы защиты информации (КСЗИ) в Системе [1,2].

Процесс создания КСЗИ включает три основных этапа [3]:

- предварительный;
- проектирования и разработки КСЗИ;
- проведение испытаний и сдача в эксплуатацию КСЗИ.

На предварительном этапе выполняются следующие основные работы:

1) классификация и описание ресурсов Системы (вычислительной системы, способов связи и коммуникаций, информации, ее категорий, вида представления, мест хранения, технологии обработки и тому подобного, обслуживающего персонала и пользователей, территории и помещений и т.п.);

2) разработка информационной модели для существующей Системы, то есть описание (формальное или неформальное) информационных потоков Системы, интерфейсов между пользователем и Системой и т. п.;

3) определение перечня угроз и возможных каналов утечки информации;

4) экспертная оценка ожидаемых потерь в случае осуществления угроз; определение услуг безопасности, которые должны предоставляться пользователям;

5) обоснование необходимости проведения спецпроверок и специсследований средств вычислительной техники (СВТ) и других технических средств, а также специального оснащения помещений;

6) определение требований к организационным, физическим и другим мероприятиям защиты, которые реализуются в дополнение к комплексу программно-технических способов защиты;

7) определение требований к метрологическому обеспечению работ;

8) определение перечня макетов, которые разрабатываются, и технологических стендов;

9) оценка стоимости и эффективности избранных способов;

10) принятие окончательного решения о составе КСЗИ Системы.

По результатам выполненных на предварительном этапе работ формируются документы: «Политика безопасности информации Системы» и «Концепция безопасности информации Системы». Кроме того, разрабатывается ТЗ на создание КСЗИ Системы и совокупность документов, в соответствии с которыми осуществляется организация защиты информации на всех этапах жизненного цикла Системы — «План защиты информации в Системе».

Следует отметить, что разработка политики и концепции безопасности Системы, должна предшествовать разработке ТЗ на создание КСЗИ Системы и Плана защиты информации в Системе.

1. Разработка Политики безопасности информации в АС

Под политикой безопасности информации в Системе (далее – политика безопасности Системы) будем понимать набор законов, нормативных документов, требований, правил, ограничений, инструкций, рекомендаций и т.п., которые регламентируют порядок обработки информации и направлены

на защиту информации от определенных угроз [3, 4]. Политика безопасности разрабатывается для отдельного компонента Системы, услуги защиты и Системы в целом. Политика безопасности информации в Системе является частью общей политики безопасности организации и должна наследовать основные ее принципы и положения.

Содержание политики безопасности Системы определяется технологией обработки информации, моделями нарушителей и угроз, особенностями вычислительной системы (ВС), физической среды и прочими факторами. Вследствие этого, если в какой-либо Системе реализуются различные технологии обработки информации, то и политика безопасности в такой Системе будет состоять из нескольких существенно отличных частей, каждая из которых будет отвечать конкретной технологии обработки информации. Как составные части общей политики безопасности Системы могут разрабатываться политики обеспечения конфиденциальности, целостности, наблюдаемости и доступности обрабатываемой информации, а также правила разграничения доступа (ПРД), которые регламентируют правила доступа пользователей и процессов к ресурсам Системы.

Политика безопасности должна предусматривать комплексное использование правовых и нравственно-этических норм, организационных (административных) мер, физических, технических (аппаратных и программных) способов и средств защиты информации, а также определять правила и порядок их применения в Системе. Политика безопасности должна базироваться на принципах системности, комплексности, непрерывности защиты, достаточности механизмов и мероприятий защиты и их адекватности угрозам, гибкости управления системой защиты, простоты и удобства ее использования, открытости алгоритмов и механизмов защиты, если другое не предусмотрено в отдельности.

Политика безопасности Системы должна доказательно давать гарантии того, что:

– в Системе (в каждой отдельной составной части, в каждой функциональной задаче и т. п..) обеспечивается адекватность уровня защиты информации уровню ее критичности;

– реализация мероприятий защиты информации является рентабельной;

– в любой среде функционирования Системы обеспечивается оцениваемость и проверяемость защищенности информации;

– обеспечивается персонификация положений политики безопасности (относительно субъектов Системы), отчетность (регистрация, аудит) для всех критичных с точки зрения безопасности ресурсов, к которым осуществляется доступ;

– персонал и пользователи обеспечены достаточно полным комплектом документации относительно порядка обеспечения защиты информации;

– все критичные с точки зрения безопасности информации технологии (функции) Системы имеют соответствующие планы обеспечения непрерывной работы и ее возобновления в случае возникновения непредвиденных ситуаций.

Методология разработки политики безопасности включает в себя следующие работы:

– разработка концепции безопасности информации в Системе;

– анализ рисков;

– определение требований к методам и средствам защиты;

– выбор основных решений по обеспечению безопасности информации;

– организация выполнения восстановительных работ и обеспечение непрерывного функционирования Системы;

– документальное оформление политики безопасности.

В общем случае документ «Политика безопасности Системы» должен включать в себя описание [3, 4]:

1) объектов (элементов ресурсов) Системы;
2) основных угроз информации;
3) требований по защите от угроз;
4) принципов управления доступом пользователей к информации;
5) правил разграничения информационных потоков;
6) правил маркирования носителей информации;
7) основных атрибутов доступа пользователей, процессов и пассивных объектов;
8) правил разграничения доступа пользователей и процессов к пассивным объектам;

9) правил администрирования КСЗИ и регистрации действий пользователей.

В разделе «Описание объектов (элементов ресурсов) Системы» на основе инвентаризации (идентификации) всех компонентов Системы, участвующих в технологическом процессе обработки информации, приводится описание критичных с точки зрения безопасности активных и пассивных компонентов Системы.

Инвентаризации (идентификации) подлежат:

- организационно-топологическая структура Системы, для которой создается КСЗИ;
- состав и назначение функциональных подсистем Системы;
- состав служб и протоколов, реализующих информационный обмен между элементами (компонентами) Системы;
- объекты защиты (виды и категории обрабатываемой информации, аппаратно-программные и информационные ресурсы на соответствующих уровнях иерархической структуры Системы);
- персонал и пользователи Системы.

При описании компонентов Системы рекомендуется составить структурную схему информационных потоков между основными компонентами Системы, а также описать (формально или неформально) технологию обработки информации. При выборе и анализе объектов Системы важным моментом является степень детализации рассматриваемых объектов. Так, для Системы 1-го класса (отдельная ПЭВМ) допустимо рассматривать всю инфраструктуру, тогда как для Системы 3-го класса (глобальная сеть) всеобъемлющая оценка может потребовать неприемлемых затрат времени и сил. В этом случае рекомендуется сосредоточиться на описании наиболее важных компонентов Системы.

В разделе «Описание основных угроз информации» на основе анализа рисков приводится перечень и классификация возможных видов угроз безопасности информации в Системе. Под угрозой безопасности понимаются какие-либо обстоятельства или действия, которые могут быть причиной нарушения политики безопасности информации и /или нанесения ущерба Системе. Ущерб заключается в нарушении качества информации пользователей (в семантическом и прагматическом смысле) путем её уничтожения, изменения или несанкционированного получения, либо в уничтожении, изменении или несанкционированном использовании ресурсов Системы. В зависимости от класса Системы анализ угроз необходимо осуществлять на уровне отдельных аппаратных, аппаратно-программных и программных средств, отдельной локальной вычислительной сети, глобальной сети. Анализ рисков предусматривает разработку модели угроз для информации и модели нарушителя, установление соответствия модели угроз и объектов защиты, оценку возможности реализации угрозы (оценка риска), количественную либо качественную оценку величины возможного ущерба вследствие реализации угроз конфиденциальности, целостности, наблюдаемости или доступности информации либо потери управляемости Системы. Для разработки модели угроз необходимо сформировать перечень основных угроз и описать возможные способы их осуществления на основе анализа объектов Системы, характеристик вычислительной системы (ВС), физической среды, персонала, особенностей функционирования Системы.

В разделе «Требования по защите от угроз» приводятся основные задачи и цели защиты информации, объекты защиты, выбранный вариант построения КСЗИ Системы. С учетом класса Системы [5, 6] для каждого компонента и /или Системы в целом перечисляются функциональные услуги безопасности и требования к уровням реализации каждой из них, уровень гарантий реализации услуг. Для каждого компонента и /или Системы в целом определяются общие подходы и требования по защите информации от утечки техническими каналами. На следующем шаге определяются механизмы безопасности, которые реализуют функциональные услуги безопасности, осуществляется выбор технических средств защиты информации от утечки техническими каналами. При необходимости определяются компоненты Системы (например, отдельная ЛВС, специализированный АРМ, Internet-узел и т. п.), для которых целесообразно разрабатывать свои собственные политики безопасности, отличные от общей политики безопасности Системы. Исходными данными для разработки требований по защите от угроз являются задачи и функции Системы, результаты анализа среды функционирование Системы, модель угроз, модель нарушителей, результаты анализа рисков.

В разделе «Описание принципов управления доступом пользователей к информации» приводятся выбранный метод управления доступом (доверительное и /или административное управление), требования к обеспечению непрерывности защиты, к набору атрибутов доступа и правилам их использования (присвоение, применение, изменение, отмена), к регистрации действий пользователей

при использовании ресурсов Системы, а также других событий, влияющих на соблюдение реализованной в Системе политики безопасности.

В разделе «Описание правил разграничения информационных потоков» приводится перечень информационных потоков, циркулирующих между компонентами Системы. В зависимости от класса Системы структурная схема информационных потоков между основными компонентами Системы может включать:

– внутренние потоки обмена между активными и пассивными объектами внутри одной ПЭВМ;

– локальные потоки обмена между рабочими станциями и серверами внутри одной ЛВС (домена);

– межсетевые потоки обмена между ЛВС (доменами), входящими в состав одной Системы;

– потоки обмена информацией с удаленными взаимодействующими объектами, не входящими в состав Системы.

Правила разграничения информационных потоков формулируются на основе анализа области (границы) существования, направленности (входные или выходные), источников и приемников, функционального назначения потоков, требований по обеспечению конфиденциальности, целостности, наблюдаемости и доступности. Правила должны определять, где и на каких уровнях взаимодействия систем должно осуществляться разграничение информационных потоков и с использованием каких атрибутов и механизмов (идентификаторов безопасности, сетевых портов, ключей аутентификации, ключей направлений и сетевых ключей шифрования и т. п.). Правила должны также определять условия и ограничения по инициированию и завершению процессов информационного обмена, например, в виде ассоциации безопасности [7].

В разделе «Описание правил маркирования носителей информации» приводятся правила, регламентирующие порядок учета, хранения, копирования, использования и уничтожения носителей информации. Правила формулируются на основе изучения форм существования критичной информации на всех этапах жизненного цикла Системы, среды функционирование Системы, модели угроз для информации и модели нарушителей, результатов анализа рисков, требований по обеспечению конфиденциальности, целостности наблюдаемости и доступности информации

В разделе «Описание основных атрибутов доступа пользователей, процессов и пассивных объектов» приводятся состав атрибутов доступа (идентификационные имена, индивидуальные и групповые идентификаторы безопасности, пароли, метки и /или маркеры доступа, списки контроля доступа и т. п.), требования к характеристикам атрибутов доступа (принадлежность, уникальность, размерность, сроки действия и т. п.) и правила работы с ними (присвоение, использование, модификация, отмена)

В разделе «Описание правил разграничения доступа пользователей и процессов к пассивным объектам» содержится набор правил определяющих состав лиц, которым разрешен доступ к ресурсам Системы, порядок правильного использования ресурсов Системы, статус, права и привилегии администратора безопасности Системы, статус, права и привилегии пользователей Системы.

В разделе «Описание правил администрирования КСЗИ и регистрации действий пользователей» приводится порядок администрирования учетных записей пользователей, профилей пользователей, групп пользователей, общих ресурсов и аудита.

2. Разработка концепции безопасности информации в АС

Концепция безопасности информации в Системе представляет собой совокупность взглядов, общих принципов и определяет основные положения и направления обеспечения безопасности информации в Системе, а также целенаправленной организации всех работ по созданию КСЗИ. Разработка документа «Концепция безопасности информации Системы» осуществляется на основе концепции построения Системы.

Методология разработки концепции безопасности включает в себя следующие работы:

1) анализ правовых и договорных основ создания Системы;

2) изучение архитектуры создаваемой Системы и технологических процессов обработки информации с целью определения активных и пассивных компонентов, влияющих на безопасность информации;

3) определение совокупности угроз и степени уязвимости ресурсов Системы (включая передаваемую, обрабатываемую и хранимую информацию);

4) определение требуемого уровня обеспечения безопасности информации;

5) определение множества средств, методов и мероприятий защиты.

По результатам выполнения работ должны быть сформулированы общие положения безопасности, которые определяют:

– цель и приоритеты, которых необходимо придерживаться в Системе во время обеспечения безопасности информации;

– общие направления деятельности, необходимые для достижения этой цели;

– аспекты деятельности в области безопасности информации, которые должны учитываться на уровне организации в целом;

– ответственность должностных лиц и других субъектов взаимоотношений в Системе, их права и обязанности относительно реализации задач безопасности информации.

В общем случае документ «Концепция …» должен включать следующие разделы:

1) область применения;
2) общие положения;
3) основные понятия;
4) цели обеспечения безопасности информации в Системе;
5) цель создания КСЗИ в Системе;
6) угрозы, объекты и задачи защиты информации в Системе;
7) концептуальные принципы и основные положения по обеспечению безопасности информации в Системе;
8) основные направления решения проблем обеспечения безопасности информации в Системе;
9) требования к подсистеме криптографической защиты информации КСЗИ Системы;
10) требования к архитектуре КСЗИ Системы;
11) этапы развития КСЗИ Системы.

В зависимости от конкретного класса Системы [6], архитектуры и условий функционирования Системы, исходных требований собственника и возможных пользователей Системы по обеспечению безопасности информации (в т. ч. по криптографической защите), допускается, в случае необходимости, объединять отдельные разделы в один, вводить новые разделы (подразделы) либо исключать неактуальные разделы.

В разделе «Область применения» излагается назначение и предметная область документа.

В разделе «Общие положения» излагаются результаты системного анализа архитектуры построения и состава пользователей Системы, исходных требований собственника и возможных пользователей Системы по обеспечению безопасности информации (в т. ч. по криптографической защите), состава нормативно-правовой базы создания концепции безопасности информации Системы. В случае отсутствия исходных данных и требований возможных пользователей Системы по обеспечению безопасности информации системный анализ возможной архитектуры построения Системы рекомендуется проводить, исходя из следующих предпосылок:

1) пользователи и информационные процессы, являющиеся получателями и источниками информации с ограниченным доступом, осуществляют взаимодействие посредством службы 7 (прикладного) уровня Системы;

2) по отношению к анализируемой Системе сторонние пользователи не обладают правами владения, эксплуатации и распоряжения её элементами;

3) пользователи Системы (в т. ч. сторонние пользователи) в соответствии с законами Украины «Об информации» и «О защите информации в автоматизированных системах» и др. могут предъявлять требования по обеспечению их права собственности на информацию;

4) система криптографической защиты информации каждого пользователя (группы пользователей) относительно систем других пользователей (групп пользователей) в сетях Системы должна быть выделенными и включать только объекты данного пользователя (группы пользователей);

5) информация, зашифрованная пользователем, должна передаваться по сетям Системы без её расшифрования на ретрансляционных узлах (станциях, центрах);

6) эксплуатация систем защиты информации (в т. ч. криптографической защиты информации) должна осуществляться только соответствующими службами пользователей и /или владельца (оператора) Системы;

7) использование зарубежных не сертифицированных в Украине аппаратных, аппаратно-программных и программных средств защиты информации неприемлемо;

– связанные участием в обеспечении качества информации с другими функциональными подсистемами как в каждой составной части, так и Системы в целом;

6) КСЗИ Системы должна представлять собой совокупность правовых (законодательных), организационных и технических мер, средств и норм, направленных на предотвращение или существенное затруднение нанесения ущерба интересам владельца (оператора) Системы, а также её пользователям (юридическим и физическим лицам, являющимися собственниками передаваемой информации);

7) политику безопасности информации в Системе и её составных частях определяют владельцы информации (пользователи Системы) в порядке, установленном соответствующими органами государственного управления;

8) объектами защиты в Системе и её составных частях являются:

– информация пользователей и информационные процессы в системах управления (информационных системах) пользователей;

– информация о Системе и её элементах;

– управляющие процессы в Системе, включая служебную управляющую информацию;

– ресурсы Системы и ее составных частей (средства связи и управления, информационно-программное обеспечение, линии связи и т. п.);

9) циркулирующая в Системе информация пользователей и служебная информация по обеспечению функционирования системы (подсистемы, элемента), как объекты защиты, имеют различный статус :

– информация с ограниченным доступом (ИсОД), включающая сведения, составляющие государственную тайну, и конфиденциальные сведения;

– открытая информация;

10) КСЗИ Системы должна обеспечивать защиту от следующих видов угроз: …. (перечень угроз определяется моделью угроз и моделью нарушителя, разработанных для конкретной Системы);

11) по проявлению и положению источников возникновения угроз относительно составных частей Системы и её элементов угрозы группируются на классы. В свою очередь классы угроз включают совокупности угроз, сгруппированные по каналам реализации: ….. (классы и группы угроз определяются для конкретной Системы);

12) принимая во внимание широкий круг пользователей, возможность размещения оборудования Системы на территориях, неконтролируемых владельцем Системы, выход на зарубежные системы телекоммуникаций, широкое использование зарубежной техники и технологий потенциально возможные угрозы безопасности информации пользователей и безопасности Системы могут проявляться в каждой составной части и Системе в целом на уровне:

– отдельных аппаратно – программных устройств (ПЭВМ, модем, операционная система, база данных, телефонный/ телеграфный аппарат и т.п.);

– отдельных локальных сетей и подсистем (комплекс, аппаратная, центр и т.п.);

– отдельной сети и Системы в целом (при работе по наземным и спутниковым магистральным, зоновым и внутризоновым каналам и трактам передачи);

13) исходя из перечня возможных угроз, КСЗИ Системы должна обеспечивать:

– целостность пользовательской и служебной информации на всех этапах её обращения при любых угрозах;

– подтверждение подлинности пользовательской и служебной информации на всех этапах её обращения при любых угрозах;

– юридическую защиту взаимодействующих сторон (пользователей, станций) на основе модели взаимного недоверия между ними;

– информационную скрытность информации с ограниченным доступом, циркулирующей в Системе;

– защиту от несанкционированного доступа к защищаемой информации и ресурсам Системы;

– юридическую ответственность взаимодействующих сторон за сформированные, переданные, принятые и обработанные сообщения (данные);

– организационно – технические и юридические меры защиты информации с ограниченным доступом от утечки.

14) реализация функций и задач КСЗИ Системы должна обеспечиваться комплексным использованием методов и средств криптографической и технической защиты информации;

15) функциональная структура КСЗИ составных частей и Системы в целом должна включать следующие функции безопасности: ….. (перечень функциональных услуг определяется для конкретной Системы по результатам анализа рисков);

16) реализация функций КСЗИ Системы должна обеспечиваться внедрением следующих средств (механизмов) защиты: …. (состав механизмов защиты определяется для конкретной Системы по результатам формирования функциональной структуры КСЗИ и модели угроз);

17) комплекс правовых (законодательных), организационных (административных), технических и физических мероприятий и средств, реализующий функции и задачи КСЗИ Системы, должен обеспечивать:

– предупреждение условий, порождающих угрозы;

– предупреждение проявления угроз;

– обнаружение проявления угроз;

– предупреждение воздействия угроз на объекты защиты;

– обнаружение воздействия угроз на объекты защиты;

– локализацию воздействия угроз на объекты защиты;

– ликвидацию последствий воздействия угроз на объекты защиты;

18) обеспечение безопасности информации в составных частях и Системе в целом должно осуществляться комплексом правовых (законодательных), организационных (административных), морально-этических, физических и технических мер, реализующих следующие методы обеспечения безопасности:

– препятствия (физическое преграждение пути);

– управление доступом (регулирование использования всех ресурсов Системы);

– маскировка (скрытие содержания информации криптографическими методами);

– регламентация (создание условий, минимизирующих возможность несанкционированного доступа к информации);

– принуждение (соблюдение установленных правил использования ресурсов системы, в т.ч. информации под угрозой наступления ответственности);

– побуждение (выполнение установленных правил использования ресурсов системы, в т. ч. информации за счет соблюдения сложившихся моральных и этических норм);

19) обеспечение безопасности информации в Системе может быть эффективным только в том случае, если оно будет представлять собою непрерывный и целенаправленный процесс, регулярно осуществляемый на всех этапах жизненного цикла. Поэтому КСЗИ Системы помимо функций, задач и средств собственно обеспечения защиты информации должна включать функции, задачи и средства управления в качестве одного из основных компонент;

20) КСЗИ Системы и её составных частей должна реализовываться на принципах:

– системности (учёт с позиций системного подхода всех взаимосвязанных и изменяющихся во времени элементов, условий, обстоятельств и факторов, существенно значимых для обеспечения безопасности информации в Системе;

– комплексности (согласованное использование различных средств и методов защиты при построении целостной системы защиты, перекрывающей все существенные каналы реализации угроз);

– непрерывности (обеспечение защиты на всех этапах жизненного цикла Системы и её составных частей);

– адекватности требованиям (построение КСЗИ Системы должно осуществляться в соответствии с требованиями политики безопасности);

– адаптируемости (способность к целенаправленному приспособлению при изменении архитектуры или условий функционирования Системы, в т.ч. видов и классов угроз);

– функциональной самостоятельности (при осуществлении функций защиты не зависеть от других функциональных подсистем Системы);

– удобства использования (не должна создавать дополнительных неудобств для пользователей и персонала Системы);

– минимизации предоставляемых прав (каждому пользователю и каждому лицу из состава персонала Системы должны предоставляться только те полномочия на доступ к услугам и ресурсам составных частей и Системы в целом (в т.ч. к информации), которые действительно необходимы для выполнения своих функций);

– полноты контроля (должны контролироваться все каналы уязвимости информации пользователей и ресурсов составных частей и Системы в целом);

– активности и своевременности реагирования (должны быть своевременные реакции на любые атаки);

– управляемости процессами обеспечения безопасности информации в Системе;

– экономичности (расходы на КСЗИ Системы должны быть минимальными);

21) помимо этого КСЗИ Системы, как функциональные подсистемы каждой составной части и Системы в целом, должны отвечать требованиям:

– функциональным (обеспечение реализации требуемой совокупности функций и задач защиты, удовлетворение всем требованиям защиты);

– организационным (структурированность всех компонент, простота эксплуатации);

– техническим (оптимальность архитектуры, комплексное использование средств);

– экономическим (максимальное использование серийных средств, минимизация затрат на систему);

– эргономическим (минимизация помех пользователям, удобство для персонала системы связи);

– открытости (способность к развитию без нарушения функционирования);

22) при работе пользователей по постоянным виртуальным каналам или коммутируемым виртуальным соединениям через ретрансляционные узлы (центры, станции), не принадлежащие пользователям:

– зашифрование информации пользователей целесообразно осуществлять на 7 уровне с использованием всех служб зашифрование [8]. Осуществлять зашифрование информации пользователей с использованием служб зашифрование соединения на 1 – 4 уровнях, зашифрование без соединения на 2 – 4 уровнях, засекречивания потока данных на 1 уровне нецелесообразно, т. к. для нормального функционирования служб данных уровней требуется расшифрование данных пользователей на ретрансляционных узлах (станциях, центрах) общего пользования в процессе передачи;

– аутентификацию пользователей целесообразно осуществлять на 7 уровне с использованием служб аутентификации одноуровневых объектов или источников данных [8]. Осуществлять аутентификацию пользователей с использованием служб аутентификации одноуровневых объектов или источников данных на 3 и 4 уровнях (с применением механизмов симметричного шифрования, формирования цифровой подписи и обеспечения аутентификации) нецелесообразно, т. к. для нормального функционирования служб данных уровней в процессе установления соединения (или периодически в течение фазы передачи данных) в общем случае требуется рассекречивание аутентифицирующей пользователей информации на ретрансляционных узлах (станциях, центрах) общего пользования;

– реализация службы контроля доступа к ресурсам Системы на 3, 4 и 7 уровнях [8] аналогична вышеизложенному, т. к. содержит процедуры аутентификации;

23) при работе пользователя по выделенным либо по коммутируемым каналам связи, соединяющим узлы (станции) данного пользователя в качестве ретрансляционных:

– зашифрование информации пользователя с использованием служб засекречивания соединения на 1 – 4 уровнях, зашифрование потока данных на 1 и 3 уровнях, засекречивания без соединения и зашифрование выборочных полей на 2 – 4 уровнях возможна, т. к. объекты будут контролироваться его соответствующей службой;

– аутентификация пользователя с использованием служб аутентификации одноуровневых объектов или источников данных на 3 и 4 уровнях (с применением механизмов симметричного шифрования, формирования цифровой подписи и обеспечения аутентификации) возможна, т. к. объекты будут контролироваться соответствующей службой пользователя;

– реализация службы контроля доступа к ресурсам Системы на 3, 4 и 7 уровнях аналогична вышеизложенному т. к. содержит процедуры аутентификации;

24) для доставки пакетов сообщения блоки данных пользователей, в т. ч. содержащиеся в служебной части сообщения истинные адреса (нумерация, наименования) источника и получателя, должны быть оформлены в форме блоков данных (N) – протокола и засекречены. Для организации информационного обмена в телеинформационной системе блоки данных (N) – протокола пользователей должны представляться в форме блока данных (N) – службы и передаваться в форме блока данных (N) – интерфейса с присвоением пользователям в управляющей информации (N) – протокола (интерфейса) сетевых (условных) адресов;

25) реализация служб целостности (соединения с восстановлением, соединения без восстановления, выборочных полей соединения, без соединения, выборочных полей без соединения) на 3, 4 и 7 уровнях протокола информационного обмена передающего и принимающего объектов пользователя возможна как при работе пользователя по постоянным виртуальным каналам или коммутируемым виртуальным соединениям через не принадлежащие пользователю ретрансляционные узлы (центры, станции), так и при работе по выделенным либо коммутируемым каналам связи, соединяющим узлы (станции) данного пользователя в качестве ретрансляционных;

26) реализация служб защиты от отказов с подтверждением источника и защиты от отказов с подтверждением доставки на 7 уровне протокола информационного обмена передающего и принимающего объектов пользователя возможна как при работе пользователя по постоянным виртуальным каналам или коммутируемым виртуальным соединениям через не принадлежащие пользователю ретрансляционные узлы (центры, станции), так и при работе по выделенным либо коммутируемым каналам связи, соединяющим узлы (станции) данного пользователя в качестве ретрансляционных;

27) если безопасность информации обеспечивается пользователем собственными средствами на уровне приложения (информационного процесса), то допускается работа пользователя по постоянным виртуальным каналам или коммутируемым виртуальным соединениям через ретрансляционные узлы (станции), не принадлежащие пользователю, либо по выделенным или по коммутируемым каналам связи, через ретрансляционные узлы (станции) пользователя, с использованием всех служб засекречивания на 1 – 4 и 7 уровнях, служб аутентификации на 3, 4 и 7 уровнях и служб контроля доступа к ресурсам Системы на 3, 4 и 7 уровнях;

28) так как КСЗИ являются функциональными подсистемами как в каждой составной части, так и Системы в целом, их архитектура должна быть аналогична архитектуре соответствующей составной части и архитектуре Системы в целом.

В разделе «Основные направления решения проблем обеспечения безопасности информации в Системе» излагается перечень основных направлений построения КСЗИ Системы, реализация которых позволит решить поставленные задачи защиты информации в Системе, например:

1) базирование КСЗИ Системы преимущественно на национальных средствах защиты. В переходный период допускается использование зарубежных средств при условии обязательной сертификации и аттестации их национальными службами;

2) внедрение систем многоуровневой защиты, которые включают оперативные пункты (центры) управления (администраторов безопасности), службы и механизмы безопасности в соответствии с рекомендациями нормативных документов ДСТСЗИ СБУ, стандартов ISO 7498-2, ISO 15408;

3) использование доменов (модель «несколько мастер – доменов» или модель «домены с целиком доверительными отношениями») для организации локальных и распределенных сетей Системы;

4) применение систем защиты типа FireWall (фильтрующие маршрутизаторы, программные фильтры, шлюзы приложений) в вариантах организации схем защиты «демилитаризованная зона»

(demilitarized zone) или «бастиона серверов» (bastion servers) для защиты информационных ресурсов Системы в доменах локальных сетей от возможных угроз со стороны Internet или удаленных пользователей;

5) широкое применение криптографических аппаратных, аппаратно – программных и программных средств для защиты информации от несанкционированного доступа, нарушения целостности, модификации и хищений, включая:

- комплексное использование канального и межконцевого (абонентского) методов криптографической защиты;
- переход от центров распределения ключей к центрам управления ключами с использованием многоуровневой иерархии ключей;
- создание электронных систем генерации, учета и распределения ключей (включая разработку состоятельных протоколов передачи и подтверждения подлинности ключей, источника и получателя);
- использование методов цифровой подписи на базе несимметричных криптосистем;
- преимущественно аппаратная реализация криптоалгоритмов в виде модулей (блоков), которые встраиваются в аппаратуру управления и связи (коммутаторы, модемы, телефонные аппараты, ПЭВМ, радиостанции и т.п.);
- уменьшение энергопотребления и массогабаритных характеристик за счет внедрения микропроцессорной техники и современных технологий;
- повышение скорости шифрования до 10 –100 Мбит/с;

6) внедрение многоуровневых систем паролирования для аутентификации и контроля полномочий объектов (пользователей, терминалов и т.п.), в т.ч. с применением интеллектуальных карточек различных типов (в т.ч. с криптографическими ключами);

7) широкое использование базового программного обеспечения (ОС, СУБД) с встроенными сертифицированными системами защиты информации;

8) унификация и стандартизация способов и методов защиты;

9) реализация моделей взаимного недоверия и взаимной защиты всех участников информационного обмена, включая арбитра и злоумышленника (криптоаналитика), в применяемых механизмах защиты информационного обмена (цифровой подписи, аутентификации, контроля целостности, шифрования, управления доступом и маршрутизации, автоматического протоколирования и аудита);

10) внедрение автоматизированной системы контроля и управления информационной безопасностью Системы;

11) внедрение интеллектуальной системы поддержки принятия решений по управлению безопасностью информации, которая содержит в себе свойства экспертной системы реального времени и интегрированной программной среды.

В разделе «Требования к подсистеме криптографической защиты информации КСЗИ Системы» приводится основные требования к составу и характеристикам применяемых криптосистем, криптопротоколов, аппаратных, аппаратно-программных и программных средств криптографической защиты. Кроме этого, приводятся требования к функциональной и организационно-топологической структурам подсистемы криптографической защиты информации.

В разделе «Требования к архитектуре КСЗИ Системы» излагаются основные требования к организационной, функциональной и топологической структурам КСЗИ, составу информационного, программного и технического обеспечения. Основные требования разрабатываются с учетом концепции построения Системы, выбранной архитектуры Системы, политики безопасности информации в Системе.

В разделе «Этапы развития КСЗИ Системы» излагаются этапы разработки и внедрения КСЗИ, которые определяются в соответствии с правовыми и договорными документами, регламентирующими создание Системы.

Заключение

Материалы, помещенные в статье, разработаны с использованием действующих на момент написания статьи отечественных и зарубежных нормативных документов по обеспечению безопасности информации. Представленные в статье методологические основы концепции и политики безопасности информационных технологий отражают опыт авторов по разработке комплексных систем защиты информации различного назначения.

Список литературы: 1. *Положення про порядок здійснення криптографічного захисту інформації в Україні*. Указ Президента України від 22.05. 1998. N 505/*98* 2. *НД ТЗИ 1.1-003-99*. Термінологія в галузі захисту інформації в комп'ютерних системах від несанкціонованого доступу. ДСТСЗІ СБУ. 1999. 3. *НД ТЗІ 3.7-001-99*. Методичні вказівки щодо розробки технічного завдання на створення комплексної системи захисту інформації в автоматизованій системі. ДСТСЗІ СБУ. 1999. 4. *НД ТЗІ 1.1-002-99*. Загальні положення щодо захисту інформації в комп'ютерних системах від несанкціонованого доступу. ДСТСЗІ СБУ. 1999. 5. *НД ТЗІ 2.5-004-99*. Критерії оцінки захищеності інформації в комп'ютерних системах від несанкціонованого доступу. ДСТСЗІ СБУ. 1999. 6. *НД ТЗІ 2.5-005 –99*. Класифікація автоматизованих систем і стандартні функціональні профілі захищеності оброблюваної інформації від несанкціонованого доступу. ДСТСЗІ СБУ. 1999. 7. *Романец Ю. В., Тимофеев П. А., Шаньгин В. Ф.* Защита информации в компьютерных системах и сетях/под редакцией В. Ф. Шаньгина. М.: Радио и связь, 1999. 328 с. 8. *ISO/DIS 7498/2*. Information Processing System – OSI. Reference Model. Part 2: Security Architecture. ISO. 1988. 32 p.

Харьковский государственный университет радиоэлектроники

Поступила в редколлегию 5.04.2001

УДК 681.3.06:519.248.681

Л. В. СКРЫПНИК, д-р. техн. наук, А. В. ПОТИЙ, канд. техн. наук.

ГИБКОСТЬ И СПЕЦИАЛИЗАЦИЯ ПРОФИЛЯ ЗАЩИТЫ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ

Введение

На сегодняшний день в Украине окончательно сформировалась точка зрения, заключающаяся в том, что обеспечение безопасности информации при её обработке в автоматизированных системах (АС) осуществляется на всех стадиях жизненного цикла АС, на всех технологических этапах обработки информации и во всех режимах функционирования АС путем создания и внедрения в АС комплексной системы защиты информации (СЗИ) как составной и неотъемлемой части АС [1]. Анализ действующей нормативной базы, как отечественной, так и международной, показал, что наряду с задачей разработки системы защиты информации, интегрированной в АС, актуальной является задача оценки защищенности информации в АС. Причем последняя задача должна решаться как для уже существующих, так и для проектируемых СЗИ. Доминирующим подходом при решении этой задачи на сегодняшний день является метод, основанный на разработке профиля защиты АС с последующей его оценкой. *Профиль защиты* представляет собой реализационно-независимую совокупность требований безопасности информационных технологий (ИТ-безопасности), которые включают в себя: 1) функциональные требования безопасности, определяющие набор функций безопасности, направленных на решение задач защиты и обеспечения требуемого уровня защиты и 2) требования адекватности, обеспечивающие высокую степень уверенности в правильности выбора и надежности функционирования механизмов безопасности. (Здесь и далее используется терминология международного стандарта ISO/IEC 15408 [5]. Причем термины «автоматизированная система» [1, 3, 4], «компьютерная система» [2] являются синонимами термина «объект оценки (ТОЕ)» [5]. Все эти объекты являются системами информационных технологий (ИТ-системы).)

При анализе и оценке профиля защиты эксперт устанавливает, насколько рассматриваемый профиль отвечает предъявленным к нему требованиям функциональной полноты, непротиворечивости, реализуемости, адекватности. По нашему мнению, помимо этих требований к профилю защиты необходимо предъявить новое требование – гибкость. Обоснование необходимости введения данного требования и является целью данной статьи.

1. Требования к профилю защиты

К профилю защиты предъявляют следующие требования.

1. *Функциональная полнота.* Профиль защиты должен содержать функционально полный набор функций безопасности, что обеспечивает решение всех задач защиты и предотвращение выявленных угроз безопасности.

2. *Непротиворечивость (согласованность).* Функция безопасности характеризует проявление свойств системы защиты информации в данной совокупности соотношений и представляет собой способ действия системы при взаимодействии элементов системы между собой, а также при взаимодействии системы с внешней средой (средой эксплуатации). В связи с этим требование непротиворечивости направлено на учет взаимозависимости функциональных требований между собой и согласованности функциональных требований и требований адекватности. Последнее означает учет влияния внешних факторов на профиль защиты. В общем, соблюдение требования непротиворечивости направлено не обеспечение целостности СЗИ как системы.

3. *Реализуемость* (техническая реализуемость). Требование отражает техническую возможность реализации определенного согласованного набора функций безопасности, с учетом имеющихся ограничений на ресурсы, выделяемых на реализацию системы.

4. *Адекватность.* Наиболее ёмкое требование, направленное на обеспечение требуемого уровня безопасности информации, т.е. адекватности СЗИ угрозам безопасности информации, существующим в конкретной среде эксплуатации и в конкретной ИТ-системе. Адекватность требует осуществления оценки эффективности функций безопасности и оценки корректности реализации функций безопасности.

Поскольку профиль защиты, по сути, представляет собой функциональную модель СЗИ, то перечисленные требования профиля преобразуются в свойства системы, которые будут характеризовать эффективность функционирования СЗИ. Однако, перечисленные требования направлены на обеспечение статических свойств СЗИ, в то время как процесс защиты информации является непрерывным

динамическим процессом. Как антропогенная система, СЗИ в числе многих свойств, таких как целенаправленность, целостность, управляемость и других, должна обладать свойством непрерывности развития. Игнорирования концепции развития в процессах проектирования, производства и эксплуатации систем приводит к созданию нежизнеспособных систем. Обеспечение свойства непрерывности требует учета динамической природы функционирования СЗИ, совершенствования, как технологий защиты информации, так и технологий, методов и способов осуществления несанкционированного доступа к защищаемой информации. В зависимости от складывающейся обстановки, с точки зрения состава и возможностей реализации тех или иных угроз безопасности, может изменяться состав решаемых системой задач защиты. Причем это изменение носит не только количественный характер (появление новых задач защиты, отпадение необходимости решения существующих задач защиты), но и качественный характер (изменение степени риска той или иной угрозы приводит к изменению приоритетности задачи защиты и, как следствие, к изменению уровня адекватности профиля защиты). С этой точки зрения непрерывность защиты информации заключается в постоянном анализе среды эксплуатации АС, оперативном анализе (оценке) рисков и выработке управляющих воздействий, направленных на коррекцию задач защиты. Изменение внешних условий влечет за собой изменение способа действия системы при её взаимодействии с внешней средой, то есть приводит к изменению функций безопасности. В более широком смысле это приводит к изменению профиля защиты и, в конечном счете, к необходимости изменения состава средств и механизмов защиты или, как минимум, к изменению режимов их работы. Таким образом, функциональная структура СЗИ не является статичной, раз и навсегда разработанной. Профиль защиты в процессе функционирования системы будет изменяться, модифицироваться. Из выше сказанного следует, что при разработке профиля защиты к нему необходимо предъявлять еще одно требование – требование *гибкости*. Причиной, объективно побуждающей введение такого требования, является усиливающееся, особенно в современных АС и СЗИ, противоречие между необходимостью решения быстроизменяющихся задач защиты и сложностью обеспечения при этом высокой эффективности СЗИ с помощью существующих средств защиты. Ясно, что высокую гибкость системы и её высокую эффективность в любых условиях изменяющейся обстановки обеспечить достаточно сложно и дорого. С одной стороны, наибольшая эффективность защиты может быть достигнута лишь при решении относительно узкого класса задач защиты путем создания «жесткой», узкопрофильной, специализированной СЗИ. Однако, при работе СЗИ в условиях неопределенности относительно состояния среды эксплуатации, быстроизменяющегося состава и качества задач защиты наибольшей эффективностью будут обладать широкопрофильные, гибкие СЗИ, с большим набором функций безопасности. Возникает ***проблема нахождения оптимального состояния гибкости и эффективности системы защиты информации.***

2. Гибкость как экономическая категория

Гибкость системы это свойство системы, состоящее в возможности её совершенствования, расширения и придания ей новых качеств [6]. Гибкость предполагает легкое, то есть без особых усилий и затрат, и достаточно быстрое изменение того или иного состояния системы. Применительно к СЗИ можно сформулировать следующее определение. ***Гибкая система защиты информации это совокупность в разных сочетаниях механизмов безопасности, средств защиты информации, программного и аппаратного обеспечения, системы анализа (оценки) рисков и системы управления, обеспечивающих функционирование СЗИ в автоматизированном режиме, обладающая свойством автоматизированной перестройки при решении задач защиты различного состава в пределах своих технических характеристик.***

Интенсивное изменение среды эксплуатации приводит к тому, что система защиты вынуждена приспосабливаться (адаптироваться) к решению новых задач защиты, обладающих определенной степенью разнообразия. Однако не всякая способность адаптации системы к быстрой смене задач защиты, быстрой смене и модификации функциональных требований может квалифицироваться как гибкость. Адаптация, достигаемая «любой ценой» с точки зрения затрачиваемых ресурсов, не может признаваться гибкостью. Отсюда, *гибкость как экономическая категория отражает способность СЗИ к эффективной адаптации, то есть рассматривается, как способность системы изменять свои цели функционирования без существенных затрат.*

Изменение целей функционирования жестких систем связано, как правило, либо с техническим перевооружением, либо с реконструкцией СЗИ. Тогда как гибкие системы могут длительно и эффективно удовлетворять постоянно изменяющиеся потребности собственника защищаемых ресурсов без технического перевооружения и модификации системы ЗИ.

Каким образом охарактеризовать гибкость профиля АС и, следовательно, гибкость СЗИ? По аналогии с решением подобной задачи для производственных систем [7] введем показатель

$$G(t_к, t_н) = J_{из}(t_к, t_н)/J_з(t_к, t_н),$$

где $G(t_к, t_н)$ – гибкость профиля защиты в период $(t_к, t_н)$; $J_{из}(t_к, t_н)$ – индекс, характеризующий изменение степени разнообразия задач защиты в системе в период $(t_к, t_н)$; $J_з(t_к, t_н)$ – индекс, характеризующий изменение приведенных динамических затрат, связанных с функционированием системы в период $(t_к, t_н)$.

Чем больше значение G, тем система более гибкая.

Значение G зависит от величины интервала $(t_к, t_н)$. Прежде чем изложить возможный подход к оценке этого интервала рассмотрим понятия техническое перевооружение и реконструкция относительно системы ЗИ. *Техническим перевооружением* будем называть комплекс мероприятий по повышению до современного уровня технического (программного, аппаратного, алгоритмического, математического) уровня элементов и подсистем СЗИ путем внедрения новых технологий обработки информации, программного и аппаратного обеспечения и других мер, направленных на обеспечение и повышение адаптивности функционирования СЗИ. Техническое перевооружение обусловлено развитием методов несанкционированного доступа к защищаемой информации, изменением возможностей злоумышленника по осуществлению атак, качественным изменением защищаемых ресурсов и другими условиями. *Реконструкция* системы ЗИ предполагает полную перестройку системы ЗИ по новым принципам. Необходимость реконструкции чаще всего вызывается физическим и моральным старением СЗИ.

Чем продолжительнее период между двумя последовательными техническими перевооружениями или реконструкциями, тем большей гибкостью при условии стабильной эффективности функционирования обладает СЗИ. С другой стороны, чем продолжительнее этот период, тем возможен больший потенциальный ущерб от нарастающего физического износа и морального старения механизмов и средств защиты информации. Особенно это зависит от продолжительности межреконструкционного периода. Это противоречие порождает необходимость решения следующих задач:

– нахождение оптимальной продолжительности межреконструкционного периода;

– выбор рациональных сроков начала и окончания работ по подготовке и проведению перевооружения и реконструкции СЗИ.

Решение этих задач направлено на предотвращение экономического ущерба, вызываемого использованием изношенных (устаревших) и преждевременной ликвидации недоамортизированных элементов СЗИ, а также «замораживанием» различного рода ресурсов. На основе решения этих задач оптимизации могут быть найдены значения периода $(t_к, t_н)$, для которого определяются значения $G(t_к, t_н)$. Полученные значения $G(t_к, t_н)$ могут рассматриваться как количественные характеристики оптимальной гибкости СЗИ.

Показатель G также зависит от индекса изменения степени разнообразия задач защиты. Среди факторов, оказывающих влияние на разнообразие задач защиты можно выделить следующие:

– степень вариативности целей функционирования АС (надсистемы), интегрированной частью которой является СЗИ. АС и СЗИ по своей сути являются многоцелевыми системами. В таких системах цели либо задаются с вышестоящих уровней иерархии, либо определяется из сложившейся ситуации лицом, принимающим решение;

– вариативность среды эксплуатации. Изменчивость среды эксплуатации складывается из возможности изменения ресурсов злоумышленника, что в свою очередь приводит к изменению количественного и качественного состава угроз безопасности, изменения защищаемых ресурсов, совокупность которых определяется изменчивостью потребностей собственника ресурсов.

Если степень изменения задач защиты, как потребностей собственника защищаемых ресурсов, позволяет в течение длительного периода эффективно использовать для их решения жесткую систему, то оптимальной будет невысокая степень гибкости. Это характерно для однообъектных систем обработки информации, решающих узкий класс задач защиты. Для распределенных АС характерно быстрое изменение внешних условий, постоянная модификация задач защиты, что приводит к изменению представлений об оптимальной гибкости. Если жесткая система начинает входить в противо-

речие с необходимостью быстрого и эффективного изменения функционального состава системы, оптимальной становится более гибкая система.

Таким образом, существуют области эффективного функционирования СЗИ различной степени гибкости. Поэтому стремление в разных условиях создавать СЗИ на основе одинаковых по степени гибкости профилей защиты экономически нецелесообразно.

3. Специализация профиля защиты

В вопросе гибкости СЗИ необходимо обратить внимание на такой аспект как специализация системы. *Специализация* – это концентрация, сосредоточение деятельности на решение какой-либо задачи, на выполнение какой-либо операции или функции. Не возникает сомнений, что в зависимости от совокупности решаемых задач защиты системы ЗИ будут иметь различную степень специализации. Различные степени специализации создают различные условия для гибкости профиля защиты.

Имеющиеся на сегодняшний день нормативные документы по оценки защищенности информационных технологий, как международные [5], так и национальные [1-4] позволяют сформулировать следующий взгляд на специализацию СЗИ.

Специализация СЗИ зависит от типа задач защиты, на решение которых направлены функции безопасности. Тип задач определяется в зависимости от класса угроз безопасности, на предотвращение которых направлено функционирование СЗИ. Согласно нормативным документам, профиль защиты должен содержать функциональные требования (ФТ), которые определяют набор функций безопасности СЗИ. Причем множество функциональных требований декомпозируются на классы по целевому признаку. Так, в международном стандарте ISO/IEC 15408 определено тринадцать классов функциональных требований, в отечественном нормативном документе НД ТЗИ 2.5.–004–99 – пять классов. Анализ содержания функциональных требований (то есть, семейств и компонент) позволят разбить их на три основные категории.

1. Категория специальных функциональных требований, в которую входят классы ФТ, целевой направленностью которых является решение задач защиты по непосредственному предотвращению конкретных типов угроз безопасности. В самом общем случае в данную категорию можно отнести классы ФТ, предотвращающие угрозы конфиденциальности, целостности и доступности информации.

2. Категория общих функциональных требований объединяет классы ФТ, целевой направленностью которых является решение общих задач, не зависящих от конкретного типа или набора предотвращаемых системой угроз безопасности. К этой категории можно отнести классы ФТ, направленные на решение задач управления системой защиты информации, аудита безопасности, адекватности (гарантированности) защиты и т.д. Отметим, что данные ФТ практически всегда включаются в систему защиты информации.

3. Категория вспомогательных функциональных требований. В данную категорию входят ФТ, целевой направленность которых является решение задач по обеспечению работы, расширению и усилению функциональных возможностей механизмов безопасности, которые реализуют функции безопасности, определяемые специальными требованиями.

Целевую специализацию будет определять совокупность специальных функциональных требований. Поскольку в настоящее время преобладают взгляды построения комплексных систем защиты информации с непрерывным процессом обеспечения безопасности данных, то современные СЗИ будут обладать специализацией по нескольким классам специальных ФТ, с выделением приоритетности отдельного класса функциональных требований в зависимости от целевого предназначения АС (например, в системах электронных платежей предъявляются усиленные требования к обеспечению целостности данных, в военных системах управления и связи – к конфиденциальности, в системах хранения данных – к доступности). Приоритетность того или иного класса ФТ может быть определена через количество функциональных компонент, включаемых в профиль защиты из конкретного класса ФТ. Специализация по классам ФТ будет отличаться разнообразием решаемых задач защиты, что создает благоприятные предпосылки для эффективного использования сложных , комплексных и, в конечном итоге, гибких СЗИ.

В заключение сформулируем главную цель СЗИ, которая позволяет определить критерий предпочтительности параметров профиля защиты. Такой целью можно считать **эффективное (т.е. своевременное, оперативное, адекватное) решение заданной совокупности задач защиты (в плане удовлетворения потребностей собственника защищаемых ресурсов) при ограничениях, накладываемых на используемые ресурсы, возрастании потенциальных возможностей злоумышлен-**

ника по несанкционированному доступу и использованию защищаемых ресурсов и одновременном совершенствовании среды эксплуатации и не ухудшении свойств и характеристик АС (надсистемы).

Заключение

1. Профиль защиты АС должен отражать не только статические свойства СЗИ, но и ее динамические свойства, отражать свойство развития системы и непрерывности процесса защиты информации. В связи с этим помимо известных требований функциональной полноты, непротиворечивости, реализуемости, адекватности, к профилю защиты автоматизированной системы необходимо предъявлять требование гибкости.

2. Гибкость, как экономическая категория, отражает способность СЗИ к эффективной адаптации, то есть рассматривается как способность системы изменять свои цели функционирования без существенных затрат. Требования гибкости позволит учитывать вариативность задач защиты, изменчивость среды эксплуатации, которая характерна для современных автоматизированных систем. В зависимости от степени варитативности задач защиты создаются системы с различной степенью гибкости, что позволяет сделать вывод о существовании различных областей эффективного функционирования гибких СЗИ.

3. Функциональные требования могут быть разбиты на три основных категории: специальные, общие и вспомогательные. В современных автоматизированных системах профили защиты будут специализироваться по нескольким классам специальных требований, что создает условия разработки гибкой системы защиты с «широким» профилем.

Список литературы. 1. *Общие положения по защите информации в компьютерных системах от несанкционированного доступа.* НД ТЗИ 1.1.–002–99. ДСТСЗИ СБ Украины. Киев, 1999. **2.** *Критерии оценки защищенности информации в компьютерных системах от несанкционированного доступа.* НД ТЗИ 2.5.–004–99.–ДСТСЗИ СБ Украины. Киев, 1999. **3.** *Классификация автоматизированных систем и стандартные функциональные профили защищенности обрабатываемой информации от несанкционированного доступа.* НД ТЗИ 2.5.–005–99. ДСТСЗИ СБ Украины. Киев, 1999. **4.** *Средства вычислительной техники.* Защита от несанкционированного доступа к информации. Показатели защищенности СВТ от НСД к информации. Руководящий документ Гостехкомиссии России. М.: ГТК РФ, 1992. **5.** *ISO/IEC 15408.* Information technology – Security techniques – Evaluation criteria for IT security. **6.** *Першиков В.И., Савинков В.М.* Толковый словарь по информатике. М.: Финансы и статистика, 1991. **7.** *Системное проектирование радиоэлектронных предприятий с гибкой автоматизированной технологией* / Под ред. В.А. Мясникова, Ф.И. Темникова. М.: Радио и связь, 1990.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 30.03.2001

УДК 681.3.06:519.248.681

И. Д. ГОРБЕНКО, д-р техн. наук, Л. В. СКРЫПНИК, д-р техн. наук,
С. А. ГОЛОВАШИЧ, Т. А. ГРИНЕНКО

СТАНДАРТ СИММЕТРИЧНОГО ШИФРОВАНИЯ 21 ВЕКА: СВОЙСТВА, РЕЖИМЫ РАБОТЫ, РЕАЛИЗАЦИЯ

Введение

Мировое информационное сообщество встретило 21 век важным событием – успешным завершением трехлетнего проекта общественного создания и принятия в качестве стандарта 21 века алгоритма симметричного шифрования. Им стал один из 15-ти кандидатов – алгоритм RIJNDAEL [1-6], авторы Joan Daemen и Vincent Rijmen. Проект создания стандарта симметричного шифрования был инициирован Национальным Институтом Стандартов и Технологий (NIST) США в 1997г. Было организовано и проведено три этапа рассмотрения, анализа и обсуждения представленных кандидатов. В 1998 г. на Первой Конференции принято решение о приеме к открытому обсуждению 15 пакетов описаний кандидатов. В результате общественного обсуждения с учетом представленных на Вторую Конференцию материалов исследований, замечаний и комментариев, число кандидатов уменьшилось до 5. В качестве кандидатов остались MARS, RC6, RIJNDAEL, SERPENT и TWOFICH. В апреле 2000 г. на Третьей Конференции были рассмотрены все поступившие материалы и проведено широкое обсуждение свойств и характеристик кандидатов. Были учтены предложения и мнения ведущих специалистов криптологов, а также результаты, полученные сотрудниками NIST США.

При отборе оценивались:

- реальная защищенность алгоритмов от криптоаналитических атак;
- статистическая безопасность криптографических алгоритмов;
- надежность математической базы криптоалгоритмов;
- вычислительная сложность (скорость) выполнения зашифрования и расшифрования;
- сложность программной, аппаратной и аппаратно-программной реализации;
- вычислительная сложность (скорость) развертывания ключей;
- возможность работы с различными длинами информационных блоков и исходных ключей;
- возможность реализации на существующем спектре программных платформ и приложений;
- возможность применения алгоритма во всех рекомендуемых режимах работы – блочного шифрования, поточного шифрования, поточного шифрования с обратной связью, выработки кодов аутентификации, хеширования, а также генератора псевдослучайных чисел;
- эквивалентность сложности программной, аппаратной и аппаратно-программной реализаций и др.

Материалы исследований, комментарии, замечания, а также мнения экспертов и организаций представлены на сайте Третьей Конференции [5]. В результате голосования участников Конференции голоса распределились следующим образом:

MARS - 13, RC6 - 23, RIJNDAEL - 86, SERPENT - 59, TWOFICH - 31.

Таким образом, специалисты-криптологи, участвующие в работе Третьей конференции, высказались за принятие в качестве стандарта симметричного шифрования алгоритма RIJNDAEL. NIST после дополнительных исследований и тестирования рекомендовал RIJNDAEL в качестве стандарта 21века.

Необходимо подчеркнуть важность и отметить высокую эффективность реализации проекта AES. По существу за три года выполнения проекта сделан прорыв в области проектирования, тестирования и методик анализа свойств и характеристик блоковых симметричных криптоалгоритмов. Необходимо отметить важность этого проекта для Украины. Специалисты получили доступ к методическому аппарату универсальных методик и тестов, средств их реализации. Безусловно, что как сами алгоритмы, так и научно-методический аппарат, созданный в процессе их создания, еще долго будут в поле зрения ученых-криптологов и практиков. В настоящей статье ставится, прежде всего, задача ознакомления специалистов в области информационной безопасности, а также практиков с криптографическим алгоритмом блочного симметричного шифрования, рекомендованного в качестве стандарта 21 века. Вместе с тем в статье приводится ряд оценок и результатов анализа свойств алгоритма.

1. Общая характеристика криптоалгоритма

Алгоритм RIJNDAEL (RD) является блоковым симметричным криптоалгоритмом. Длина информационного блока l_n может быть равной 128, 192 или 256 бит. Криптографические преобразования в алгоритме осуществляются за счет преобразования блоков информации с использованием ключевых данных. Ключ, вводимый в средства реализации RIJNDAEL, называют исходным ключом K_u. Разрешенными длинами исходного ключа есть l_{k_u}=128, 192 и 256 бит. Ключи, используемые в циклах преобразования, формируются из исходного K_u. Для этого выполняется процедура развертывания из исходного цикловых ключей. Число развернутых ключей N_p определяется как $N_p = n_ц + 1$, а длина l_p развернутого ключа как $l_p = (n_ц + 1)l_u$, где $n_ц$ – есть число циклов преобразования, выполняемых в алгоритме, а l_u – длина информационного блока.

Алгоритм RD может применяться в пяти режимах:

- блочного шифрования;
- поточного шифрования;
- поточного шифрования с обратной связью;
- выработки ключевой хеш-функции (кода-аутентификации);
- генератора псевдослучайных последовательностей.

2. Представление преобразуемой информации и ключей, цикл преобразования

В алгоритме RD преобразования выполняются при задании длины блоков информации l_u=128, 192, 256 бит и длины исходных ключей l_{k_u}=128, 192 и 256 бит. Необходимая стойкость в алгоритме обеспечивается за счет многоциклового преобразования, причем, в каждом из циклов применяются табличные и ключевые преобразования, т.е. преобразования по фиксированным таблицам и развернутым ключам. Пусть необходимо зашифровать A_i блок информации длиной соответственно 128 (192 или 256 бит). Назовем значение $A_i = M_i$, где M_i – значение зашифровываемого блока, начальным состоянием. Далее в зависимости от номера цикла j преобразования состояние (state) будем обозначать как A_{ij}, представляя его в виде матрицы байтов. Входные данные A_i и выходные C_i рассматриваются как одномерные массивы из 8-битовых слов (байтов), пронумерованные по столбцам от 0 до $4n_c - 1$, где n_c есть количество столбцов.

Для l_u=128 бит (16 байтов)
$$A_{ij} = \begin{vmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{vmatrix} \quad (1)$$

Для l_u=192 бита (24 байтов)
$$A_{ij} = \begin{vmatrix} a_{00} & a_{01} & a_{02} & a_{03} & a_{04} & a_{05} \\ a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{20} & a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{30} & a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{vmatrix} \quad (2)$$

Для l_u=256 бит (32 байтов)
$$A_{ij} = \begin{vmatrix} a_{00} & a_{01} & a_{02} & a_{03} & a_{04} & a_{05} & a_{06} & a_{07} \\ a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ a_{20} & a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ a_{30} & a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \end{vmatrix} \quad (3)$$

Таким образом, состояние A_{ij} описывается матрицей байтов, в которой четыре строки и разное число столбцов – 4, 6 и 8.

Аналогично состоянию информационного блока задается и исходный ключ k_u, например, для длины l_{k_i} =128 бит

$$k_u = \begin{vmatrix} k_{00} & k_{01} & k_{02} & k_{03} \\ k_{10} & k_{11} & k_{12} & k_{13} \\ k_{20} & k_{21} & k_{22} & k_{23} \\ k_{30} & k_{31} & k_{32} & k_{33} \end{vmatrix} \qquad (4)$$

Для длин ключей k_u, равных 192 и 256 бит, исходный ключ задается аналогично (4).

Таблица 1

$l_k \backslash l_м$	128	192	126
128	10	12	14
192	12	12	14
256	14	14	14

В алгоритме RD число циклов $n_ц$ преобразования зависит от длины информационного блока $l_м$ и длины исходного ключа l_{k_u}. В табл. 1 приведено значение $n_ц$ циклов преобразования как функции $l_м$ и l_{k_u}.

В каждом из циклов преобразования (за исключением последнего) выполняется три табличных преобразования и одно криптографическое. В процессе преобразования байты считываются по столбцам, т.е. $a_{00}, a_{10}, a_{20}, a_{30}, a_{01}, a_{11}, \ldots$ и т.д. На языке псевдо-Си один цикл имеет вид

```
Round(State); //циклы 1, n_ц − 1
{
  Bytesub(State); //замена байт
  ShiftRow(State); //сдвиг строчек
  MixColumn(State); //перемешивание в столбцах
  AddRoundKey(State, RoundKey); //сложение с ключом
}
```

На последнем цикле преобразование MixColumn отсутствует, т.е. перемешивание в столбцах не выполняется. Алгоритм RD блочного зашифрования на $n_ц$ представлен на рис. 1.

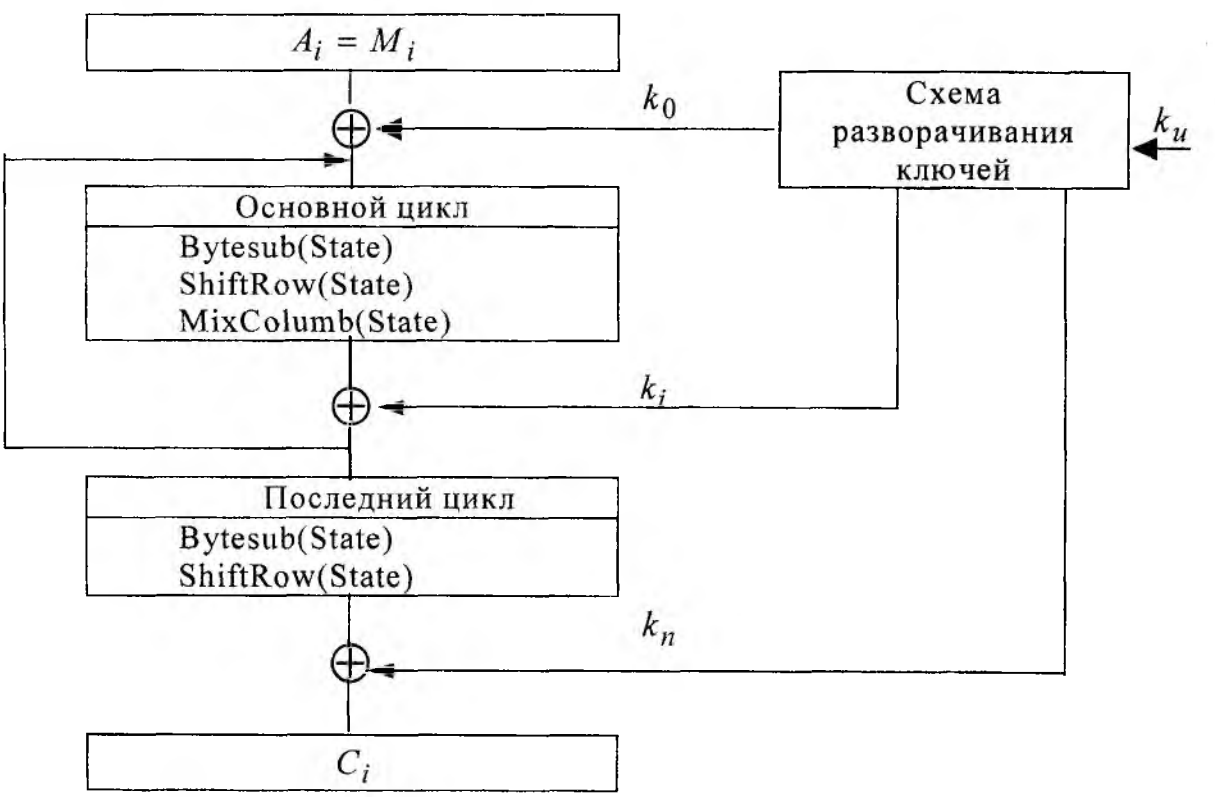


Рис. 1

3. Табличные и криптографические преобразования

Преобразование Bytesub (замена байт) математически может быть представлено в виде двух преобразований. При первом преобразовании каждый байт состояния a_{ij} заменяется мультипликативно обратным элементом a_{ij}^{-1} в поле Галуа GF(2^8), т.е.

$$a_{ij} \cdot a_{ij}^{-1} \equiv 1 \left(\mathrm{mod}\, f(x) = x^8 + x^4 + x^3 + x + 1\right) \tag{5}$$

Второе преобразование обеспечивает аффинное преобразование $a_{ij}^{-1}\left(x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 = x\right)$ по правилу:

$$Y = C \cdot x + C_1 \left(\mathrm{mod}\, x^8 + 1\right) \tag{6}$$

В матричном представлении (6) имеет вид:

$$\begin{vmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{vmatrix} \oplus \begin{vmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{vmatrix} \tag{7}$$

Преобразования (5) и (7) можно представить в виде одной таблицы подстановки типа «байт в байт». Если состояние A_i имеет вид (2), то после преобразования «замена байт» (Bytesub) получим состояние A_i', что можно представить преобразованием последовательной замены байт согласно таблице замены (S-box). На рис. 2 показано, как a_{23} байт заменяется по таблице на a'_{23}:

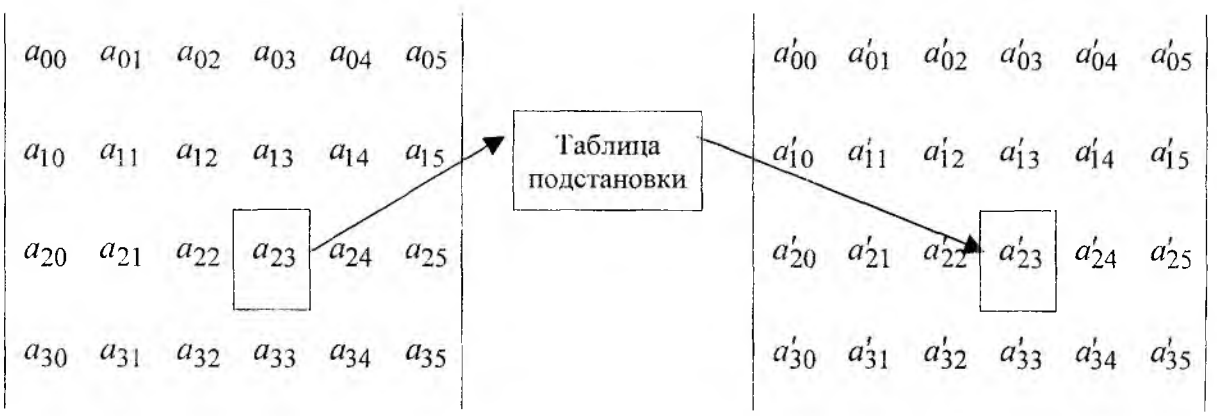


Рис. 2

Преобразование ShiftRow (Сдвиг строчек) обеспечивает циклический сдвиг строчек состояния. При этом первая строка циклически не сдвигается, а вторая, третья и четвертая циклически сдвигаются на величину, указанную в табл. 2. Причем, величины сдвига C_1, C_2, C_3 зависят от числа столбцов n_c состояния $A_i((1),(2)\,или\,(3))$.

Таблица 2

n_c	C_1	C_2	C_3
4	1	2	3
6	1	2	3
8	1	3	4

На рис. 3 приведен пример преобразования «сдвиг строчек» для состояния A_i (1), длина блока состояния 128 бит (16 байт).

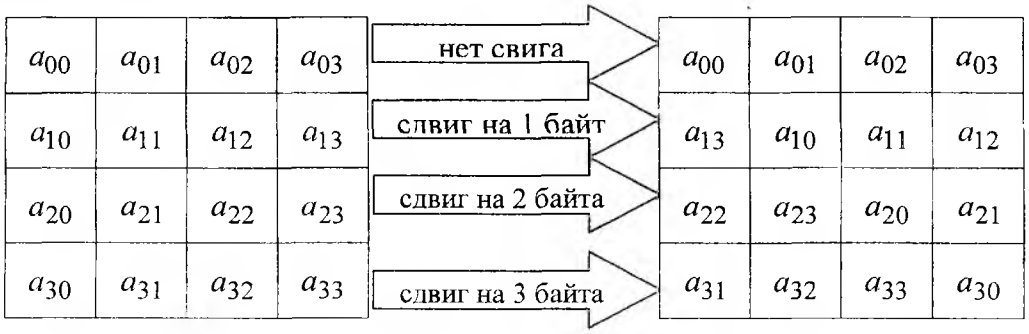


Рис. 3

Преобразование типа «перемешивание в столбцах» обеспечивает последовательное табличное преобразование элементов-байтов столбцов. Столбцы, всегда состоящие из четырех байт, представляются полиномами третьей степени:

$$a(x) = a_{3j}x^3 + a_{2j}x^2 + a_1 x + a_0 \left(\mathrm{mod}\, x^4 + 1\right),\ j = \overline{0, n_c}\,. \tag{8}$$

Коэффициенты $a(x)$ принимают в (8) значения в интервале от нуля (00000000) до 255 (11111111).

Перемешивание производится умножением $a(x)$ на константу

$$C(x) = '03'x^3 + '01'x^2 + '01'x + '02'\,, \tag{9}$$

т.е. преобразованный столбец $a'(x)$ есть

$$a'(x) = C(x)\cdot a(x)\left(\mathrm{mod}\, x^4 + 1\right) \tag{10}$$

В матричном виде преобразование (10) имеет вид:

$$\begin{vmatrix} a'_0 \\ a'_1 \\ a'_2 \\ a'_3 \end{vmatrix} = \begin{vmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{vmatrix} \begin{vmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{vmatrix} \tag{11}$$

На рис. 4 показано преобразование типа «перемешивание в столбцах» для A_i состояния вида (1).

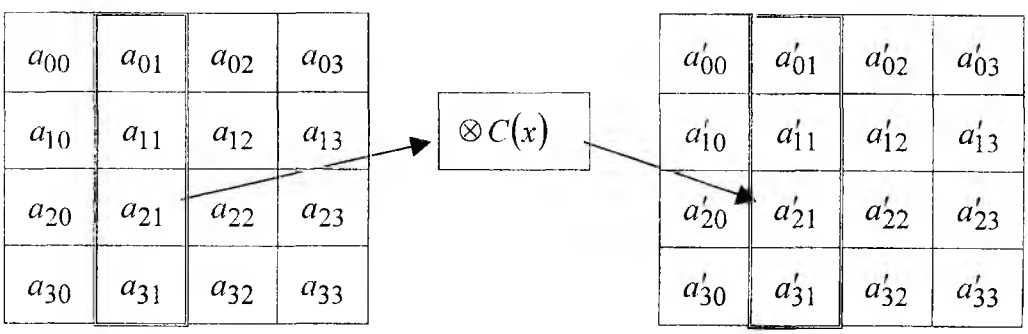


Рис. 4

Преобразование, представленное на рис. 4, может быть выполнено с использованием заранее составленной таблицы.

Криптографическое преобразование в алгоритме RD осуществляется посредством сложения по модулю 2 A_i состояния и k_ν ключа ν-го цикла преобразования. В общем виде это преобразование можно представить как:

$$A'_{ij} = A_{ij} \oplus K_{\nu j} \tag{12}$$

В матричном виде для состояния (1) его можно представить как:

$$\begin{vmatrix} a'_{00} & a'_{01} & a'_{02} & a'_{03} \\ a'_{10} & a'_{11} & a'_{12} & a'_{13} \\ a'_{20} & a'_{21} & a'_{22} & a'_{23} \\ a'_{30} & a'_{31} & a'_{32} & a'_{33} \end{vmatrix} = \begin{vmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{vmatrix} \oplus \begin{vmatrix} k_{00} & k_{01} & k_{02} & k_{03} \\ k_{10} & k_{11} & k_{12} & k_{13} \\ k_{20} & k_{21} & k_{22} & k_{23} \\ k_{30} & k_{31} & k_{32} & k_{33} \end{vmatrix} \tag{13}$$

Таким образом, зашифрование производится посредством побитного сложения бит A_i состояния и битов k_ν развернутого ключа.

После окончания процедуры зашифрования блок-криптограмма C_i представляет собой состояние, одномерный массив байтов которого формируется считыванием столбцов матрицы $C_{0,0} = a'_{00}, C_{1,0} = a'_{10}, C_{3,0} = a'_{30}, C_{0,1} = a'_{01}, C_{0,2} = a'_{02}, \ldots$.

4. Алгоритмы выработки цикловых ключей

В алгоритме RD применяется принцип разворачивания цикловых ключей $k_ц$. Сущность его заключается в том, что необходимое число цикловых ключей $k_ц$ вырабатывается из исходного ключа $K_и$. Как исходные, так и цикловые ключи представляются в виде одномерных массивов, например, $k_{00}, k_{10}, k_{20}, k_{30}, k_{01}, k_{11}, \ldots$ и т.д., причем, преобразование в одномерный массив производится последовательным считыванием по столбцам матричного представления ключа. Исходный ключ может иметь длину 128, 192 или 256 бит. Он может быть представлен соответственно 4(16), 6(24) и 8(32) столбцами (байтами). Развернутые цикловые ключи всегда имеют длину, равную длине информационного блока (13).

Для выработки цикловых ключей в алгоритме RD рекомендуется применять два алгоритма: первый, если длина исходного ключа $n_{k_и} \le 6$, и второй, если $n_k = 8$. Значение $n_{k_и}$ указано в числе столбцов (32 битных или 4 байтовых слов). Рекомендуемое число циклов преобразования как функция длины информационного блока $l_м$ и длины исходного ключа $k_и$ приведено в табл. 1.

Алгоритм 1. Исходный ключ представляет собой массив байтов размерности $4n_{k_и}$, $n_{k_и} = 4$ *или 6*. Расширенный массив развернутых ключей представляет собой массив размерности $n_м(n_ц + 1)$ 32-битных слов или $4n_м(n_ц + 1)$ байтов, где $n_м$ - число столбцов информационного блока (может быть 4, 6 или 8).

Исходный ключ задается массивом байтов $k_и$ размера $4n_{k_и}$, а развернутый – $k_р$ размера $n_м(n_ц + 1)$. Вначале исходный ключ переписывается в качестве значений первых байтов (слов) разворачиваемого. Затем производится развертывание остальных цикловых ключей.

Алгоритм развертывания ключа на языке псевдо-Си имеет вид ($n_r = n_ц$):

```
for (i = nk; i<nм×(nr+1); i++)
{
   temp = Kp[i-1];
   if (i%nk == 0)
   temp = SubByte(RotByte(temp)⊕Rcon[i/nk]);
   Kp[i] = K[i-nk]⊕temp;
}
```

В алгоритме функция SubByte($K_р$) формирует 4-байтовое слово, каждый байт которого в свою очередь формируется применением табличного преобразования ByteSub, задаваемого правилами (5)

и (7). Функция RotByte(K_p) производит циклический сдвиг входных байтов влево, например, RotByte(a,b,c,d)=(b,c,d,a).

Таким образом, при развертывании исходного ключа и формировании цикловых ключей в качестве первых n_k слов записывается исходный ключ. Каждое последующее слово $K_p(i)$ формируется на основе сложения по модулю 2 $K_p[i-n_k]$ слова и $K_p[i-1]$ слова. При этом для слов $i \equiv 0 (\mathrm{mod}\ n_k)$, т.е. кратных длине исходного ключа n_k, $K_p[i-1]$ слово преобразуется с использованием функций SubByte, RotByte, а также складывается с константой Rcon. Циклические константы не зависят от n_k и задаются как Rcon[i]=(RC[i],'00','00','00') (14), а R[i] вычисляется с использованием образующего элемента x='02', причем

$$RC[1]='01',\ a\ RC[i]=x\cdot[RC(i-1)]=x^{i-1}. \qquad (15)$$

Алгоритм 2, приведенный ниже, рекомендуется использовать при $n_k > 6$ $(n_k = 8)$.

Алгоритм 2

```
for (i = nk; i<nм×(nr+1); i++)
{
   temp = Kp[i-1];
   if (i%nk == 0)
   temp = SubByte(RotByte(temp)⊕Rcon[i/nk]);
   else if (i%nk == 4)
   temp = SubByte(temp);
   Kp[i] = K[i-nk]⊕temp;
}
```

Подчеркнем, что в алгоритме 2, как и в алгоритме 1, исходный ключ записывается в качестве первых слов расширенного ключа. Кроме того, в случае наличия ограничений на память, цикловые ключи могут формироваться динамически в каждом цикле соответственно.

Выбор цикловых ключей из массива Kp производится в каждом из циклов последовательно. При этом каждый раз выбирается цикловый ключ $l_ц$ длиной, равной длине информационного блока $l_м$.

5. Обратные преобразования (расшифрование)

Алгоритм RD в явном виде не является симметричным. Поэтому расшифрование должно производиться в обратном порядке. На рис. 5 приведена структурная схема расшифрования C_i блока.

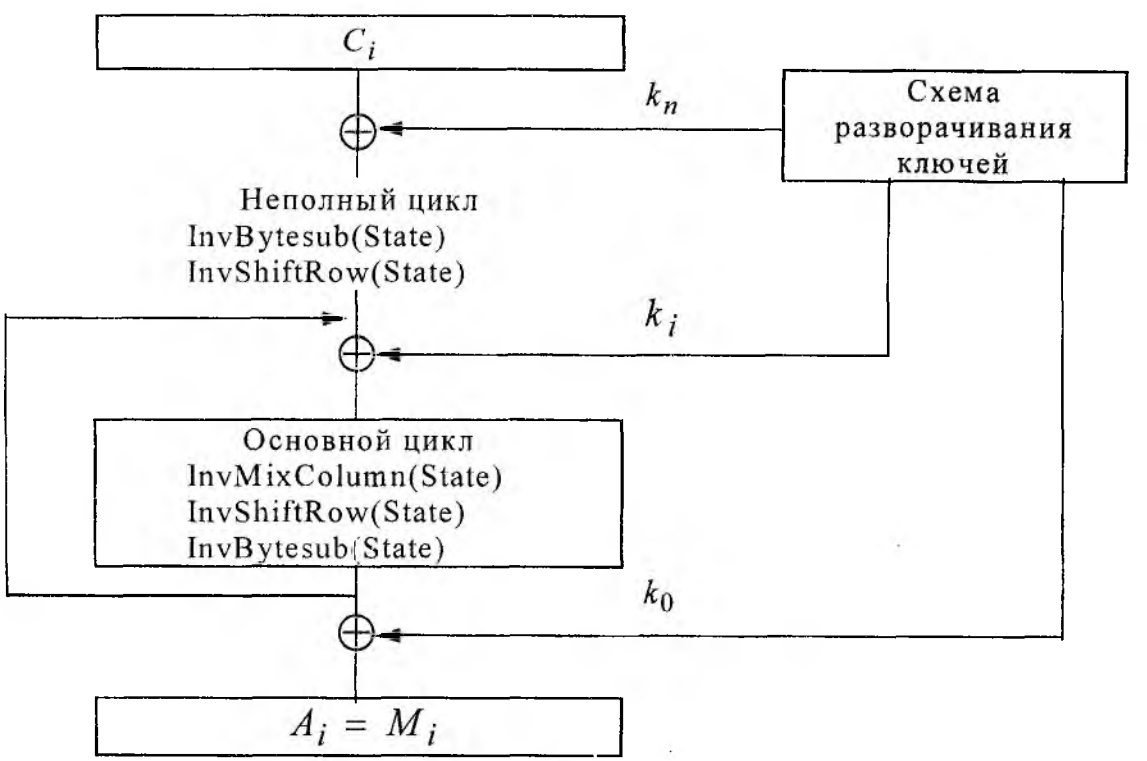


Рис. 5

В алгоритме используются инверсные преобразования Inv, отличительной особенностью которых является то, что при их выполнении используются другие таблицы преобразований.

На рис. 6 приведен процесс, поясняющий выбор цикловых ключей из развернутого ключа, при длине циклового ключа равной 256 бит.

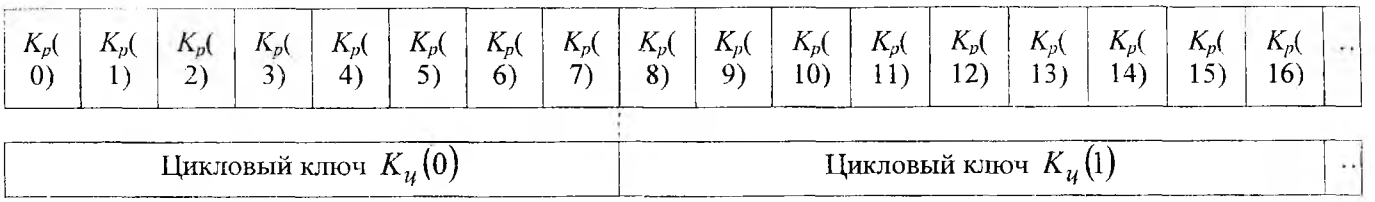


Рис. 6

Как показали исследования, на применяемые в RD исходные ключи не накладывается никаких ограничений. Поэтому можно предположить, что слабых ключей для этого алгоритма не существует. Следует заметить, что авторы RD алгоритма не представили обоснования алгоритмов разворачивания исходного ключа в цикловые. На наш взгляд, в качестве алгоритма разворачивания может использоваться любой алгоритм, который формирует $n_ц$ независимых случайных и равновероятных цикловых ключей.

В алгоритме зашифрования (рис. 1) в качестве первого используется нелинейное преобразование Bytesub(замена байт). Затем используются преобразования «сдвиг строчек» и «преобразование столбцов». При обратном преобразовании, т.е. расшифровании, порядок преобразований в цикле должен быть обратным. Вначале выполняется обратное преобразование в столбцах MixColumb, затем – обратное преобразование сдвиг строчек, последним – обратное нелинейное преобразование Bytesub. Анализ показывает, что для того, чтобы обратное преобразование (расшифрование) могло быть выполнено в той же последовательности, что и прямое (т.е. зашифрование), должна существовать таблица замены байт, реализующая сразу три шага – Bytesub, ShiftRow и MixColumb, а также обратная ей. Как удалось установить, это не так. Поэтому обратное преобразование в циклах должно выполняться в другом порядке – обратном. Следовательно, при расшифрвании нельзя использовать алгоритм, представленный на рис.1, а только алгоритм, представленный на рис.5. При этом таблицы обратных преобразований MixColumb, ShiftRow и Bytesub также должны быть другими, но однозначно связанными с прямыми.

Анализ алгоритма RD показывает, что порядок выполнения преобразований ShiftRow и Bytesub в нем может быть любой. Это объясняется тем, что преобразование ShiftRow только перемещает байты и не влияет на содержание (значение) байта. Преобразование Bytesub обеспечивает замену только отдельных байт независимо от их положения.

Кроме того, последовательность преобразования

```
AddRoundKey(State, RoundKey);      // Сложение состояния с цикловым ключом
InvMixColumn(State);               // Обратное перемешивание в столбцах
                                   может быть заменена на последовательность
InvMixColumn(State);               // Обратное перемешивание в столбцах
AddRoundKey(State, InvRoundKey);   //Сложение состояния с инверсным цикловым
                                    ключом.
```

Инверсный цикловый ключ можно получить посредством применения инверсного перемешивания в столбцах к соответствующему цикловому ключу RoundKey. Это объясняется тем, что названное A-преобразование линейное и A(x+y)=A(x)+A(y).

В целом обратное преобразование криптоалгоритма RD может быть задано следующим образом:

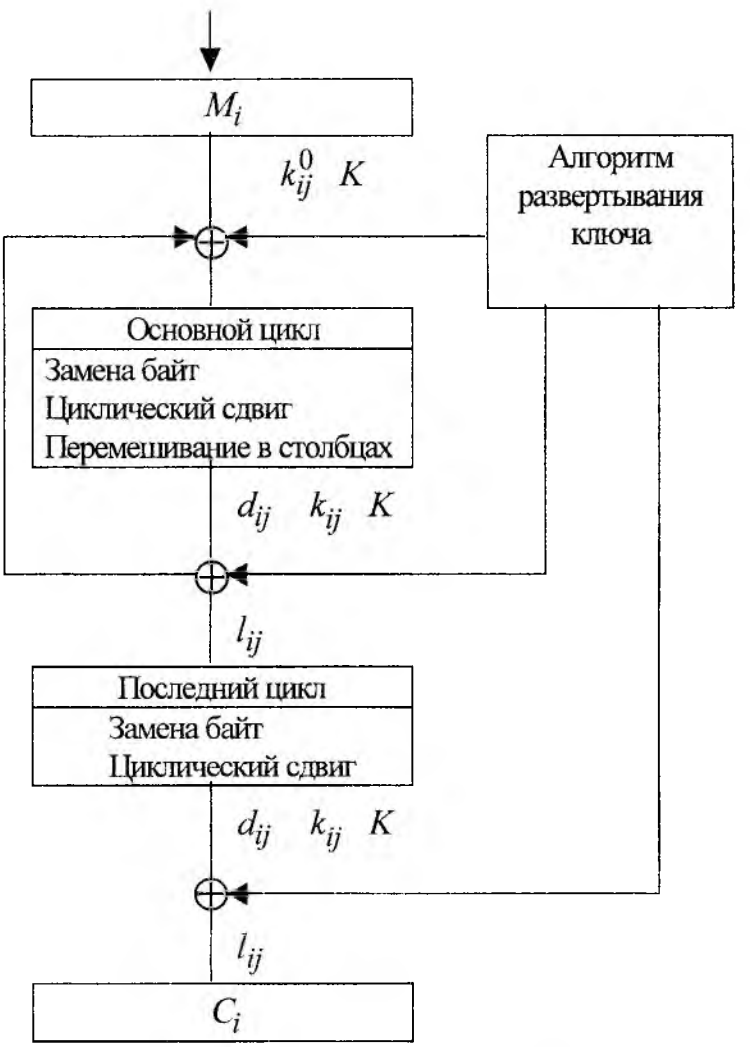
```
I_RD(State, CipherKey)                    // I_RD(состояние, исходный ключ шифрования)
{
I_KeyExprension(CipherKey, I_ExpandedKey); // I_Расширение ключа (исходный ключ,
                                           развернутые цикловые ключи)
```

```
AddRoundKey(State, I_ExpandedKey+nc*nr);        // Сложение по модулю 2 состояния с цикловым ключом

For (i = nr-1); I>0;i--)
Round(State, I_ExpandedKey+nc*i);               // Выполнение nr циклов
FinalRound(State, I_ExpandedKey);               // Выполнение последующего цикла
}
```

6. Организация табличных преобразований

При практической реализации алгоритма шифрования одним из наиболее важных требований является требование минимизации сложности зашифрования/расшифрования. На рис. 7 приведена схема алгоритма зашифрования с более подробным, чем на рис. 1, описанием преобразований. Здесь также рассмотрим циклы преобразования, основное внимание уделив на организацию идентичных операций в циклах.



$$a_{ij} = M_{ij} \oplus K_{ij}^0$$
$$b_{ij} = S\lfloor a_{ij} \rfloor$$
$$C_{ij} = b_{ij}'$$
$$d_{ij} = c_{ij} \cdot C$$
$$l_{ij} = d_{ij} \oplus K_{ij}$$
$$b_{ij} = S\lfloor a_{ij} \rfloor$$
$$C_{ij} = b_{ij}'$$
$$l_{ij} = C_{ij} \oplus K_{ij}^n$$
$$C_j = l_{ij}$$

Рис. 7

Рассмотрим основной цикл в обратном порядке преобразования, опираясь на уже приведенный выше материал. Байты состояния будем задавать a_{ij} (байт i-й строки и j-го столбца). Соотношение $l_{ij} = d_{ij} \oplus K_{ij}$ представим в развернутом виде – столбцами состояний, начиная с l_{ij} :

$$\begin{bmatrix} l_{0j} \\ l_{1j} \\ l_{2j} \\ l_{3j} \end{bmatrix} = \begin{bmatrix} d_{0j} \\ d_{1j} \\ d_{2j} \\ d_{3j} \end{bmatrix} \oplus \begin{bmatrix} k_{0j} \\ k_{1j} \\ k_{2j} \\ k_{3j} \end{bmatrix}, j = \overline{0, n_c - 1}. \tag{16}$$

Далее d_{ij} состояние представим как результат выполнения преобразования «перемешивание в столбцах», причем, входным является состояние C_{ij}:

$$\begin{bmatrix} d_{0j} \\ d_{1j} \\ d_{2j} \\ d_{3j} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \cdot \begin{bmatrix} C_{0j} \\ C_{1j} \\ C_{2j} \\ C_{3j} \end{bmatrix}, \quad j = \overline{0, n_c - 1}. \tag{17}$$

В (17) матрица перемешивания представлена константой C. Циклический сдвиг строк, т.е. состояние C_{ij}, положив, что входом циклического преобразования есть состояние a_{ij}, представим в виде:

$$\begin{bmatrix} C_{0j} \\ C_{1j} \\ C_{2j} \\ C_{3j} \end{bmatrix} = \begin{bmatrix} b_{0,j} \\ b_{1,j-C_1} \\ b_{2,j-C_2} \\ b_{3,j-C_3} \end{bmatrix}, \quad j = \overline{0, n_c - 1}. \tag{18}$$

Константы C_1, C_2 и C_3 выбираются из табл. 1.

Табличное преобразование «замена байт» представим в виде последовательной замены a_{ij} байт по таблице подстановки S, при этом выходом является состояние b_{ij}, причем:

$$b_{ij} = S[a_{ij}], \quad i = \overline{0,3}, \; N = \overline{0, n_c - 1}. \tag{19}$$

Преобразование (19) представлено одной таблицей подстановки, в которой учитываются два преобразования – замена a_{ij} байта на обратный элемент и аффинное преобразование вида (7). Возможность и алгоритм построения S-таблицы рассмотрим отдельно.

Последовательность преобразований (16) – (19) можно представить в виде одного, заменив состояние d_{ij} :

$$\begin{bmatrix} l_{0j} \\ l_{1j} \\ l_{2j} \\ l_{3j} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \cdot \begin{bmatrix} S[a_{0j}] \\ S[a_{1j-C1}] \\ S[a_{2j-C2}] \\ S[a_{3j-C3}] \end{bmatrix} \oplus \begin{bmatrix} k_{0j} \\ k_{1j} \\ k_{2j} \\ k_{3j} \end{bmatrix}. \tag{20}$$

Далее, учитывая то, что умножение производится по правилу «строка на столбец» (20) представим в виде:

$$\begin{bmatrix} l_{0j} \\ l_{1j} \\ l_{2j} \\ l_{3j} \end{bmatrix} = \begin{bmatrix} S[a_{0j}] \begin{bmatrix} 02 \\ 01 \\ 01 \\ 03 \end{bmatrix} \oplus S[a_{1j-C1}] \begin{bmatrix} 03 \\ 02 \\ 01 \\ 01 \end{bmatrix} \oplus S[a_{2j-C2}] \begin{bmatrix} 01 \\ 03 \\ 02 \\ 01 \end{bmatrix} \oplus S[a_{3j-C3}] \begin{bmatrix} 01 \\ 01 \\ 03 \\ 02 \end{bmatrix} \end{bmatrix} \oplus \begin{bmatrix} k_{0j} \\ k_{1j} \\ k_{2j} \\ k_{3j} \end{bmatrix}. \tag{21}$$

В выражении введены скобки [ ], которые указывают на необходимость выполнения операций сначала в квадратных скобках. Вычисления значения $[l_{ij}]$ можно свести к построению четырех таблиц.

$$T_0[a]=\begin{bmatrix} S[a]\cdot 02 \\ S[a] \\ S[a] \\ S[a]\cdot 03 \end{bmatrix}; T_1[a]=\begin{bmatrix} S[a]\cdot 03 \\ S[a]\cdot 02 \\ S[a] \\ S[a] \end{bmatrix}; T_2[a]=\begin{bmatrix} S[a] \\ S[a]\cdot 03 \\ S[a]\cdot 02 \\ S[a] \end{bmatrix}; T_3[a]=\begin{bmatrix} S[a] \\ S[a] \\ S[a]\cdot 03 \\ S[a]\cdot 02 \end{bmatrix}. \quad (22)$$

В выражениях (22) умножения выполняются побайтно и по модулю $m(x)=x^8+x^4+x^3+x+1$. Причем, вместо операции замены конкретного значения байта записана полная таблица замены байт. Это позволяет выполнить операцию умножения таблично. Всего нужно 4 таблицы T_v по 4 таблицы замены в каждой из них. Общий объем 4 таблицы. $N_T=4\cdot 4\cdot 2^8=2^{12}=4096$ *байт*.

Преобразование (21) с учетом (22) можно представить:

$$l_j=T_0(a_{0j})\oplus T_1(a_{1j-C1})\oplus T_2(a_{2j-C2})\oplus T_3(a_{3j-C3})\oplus K_j,\ j=\overline{0,n_c-1}. \quad (23)$$

Операцию нахождения T_v будем называть табличным поиском. Следует учитывать, что каждая из таблиц задается 4-х байтовым числом, результатом l_j также будет 4-х байтовое значение l_j. Всего за цикл вычисления необходимо выполнить 4 операции сложения по модулю 2 и 4 поиска по таблицам T_0, T_1, T_2 и T_3.

Анализ показывает, что таблицы T_1, T_2 и T_3 отличаются от T_0 только по циклическим сдвигам на C_1, C_2 и C_3 байт. Поэтому вычисления (23) можно свести к рекуррентному определению согласно выражению:

$$l_j=k_j\oplus T_0\lfloor b_{0j}\rfloor\oplus R_b\left(T_0\lfloor b_{1,j-C1}\rfloor\oplus R_b\left(T_0\lfloor b_{2,j-C2}\rfloor\oplus R_b\left(T_0\lfloor b_{3,j-C3}\rfloor\right)\right)\right). \quad (24)$$

В (24) используются 4-х байтовые слова, полученные после циклического сдвига на величины C_1, C_2 и C_3 соответственно b_{1j}, b_{2j} и b_{3j}.

При этом вычисление (24) по сравнению с (23) требует трех дополнительных операций циклического сдвига. Но в этом случае достаточно одной таблицы преобразования длиной $N_T=4\cdot 2^8=2^{10}=1024$ *байт*.

7. Анализ стойкости криптоалгоритма

В процессе всех этапов общественного обсуждения и анализа алгоритма RD, а также исследования и тестирования его со стороны NIST США, особое внимание уделялось анализу и оценке по критериям реальной криптозащищенности, статистической безопасности и надежности математической базы криптоалгоритма [1-3]. На сегодняшний день не известно, а возможно и не существует ни одного метода криптоанализа, сложность реализации которого была бы меньше, чем методы, основанные на «грубой силе» (переборе). Тем не менее, RD алгоритм оказался защищенным от атак, реализованных на основе:

- дифференциального криптоанализа;
- поиска наилучшей дифференциальной характеристики;
- расширения для дифференциального криптоанализа;
- линейного криптоанализа;
- вторжения с использованием связанных ключей;
- вторжения с частичным угадыванием ключа;
- вторжения на основе обработки сбоев;
- интерполяционного (алгебраического) вторжения;
- поиска лазеек.

Основу криптозащищенности в алгоритме составляют табличные и криптографические преобразования на множестве циклов. В качестве табличных преобразований используются преобразования (замена) байт, преобразование (сдвиг) строк и преобразование (перемешивание) в столбцах. При этом наибольший вклад в криптостойкость вносят преобразования вида (5) и (6): преобразование (5) ввиду нелинейности обеспечивает защиту от линейных и дифференциально-подобных атак, а (6) – от раз-

личных алгебраических вторжений. Оба преобразования, как следует из (23), обеспечивают табличную нелинейную многократную замену байт. Многократное выполнение замены байт совместно с преобразованием строк и столбцов, а также зашифрование с использованием цикловых ключей, дают определенную уверенность в стойкости криптоалгоритма RD.

Что касается поиска лазеек, то пока в самом алгоритме их обнаружить не удалось. На наш взгляд, лазейки могут быть встроены в алгоритм формирования исходных ключей или в алгоритм развертывания ключей, т.е. формирования цикловых ключей.

Статистическая безопасность проверялась по методике, изложенной в [6]. В табл. 3 в качестве примера приведены статистики (не лучшие) связанности C_i и M_i (функция F_1), а также зависимости C_i и C_k (функция F_2) при различных фактических значениях данных M_i и исходных ключей k_u. Причем, $\overline{m}(F_r)$ - среднее значение, а $m^2(F_r)$ - дисперсия, вычисляемые в метрике Хэмминга (по числу совпадающий битов).

Таблица 3

Варианты	$m(F_1)$	$m^2(F_1)$	$m(F_2)$	$m^2(F_2)$	min F_1/max F_1	min F_1/max F_1
1. M_i – случайные, k_u - случайный постоянный	64,16	32,25	63,95	32,12	44/90	44/85
2. M_i – случайные, k_u - единичный постоянный	63,79	31,86	64,03	31,07	39/90	44/85
3. M_i – единичное, k_u - случайный постоянный	63,99	32,40	63,97	32,30	39/90	42/86
4. M_i – нулевые, k_u - случайный изменяемый	63,98	33,47	64,05	31,92	39/90	42/86
5. M_i – случайные, k_u - случайный изменяемый	63,88	31,70	64,00	30,67	39/90	42/86

Анализ данных таблицы позволяет высказать гипотезу о биноминальном законе распределения функций F_1 и F_2, так как биномиального закона распределения математическое ожидание $\overline{m}(F_v)$ и $m^2(F_v)$ соответственно равны 64 и 32. В то же время величины, подчиняющиеся биномиальному закону распределения, являются случайными и равновероятными.

Таким образом, алгоритм RD можно с определенной уверенностью считать статистически безопасным.

Что касается доказательства надежности математического аппарата, то это сложная и отдельная проблема. Ее надо решать, начиная с обоснования терминологии, понятий, критериев, методов и методик и т.д.

8. Анализ алгоритмов (схем) развертывания ключей

К алгоритмам развертывания ключей предъявляются требования по двум критериям:

- случайность, равновероятность и независимость цикловых ключей;
- минимизация сложности развертывания цикловых ключей.

В табл. 4 приведены значения сложности развертывания цикловых ключей для различных длин исходных ключей.

Таблица 4

Длина ключа	128	192	256
Количество тактов	2066	2418	2937

Важной характеристикой, которая существенно влияет на выбор того или иного криптоалгоритма, является сложность (скорость) прямых и обратных криптографических преобразований. При выборе алгоритма эта характеристика может быть решающей. Результаты исследований в этом направлении представлены отдельной статьей.

В процессе анализа AES мы пришли к выводу, что определенные требования должны быть предъявлены и к развернутым ключам. В частности, желательно, чтобы изменение на каждой из позиции исходного ключа бита приводило к равномерному изменению битов в развернутом ключе. Были проведены экспериментальные исследования такого эффекта размножения ошибок для RD. По существу изучался эффект размножения ошибок в развернутом K^p-ключе при изменении каждого бита исходного K_j - ключа. Для всех алгоритмов процедуру разворачивания ключа можно представить в виде:

$$K_j^p = X\left(K_j, C_V\right), \tag{25}$$

где C_V – V-я константа алгоритма развертывания ключа.

Число измененных n_i битов в развернутом ключе можно представить как функцию φ номера измененного бита исходного ключа и констант, т.е.

$$n_i = \varphi\left(i, c_V\right). \tag{26}$$

Таким образом, в эксперименте изменялся i-й бит исходного ключа и определялось количество измененных битов по следующей методике:

1. Формировался K_j исходный ключ, а затем согласно (25) – соответствующий ему развернутый K_j^p ключ.
2. В каждом исходном K_j ключе изменялся i-й бит, $i = \overline{1, l_p}$, а затем, согласно (25), формировался соответствующий ему развернутый K_{j1}^p ключ, где l_p - длина развернутого ключа.
3. Вычислялось расстояние Хэмминга n_i между развернутыми ключами K_j^p для каждой i-й позиции и различных ключей.

На наш взгляд, лучшим алгоритмом развертывания ключа может быть тот, для которого, во-первых, распределение $\overline{n_i}$, $i = \overline{1, l_p}$ подчиняется равномерному закону, а во-вторых, математическое ожидание распределения $\bar{n}$ ближе к $l_p/2$. Физически в первом случае это означает, что изменение каждого бита приводит к лавинному эффекту с равномерным размножением ошибок, а во втором, что изменение одного бита приводит в среднем к изменению половины битов развернутого ключа. В целом при выполнении обоих условий развернутые ключи отличаются в половине бит, поэтому их можно считать независимыми.

На рис. 8 и 9 приведены гистограммы коэффициентов размножения ошибок в развернутом ключе в зависимости от номера позиции измененного бита в исходном ключе для алгоритмов RD и RC6 соответственно (длина ключа 128 бит).

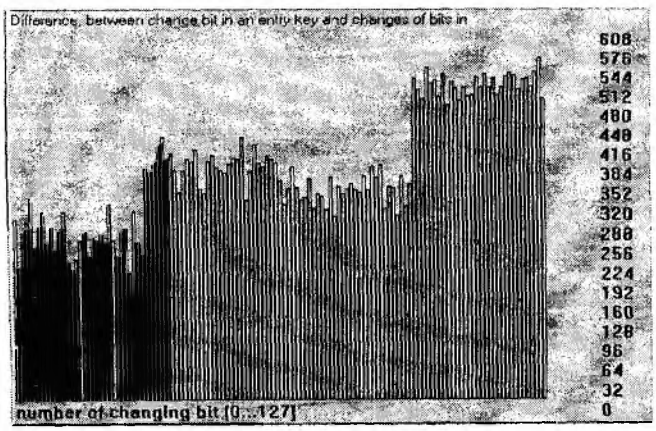


Рис. 8

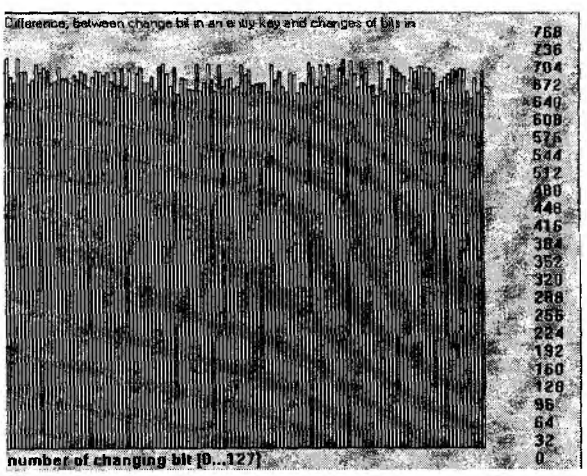


Рис. 9

Анализ данных рис. 8 показывает, что для алгоритма RIJNDAEL коэффициент размножения «ошибки» в зависимости от позиции изменяемого бита является существенно неравномерным. Природа неравномерности объясняется особенностью алгоритма разворачивания ключа.

Как следует из рис. 9, коэффициент размножения ошибок для алгоритма RC6 [3] близок к равномерному и имеет один и тот же характер для различных длин исходных ключей.

На наш взгляд, выявленный факт неравномерности размножения «ошибок» является слабой стороной схемы развертывания ключей, применяемого в RD. Если лазейка в этом алгоритме имеется, то начало ее может начинаться со значительной вероятностью с указанного недостатка.

Из рис. 9 следует, что алгоритм разворачивания ключей RC6 по статистике «размножения» ошибок лучше.

Заключение

В практической криптографии на мировом уровне успешно проведен важный научно-практический эксперимент "общественного" создания и обсуждения претендентов на стандарт симметричного шифрования 21 века. Ему предшествовала по существу неудачная попытка создания стандарта США в виде различных версий Skipjack. Явно понятно, что заказчиком такого эксперимента были США в лице NIST и других структур. Эксперимент закончен успешно. Лучшие криптологи участвовали в эксперименте, дали на обсуждение свои лучшие разработки, методики и знания. Эксперимент позволил специалистам освоить многогранные знания, методы и методики, используемые в передовых странах.

Блоковый криптоалгоритм, вернее его разработчики, обоснованно победили в конкурсе на лучший алгоритм среди кандидатов на стандарт 21 века. При современном уровне знаний алгоритм RD является защищенным от существующих аналитических атак, наименее опасной для него является атака типа "грубая сила". Алгоритм обеспечивает приемлемую скорость зашифрования/расшифрования, может быть реализован на процессорах различной разрядности.

С определенной вероятностью можно предположить, что в нем отсутствуют лазейки. Вместе с тем необходимо отметить недостаточную обоснованность используемой схемы развертывания ключа. По сравнению со схемами других кандидатов схема развертывания цикловых ключей RD уступает им. Следует предположить, что эта схема будет или должна быть усовершенствована или заменена по ряду причин.

Безусловно, что алгоритм RD еще многие годы будет подвергаться анализу и совершенствованию. Выражаем признательность специалистам и всем интересующимся практической криптографией за внимание, уделенное нашей статье.

Список литературы: 1. *Announcing Development of Federal Information Processing Standard for Advanced Encryption Standard.* Federal Register. 1997. Vol. 62, №1. Pp. 93-94. 2. *Announcing Reguest for Candidate Algorithm Nomination for Advanced Encryption Standard (AES).* Federal Register. 1997. Vol. 62, №177. Pp. 48051-48058. 3. *М.Ф. Бондаренко, И.Д. Горбенко, А.В. Потий.* Улучшеный стандарт симметричного шифрования XXI века: концепция создания и свойства кандидатов. Радиотехника. 2000. Вып. 114. С.5-15. 4. *Status report on the 2-nd round of the Development of the Advanced Encryption Standard* / AES home page. Http://www.nist.gov/aes. 5. *Status report on the 3-th round of the Development of the Advanced Encryption Standard* / AES home page. http://www.nist.gov/aes. 6. *Joan Daemen, Vincent Rijmen.* The Rijndael Block Cipher. AES Proposal: Rijndael, Document version 2. 3.09.99.

Харьковский государственный технический университет радиоэлектроники
Служба безопасности Украины

Поступила в редколлегию 18.01.2001

УДК 681.3.06:519.248.681

И. Д. ГОРБЕНКО, д-р техн. наук, Д. А. ЧЕКАЛИН

СВОЙСТВА И ВОЗМОЖНОСТИ ОПТИМИЗАЦИИ КРИПТОГРАФИЧЕСКИХ ПРЕОБРАЗОВАНИЙ В AES – RIJNDAEL

Введение

В конце 90-х годов в мире, в кругу специалистов-криптологов, сформулирована и решена задача создания стандарта блочного симметричного шифрования XXI века. В течение трех лет проводились три конгресса по прикладной криптологии, на которых из 15 алгоритмов к сегодняшнему дню выбрали Rijndael [1].

В соответствии с минимальными и общими требованиями к блочным симметричным алгоритмам проведены исследования, направленные на минимизацию вычислительной сложности уже существующих программных реализаций алгоритма Rijndael. Среди существующих реализаций нами была выбрана программа, написанная Брайаном Гладманом [1], обеспечивающая, по мнению авторов указанной реализации, максимальную скорость прямого и обратного криптографических преобразований, и, соответственно, имеющая минимальную вычислительную сложность.

Целью настоящей статьи является рассмотрение возможностей усовершенствования названной программной реализации с целью повышения скорости прямых и обратных преобразований.

1. Сущность цикловых табличных и криптографических преобразований

Rijndael является цикловым блочным симметричным криптоалгоритмом. Длины ключа и блока могут иметь независимо друг от друга значения 128, 192 и 256 бит. Количество циклов в алгоритме зависит от длин ключа и блока. Их зависимость отображается в табл. 1. Длины ключа и блока берутся деленными на 32.

Таблица 1

N_r	$N_b = 4$	$N_b = 6$	$N_b = 8$
$N_k = 4$	10	12	14
$N_k = 6$	12	12	14
$N_k = 8$	14	14	14

Обозначим промежуточный результат шифрования, согласно [1], как **State** (состояние). Состояние можно представить в виде прямоугольного массива байтов. Этот массив имеет 4 строки, а число столбцов обозначено как N_b и равно длине блока, деленной на 32. Цикловое преобразование состоит из четырех различных преобразований: трех табличных и одного криптографического (сложение с ключом). На псевдо-Си это выглядит следующим образом:

```
Round (State, RoundKey)
 {
ByteSub(State):                  // замена байт
ShiftRow(State);                 // сдвиг строк
MixColumn(State);                // замешивание столбцов
AddRoundKey(State, RoundKey);    // добавление циклового ключа
}
```

Программная реализация позволяет свести три табличных преобразования к одному. Обозначим *e* как результат циклового преобразования (State). Для преобразований MixColumn и сложения с ключом мы имеем:

$$\begin{bmatrix} e_{0,j} \\ e_{1,j} \\ e_{2,j} \\ e_{3,j} \end{bmatrix} = \begin{bmatrix} d_{0,j} \\ d_{1,j} \\ d_{2,j} \\ d_{3,j} \end{bmatrix} \oplus \begin{bmatrix} k_{0,j} \\ k_{1,j} \\ k_{2,j} \\ k_{3,j} \end{bmatrix} u \begin{bmatrix} d_{0,j} \\ d_{1,j} \\ d_{2,j} \\ d_{3,j} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} c_{0,j} \\ c_{1,j} \\ c_{2,j} \\ c_{3,j} \end{bmatrix} \quad (1)$$

Для преобразований ShiftRow и ByteSub мы имеем:

$$\begin{bmatrix} c_{0,j} \\ c_{1,j} \\ c_{2,j} \\ c_{3,j} \end{bmatrix} = \begin{bmatrix} b_{0,j} \\ b_{1,j-C1} \\ b_{2,j-C2} \\ b_{3,j-C3} \end{bmatrix} u\ b_{1,j} = S[a_{i,j}] \tag{2}$$

Подставив выражение (2) в (1) получим (3):

$$\begin{bmatrix} e_{0,j} \\ e_{1,j} \\ e_{2,j} \\ e_{3,j} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} S[a_{0,j}] \\ S[a_{1,j-C1}] \\ S[a_{2,j-C2}] \\ S[a_{3,j-C3}] \end{bmatrix} \oplus \begin{bmatrix} k_{0,j} \\ k_{1,j} \\ k_{2,j} \\ k_{3,j} \end{bmatrix} \tag{3}$$

Перемножение матриц может быть представлено как линейная комбинация векторов:

$$\begin{bmatrix} e_{0,j} \\ e_{1,j} \\ e_{2,j} \\ e_{3,j} \end{bmatrix} = S[a_{0,j}] \begin{bmatrix} 02 \\ 01 \\ 01 \\ 03 \end{bmatrix} \oplus S[a_{1,j-C1}] \begin{bmatrix} 03 \\ 02 \\ 01 \\ 01 \end{bmatrix} \oplus S[a_{2,j-C2}] \begin{bmatrix} 01 \\ 03 \\ 02 \\ 01 \end{bmatrix} \oplus S[a_{3,j-C3}] \begin{bmatrix} 01 \\ 01 \\ 03 \\ 02 \end{bmatrix} \oplus \begin{bmatrix} k_{0,j} \\ k_{1,j} \\ k_{2,j} \\ k_{3,j} \end{bmatrix}$$

Определим четыре таблицы $T_0 - T_3$

$$T_0[a] = \begin{bmatrix} S[a] \cdot 02 \\ S[a] \\ S[a] \\ S[a] \cdot 03 \end{bmatrix} \quad T_1[a] = \begin{bmatrix} S[a] \cdot 03 \\ S[a] \cdot 02 \\ S[a] \\ S[a] \end{bmatrix} \quad T_2[a] = \begin{bmatrix} S[a] \\ S[a] \cdot 03 \\ S[a] \cdot 02 \\ S[a] \end{bmatrix} \quad T_3[a] = \begin{bmatrix} S[a] \\ S[a] \\ S[a] \cdot 03 \\ S[a] \cdot 02 \end{bmatrix}$$

Таким образом, мы получили 4 таблицы, содержащие по 256 4-х байтных слов и занимающие около 4 Кб. Полностью таблицы $T_0 - T_3$ приведены ниже в С-транскрипции. Используя эти таблицы. представим цикловое преобразование как

$$e_j = T_0[a_{0,j}] \oplus T_1[a_{1,j-C1}] \oplus T_2[a_{2,j-C2}] \oplus T_3[a_{3,j-C3}] \oplus k_j$$

Учитывая достаточную сложность получения таблиц и необходимость для разработчиков «эталона» таблиц преобразований, а также то, что алгоритм RIJNDAEL рекомендован в качестве стандарта XXI века, считаем необходимым опубликование их в настоящей статье.

```
u4byte ft_tab[4][256]={{
A56363C6, 847C7CF8, 997777EE, 8D7B7BF6,  DF2F2FF, BD6B6BD6, B16F6FDE, 54C5C591,
50303060,  3010102, A96767CE, 7D2B2B56, 19FEFEE7, 62D7D7B5, E6ABAB4D, 9A7676EC,
45CACA8F, 9D82821F, 40C9C989, 877D7DFA, 15FAFAEF, EB5959B2, C947478E,  BF0F0FB,
ECADAD41, 67D4D4B3, FDA2A25F, EAAFAF45, BF9C9C23, F7A4A453, 967272E4, 5BC0C09B,
C2B7B775, 1CFDFDE1, AE93933D, 6A26264C, 5A36366C, 413F3F7E,  2F7F7F5, 4FCCCC83,
5C343468, F4A5A551, 34E5E5D1,  8F1F1F9, 937171E2, 73D8D8AB, 53313162, 3F15152A,
 C040408, 52C7C795, 65232346, 5EC3C39D, 28181830, A1969637,  F05050A, B59A9A2F,
 907070E, 36121224, 9B80801B, 3DE2E2DF, 26EBEBCD, 6927274E, CDB2B27F, 9F7575EA,
1B090912, 9E83831D, 742C2C58, 2E1A1A34, 2D1B1B36, B26E6EDC, EE5A5AB4, FBA0A05B,
F65252A4, 4D3B3B76, 61D6D6B7, CEB3B37D, 7B292952, 3EE3E3DD, 712F2F5E, 97848413,
F55353A6, 68D1D1B9,        0, 2CEDEDC1, 60202040, 1FFCFCE3, C8B1B179, ED5B5BB6,
BE6A6AD4, 46CBCB8D, D9BEBE67, 4B393972, DE4A4A94, D44C4C98, E85858B0, 4ACFCF85,
6BD0D0BB, 2AEFEFC5, E5AAAA4F, 16FBFBED, C5434386, D74D4D9A, 55333366, 94858511,
CF45458A, 10F9F9E9,  6020204, 817F7FFE, F05050A0, 443C3C78, BA9F9F25, E3A8A84B,
F35151A2, FEA3A35D, C0404080, 8A8F8F05, AD92923F, BC9D9D21, 48383870,  4F5F5F1,
DFBCBC63, C1B6B677, 75DADAAF, 63212142, 30101020, 1AFFFFE5,  EF3F3FD, 6DD2D2BF,
4CCDCD81, 140C0C18, 35131326, 2FECECC3, E15F5FBE, A2979735, CC444488, 3917172E,
```

```
57C4C493, F2A7A755, 827E7EFC, 473D3D7A, AC6464C8, E75D5DBA, 2B191932, 957373E6,
A06060C0, 98818119, D14F4F9E, 7FDCDCA3, 66222244, 7E2A2A54, AB90903B, 8388880B,
CA46468C, 29EEEEC7, D3B8B86B, 3C141428, 79DEDEA7, E25E5EBC, 1D0B0B16, 76DBDBAD,
3BE0E0DB, 56323264, 4E3A3A74, 1E0A0A14, DB494992,  A06060C, 6C242448, E45C5CB8,
5DC2C29F, 6ED3D3BD, EFACAC43, A66262C4, A8919139, A4959531, 37E4E4D3, 8B7979F2,
32E7E7D5, 43C8C88B, 5937376E, B76D6DDA, 8C8D8D01, 64D5D5B1, D24E4E9C, E0A9A949,
B46C6CD8, FA5656AC,  7F4F4F3, 25EAEACF, AF6565CA, 8E7A7AF4, E9AEAE47, 18080810,
D5BABA6F, 887878F0, 6F25254A, 722E2E5C, 241C1C38, F1A6A657, C7B4B473, 51C6C697,
23E8E8CB, 7CDDDDA1, 9C7474E8, 211F1F3E, DD4B4B96, DCBDBD61, 868B8B0D, 858A8A0F,
907070E0, 423E3E7C, C4B5B571, AA6666CC, D8484890,  5030306,  1F6F6F7, 120E0E1C,
A36161C2, 5F35356A, F95757AE, D0B9B969, 91868617, 58C1C199, 271D1D3A, B99E9E27,
38E1E1D9, 13F8F8EB, B398982B, 33111122, BB6969D2, 70D9D9A9, 898E8E07, A7949433,
B69B9B2D, 221E1E3C, 92878715, 20E9E9C9, 49CECE87, FF5555AA, 78282850, 7ADFDFA5,
8F8C8C03, F8A1A159, 80898909, 170D0D1A, DABFBF65, 31E6E6D7, C6424284, B86868D0,
C3414182, B0999929, 772D2D5A, 110F0F1E, CBB0B07B, FC5454A8, D6BBBB6D, 3A16162C
}{
6363C6A5, 7C7CF884, 7777EE99, 7B7BF68D, F2F2FF0D, 6B6BD6BD, 6F6FDEB1, C5C59154,
30306050,  1010203, 6767CEA9, 2B2B567D, FEFEE719, D7D7B562, ABAB4DE6, 7676EC9A,
CACA8F45, 82821F9D, C9C98940, 7D7DFA87, FAFAEF15, 5959B2EB, 47478EC9, F0F0FB0B,
ADAD41EC, D4D4B367, A2A25FFD, AFAF45EA, 9C9C23BF, A4A453F7, 7272E496, C0C09B5B,
B7B775C2, FDFDE11C, 93933DAE, 26264C6A, 36366C5A, 3F3F7E41, F7F7F502, CCCC834F,
3434685C, A5A551F4, E5E5D134, F1F1F908, 7171E293, D8D8AB73, 31316253, 15152A3F,
 404080C, C7C79552, 23234665, C3C39D5E, 18183028, 969637A1,  5050A0F, 9A9A2FB5,
 7070E09, 12122436, 80801B9B, E2E2DF3D, EBEBCD26, 27274E69, B2B27FCD, 7575EA9F,
 909121B, 83831D9E, 2C2C5874, 1A1A342E, 1B1B362D, 6E6EDCB2, 5A5AB4EE, A0A05BFB,
5252A4F6, 3B3B764D, D6D6B761, B3B37DCE, 2929527B, E3E3DD3E, 2F2F5E71, 84841397,
5353A6F5, D1D1B968,        0, EDEDC12C, 20204060, FCFCE31F, B1B179C8, 5B5BB6ED,
6A6AD4BE, CBCB8D46, BEBE67D9, 3939724B, 4A4A94DE, 4C4C98D4, 5858B0E8, CFCF854A,
D0D0BB6B, EFEFC52A, AAAA4FE5, FBFBED16, 434386C5, 4D4D9AD7, 33336655, 85851194,
45458ACF, F9F9E910,  2020406, 7F7FFE81, 5050A0F0, 3C3C7844, 9F9F25BA, A8A84BE3,
5151A2F3, A3A35DFE, 404080C0, 8F8F058A, 92923FAD, 9D9D21BC, 38387048, F5F5F104,
BCBC63DF, B6B677C1, DADAAF75, 21214263, 10102030, FFFFE51A, F3F3FD0E, D2D2BF6D,
CDCD814C,  C0C1814, 13132635, ECECC32F, 5F5FBEE1, 979735A2, 444488CC, 17172E39,
C4C49357, A7A755F2, 7E7EFC82, 3D3D7A47, 6464C8AC, 5D5DBAE7, 1919322B, 7373E695,
6060C0A0, 81811998, 4F4F9ED1, DCDCA37F, 22224466, 2A2A547E, 90903BAB, 88880B83,
46468CCA, EEEEC729, B8B86BD3, 1414283C, DEDEA779, 5E5EBCE2,  B0B161D, DBDBAD76,
E0E0DB3B, 32326456, 3A3A744E,  A0A141E, 494992DB,  6060C0A, 2424486C, 5C5CB8E4,
C2C29F5D, D3D3BD6E, ACAC43EF, 6262C4A6, 919139A8, 959531A4, E4E4D337, 7979F28B,
E7E7D532, C8C88B43, 37376E59, 6D6DDAB7, 8D8D018C, D5D5B164, 4E4E9CD2, A9A949E0,
6C6CD8B4, 5656ACFA, F4F4F307, EAEACF25, 6565CAAF, 7A7AF48E, AEAE47E9,  8081018,
BABA6FD5, 7878F088, 25254A6F, 2E2E5C72, 1C1C3824, A6A657F1, B4B473C7, C6C69751,
E8E8CB23, DDDDA17C, 7474E89C, 1F1F3E21, 4B4B96DD, BDBD61DC, 8B8B0D86, 8A8A0F85,
7070E090, 3E3E7C42, B5B571C4, 6666CCAA, 484890D8,  3030605, F6F6F701,  E0E1C12,
6161C2A3, 35356A5F, 5757AEF9, B9B969D0, 86861791, C1C19958, 1D1D3A27, 9E9E27B9,
E1E1D938, F8F8EB13, 98982BB3, 11112233, 6969D2BB, D9D9A970, 8E8E0789, 949433A7,
9B9B2DB6, 1E1E3C22, 87871592, E9E9C920, CECE8749, 5555AAFF, 28285078, DFDFA57A,
8C8C038F, A1A159F8, 89890980,  D0D1A17, BFBF65DA, E6E6D731, 424284C6, 6868D0B8,
414182C3, 999929B0, 2D2D5A77,  F0F1E11, B0B07BCB, 5454A8FC, BBBB6DD6, 16162C3A
}
{
63C6A563, 7CF8847C, 77EE9977, 7BF68D7B, F2FF0DF2, 6BD6BD6B, 6FDEB16F, C59154C5,
30605030,  1020301, 67CEA967, 2B567D2B, FEE719FE, D7B562D7, AB4DE6AB, 76EC9A76,
CA8F45CA, 821F9D82, C98940C9, 7DFA877D, FAEF15FA, 59B2EB59, 478EC947, F0FB0BF0,
AD41ECAD, D4B367D4, A25FFDA2, AF45EAAF, 9C23BF9C, A453F7A4, 72E49672, C09B5BC0,
B775C2B7, FDE11CFD, 933DAE93, 264C6A26, 366C5A36, 3F7E413F, F7F502F7, CC834FCC,
34685C34, A551F4A5, E5D134E5, F1F908F1, 71E29371, D8AB73D8, 31625331, 152A3F15,
 4080C04, C79552C7, 23466523, C39D5EC3, 18302818, 9637A196,  50A0F05, 9A2FB59A,
 70E0907, 12243612, 801B9B80, E2DF3DE2, EBCD26EB, 274E6927, B27FCDB2, 75EA9F75,
 9121B09, 831D9E83, 2C58742C, 1A342E1A, 1B362D1B, 6EDCB26E, 5AB4EE5A, A05BFBA0,
52A4F652, 3B764D3B, D6B761D6, B37DCEB3, 29527B29, E3DD3EE3, 2F5E712F, 84139784,
53A6F553, D1B968D1,        0, EDC12CED, 20406020, FCE31FFC, B179C8B1, 5BB6ED5B,
6AD4BE6A, CB8D46CB, BE67D9BE, 39724B39, 4A94DE4A, 4C98D44C, 58B0E858, CF854ACF,
D0BB6BD0, EFC52AEF, AA4FE5AA, FBED16FB, 4386C543, 4D9AD74D, 33665533, 85119485,
```

```
458ACF45, F9E910F9,  2040602, 7FFE817F, 50A0F050, 3C78443C, 9F25BA9F, A84BE3A8,
51A2F351, A35DFEA3, 4080C040, 8F058A8F, 923FAD92, 9D21BC9D, 38704838, F5F104F5,
BC63DFBC, B677C1B6, DAAF75DA, 21426321, 10203010, FFE51AFF, F3FD0EF3, D2BF6DD2,
CD814CCD,  C18140C, 13263513, ECC32FEC, 5FBEE15F, 9735A297, 4488CC44, 172E3917,
C49357C4, A755F2A7, 7EFC827E, 3D7A473D, 64C8AC64, 5DBAE75D, 19322B19, 73E69573,
60C0A060, 81199881, 4F9ED14F, DCA37FDC, 22446622, 2A547E2A, 903BAB90, 880B8388,
468CCA46, EEC729EE, B86BD3B8, 14283C14, DEA779DE, 5EBCE25E,  B161D0B, DBAD76DB,
E0DB3BE0, 32645632, 3A744E3A,  A141E0A, 4992DB49,  60C0A06, 24486C24, 5CB8E45C,
C29F5DC2, D3BD6ED3, AC43EFAC, 62C4A662, 9139A891, 9531A495, E4D337E4, 79F28B79,
E7D532E7, C88B43C8, 376E5937, 6DDAB76D, 8D018C8D, D5B164D5, 4E9CD24E, A949E0A9,
6CD8B46C, 56ACFA56, F4F307F4, EACF25EA, 65CAAF65, 7AF48E7A, AE47E9AE,  8101808,
BA6FD5BA, 78F08878, 254A6F25, 2E5C722E, 1C38241C, A657F1A6, B473C7B4, C69751C6,
E8CB23E8, DDA17CDD, 74E89C74, 1F3E211F, 4B96DD4B, BD61DCBD, 8B0D868B, 8A0F858A,
70E09070, 3E7C423E, B571C4B5, 66CCAA66, 4890D848,  3060503, F6F701F6,  E1C120E,
61C2A361, 356A5F35, 57AEF957, B969D0B9, 86179186, C19958C1, 1D3A271D, 9E27B99E,
E1D938E1, F8EB13F8, 982BB398, 11223311, 69D2BB69, D9A970D9, 8E07898E, 9433A794,
9B2DB69B, 1E3C221E, 87159287, E9C920E9, CE8749CE, 55AAFF55, 28507828, DFA57ADF,
8C038F8C, A159F8A1, 89098089,  D1A170D, BF65DABF, E6D731E6, 4284C642, 68D0B868,
4182C341, 9929B099, 2D5A772D,  F1E110F, B07BCBB0, 54A8FC54, BB6DD6BB, 162C3A16
}
{
C6A56363, F8847C7C, EE997777, F68D7B7B, FF0DF2F2, D6BD6B6B, DEB16F6F, 9154C5C5,
60503030,  2030101, CEA96767, 567D2B2B, E719FEFE, B562D7D7, 4DE6ABAB, EC9A7676,
8F45CACA, 1F9D8282, 8940C9C9, FA877D7D, EF15FAFA, B2EB5959, 8EC94747, FB0BF0F0,
41ECADAD, B367D4D4, 5FFDA2A2, 45EAAFAF, 23BF9C9C, 53F7A4A4, E4967272, 9B5BC0C0,
75C2B7B7, E11CFDFD, 3DAE9393, 4C6A2626, 6C5A3636, 7E413F3F, F502F7F7, 834FCCCC,
685C3434, 51F4A5A5, D134E5E5, F908F1F1, E2937171, AB73D8D8, 62533131, 2A3F1515,
 80C0404, 9552C7C7, 46652323, 9D5EC3C3, 30281818, 37A19696,  A0F0505, 2FB59A9A,
 E090707, 24361212, 1B9B8080, DF3DE2E2, CD26EBEB, 4E692727, 7FCDB2B2, EA9F7575,
121B0909, 1D9E8383, 58742C2C, 342E1A1A, 362D1B1B, DCB26E6E, B4EE5A5A, 5BFBA0A0,
A4F65252, 764D3B3B, B761D6D6, 7DCEB3B3, 527B2929, DD3EE3E3, 5E712F2F, 13978484,
A6F55353, B968D1D1,        0, C12CEDED, 40602020, E31FFCFC, 79C8B1B1, B6ED5B5B,
D4BE6A6A, 8D46CBCB, 67D9BEBE, 724B3939, 94DE4A4A, 98D44C4C, B0E85858, 854ACFCF,
BB6BD0D0, C52AEFEF, 4FE5AAAA, ED16FBFB, 86C54343, 9AD74D4D, 66553333, 11948585,
8ACF4545, E910F9F9,  4060202, FE817F7F, A0F05050, 78443C3C, 25BA9F9F, 4BE3A8A8,
A2F35151, 5DFEA3A3, 80C04040,  58A8F8F, 3FAD9292, 21BC9D9D, 70483838, F104F5F5,
63DFBCBC, 77C1B6B6, AF75DADA, 42632121, 20301010, E51AFFFF, FD0EF3F3, BF6DD2D2,
814CCDCD, 18140C0C, 26351313, C32FECEC, BEE15F5F, 35A29797, 88CC4444, 2E391717,
9357C4C4, 55F2A7A7, FC827E7E, 7A473D3D, C8AC6464, BAE75D5D, 322B1919, E6957373,
C0A06060, 19988181, 9ED14F4F, A37FDCDC, 44662222, 547E2A2A, 3BAB9090,  B838888,
8CCA4646, C729EEEE, 6BD3B8B8, 283C1414, A779DEDE, BCE25E5E, 161D0B0B, AD76DBDB,
DB3BE0E0, 64563232, 744E3A3A, 141E0A0A, 92DB4949,  C0A0606, 486C2424, B8E45C5C,
9F5DC2C2, BD6ED3D3, 43EFACAC, C4A66262, 39A89191, 31A49595, D337E4E4, F28B7979,
D532E7E7, 8B43C8C8, 6E593737, DAB76D6D,  18C8D8D, B164D5D5, 9CD24E4E, 49E0A9A9,
D8B46C6C, ACFA5656, F307F4F4, CF25EAEA, CAAF6565, F48E7A7A, 47E9AEAE, 10180808,
6FD5BABA, F0887878, 4A6F2525, 5C722E2E, 38241C1C, 57F1A6A6, 73C7B4B4, 9751C6C6,
CB23E8E8, A17CDDDD, E89C7474, 3E211F1F, 96DD4B4B, 61DCBDBD,  D868B8B,  F858A8A,
E0907070, 7C423E3E, 71C4B5B5, CCAA6666, 90D84848,  6050303, F701F6F6, 1C120E0E,
C2A36161, 6A5F3535, AEF95757, 69D0B9B9, 17918686, 9958C1C1, 3A271D1D, 27B99E9E,
D938E1E1, EB13F8F8, 2BB39898, 22331111, D2BB6969, A970D9D9,  7898E8E, 33A79494,
2DB69B9B, 3C221E1E, 15928787, C920E9E9, 8749CECE, AAFF5555, 50782828, A57ADFDF,
 38F8C8C, 59F8A1A1,  9808989, 1A170D0D, 65DABFBF, D731E6E6, 84C64242, D0B86868,
82C34141, 29B09999, 5A772D2D, 1E110F0F, 7BCBB0B0, A8FC5454, 6DD6BBBB, 2C3A1616
}};
```

Расшифрование в алгоритме Rijndael осуществляется с использованием таблицы расшифрования, согласованной с таблицей зашифрования.

2. Оптимизация программной реализации

Проведенный анализ наилучшей программной реализации Rijndael показал, что существует возможность уменьшения вычислительной сложности даже в этой реализации. В известной наиболее быстрой реализации значительной сложностью обладает операция вычисления адреса. Этих затрат

можно избежать, учитывая малую длину цикла. Сущность оптимизации заключается в «разворачивании» цикла и прямого указания адресов (вместо их вычисления) в функциях зашифрования-расшифрования.

Рассмотрим известную [1] функцию зашифрования с точки зрения ее программной реализации:

```
void encrypt(const u4byte in_blk[4], u4byte out_blk[4])
{   u4byte  i, b0[4], b1[4];

    b0[0] = in_blk[0] ^ e_key[0];
    b0[1] = in_blk[1] ^ e_key[1];
    b0[2] = in_blk[2] ^ e_key[2];
    b0[3] = in_blk[3] ^ e_key[3];

    for(i = 1; i < k_len + 6; ++i)
    {
        f_rn(b1, b0, 0); f_rn(b1, b0, 1);
        f_rn(b1, b0, 2); f_rn(b1, b0, 3);

        b0[0] = b1[0] ^ e_key[4 * i];
        b0[1] = b1[1] ^ e_key[4 * i + 1];
        b0[2] = b1[2] ^ e_key[4 * i + 2];
        b0[3] = b1[3] ^ e_key[4 * i + 3];
    }

    f_rl(b1, b0, 0); f_rl(b1, b0, 1);
    f_rl(b1, b0, 2); f_rl(b1, b0, 3);

    out_blk[0] = b1[0] ^ e_key[4 * k_len + 24];
    out_blk[1] = b1[1] ^ e_key[4 * k_len + 25];
    out_blk[2] = b1[2] ^ e_key[4 * k_len + 26];
    out_blk[3] = b1[3] ^ e_key[4 * k_len + 27];
};
```

Видно, что операция вычисления адреса производится достаточно часто. При вычислении используются операции умножения и сложения. Цикл в оригинальной функции зашифрования используется для реализации циклового преобразования. Оптимизация производится путем записи функции зашифрования в виде макросов, так как макросы подставляются в программный код, и все необходимые предвычисления выполняются компилятором. Разобьем программу на три части: начального, циклового преобразования и завершающую согласно тексту программы, приведенному выше.

Обозначим начальную часть (операцию сложения с ключом) в виде макроса **e_rn**:

```
#define e_rn(bo, bi, n)           \
  bo[0] = bi[0] ^ e_key[n];       \
  bo[1] = bi[1] ^ e_key[n + 1];   \
  bo[2] = bi[2] ^ e_key[n + 2];   \
  bo[3] = bi[3] ^ e_key[n + 3]
```

Обозначим цикловое преобразование в виде макроса **e_round**:

```
#define e_round(bo, bi, n)        \
  f_rn(bo, bi, 0);                \
  f_rn(bo, bi, 1);                \
  f_rn(bo, bi, 2);                \
  f_rn(bo, bi, 3);                \
  e_rn(bi, bo, n)
```

Обозначим завершающую часть в виде макроса **e_final**:

```
#define e_final(bo, bi, n)    \
  f_rl(bo, bi, 0);            \
  f_rl(bo, bi, 1);            \
  f_rl(bo, bi, 2);            \
  f_rl(bo, bi, 3);            \
  e_rn(out_blk, bo, n)
```

Полностью предлагаемая альтернативная функция зашифрования имеет вид:

```
void encrypt(const u4byte in_blk[4], u4byte out_blk[4])
{   u4byte  b0[4], b1[4];

  e_rn(b0, in_blk, 0);     e_round(b1, b0,  4);
  e_round(b1, b0,  8);  e_round(b1, b0, 12);
  e_round(b1, b0, 16);  e_round(b1, b0, 20);
  e_round(b1, b0, 24);  e_round(b1, b0, 28);
  e_round(b1, b0, 32);  e_round(b1, b0, 36);

    switch(k_len)
    {
  case 4:       e_final(b1, b0, 40);        break;

  case 6:       e_round(b1, b0, 40);        e_round(b1, b0, 44);
                e_final(b1, b0, 48);        break;

  case 8:       e_round(b1, b0, 40);        e_round(b1, b0, 44);
                e_round(b1, b0, 48);        e_round(b1, b0, 52);
                 e_final(b1, b0, 56);
    }
};
```

Таким образом, оптимизированная функция зашифрования имеет явные преимущества перед ригинальной: отсутствует цикл, не выполняются операция вычисления адреса (указывается явно).

3. Оценка вычислительной сложности

Произведем оценку вычислительной сложности. В качестве показателя вычислительной сложности функций зашифрования-расшифрования используется количество команд (тактов), необходимых для выполнения функций с учетом вероятности выполнения каждой операции. Сразу заметим, что асчет будет учитывать элементарные операции и не будет учитывать затраты на реализацию цикла, ледовательно, будет приблизительным. В данном случае такое упрощение допустимо, так как при ычете оптимизированной версии алгоритма такое упрощение также будет иметь место. Вычислительная сложность для функций зашифрования-расшифрования в обеих реализациях одинакова.

Проведем непосредственный анализ предлагаемой программной реализации вычислительной ожности цикловых преобразований. При этом вычислительная сложность известных функций зашифрования-расшифрования составляет:

$$I_1 = 41 + 35 \bullet N_r \text{ (тактов)}, \tag{4}$$

N_r – число циклов, зависящее от значений длины блока N_b и длины ключа N_k. Их зависимость показывается в таблице 1.

Вычислительная сложность оптимизированных функций зашифрования-расшифрования состав-т:

$$I_2 = 40 + 28 \bullet N_r \text{ (тактов)} \tag{5}$$

Произведем теоретическую оценку сложности известного и оптимизированного алгоритмов. Результаты оценки приведены в табл. 2 при длине блока (N_b), равной 128 бит, в соответствии с формулами (1) и (2).

Таблица 2

Длина ключа/алгоритм	Известный	Оптимизированный
128 бит	391 такт	320 тактов
192 бит	461 такт	376 тактов
256 бит	531 такт	432 тактов

4. Экспериментальная оценка скорости преобразований

Полученные результаты свидетельствуют о наличии 18% повышения скорости работы алгоритма. Для проверки теоретических данных были проведены экспериментальные исследования. Сущность исследований заключается в осуществлении большого количества операций зашифрования-расшифрования (в нашем случае 100 000 000) при разной длине ключа и оценке времени, за которое эти операции были осуществлены. Исследования проводились в среде Visual C 6.0 с включенной оптимизацией по скорости. Результаты тестирования при количестве итераций 1000000 приведены в табл.3. В табл.4 приведены результаты оценки выигрыша в скорости оптимизированного алгоритма по сравнению с оригинальным. Эти результаты были получены на компьютере с процессором Celeron 600 Мгц, 128 Мб в операционной системе Windows 98 с минимально возможным числом запущенных процессов.

Таблица 3

Длина ключа	Время зашифр.(сек)		Время расшифр.(сек)	
	Известный алг.	Оптимизир. алг.	Известный алг.	Оптимизир. алг.
128	75,47	**66,63**	73,54	**64,37**
192	87,12	**77,77**	86,01	**75,85**
256	99,19	**89,26**	99,14	**86,89**
Длина ключа	Скорость зашифр.(Мбит/с)		Скорость расшифр.(Мбит/с)	
	Известный алг.	Оптимизир. алг.	Известный алг.	Оптимизир. алг.
128	169,6	**192,1**	174,05	**198,85**
192	146,92	**164,58**	148,81	**168,75**
256	129,04	**143,4**	129,11	**147,31**

Таблица 4

Длина ключа	Процент выигрыша (зашифр.)	Процент выигрыша (расшифр.)
128	**11,71%**	**12,47%**
192	**10,73%**	**11,81%**
256	**10,01%**	**12,36%**

Заключение

Таким образом, полученные теоретические и экспериментальные сложности (скорости) прямых и обратных криптографических преобразований в алгоритме Rijndael с использованием предложенных усовершенствований позволяют сделать вывод о его предпочтительности по сравнению с известным. Причем выигрыш достигается не менее, чем на 11%. Очевидно, что дальнейшее повышение скорости преобразований возможно за счет применения процессоро-зависимой реализации преобразований на ассемблере.

Список литературы: 1. *J. Daemen, V. Rijemen*, The Rijndael block cipher. AES Proposal. Http://csrc.nist.gov/encryption/aes/rijndael/Rijndael.pdf

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 16.03.2001

УДК 681.3.06:519.248.681

И. Д. ГОРБЕНКО, д-р техн. наук, С. И. ЗБИТНЕВ, А. А. ПОЛЯКОВ

КРИПТОАНАЛИЗ КРИПТОГРАФИЧЕСКИХ ПРЕОБРАЗОВАНИЙ В ГРУППАХ ТОЧЕК ЭЛЛИПТИЧЕСКИХ КРИВЫХ МЕТОДОМ ПОЛЛАРДА.

Введение

В последние годы для криптографических преобразований начали использоваться эллиптические кривые (ЭК) над полем Галуа $GF(p^n)$, где p – простое, а n – целое. Появились рабочие версии стандарта цифровой подписи X9.62[1] и протоколов Диффи – Хелмана X9.63[2], которые, по мнению разработчиков и специалистов, имеют ряд преимуществ и разрешают ряд противоречий, которые в последние годы явно появились в криптографии с открытыми ключами и открытым распространением ключей. Суть противоречий заключается в том, что в условиях интенсивного развития и создание нового математического аппарата, а также высокопроизводительных систем и средств криптоанализа, для обеспечения требуемого уровня стойкости необходимо непрерывно увеличивать длины параметров (модулей) преобразований и ключей, на сегодня до единиц тысяч и более битов. Это противоречие присуще криптографическим преобразованиям в кольцах и полях. Уменьшение длин параметров и ключей, по мнению специалистов, может быть достигнуть за счет использование при криптографических преобразованиях групп точек на эллиптической кривой над полем Галуа $GF(2^m)$ и $GF(p^n)$ [3]. Использование преобразований в группах точек на ЭК предположительно позволяет реализовать вероятно-стойкие или по другой терминологии доказуемо стойкие криптоалгоритмы. При этом доказательство стойкости в наиболее общем случае сводят к доказательству сложности решение дискретного сравнения в группе точек ЭК над полем Галуа $GF(p^n)$

$$Q = d \cdot G(\mathrm{mod}\, f(x), p), \tag{1}$$

относительно d, где G – базовая точка на ЭК порядка n, d – личный ключ (целое число, $1 \le d \le n-1$). Q – открытый ключ, $f(x)$ – примитивный полином, p – простое число.

Более частной является задача Диффи – Хелмана, которое формируется в следующем виде. Известны открытые ключи

$$Q_1 = d_1 G(\mathrm{mod}\, f(x), p) \text{ и } Q_2 = d_2 G(\mathrm{mod}\, f(x), p).$$

Необходимо найти значение общего секрета.

$$K_{21} = K_{21} = d_1 d_2 G(\mathrm{mod}\, f(x), p). \tag{2}$$

На сегодня известно несколько методов решения сравнений вида (1). Получили распространения методы Полларда ρ [4] и λ – метод.

Целью настоящей статьи является разработка математического аппарата и алгоритмов криптоанализа с использованием метода ρ – Полларда. Основной причиной выбора этого метода является возможность эффективного распараллеливания процесса решения сравнения вида (1) и по взглядам на сегодняшний день меньшая по сравнению с другими методами сложность.

1. Математическая постановка задачи

Пусть задана супернесингулярная ЭК над полем $GF(2^m)$ в аффинном представлении

$$y^2 + xy = x^3 + ax^2 + b(\mathrm{mod}\, f(x), 2), \tag{3}$$

где a и b – параметры ЭК.

Она определена множеством точек $(x, y) \in GF(2^m) \times GF(2^m)$, a и $b \in GF(2^m)$, $b \neq 0(\mathrm{mod}\, f(x), 2)$. Точки на ЭК, включая точку бесконечности O, образуют группу с операцией сложения.

Если точка $P_1 = (x_1, y_1)$ и $P_2 = (x_2, y_2)$ принадлежат эллиптической кривой, т.е. $P_i \in E(GF(2^m))$, то для каждой из них существует обратная точка, соответственно $-P_1 = (x_1, x_1 + y_1)$ и

$-P_2=(x_2,x_2+y_2)$, а также точка $P_3=(x_3,y_3)$, такая что $P_1+P_2=P_3$. Координаты точки $P_3=(x_3,y_3)$ определяются с использованием соотношений

$$x_3=\lambda^2+\lambda+x_1+x_2+a \pmod{f(x),2}; \tag{4}$$

$$y_3=\lambda(x_1+x_3)+x_3+y_1 \pmod{f(x),2}; \tag{5}$$

$$\lambda=\begin{cases}\dfrac{y_1+y_2}{x_1+x_2}\pmod{f(x),2}, & \text{если} \quad P_1\neq P_2;\\ \dfrac{x_1^2+y_1}{x_1}\pmod{f(x),2}, & \text{если} \quad P_1=P_2,\end{cases} \tag{6}$$

скалярное умножение определяется для нескольких точек $G\in E(GF(2^m))$ как

$$d\cdot G\pmod{f(x),2}=\underbrace{G+G+G+\ldots+G}_{d\text{ раз}}\pmod{f(x),2}. \tag{7}$$

Операция (7) реализуется за счет применения операций сложения $(P_1\neq P_2)$ или удвоения $(P_1=P_2)$.

Точка G имеет порядок n на ЭК, если

$$n\cdot G\pmod{f(x),2}=O; \tag{8}$$

где O – точка на бесконечности (ноль).

Важнейшей задачей при выполнении операций (4) – (8) является минимизация сложности. К сожалению продуктивных методов и алгоритмов выполнения аффинных преобразований, особенно в (6), где требуется выполнять деление по модулю, неизвестно или нет. При выполнении криптоанализа задача минимизация сложности вычислений (4) – (8) становиться особенно актуальной и требует особого внимания.

Одним из возможных методов уменьшения сложности преобразований при решении задач вида (1) и (2) является использование проективного представления точек на эллиптической кривой и выполнение операций (4) – (8) в проективном базисе. Существует ряд проблемных вопросов реализации проективного представления вычислений при криптоанализе.

При переходе от аффинного представления к проективному используют следующее преобразование [3]

$$x=\frac{X}{Z^2};\ y=\frac{Y}{Z^3}.$$

При этом точка в аффинном представлении $Q_{afin}=(x,y)$ отображается в точку проективного представления $Q_{project}=(x,y,1)$, а обратный переход выполняется в виде $Q_{afin}=\left(\dfrac{x}{z^2},\dfrac{y}{z^3}\right)$.

Проективным аналогом сравнения (3) является сравнение

$$Y^2+XYZ=X^3+aX^2Z^2+bZ^6\pmod{f(x),2}. \tag{9}$$

Если $P_1=(X_1,Y_1,Z_1)$ и $P_2=(X_2,Y_2,Z_2)$ и $P_1\neq P_2$, то суммой двух точек P_1+P_2 является точка $P_3=(X_3,Y_3,Z_3)$, координаты которой определяются с использованием формул [3] по модулю $(f(x),2)$.

$$X_3=SU; \tag{10}$$

$$Y_3=T(U+S^2X_1Z_2)+S^3Y_1Z_2+SU; \tag{11}$$

$$Z_3=S^3Z_1Z_2, \tag{12}$$

где $S=X_2Z_1+X_2Z_2;\ T=Y_2Z_1+Y_1Z_2;\ U=(T^2+TS+aS^2)Z_1Z_2+S^3.$ (13)

Для случая, когда $P_1 = P_2$ с использованием формул

$$X_3 = ST; \tag{14}$$

$$Y_3 = X^4 S + T\left(S + YZ + X^2\right); \tag{15}$$

$$Z_3 = S^3, \tag{16}$$

где $S = XZ;\ T = bZ^4 + X^4$. (17)

Причем математические операции в (14) – (17) выполняется по модулю $\left(f(x),2\right)$.

Подробные сведения, методы и алгоритмы выполнение операций сложения и удвоения точек эллиптической кривой приведены в [3].

2. Характеристики известных методов криптоанализа

Основными среди известных методов решения сравнения вида (1) являются методы Полларда ρ и λ – метод [4]. Сложность ρ – Полларда метода можно оценить с использованием соотношения [4]

$$I_\rho = \sqrt{\frac{\pi n}{2}}, \tag{18}$$

где n – порядок базовой точки на эллиптической кривой. В [4] показано, что ρ – Поллард метод может быть ускорен в $\sqrt{2}$ раза, в этом случае

$$I_\rho{}' = \sqrt{\frac{\pi n}{4}}. \tag{19}$$

Одним из преимуществ методов ρ – Полларда является то, что он допускает распараллеливание на r независимых процессов. В этом случае сложность реализации каждого из процессов можно оценить как

$$I_{\rho_1} = \frac{\sqrt{\frac{\pi n}{2}}}{r} = \sqrt{\frac{\pi n}{2r^2}}; \tag{20}$$

$$I_{\rho_1}{}' = \frac{\sqrt{\frac{\pi n}{4}}}{r} = \sqrt{\frac{\pi n}{4r^2}}. \tag{21}$$

Сложность при использовании метода λ – Полларда может быть оценена как [2]

$$I_\lambda = 2\sqrt{n}; \tag{22}$$

а при распараллеливании

$$I_{\lambda_1} = \frac{2\sqrt{n}}{r}. \tag{23}$$

В (18) – (23) сложность оценивается количеством операций сложений на эллиптической кривой.

Проведем сравнительный анализ по сложности названных методов. Для этого найдем отношение

$$\frac{I_\lambda}{I_\rho} = \frac{2\sqrt{n}}{\sqrt{\frac{\pi n}{2}}} = \frac{2\sqrt{2}}{\sqrt{\pi}} = \frac{\sqrt{8}}{\sqrt{\pi}} \approx 1.558. \tag{24}$$

Из отношения (24) вытекает, что λ – метод более сложен чем даже не оптимизированный ρ – метод. Поэтому в дальнейшем сосредоточим внимание на методах ρ – Полларда.

3. Алгоритм решения сравнения по основе ρ – метода Полларда

Непосредственно из (1) следует, что используя ρ – метод необходимо найти число d, такое что $1 \le d \le n-1$. Будем формировать последовательность предполагаемых решений d_i сравнения (1), используя функцию $f(Z_i)$, $i = 0,1,3,...,n$, которая в соответствии с ρ – методом обеспечивает (может обеспечивать) нахождение пары значений $f(Z_i)$ и $f(Z_j)$, таких, что

$$f(Z_i) = f(Z_j), \ (i \ne j). \tag{25}$$

Одним из предпочтительных представлений точек Z_i является

$$Z_i = A_i G + B_i Q, \tag{26}$$

где G – базовая точка на эллиптической кривой, а Q – открытый ключ [3].

Выберем случайным образом два числа A_0 и $B_0 \in [0, n-1]$, но так чтобы одновременно A_0 и B_0 не были равны нулю.

$$f(Z_0) = A_0 G + B_0 Q. \tag{27}$$

Далее будем искать значения $f(Z_i)$ и $f(Z_j)$, удовлетворяющие условию (25). В этом случае получим, что

$$A_j G + B_j Q \equiv A_i G + B_i Q (\mathrm{mod}\, f(x), p), \ (i \ne j). \tag{28}$$

где наличие $f(x)$ и p означает, что кривая рассматривается над полем $GF(p^m)$, m – порядок расширения поля. После преобразований (28) имеем

$$(B_i - B_j) Q = (A_j - A_i) G (\mathrm{mod}\, f(x), p), \ (i \ne j),$$

или

$$Q = \frac{A_j - A_i}{B_i - B_j} G (\mathrm{mod}\, f(x), p), \ (B_i \ne B_j), \tag{29}$$

сравнивая (1) и (29) имеем, что

$$d = \frac{A_j - A_i}{B_i - B_j} (\mathrm{mod} n), \ (B_i \ne B_j). \tag{30}$$

Выясним, каким образом нужно формировать пары коэффициентов (A_i, B_i) для всех $i \le n$. Наиболее простым алгоритмом формирования коэффициентов является следующим [4].

Разобьем точки на ЭК на три равных множества S_1, S_2 и S_3 и вычислим рекуррентно по правилу

$$Z_{i+1} = f(Z_i) = \begin{cases} 2Z_i, & если \quad Z_i^x \in S_1; \\ Z_i + G, & если \quad Z_i^x \in S_2; \\ Z_i + Q, & если \quad Z_i^x \in S_3; \end{cases} \tag{31}$$

где Z_i^x означает x – координату точки на ЭК.

Значение Z_0 формируется по (27), в простейшем случае $Z_i = G$ или $Z_i = Q$, или $Z_i = G + Q$. Выполняя последовательно вычисления по правилу (31) мы, по существу, будем изменять и коэффициенты A_i, B_i. Найдя Z_i и Z_j, удовлетворяющие условию (25), мы получим решение в виде (30).

Пример 1. Пусть имеется точка $G=(13,7)$, которая принадлежит кривой

$$y^2 = x^3 + x + 1 (\mathrm{mod}\, 23)$$

и $Q=(17,20)$. Можно показать, что G имеет порядок $n=7$. Необходимо решить сравнение (1) вида

$$(17,20) = d(13,17) (\mathrm{mod}\, 23),\ d \in [1, n-1].$$

Разобьем $p-1=22$ на три области $S_1 \in [0,7]$, $S_2 \in [8,15]$, $S_3 \in [16,22]$ и выберем $Z_0 = G = (13,7)$.

Вычислим последовательность Z_{i+1} используя (31). Поскольку $Z_0^{\,x} \in S_2$, то $Z_1 = Z_0 + G = (13,7) + (13,7) = (5,4)$. Далее $Z_1^{\,x} \in S_1$, поэтому $Z_2 = 2Z_1 = 2(5,4) = (17,20)$. $Z_2 \in S_3$, поэтому $Z_3 = Z_2 + Q = (17,20) + (17,20) = (13,7)$.

Мы получили значения Z_3 такое, что

$$Z_3 = Z_0.$$

Теперь необходимо определить пары коэффициентов (A_i, B_i) и (A_j, B_j), для которых выполняется условие (28). Результаты вычислений приведены в табл. 1.

Таблица 1

Шаг	Z_i	A_i	B_i	$A_iG + B_iQ$
0	$(13,7)$	1	0	$1\cdot(13,7) + 0\cdot(17,20)$
1	$(5,7)$	2	0	$2\cdot(13,7) + 0\cdot(17,20)$
2	$(17,20)$	4	0	$4\cdot(13,7) + 0\cdot(17,20)$
3	$(13,7)$	4	1	$4\cdot(13,7) + 1\cdot(17,20)$

Из таблицы определяем, что

$$A_0 = 1;\ B_0 = 0;\ A_3 = 4;\ B_3 = 1.$$

По правилу (30) имеем

$$d = \frac{1-4}{1-0} (\mathrm{mod}\, 7) = -3 (\mathrm{mod}\, 7) = 4.$$

Проверкой убедимся, что $d = 4$ является решением примера.

Алгоритм (31) при больших значениях n становится не эффективным. Как показали исследования, он может быть улучшен [4], если все множество точек эллиптической кривой, включая точку бесконечности, разбивать на $r > 3$ не пересекающихся множеств. В этом случае функция $f(Z_i)$ задается в виде

$$Z_{i+1} = f(Z_i) = \begin{cases} (Z_i + c_1 G + d_1 Q), & \text{если} \quad Z_i \in S_1; \\ (Z_i + c_2 G + d_2 Q), & \text{если} \quad Z_i \in S_2; \\ \ldots\ldots\ldots\ldots & \ldots\ldots \quad \ldots\ldots\ldots \\ (Z_i + c_r G + d_r Q), & \text{если} \quad Z_i \in S_r; \end{cases} \qquad (32)$$

где c_v, d_v – случайные целые числа из интервала $[0, n-1]$.

Установлено [4], что при использовании итерационных формул вида (32) можно уменьшить сложность (стоимость времени) процедуры криптоанализа. Кроме того, применение правил (32) позволяет эффективно распараллелить процесс нахождения коэффициентов A_i и B_i, для которых выполняется условие (28), по крайней мере, на r процессов.

Если числа c_v и d_v выбирать таким образом, чтобы пересечения значений Z_i в областях не было, то и каждую область, можно разбить на необходимое число подобластей, Это позволяет распараллелить задачу на необходимое количество независимых процессов. Результаты работы каждого из процессов в виде троек чисел (Z_i, A_i, B_i) заносятся в общую базу данных. Центральный процессор осуществляет параллельный поиск значений

$$Z_i = Z_j,\ (i \neq j),$$

и после этого с использованием выражения (30) вычисляют личный ключ d.

На практике процесс нахождения d может быть ускорен, если воспользоваться взаимосвязью основной и обратной точек. Если существует точка $P=(x,y)$, то обратной точкой, над полем $GF(2^m)$, является $(-P)=(x,x+y)$. Это означает, что у основной и обратной точек координаты x одинаковые. Поэтому для алгоритма (32) необходимо искать только точки у которых x – координаты одинаковые. В этом случае возможны два исхода:

1) координаты y у обеих точек одинаковы;
2) координаты y у обеих точек разные.

В первом случае решение находится так же как уже описывалось выше. Во втором случае правило неприменимо. Но если учесть, что основная и обратная точки связаны соотношением

$$Z_j + Z_k = O, \tag{33}$$

то истинные значения координат точек, для которых совпадают обе координаты точек, можно пересчитать используя то, что [4]

$$A_k + A_j = n \text{ и } B_k + B_j = n.$$

Получим, что

$$A_j = n - A_k \text{ и } B_j = n - B_k,$$

где положено, что A_k и B_k – коэффициенты точки у которой не совпадают y координаты. Использование отмеченного свойства позволяет уменьшить сложность нахождения личного ключа не менее чем в 2 раза

4. Примеры нахождения личного ключа на ЭК над полем $GF(2^m)$.

Пусть имеется ЭК $y^2+xy=x^3+ax^2+b$ над полем $GF(2^m)$ $(m=33)$, параметры ЭК: $a=571EF7A8$ и $b=39A75A68$, а полином над $GF(2^{33})$ $f(x)=200000401$. В качестве базовой выберем точку $G=(030D5A653,1B081C3F)$. Эта точка на кривой имеет порядок $n=7FFF9BEF$, открытый ключ $Q=(\text{C69A56D4,1955247BB})$.

Составим аналогично (1) уравнение

$$(\text{C69A56D4,1955247BB})=d(030D5A653,1B081C3F).$$

В таблице 2 приведены значения c_v и d_v. В таблице 3 приведены некоторые значения коэффициентов и процесс поиска коэффициентов. Он аналогичен работе при ручном поиске (пример 1).

Таблица 2

Интервал	Z_i	A_i	B_i	Интервал	Z_i	A_i	B_i
0	(0 17db318b,0 17db318b)	2ed6d0f2	7149f463	10	(1 ae1c4f64,0 ae1c4f64)	5066a455	3731287c
1	(1 1c1b66d4,0 1c1b66d4)	6bea2002	3cdba6b3	11	(1 71c072f7,1 71c072f7)	22a13b32	1d918707
2	(0 28db3c5d,1 28db3c5d)	193e7098	1c0d5e24	12	(1 77eab9f7,0 77eab9f7)	5a1feff2	299cab0
3	(0 da315ef2,0 da315ef2)	5e481b17	484d76c4	13	(1 830e176e,0 830e176e)	499d72b9	457ee43d
4	(1 a58b7411,0 a58b7411)	268c1f54	7d5a031f	14	(0 3a3b41a8,0 3a3b41a8)	ea39aa	4cca4ec3
5	(0 db9be265,1 db9be265)	1f253809	7ccc386e	15	(0 2d17b52e,1 2d17b52e)	56682364	7ed614bf
6	(0 668ecf21,1 668ecf21)	3b45f596	a890d13	16	(0 6fecbfbb,1 6fecbfbb)	668edc2d	4957ef83
7	(0 3b7b8cf6,1 3b7b8cf6)	32aedb1c	6e50fe31	17	(1 a3edec8d,0 a3edec8d)	f6a6372	1ed1cd61
8	(0 14d557ad,1 14d557ad)	7d36dc51	7632ad23	18	(0 c7420d46,0 c7420d46)	72169c9d	701e4a83
9	(0 1d5d56f4,1 1d5d56f4)	5c49e1af	7214b3be	19	(0 124f81ea,1 124f81ea)	774a4f84	39e8d702

Таблица 3

Шаг	Z_i	Применяемый интервал	A_i	B_i
1	(0 124f81ea,1 124f81ea)	Интервал=10	A[1]=774a4f84	B[1]=39e8d702
2	(0 41d32423,1 41d32423)	Интервал=7	A[2]=47b157ea	B[2]=7119ff7e
3	(1 d56a393e,0 d56a393e)	Интервал=18	A[3]=7a603306	B[3]=5f6b61c0
4	(0 efb109bf,1 efb109bf)	Интервал=11	A[4]=6c7733b4	B[4]=4f8a1054
5	(0 c4ffc7df,0 c4ffc7df)	Интервал=3	A[5]=f18d2f7	B[5]=6d1b975b
	.....	.....	.....	.....
7545	(1 ccab1475,0 ccab1475)	Интервал=17	A[7545]=2cc89bd9	B[7545]=797aa370
7546	(1 a7b63df1,1 a7b63df1)	Интервал=1	A[7546]=3c32ff4b	B[7546]=184cd4e2
7547	(1 e4592c48,0 e4592c48)	Интервал=8	A[7547]=281d835e	B[7547]=55287b95
7548	(1 6e8a5c1b,1 6e8a5c1b)	Интервал=7	A[7548]=2554c3c0	B[7548]=4b5b8cc9

Проведены эксперименты по нахождению личного ключа для ЭК над полем $GF(2^m)$ при $m=64$. Среднее время решения на одном Pentium III-800 составляет 48 мин. Отрабатываются составляющие решения сравнения (1) для $m=96$ и $m=112$. Заметим, что при решении сравнения могут возникать тупиковые ситуации.

5. Оценка сложности криптоанализа практически применяемых ЭК над $GF(2^m)$.

Группа точек ЭК является циклической группой. Поэтому потенциальные оценки сложности дискретного логарифма, которая может быть обеспечена в предельном случае, необходимо искать в общем случае как для произвольной циклической группы G. Известен метод Шенкса, при использовании которого задачу дискретного логарифма в произвольной циклической группе можно решить за $\sqrt{2^m}$ операций. Решение основывается на составление двух стеков размером $t=\sqrt{2^m}$ отсортированных по вторым компонентам [1]. Например первый стек состоит из пар $(i,\theta_e^{\,it})$, $i=\overline{0,t-1}$ и отсортирован по второму компоненту, где θ – первообразный элемент группы. Второй стек состоит из пар (α,y^j), $j=\overline{0,t-1}$ и тоже отсортирован по второй компоненте. При наличии таких стеков можно найти две пары с равными вторыми компонентами, т.е. (i,θ^{it}) и $(j,y\theta^j)$, причем $\theta^{it}=y\theta^j$, и далее

$$x=(it-j)(\bmod 2^m). \tag{34}$$

Полагая, что для решения сравнения (1) используется система, выполняющая соответственно $\gamma=10^6$ и 10^8 операций сложений на ЭК получим оценки безопасного времени $t_б$ для ρ – метода Полларда (выражение (19)). В таблице 4 приведены оценки $t_б$ для $\gamma=10^6$ (для $\gamma=10^8$ $t_б'$)операций сложения на ЭК/с, полученные с использованием выражения $t_б=\dfrac{I_\rho}{\gamma k}$, где $k=3.1\cdot10^7$ с/год (кол-во секунд в год).

Таблица 4

n	96	128	160	192	224	256	512
I_ρ (оп/с)	$249\cdot10^{12}$	$16{,}3\cdot10^{18}$	10^{25}	$70{,}2\cdot10^{27}$	$4{,}6\cdot10^{33}$	$3\cdot10^{38}$	10^{77}
$t_б$ (лет)	8,04	$5{,}2\cdot10^5$	$34{,}5\cdot10^9$	$2{,}2\cdot10^{15}$	$1{,}48\cdot10^{20}$	$9{,}72\cdot10^{24}$	$3{,}31\cdot10^{63}$
$t_б'$ (лет)	0,0804	$5{,}2\cdot10^3$	$34{,}5\cdot10^7$	$2{,}2\cdot10^{13}$	$1{,}48\cdot10^{18}$	$9{,}72\cdot10^{22}$	$3{,}31\cdot10^{61}$

Заключение

В ближайшие годы следует ожидать сосредоточение особого внимания и интеллектуальных ресурсов на решение задач криптоанализа в группах точек эллиптических кривых. По видимому основные усилия будут направлены на развитие математического аппарата решения сравнения вида (1) и (2), а также создания распараллеленых криптоаналитических систем . Если будут серьезные достижения в минимизации сложности вычислений в группах точек эллиптических кривых, то это приведет к необходимости увеличению параметров эллиптической кривой. Поэтому во внедряемых

стандартах необходимо предусматривать широкий спектр значений параметров. Так по нашим оценкам в цифровых подписях и протоколов аутентификации и управление ключами необходимо было бы предусмотреть преобразования с порядками базовых точек $n = 2^{160}, 2^{224}, 2^{256}, 2^{512}$ и более. Использование такого спектра значений порядка базовой точки с одной стороны может обеспечивать при необходимости требуемый уровень стойкости, а с другой допустимый уровень вычислительной сложности.

Список литературы: 1. *X9.62 Public Key Cryptography For The Financial Services Industry*: The Elliptic Curve Digital Signature Algorithm (ECDSA), 1998, 87 с. 2. *X9.63 Public Key Cryptography For The Financial Services Industry*: Key Agreement and Key Transport Using Elliptic Curve Cryptography, 1999. 207 с. 3. *IEEE P1363 / D11 (Draft Version 11).* Standard Specifications for Public Key Cryptography. Annex A (Informative). Number-Theoretic Background, 1999. 91 с. 4. *Michael J. Wiener, Robert J. Zuccherato* Faster Attacks on Elliptic Curve Cryptosystems, 1998, 8 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 19.03.2001

УДК 681.3.06:519.248.681

И. Д. ГОРБЕНКО, д-р техн. наук, С. И. ЗБИТНЕВ, А. А. ПОЛЯКОВ

СЛОЖНОСТЬ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ В ГРУППАХ ТОЧЕК ЭЛЛИПТИЧЕСКИХ КРИВЫХ ДЛЯ КРИПТОГРАФИЧЕСКИХ ОПЕРАЦИЙ

Введение

В 90-е годы разработаны и нашли широкое распространение методы и средства обеспечения целостности и подлинности информации, которые базируются на несимметричной криптографии. Необходимый уровень стойкости в них обеспечивается за счет выполнения арифметических операций над целыми числами в кольцах, полях и подполях (подгруппах). В то же время непрерывное развитие математических методов и средств криптоанализа требует постоянного увеличения величин модуля преобразования. На сегодня безопасным считается длина модуля 2048 и более битов. Однако увеличение длин преобразуемых чисел приводит к увеличению сложности преобразований, а также длин ключей и параметров. Разрешение этого противоречия наметилось за счет выполнения криптографических преобразований в группе точек эллиптической кривой (ЭК). Разработаны и нашли применения стандарты цифровой подписи X9.62 [1] и распределение ключей X9.63 [2]. При их использовании появилась возможность уменьшения длины параметров и ключей, а также уменьшения вычислительной сложности преобразований и, как следствие, повышение скорости преобразований. В то же время проблема дальнейшего уменьшения сложности криптографических преобразований в группах точек эллиптических кривых остается весьма актуальной.

Проведенный анализ показал, что выполнение операций сложения и умножения в группе точек эллиптической кривой может выполняться с использованием аффинных или проективных координат[3]. Представляется необходимым проведение подробных исследований и сравнительного анализа сложности арифметических операций (преобразований) с целыми числами большой разрядности (160 и более) с использованием аффинных и проективных координат. При этом первоочередным является проведение сравнительного анализа сложности операций и умножения в группе точек эллиптических кривых над полем $GF(2^m)$.

Целью настоящей статьи и является проведение сравнительного анализа сложности выполнения арифметических операций в группах точек эллиптических кривых в поле $GF(2^m)$ для $m = 160$ и более битов.

1. Сложность арифметических операций над элементами в поле $GF(2^m)$

Наиболее распространенные представления элементов в полях $GF(2^m)$ – полиномиальное и нормальное [3], но наиболее эффективным, позволяющим достичь большей производительности, является полиномиальное. Подробное сравнение проведено в [4] . Далее мы будем рассматривать только полиномиальное представление элементов поля $GF(2^m)$.

Основными операциями над элементами в поле $GF(2^m)$ являются операции сложения по модулю *2*, умножения по модулю $(f(x),2)$, вычитание (идентично сложению), возведение в квадрат (частный случай умножения), нахождение обратного элемента в поле. Ниже приведена сложность алгоритмов из [3] с учетом представления чисел в ЭВМ.

Сложение по модулю 2 (I_{sum}):

$$I_{sum}[x] = bl[x], \tag{1}$$

где x – дли блока преобразования в битах; $bl[x] = \left\lceil \frac{x}{32} \right\rceil$.

Умножение по модулю $(f(x),2)$ (I_{mul}):

$$I_{mul} = (5 + 1984 \cdot bl[x]) \cdot x \tag{2}$$

Возведение в квадрат по модулю $(f(x),2)$ (I_{sqr}):

$$I_{sqr} = x \cdot (5.5 + 993 \cdot bl[x]) + 468 \cdot bl[x] \tag{3}$$

Нахождение обратного элемента по модулю $(f(x),2)$ (I_{inv}):

$$I_{inv} = (x-1)\cdot I_{sqr} + 0.5\cdot x\cdot I_{mul} \tag{4}$$

2. Сущность и анализ сложности преобразований в аффинных координатах

Пусть в группе точек эллиптических кривых над полем $GF(2^m)$, вида $y^2 + xy = x^3 + ax^2 + b \bmod(f(x),2)$ заданы точки $P_1 = (x_1, y_1)$ и $P_2 = (x_2, y_2)$. Суммой двух точек $P_1 + P_2$ называется точка, имеющая координаты (x_3, y_3), причем:

$$x_3 = \lambda^2 + \lambda + x_1 + x_2 + a \bmod(f(x),2), \tag{5}$$

$$y_3 = \lambda(x_1 + x_3) + x_3 + y_1 \bmod(f(x),2), \tag{6}$$

где $\lambda = \dfrac{y_1 + y_2}{x_1 + x_2}$.

Подсчитав общее количество сложений, умножений и инверсий и подставив сложность соответствующих арифметических операций, получим:

$$I^{af}{}_{add} = (8 + x)\cdot I_{mul} + (x-1)\cdot I_{sqr} + 9I_{sum} \tag{7}$$

Существует частный случай сложение двух точек – удвоение. Пусть в группе точек эллиптических кривых над полем $GF(2^m)$ вида $y^2 + xy = x^3 + ax^2 + b \bmod(f(x),2)$ задана точка $P_1 = (x_1, y_1)$. Удвоением точки $2\cdot P_2$ называется точка, имеющая координаты (x_2, y_2), причем:

$$x_3 = \lambda^2 + \lambda + a \bmod(f(x),2), \tag{8}$$

$$y_3 = x_1^{\,2} + (\lambda + 1)\cdot x_3 \bmod(f(x),2), \tag{9}$$

где $\lambda = x_1 + \dfrac{y_1}{x_1}$.

Подсчитав общее количество сложений, умножений и инверсий и подставив сложность соответствующих арифметических операций, получим:

$$I^{af}{}_{double} = (5 + x)\cdot I_{mul} + x\cdot I_{sqr} + 5I_{sum} \tag{10}$$

3. Сущность и анализ сложности преобразований в проективных координатах

Пусть в группе точек эллиптических кривых над полем $GF(2^m)$ вида $Y^2 + XYZ = X^3 + aX^2Z^2 + bZ^6 \bmod(f(x),2)$ заданы точки $P_1 = (X_1, Y_1, Z_1)$ и $P_2 = (X_2, Y_2, Z_2)$. Суммой двух точек $P_1 + P_2$ называется точка, имеющая координаты (X_3, Y_3, Z_3), причем:

$$U_0 = X_1\cdot Z_2^{\,2}, \quad U_1 = X_2\cdot Z_1^{\,2}, \quad S_0 = Y_1\cdot Z_2^{\,3}, \quad S_1 = Y_2\cdot Z_1^{\,3}, \quad W = U_0 + U_1, \quad R = S_0 + S_1;$$
$$L = Z_1\cdot W, \qquad V = R\cdot X_2 + L\cdot Y_2,$$

$$Z_3 = L\cdot Z_2;\ T = R + Z_3, \qquad X_3 = a\cdot Z_3^{\,2} + T\cdot R + W^3, \qquad Y_3 = T\cdot X_3 + V\cdot L^2 \tag{11}$$

Подсчитав общее количество сложений, умножений и инверсий и подставив сложность соответствующих арифметических операций, получим:

$$I^{pr}{}_{add} = 22\cdot I_{mul} + 5\cdot I_{sqr} + 7I_{sum} \tag{12}$$

В проективных координатах также существует частный случай сложение двух точек – удвоение. Пусть в группе точек эллиптических кривых над полем $GF(2^m)$ вида

$Y^2 + XYZ = X^3 + aX^2Z^2 + bZ^6 \bmod\left(f(x),2\right)$ задана точка $P_1 = (X_1, Y_1, Z_1)$. Удвоением точки $2 \cdot P_2$ называется точка, имеющая координаты (X_3, Y_3, Z_3). причем:

$$c = b^{2^{m-2}}, \qquad Z_2 = X_1 \cdot Z_1^{\,2}, \qquad X_2 = \left(X_1 + c \cdot Z_1^{\,2}\right)^4,$$
$$U = Z_2 + X_1^{\,2} + Y_1 \cdot Z_1, \qquad Y_2 = X_1^{\,4} \cdot Z_2 + UX_2. \tag{13}$$

Подсчитав общее количество сложений, умножений и инверсий и подставив сложность соответствующих арифметических операций, получим:

$$I^{pr}{}_{double} = 9 \cdot I_{mul} + 5 \cdot I_{sqr} + 4I_{sum} \tag{14}$$

С использованием выражений (7) и (12) определяем сложность сложения точек на ЭК, а с использованием выражений (10) и (14) - сложность удвоения на ЭК. При этом перевод из аффинных координат в проективные выполняется в виде [3]

$$X = x \ , \ Y = y, \ Z = 1. \tag{15}$$

Преобразование из проективных координат в аффинные выполняется в виде [3]

$$x = \frac{X}{Z^2}, \ y = \frac{Y}{Z^3} \tag{16}$$

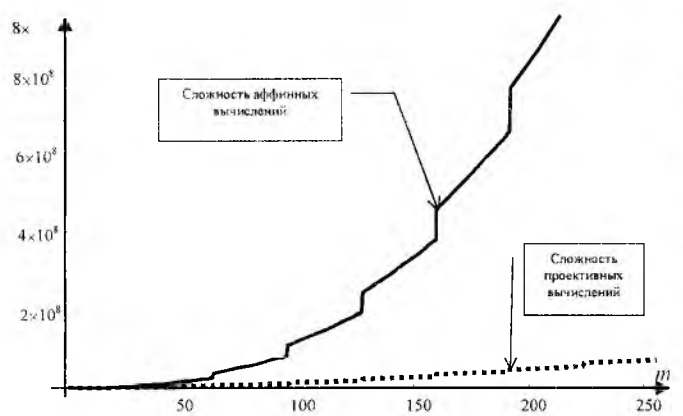


Рис. 1

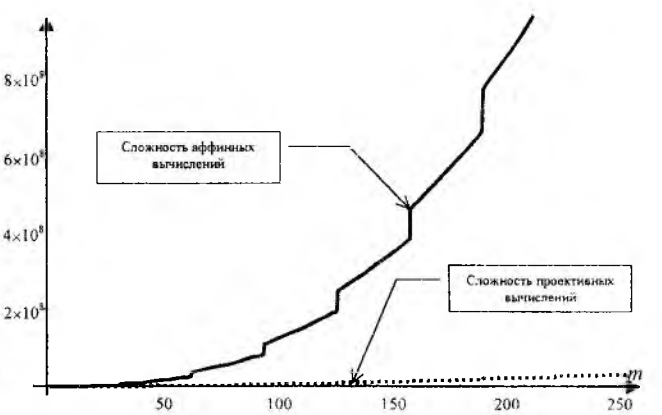


Рис. 2

В приведенных выражениях x, y - аффинные координаты, X, Y, Z – проективные координаты точек ЭК.

Графики сложности сложения и удвоения на эллиптической кривой в аффинных и проективных координатах (в зависимости от порядка расширенного поля m) приведены на рис.1, 2 соответственно.

Сложность сложения и удвоения точек на ЭК вычислена в числе тактов, необходимых для выполнения операций с длиной чисел в m бит. Зная число тактов выполнения операций сложения и удвоения можно определить время выполнения каждой из операций.

На рис. 3 и 4 приведены графики сложности сложения и удвоения для небольших значений m. Из графиков следует, что сложность вычислений в аффинных координатах меньше сложности вычислений в проективных координатах при $m \le 11$, а удвоение $m \le 5$.

Следует отметить, что на рис. 1-4 приведены значения сложности сложения и удвоения без учета преобразований из одних координат в другие. На практике выполняются операции умножения большого целого числа d на базовую точку G с координатами x_G и y_G, так, что

$$Q = d \cdot G(\bmod f(x), 2) \tag{17}$$

где Q - точка на ЭК (открытый ключ). Причем, значение (точка) Q вычисляется посредством многократного выполнения операций сложения и удвоения. После нахождения Q в проективных координ-

натах необходимо преобразовать его в аффинных координатах, что выполняется с использованием (16). Эта операция выполняется один раз.

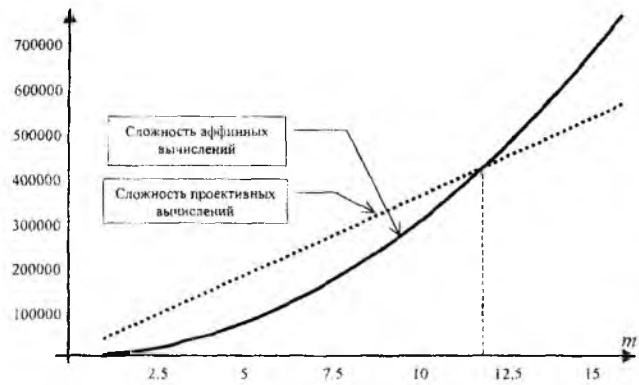


Рис. 3

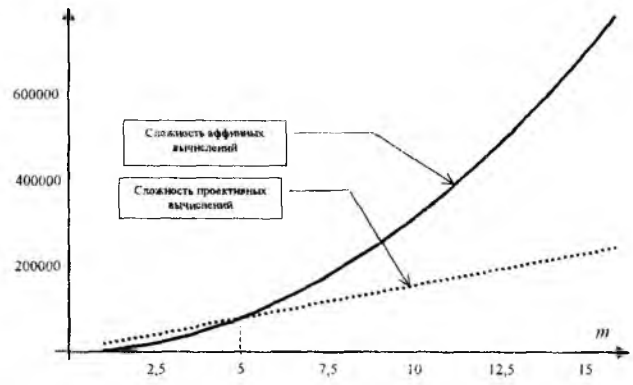


Рис. 4

В связи с тем, что основной операцией в криптографических преобразованиях является операция умножения (17), исследовалась сложность ее выполнения в аффинных и проективных координатах. На рис. 5-8 приведены зависимости сложности выполнения операции умножения (17) в зависимости от величины d при значения порядка расширения поля $m = 16, 32, 128, 256$, битов соответственно, причем, d изменялось в интервале от 1 до 2^m.

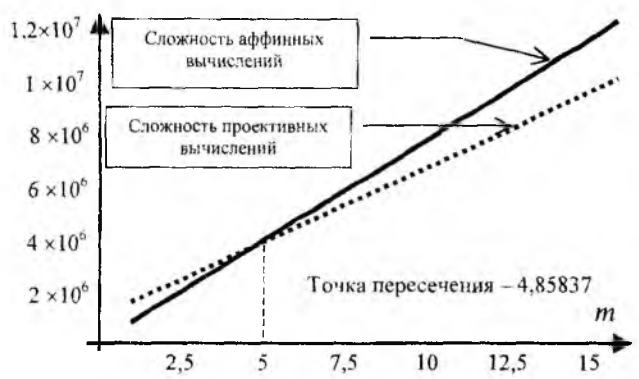


Рис. 5

Рис. 6

Рис. 7

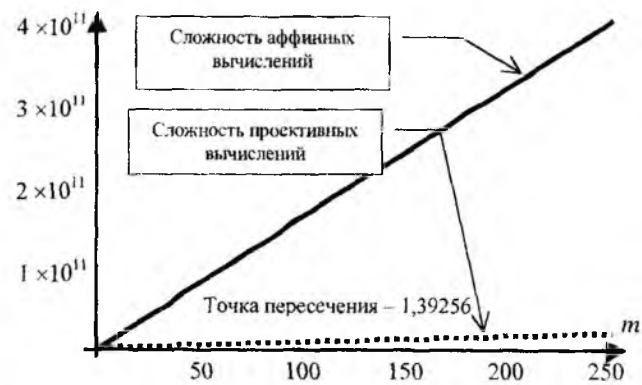


Рис.8

На рис. 9 и 10 приведены графики зависимости сложности умножений в аффинных и проективных координатах как функция величин d и m.

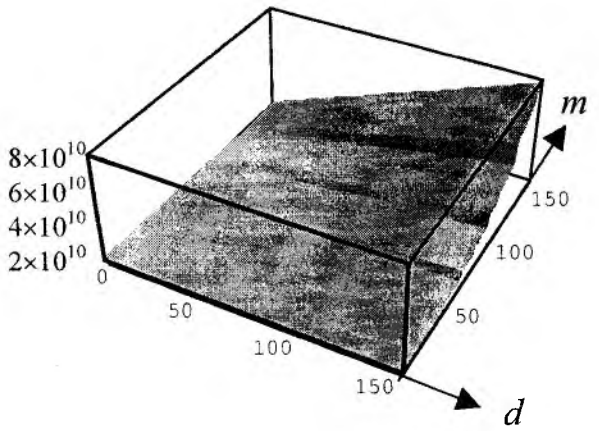


Рис. 9

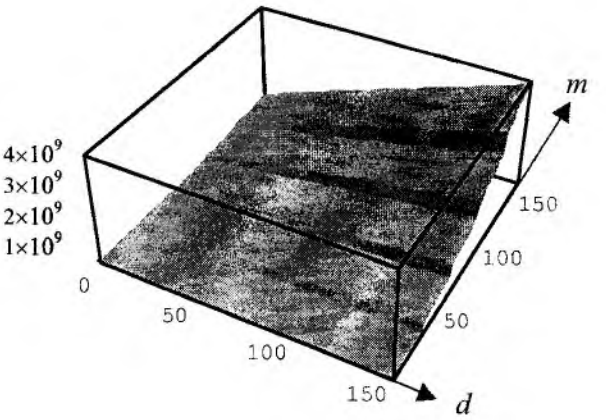


Рис. 10

Таблица 1

M	160	192	256
t_a	0,494	2,829	6,712
t_n	0,010	0,108	0,229

Для проверки сложности выполнения операции умножения (17) выполнены экспериментальные измерения с использованием стандартных библиотек, реализованных в аффинных и проективных координатах. В табл. 1 приведены значения среднего времени выполнения операции (17) в секундах при m = 160, 192 и 256, для аффинных t_a и проективных t_n координат с использованием ЭВМ на процессоре Celeron 600 MHz.

Экспериментальные результаты подтверждают теоретические.

Выводы

Анализ полученных результатов позволяет сделать вывод, что использование проективных координат при умножении является более предпочтительным, учитывая то, что минимальная длина поля, для обеспечения приемлемого уровня стойкости [3], должна быть не менее 160 бит. Меньшая сложность достигается при использовании проективных координат. Величина выигрыша в зависимости от значений d и m может быть определена из рис. 9 и 10.

Список литературы: 1. *X9.62 Public Key Cryptography For The Financial Services Industry*: The Elliptic Curve Digital Signature Algorithm (ECDSA). 1998. 87 с. 2. *X9.63 Public Key Cryptography For The Financial Services Industry*: Key Agreement and Key Transport Using Elliptic Curve Cryptography. 1999. 207 с. 3. *IEEE P1363/D11 (Draft Version 11).* Standard Specifications for Public Key Cryptography. Annex A (Informative). Number-Theoretic Background. 1999. 91 с. 4. *И.Д. Горбенко, С.И. Збитнев* Расширенное поле Галуа $GF(2^m)$. Вычислительная сложность простейших операций над расширенным полем $GF(2^m)$ // Радиотехника. 2000. Вып. 114. 10 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 19.03.2001

УДК 681.3.06

Д. И. ЛАВРИНЕНКО

ОПТИМАЛЬНО РАСШИРЕННЫЕ ПОЛЯ В АЛГОРИТМАХ ДЛЯ ЭЛЛИПТИЧЕСКИХ КРИВЫХ

Введение

Математические вычисления в конечных полях (полях Галуа) являются неотъемлемой частью всех криптографических систем, использующих несимметричные алгоритмы шифрования [1]. К таким системам относятся, в том числе, системы, основанные на решении дискретного логарифма [2], на эллиптических [3] и гиперэллиптических кривых [4]. К таким системам предъявляются жесткие требования по производительности, которые напрямую зависят от возможности быстро выполнять математические преобразования в полях Галуа. Особенностью таких систем является то, что элементами этих полей являются числа большой разрядности, которая в несколько раз превышает разрядность современных ЭВМ. Базовой и, следовательно, наиболее важной операцией в таких системах является операция модульного умножения. Следовательно, уменьшение вычислительной сложности данной операции приводит к пропорциональному росту производительности криптографической системы в целом.

В дальнейшем в статье внимание будет сконцентрировано на алгоритмах эллиптических кривых, как наиболее перспективных с точки зрения применения подхода, описанного в статье. Но ничего не мешает применить данный подход и методы в любых других криптографических алгоритмах, использующих вычисления в полях Галуа.

1. Проблемная область и существующие решения

Общий вид записи поля Галуа, в котором производятся преобразования, имеет вид $GF(p^m)$, где p – простое, а m – положительное целое. В криптографии обычно используются частные специальные случаи таких полей с целью уменьшения вычислительной сложности модульных преобразований в таких полях.

Случай $p = 2$ особенно привлекателен с точки зрения аппаратной реализации умножения в поле Галуа, поскольку элементы поля могут быть однозначно представлены сигналами логических «0» и «1». Однако с точки зрения программной реализации умножения случай $GF(2^m)$ не так привлекателен, как кажется на первый взгляд. Причина в том, что современные вычислительные системы оперируют числами большей разрядности, известными как слова. Для достаточно больших m, которые требуются при проектировании реальных криптографических систем, реализация умножения будет иметь большую вычислительную сложность. Существуют методы [3], позволяющие заметно повысить производительность операции умножения в полях $GF(2^m)$. Большинство из них основано на приведении $GF(2^m)$ к виду $GF((2^n)^m)$, где n максимально приближено к разрядности вычислительной сетки ЭВМ. Это позволяет разрешить вышеуказанную проблему, но здесь снижение вычислительной сложности происходит за счёт увеличения сложности пространственной. К тому же достигнутые результаты проигрывают другим существующим в настоящее время методам.

Аналогично, случай $GF(p)$ также имеет трудности реализации на современных ЭВМ. Для представления элементов поля в разрядной сетке ЭВМ необходимо несколько машинных слов, и здесь имеются две основные трудности при реализации операции умножения. Во-первых, при выполнении собственно умножения необходимо учитывать переносы между отдельными словами. Во-вторых, после умножения нужно выполнить модульное преобразование, которое имеет те же трудности реализации и такую же, если не большую, вычислительную сложность. В настоящее время существуют достаточно эффективные методы решения этих проблем, например, модульное преобразование Монтгомери [6]. Другой метод решения проблемы – подбор p специального вида для уменьшения вычислительной сложности собственно операции умножения, как описано в [7].

2. Оптимально расширенное поле и операции в нём

Объединим все вышеуказанные методы оптимизации с целью достижения максимальной производительности в $GF(p^m)$ реализации эллиптических и гиперэллиптических криптосистем. Чтобы

оптимизировать арифметику в поле, ограничим выбор параметров поля p и m следующими требованиями:

1. Выберем p возможно большим, но таким, чтобы оно помещалось в разрядную сетку машинного слова. Это требование позволит с максимальной эффективностью использовать быстрые процессорные команды.

2. Выберем p таким, чтобы оно было псевдопростым числом Мерсенна, то есть вида $2^n \pm c$, при $\log_2 c \le \frac{1}{2} n$. Это позволит выполнять операцию модульного преобразования над элементами поля с минимальной вычислительной сложностью.

3. Выберем m таким, чтобы в поле существовал несократимый бином вида $x^m - \omega$ для эффективного модульного преобразования в расширенном поле. Степень расширенности m тем меньше, чем большим может быть выбрано p.

Поле, выбранное таким образом, назовём оптимально расширенным полем (ОРП). Операции умножения в таком поле, как уже было указано ранее, позволяют совместно применять виды оптимизации, описанные в [7].

Для такого поля $GF(p)$ является подполем, m – степенью расширения, так что поле можно обозначить как $GF(p^m)$. Это поле является изоморфным по отношению к $GF(p)[x]/(P(x))$, где $P(x) = x^m + \sum_{i=0}^{m-1} p_i x^i$, $p_i \in GF(p)$ является нормированным несократимым полиномом степени m над $GF(p)$. Элемент поля $A \in GF(p^m)$ в каноническом (полиномиальном) виде можно записать как

$$A(x) = a_{m-1} x^{m-1} + \ldots + a_1 x + a_0,$$

где $a_i \in GF(p^m)$. Так как p выбиралось таким образом, чтобы не превышать разрядность машинного слова, то $A(x)$ можно представить с помощью m регистров или машинных слов.

Все математические преобразования в поле выполняются по модулю полинома поля. Выбор полинома поля определяет сложность операции модульного преобразования. В дальнейшем здесь будут описываться только операции сложения, умножения и возведения в квадрат и оценена их вычислительная сложность. Введём следующие обозначения для определения вычислительной сложности операций в подполе и расширенном поле:

I_{a0} – вычислительная сложность машинной команды сложения, вычитания или пересылки,

I_{l0} – вычислительная сложность машинной команды цикла или перехода,

I_{m0} – вычислительная сложность машинной команды умножения.

2.1 Сложение и вычитание

Сложение или вычитание двух элементов поля реализуется как последовательное сложение или вычитание коэффициентов соответствующих степеней полинома и, если необходимо, выполнение модульного преобразования по модулю p. Если сравнивать данную реализацию по отношению к операции сложения или вычитания в $GF(p)$, то можно заметить, что здесь не требуется учитывать переносы между машинными словами в процессе вычисления. Это является небольшим преимуществом по отношению к $GF(p)$. Приведём схему алгоритма сложения (схема алгоритма вычитания отличается незначительно и имеет аналогичную вычислительную сложность):

Алгоритм 1. Сложение в расширенном поле $GF(p^m)$

Исходные данные: $A(x) = a_{m-1} x^{m-1} + \ldots + a_1 x + a_0$, $B(x) = b_{m-1} x^{m-1} + \ldots + b_1 x + b_0$, $A(x), B(x) \in GF(p^m)$

Результат: $A(x) + B(x) = C(x) \in GF(p^m)$

for $i = 0$ to $m - 1$
$c_i = a_i + b_i$
if $c_i \geq p$
$c_i = c_i - p$
end if
end for

Вычислительная сложность алгоритма составит:

$$I_{addOEF} = m \cdot (I_{a0} + I_{l0} + \frac{1}{2} I_{a0}) = m \cdot I_{addsf} \quad (1)$$

2.2 Умножение

Операция умножения в ОРП выполняется в два этапа: собственно умножение и модульное преобразование. Так как операция умножения в поле является базовой операцией и, следовательно, самой критичной по вычислительной сложности для криптосистем с открытым ключом, то в дальнейшем мы уделим максимальное внимание всем аспектам реализации данной операции.

Первый этап представляет собой обычное полиномиальное умножение:

$$C'(x) = A(x) \times B(x) = c'_{2m-2} x^{2m-2} + \ldots + c'_1 x + c'_0 ; c'_i \in GF(p)$$

Данный метод требует m^2 умножений и $(m-1)^2$ сложений в подполе $GF(p)$. Следовательно, вычислительная сложность данной операции составит:

$$I_{mul} = I_{mulsf} \cdot m^2 + I_{addsf} \cdot (m-1)^2, \quad (2)$$

где I_{mulsf} и I_{addsf} обозначают вычислительную сложность умножения и сложения в подполе $GF(p)$ соответственно. Операция умножения в подполе $GF(p)$ является здесь ключевой и будет подробно рассмотрена ниже, в разделе 2.2.1.

Второй этап – модульное преобразование в поле $GF(p^m)$:

$$C(x) = C'(x) \bmod P(x); C(x) \in GF(p^m)$$

Обозначим вычислительную сложность данного этапа как I_{red}. Ниже, в разделе 2.2.2, будет представлен эффективный алгоритм для выполнения такого преобразования и оценена его вычислительная сложность.

Вычислительная сложность операции умножения в ОРП будет равна сумме вычислительных сложностей операций умножения и модульного преобразования, т.е.

$$I_{mulOEF} = I_{mul} + I_{red} \quad (3)$$

2.2.1 Умножение в подполе *GF(p)*

Как уже было упомянуто выше, производительность операции умножения в подполе $GF(p)$ является одним из определяющих факторов для производительности умножения в поле $GF(p^m)$.

Умножение двух элементов $a, b \in GF(p)$ выполняется как $a \times b \bmod p$. Благодаря тому, что разрядность элемента $GF(p)$ не превышает разрядности машинного слова, умножение можно выполнить непосредственно одной машинной инструкцией. Например, для 32-разрядных машин наибольшим простым числом, которое помещается в разрядную сетку, является $2^{32} - 5$. Следовательно. для такой ЭВМ нужно выбирать $p \leq 2^{32} - 5$. Подытоживая, можно рекомендовать использование p как можно большей величины, пока оно остаётся в пределах разрядной сетки.

При умножении двух целых чисел размером в машинное слово результатом, в общем случае, является целое размером в двойное машинное слово. Чтобы завершить вычисление результата, нужно

выполнить модульное преобразование. Тривиальное решение этой проблемы – получение остатка при делении двух целых чисел. Однако инструкция деления в современных ЭВМ имеет самую большую вычислительную сложность среди арифметических инструкций. Выбор параметра p в соответствии с требованием 2 (см. раздел 2) позволяет применить гораздо более эффективный алгоритм.

Хорошо известно, что возможно быстрое модульное преобразование по модулю вида $2^n - c$, где c является «малым» целым числом. Модуль такого вида позволяет выполнить модульное преобразование без операции деления. Здесь будет представлен алгоритм, базирующийся на алгоритме из [9]. Здесь будут рассматриваться только модули вида $2^n - c$, хотя незначительная модификация алгоритма позволяет использовать вид $2^n + c$. В приведённом ниже алгоритме операторы << и >> обозначают операции сдвигов влево и вправо соответственно.

Алгоритм 2. Модульное преобразование в подполе $GF(p)$

Исходные данные: $p = 2^n - c, \log_2 c \le \frac{1}{2} n,\ x < p^2$

Результат: $r \equiv x \bmod p$

$q = x >> n$
$res = x - q << n$
$r = res$
while $q > 0$
$t = qc$
$q = t >> n$
$res = t - q << n$
$r = r + res$
end while
while $r \ge p$
$r = r - p$
end while

В соответствии с выбранными требованиями главный цикл алгоритма будет выполнен, как максимум, два раза, т.е. вычислительная сложность алгоритма составит:

$$I_{redsf} = 2(I_{m0} + I_{l0}) + 12I_{a0} \tag{4}$$

Если n в точности соответствует разрядности машинного слова (а в практической реализации такой случай будет преимущественно использоваться), то операции сдвига можно выполнить простым присвоением, что позволит дополнительно снизить вычислительную сложность:

$$I_{redsf} = 2(I_{m0} + I_{l0}) + 6I_{a0} \tag{4'}$$

Полная вычислительная сложность умножения в подполе составляет:

$$I_{mulsf} = I_{m0} + I_{redsf} \tag{5}$$

При практической реализации данного алгоритма получается значительный выигрыш в производительности по отношению к выполнению непосредственного деления.

2.2.2 Модульное преобразование в расширенном поле $GF(p^m)$

После выполнения умножения элементов поля $GF(p^m)$ в полиномиальном представлении мы получаем промежуточный результат $C'(x)$. В общем случае степень $C'(x)$ больше или равна m. В этом случае требуется выполнить модульное преобразование. Канонический метод выполнения такого преобразования – полиномиальное деление с остатком на полином поля $P(x)$. Вычислительная сложность этого преобразования является одним из самых высоких среди операций арифметики многократной точности [6], и она тем выше, чем больше число членов в полиноме поля. Но если использовать полиномы специального вида, в частности, с малым количеством членов и малым весом коэф-

фициентов, то вычислительную сложность можно значительно снизить. Это обеспечивается требованием 3 (см. раздел 2) о том, чтобы полином поля имел вид:

$$P(x) = x^m - \omega .$$

По определению, $C'(x)$ имеет вид:

$$C'(x) = c'_{2m-2}\, x^{2m-2} + \ldots + c'_m\, x^m + c'_{m-1}\, x^{m-1} + \ldots + c'_1\, x + c'_0$$

Только члены $c'_{m+i}\, x^{m+i}, i \geq 0$ должны быть преобразованы по модулю $P(x)$. Заметим, что

$$c'_{m+i}\, x^{m+i} \equiv \omega c'_{m+i}\, x^i \bmod P(x);\ i = 0,1,\ldots,m-2$$

Так как степень $C'(x) \leq 2m-2$, требуется максимум $m-1$ умножения на ω и $m-1$ операций сложения для редуцированных членов.

Общий вид выражения полиномиального преобразования по модулю:

$$C(x) \equiv c'_{m-1}\, x^{m-1} + [\omega c'_{2m-2} + c'_{m-2}]x^{2m-2} \ldots + [\omega c'_{m+1} + c'_1]x + [\omega c'_m + c'_0] \bmod P(x)$$

Вычислительная сложность преобразования составляет:

$$I_{red} = (m-1)(I_{mulsf} + I_{addsf}) \qquad (6)$$

2.3 Возведение в квадрат

Операцию возведения в квадрат можно выполнить, используя метод операции умножения. Но хорошо известен метод, позволяющий сократить количество требуемых операций умножения в подполе почти в два раза. Этот метод требует $m(m+1)/2$ операций умножения в подполе вместо m^2 в операции умножения. Пропорционально уменьшается также и число операций сложения в подполе, а модульное преобразование будет точно таким же, как и для операции умножения. Следовательно, вычислительная сложность данной операции составит:

$$I_{sqrOEF} = (m(m-1)/2)(I_{mulsf} + I_{addsf}) + I_{res} \qquad (7)$$

2.4 Дополнительная оптимизация и улучшения

Обратим внимание на то, что два особых частных случая предоставляют возможность дополнительной оптимизации. Назовём их типами ОРП I и II.

Тип I имеет p вида $2^n \pm 1$. Такой случай позволяет выполнять модульное преобразование над элементами подполя с минимальной вычислительной сложностью. Тогда главный цикл алгоритма 2 будет выполняться только единожды, и в цикле не требуются операции умножения. В этом случае вычислительная сложность модульного преобразования в подполе будет равна

$$I_{redsfI} = 8I_{a0}, \qquad (8)$$

а вычислительная сложность всей операции умножения в расширенном поле

$$I_{mulOEFI} = I_{mul} + I_{red} = (m^2 + m - 1)I_{m0} + (9.5m^2 + 6.5m - 8)I_{a0} \qquad (9)$$

Тип II имеет несократимый бином вида $x^m - 2$. Полином такого вида позволяет ускорить модульное преобразование в расширенном поле. В этом случае операции умножения на ω можно заменить на операции сдвигов. Вычислительные сложности модульного преобразования и умножения в расширенном поле будут равны соответственно:

$$I_{redII} = (m-1)(\frac{3}{2}I_{a0} + I_{addsf}) \qquad (10)$$

$$I_{mulOEFII} = 2m^2 I_{m0} + (7{,}5m^2 - 1.5)I_{a0} \qquad (11)$$

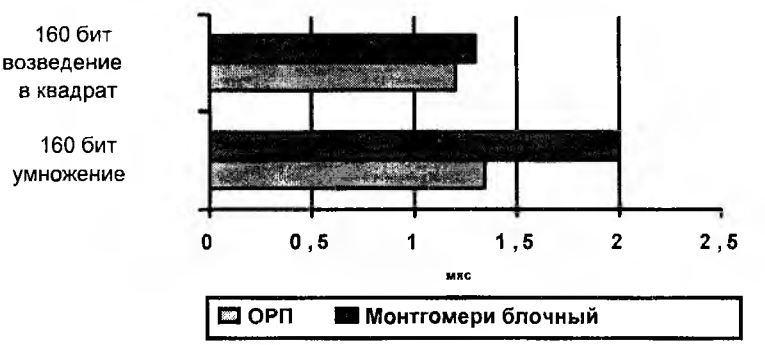
3 Экспериментальные результаты

Для получения экспериментальных результатов реализуем алгоритмы на наиболее распространённой в настоящее время архитектуре процессоров Intel Pentium II. Поскольку разрядность процессора – 32 бита, то выбор параметров был произведен следующим образом: $p = 2^{32} - 5$, $m = 5$, $\omega = 2$. Данный случай соответствует ОРП II и позволяет применить дополнительную оптимизацию.

Производительность операций умножения и возведения в квадрат в ОРП будем сравнивать с производительностью аналогичных операций в $GF(p)$, $p \le 2^{160}$ с применением блочной арифметики Монтгомери [6] как с наиболее производительной на текущий момент и аналогичной по криптостойкости. Теоретический расчёт вычислительной сложности будем производить, установив $I_{m0} = 4$, $I_{l0} = 2$, $I_{a0} = 1$. С расчётом вычислительной сложности операций в $GF(p)$ подробно можно ознакомиться в [6].

Вместе с теоретическим расчётом проведём вычислительный эксперимент оценки вычислительной сложности. Результаты были получены на ЭВМ с процессором Pentium II 366 MHz. Результаты показаны на графике.

Теоретические и экспериментальные данные сведены в таблицу.



Метод	Выч. сложность, квадрат	Выч. сложность, умножение	Время выполнения, квадрат	Время выполнения, умножение
Блочн. Монтгомери	485 ед.	825 ед.	1,304 мкс	1,983 мкс
ОРП	446 ед.	524 ед.	1.203 мкс	1,346 мкс

Исходя из полученных данных, можно заключить, что метод ОРП обеспечивает более высокую производительность по сравнению с блочной арифметикой Монтгомери.

Заключение

Принимая во внимание экспериментальные результаты арифметики в ОРП, можно рекомендовать данный подход для использования как в криптосистемах на базе эллиптических и гиперэллиптических кривых, так и в любых других, которые используют операции умножения и возведения в степень в полях Галуа. В ОРП можно также применить метод проективных координат [11], который позволит получить дополнительный выигрыш в производительности операции возведения в степень.

В данной статье были рассмотрены исключительно математические аспекты оптимизации математических преобразований в полях Галуа. Но ничто не мешает нам применить совместно с ними и другие методы оптимизации, например, аппаратно-зависимую для конкретного типа ЭВМ.

Список литературы 1. *Diffie W., Hellman* M.E. New directions in cryptography // IEEE Trans. on Information Theory. 1976. V.IT-22. №6. P. 644-654. 2. *Lenstra A.K., Manasse M.S.* Factoring with two large primes // Advances in cryptology – Eurocrypt '90. Berlin,1991. P. 72-82. 3. *R. Schroeppel, H. Orman, S. O'Malley, et. al.* Fast key exchange with elliptic cryptosystems // Advances in Cryptography – CRYPTO '95. 1995. 4. *Menezes A., Oorschot P., Vanstone S.* Handbook of Applied Cryptography. CRC Press, 1996 816 p. 5. *Кнут Д.* Искусство программирования для ЭВМ: В 3-х т. Т.2. Получисленные алгоритмы. М.: Мир, 1977. 387 с. 6. *Свинарёв А.В.* Методы и средства комбинированных несимметричных криптографических преобразований информации с уменьшенной вычислительной сложностью // На правах рукописи. 1998. 7. *Kenji Koyama and Yukio Tsuruoka.* Speeding up elliptic cryptosystems by used a signed binary window method // In Crypto' 92 Springer Lecture Notes in Computer Science. 1992.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 27.03.2001

УДК 681.3.06: 519.248.681

И. Д. ГОРБЕНКО, д-р техн. наук, П. В. КОЛЕСНИКОВ

ОЦЕНКА СТОЙКОСТИ RSA СИСТЕМ, В КОТОРЫХ ОТКРЫТЫЕ КЛЮЧИ ИЛИ ПАРАМЕТРЫ ЯВЛЯЮТСЯ ЛИЧНЫМИ

Введение

В 90-е годы широкое распространение получили криптографические системы, базирующиеся на преобразованиях с использованием математического аппарата колец, полей, групп, включая эллиптические группы. Необходимость таких систем можно объяснить требованиями практических задач, для которых должна быть реализована модель взаимного недоверия и взаимной защиты. Были разработаны и практически нашли применение криптографические алгоритмы под названием "алгоритмы с открытыми ключами" (RSA [1]), "с открытым распределением ключей" (Диффи-Хелмана), алгоритмы цифровой подписи с открытыми ключами, базирующиеся на RSA и преобразованиях класса Эль-Гамаля. В последние годы получили распространение криптографические алгоритмы, использующие группы точек в эллиптических кривых. Кроме названных алгоритмов определенное распространение получили и другие, но они по своей структуре базируются на описанных выше. Отличительная особенность их в том, что ключ делится на два, один из которых является личным, другой открытым. При этом принятие такого протокола существенно снижает стойкость криптозащиты, либо требует применения "усиленных" параметров криптографических преобразований. Физически криптографическое "ослабление" можно пояснить тем, что атака производится при известных открытых ключах и параметрах. Необходимость такой модели, на наш взгляд, связана с тем, что все санкционированные пользователи не доверяют друг другу и требуют защиты.

В то же время в ряде приложений можно обеспечить, с одной стороны, применение систем с открытыми ключами, а с другой – гарантировать конфиденциальность и целостность как личных, так и открытых ключей и параметров. В таких условиях криптоаналитические атаки могут осуществляться только извне информационной системы. На наш взгляд, весьма целесообразно проведение анализа и оценки криптостойкости в условиях атаки со стороны внешних пользователей, не имеющих доступа к личным и открытым ключам и параметрам. При этом чрезвычайно важным является оценка целесообразности использования такого режима. Анализ показал, что на RSA с закрытыми параметрами криптонападения возможны при следующих условиях:

1. Когда ключи E и D неизвестны, модуль N известен.
2. Для случая, когда E, D и N неизвестны.

Рассмотрим и оценим криптонападения при указанных условиях.

1. Классическая атака на системы с открытыми ключами

Цель криптоаналитической атаки на RSA-подобные системы заключается в нахождении секретного ключа E_k, и далее передаваемого сообщения M, исходя из известных открытых параметров N, D_k и криптограммы C. В ряде источников показано, что наилучшей среди известных атак на RSA является факторизация модуля преобразования N. Суть методики криптоанализа, которую может реализовать санкционированный пользователь, заключается в следующем.

1. Перехват сертификатов атакуемого объекта. N – модуль преобразования, D_k – открытый ключ.
2. Факторизация числа $N=P*Q$.
3. Вычисление функции Эйлера $\varphi(N)=(P-1)(Q-1)$.
4. Решение сравнения $E_kD_k=1 \bmod \varphi(N)$. При известных D_k и $\varphi(N)$ это уравнение решается за полиномиальное время.

Проведем анализ сложности этой атаки [2]. Факторизация N требует реализации субэкспоненциальной сложности. Известно значительное множество методов решения задачи факторизации. Основным из них является метод Полларда (ρ-метод), Ленстры с использованием аппарата эллиптических кривых, квадратичное решето числового поля, общее решето числового поля и другие. Сложность приведенных методов уменьшается в порядке, перечисленном выше. Сложность факторизации для общего решета числового поля составляет [5] :

$$I = e^{\delta \cdot (\ln N)^{\nu} \cdot (\ln \ln N)^{(1-\nu)}} \qquad (1)$$

Известно решение задачи, где параметры δ и ν равны [5] :

$$(\delta,\nu)=(1.96,\ 1/3) \qquad (2)$$

Сложность вычисления шагов 3, 4 приведенной методики криптоанализа является полиномиальной. В табл. 1 приведены оценки сложности факторизации методом обобщенного числового поля, полученные с использованием (1) и (2).

Таблица 1

Длина модуля N	Сложность факторизации I
256	$2^{47,56}$
512	$2^{65,15}$
1024	$2^{88,43}$

Второй способ криптоатаки – это атака на операцию возведения в степень по модулю и нахождение дискретного логарифма. Цель атаки состоит в нахождении передаваемого сообщения M, исходя из криптограммы C и известных открытых параметров N, D_k.

Из приведенной методики следует, что криптоанализ успешно может быть выполнен, если сертификаты N, E_k известны криптоаналитику. Представляет интерес случай, когда E_k, D_k – неизвестны. Рассмотрим его более подробно.

2. Атака на RSA-алгоритм при неизвестных открытом ключе D и известном модуле N

Напомним, что RSA-преобразования выполняются в кольце целых чисел, содержащем единицу, по модулю N. В таком кольце существует множество пар ключей (E_k, D_k), удовлетворяющих сравнению

$$E_kD_k\equiv 1 \bmod \varphi(N) \qquad (3)$$

Действительно, из ((3) следует, что E_k и D_k принимают значения, не превышающие $\varphi(N)$. Для чисел в кольце введены основные операции. Если предъявить требования, чтобы кольцо содержало "1" (т.е. существовало решение сравнения), тогда в качестве E_k могут быть выбраны такие числа, что (E_k, $\varphi(N)$)=1, поэтому число решений сравнения можно определить как [4]

$$n_{реш}=\varphi(\varphi(N))=\varphi((P-1)(Q-1))=\varphi(P-1)\,\varphi(Q-1) \qquad (4)$$

Предварительный анализ соотношения показывает, что количество решений зависит от вида разложения P-1 и Q-1. Например, если P-1 содержит большие числа, то $n_{реш}$ будет большим числом.

Вероятность подбора пары (E_k, D_k) равна

$$Вер(E_k, D_k)=Вер(E_k) \qquad (5)$$

В табл. 2 приведены вероятности подбора пары (E_k, D_k) при известном N.

Таблица 2

Длина числа N, (бит)	Для простых чисел	Для сильных простых чисел вида R|P-1	Для сильно простых чисел вида R|P-1, S|P+1
511	$2^{-508,72}$	$2^{-508,86}$	$2^{-508,86}$
1024	$2^{-1020,45}$	$2^{-1020,72}$	$2^{-1020,72}$
2048	$2^{-2044,22}$	$2^{-2044,45}$	$2^{-2044,45}$

Число попыток, которые необходимо выполнить для удачного подбора N, можно оценить как

$$n_N=P_{треб}*1/Вер(P,Q)\ , \qquad (6)$$

где $P_{треб}$- требуемая вероятность подбора N.

3. Атака на RSA алгоритм при неизвестных открытых модуле N и ключе D

Проведем анализ возможных атак и дадим оценку сложности их выполнения для случая, когда один или несколько параметров используются в режиме конфиденциальных параметров.

Известно, что, используя метод грубой силы, можно подобрать модуль N и ключи D_k, E_k. Выполняя атаку по этой схеме, нужно подобрать сначала модуль N. Определим количество вариантов перебора n_N. Величину n_N можно определить, зная количество простых чисел P,Q – n_P и n_Q.

Если P есть простое, то, используя теорему Чебышева [3], число простых чисел, находящихся в интервале [1, P-1], можно определить как:

$$n_p=\frac{P}{\ln P} \qquad (7)$$

В интервале $[P_1, P_2]$ число простых чисел Δn_P можно оценить как

$$\Delta n_P = \frac{P_2}{\ln P_2} - \frac{P_1}{\ln P_1} \tag{8}$$

где $[P_1, P_2]$ – заданный интервал.

Для Q аналогично находим:

$$\Delta n_Q = \frac{Q_2}{\ln Q_2} - \frac{Q_1}{\ln Q_1} = \Delta n_P \tag{9}$$

Если P простое число общего вида, то вероятность подбора числа N можно определить через вероятности подбора чисел P и Q:

$$\text{Вер}(P, Q) = \text{Вер}(P)\ \text{Вер}(Q) = \frac{1}{\Delta n_P} \cdot \frac{1}{\Delta n_Q} =$$

$$= \frac{\ln P_1 \cdot \ln P_1}{P_2 \cdot \ln P_1 - P_1 \cdot \ln P_2} \cdot \frac{\ln Q_1 \cdot \ln Q_1}{Q_2 \cdot \ln Q_1 - Q_1 \cdot \ln Q_2} \tag{10}$$

Для случая многократного подбора:

$$\text{Вер}(N) = n * \text{Вер}(P, Q), \tag{11}$$

где n – число попыток

В табл. 3 приведены оценки вероятности подбора числа N для случая, когда l_N=256, 512, 1024, 2048 бит.

Таблица 3

Длина числа N, (бит)	Для простых чисел (экспериментально)	Для простых чисел (по формуле)	Для простых чисел R|P-1	Для простых чисел R|P-1, R|P+1
511	$2^{-495,08}$	$2^{-495,05}$	$2^{-493,72}$	$2^{-492,69}$
1024	$2^{-1005,0}$	$2^{-1005,05}$	$2^{-1003,58}$	$2^{-1002,55}$
2048	$2^{-2027,19}$	$2^{-2027,05}$	$2^{-2025,44}$	$2^{-2024,41}$

Рассмотрим дальше задачу "взлома" RSA – систем при неизвестных открытых параметрах E, D и N. Безопасное время равно:

$$t_{без} = P_{треб} * n_N / (\gamma K) \tag{12}$$

Очевидно, что полная вероятность успешного криптоанализа, ввиду независимости событий равна:

$$\text{Вер}(N, (E, D)) = \text{Вер}(P,Q) * \text{Вер}(E_k) = (\ln P_1 \ln P_2)/(P_1 \ln P_2 - P_2 \ln P_1) * 1/(\varphi(P-1)\ \varphi(Q-1)) \tag{13}$$

В табл. 4 приведены значения вероятностей подбора N при неизвестных открытых параметрах E, D, N для прямой атаки при использовании простых чисел P, простых чисел в узком смысле и широком смысле.

Таблица 4

Длина числа N, (бит)	Для простых чисел P и Q	Для сильных простых чисел R|P-1	Для простых чисел R|P-1, S|P+1
511	$2^{-1003,77}$	$2^{-1002,58}$	$2^{-1001,55}$
1024	$2^{-2025,50}$	$2^{-2024,30}$	$2^{-2023,27}$
2048	$2^{-4071,27}$	$2^{-4069,89}$	$2^{-4068,86}$

Для случая, когда N – известно, E_k, D_k – неизвестны, необходимо сначала выполнить факторизацию N, а затем найти E_k и D_k.

4. Экспериментальная оценка количества сильных простых чисел

Если P, Q – ограниченные числа, например сильные простые числа в узком или широком смысле, то оценку для количества простых чисел P заданной длины можно получить экспериментально.

Суть эксперимента заключается в определении плотности распространения простых чисел на оси натуральных чисел N. Для обычных простых чисел [4]:

$$n_p = \frac{P}{\ln P} \qquad (14)$$

Плотность распределения простых чисел можно определить, взяв производную по P

$$n_p' = \frac{P' \cdot \ln P - P \cdot \frac{1}{P}}{\ln^2 P} = \frac{\ln P - 1}{\ln^2 P} = \frac{1}{\ln P} - \frac{1}{\ln^2 P} \qquad (15)$$

Если пренебречь $\frac{1}{\ln^2 P}$, то среднее расстояние Δr между простыми числами P_i и P_{i+1} на оси N можно оценить как:

$$\Delta r = \frac{1}{\ln P} \qquad (16)$$

Цель эксперимента состоит в оценке количества простых чисел из множества допустимых с длиной 256, 512, 1024 бит. Для сильно простого числа в широком смысле справедливы соотношения:

$$R_1 | P-1,\ R_2 | P+1,$$

$$l_{R1} \geq l_P/2,\ l_{R2} \geq l_P/2 \qquad (17)$$

1. Найдем сначала количество простых чисел:

а) в заданном диапазоне его можно найти путем определения количества таких чисел исходя из малого диапазона чисел длины 256, 512, 1024 бит;

б) далее, зная распределение простых чисел в малом диапазоне, сделаем оценку их количества во всем диапазоне.

Результаты эксперимента приведены в табл. 5. Результаты эксперимента можно сравнить с теоретическими данными, полученными с использованием формулы (16)

2. Оценим количество сильных простых чисел в широком и узком смысле. Для чисел большой длины эта задача не имеет простого решения, т.к. для этого необходимо разложить большие P-1 и P+1 на множители. Найдем оценки, следующим образом:

а) рассмотрим все числа длиной: 11, 12,13,...,32 битов;

б) для каждой длины найдем среднее значение количества простых чисел из 1000 рассмотренных с данной длиной;

в) найдем зависимость количества простых чисел в диапазоне длиной 1000 от длины рассматриваемого числа. Используя (16) получим, что количество простых чисел среди $k_{расам}$ чисел, сравнимых по длине с P, равно:

$$n_P = \frac{k_{рассм}}{\ln P} \qquad (18)$$

Если учесть тот факт, что длина числа P в битах равна l_P, перепишем (18) в следующем виде:

$$n = k_{рассм} \cdot \frac{\log_2 e}{\log_2 P} = k_{рассм} \cdot \frac{1{,}44}{l_P} \qquad (19)$$

где n – количество простых чисел среди $k_{рассм}$ чисел длины lP;

г) по экспериментальным данным найдем закон зависимости количества простых чисел в интервале длиной 1000 от длины простого числа. Далее обобщим полученные результаты на числа большой длины: 256, 512, 1024 бит.

Для обыкновенных простых чисел зависимость их количества от длины числа подчиняется закону (20). Для определения количества сильных простых чисел воспользуемся модифицированной формулой:

$$n = k_{рассм} \cdot c_1 \cdot \frac{1{,}44}{l_P^{c_2}}, \qquad (20)$$

где коэффициенты c_1 и c_2 будут подбираться экспериментально. В табл. 5 Количество простых и сильно простых чисел заданной длины.

Таблица 5

Длина числа P, бит	Количество простых чисел (экспериментально)	Количество сильных простых чисел R|P-1 (экспериментально)	Количество сильных простых чисел R|P-1, S|P+1 (экспериментально)	Количество простых чисел по формуле (19)	Количество простых чисел по формуле (20) c1=1,07 c2=0,93	Количество простых чисел по формуле (20) c1=1,07 c2=0,65
11	134	112	83	131	103	72
12	124	87	55	120	94	66
13	113	90	65	111	86	60
14	107	77	52	104	80	56
15	99	76	58	96	74	52
16	93	66	45	90	69	48
17	87	68	51	85	65	45
18	82	59	41	80	61	43
19	78	59	45	75	57	40
20	75	52	35	72	54	38
21	72	53	39	69	52	36
22	68	48	32	66	49	34
23	65	48	35	63	47	33
24	63	44	30	60	45	31
25	60	44	32	58	43	30
26	58	40	27	55	41	29
27	56	40	29	53	39	28
28	54	37	26	52	38	27
29	52	38	27	50	37	26
30	50	34	24	48	35	25
31	48	35	25	47	34	24
32	46	32	22	45	33	23

На рисунке показаны результаты экспериментальной и теоретической оценки числа простых чисел на 1000 чисел.

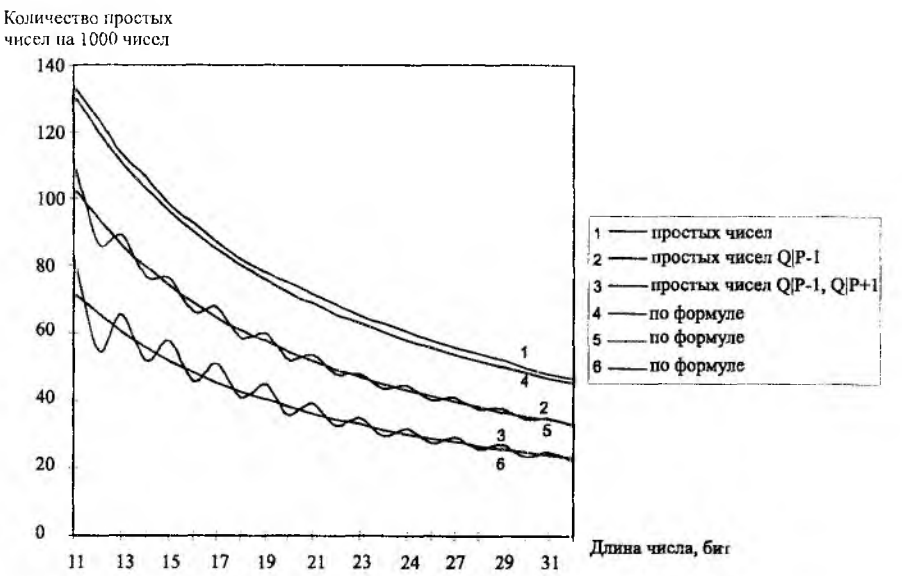


Рис. 1

На рис. 1 показана зависимость количества простых чисел на 1000 исследуемых от длины числа. Здесь объединены теоретические и практические результаты. В таблице 6 даны оценки количества простых чисел во всем диапазоне чисел заданной длины для реальных значений P и Q.

Таблица 6

Длина выборки	Длина числа P, Q, (бит)	Количество простых чисел в диапазоне $[1, 2^n]$ (экспериментально)	Количество простых чисел по формуле (16)	Количество простых чисел по формуле (20) R|P-1	Количество простых чисел по формуле (20) R|P-1, R|P+1
2^{255}	256	$2^{247,54}$	$2^{247,52}$	$2^{246,86}$	$2^{246,34}$
2^{511}	512	$2^{502,51}$	$2^{502,52}$	$2^{501,79}$	$2^{501,27}$
2^{1023}	1024	$2^{1013,59}$	$2^{1013,52}$	$2^{1012,72}$	$2^{1012,20}$

5. Экспериментальная оценка значения функции Эйлера

Из теории чисел известно [3], что значение функции Эйлера для любого целого числа n, имеющего каноническое разложение вида

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots \cdot p_n^{\alpha_n}$$

можно определить как

$$\varphi(n) = p_1^{\alpha_1 - 1} \cdot (p_1 - 1) \cdot \ldots \cdot p_n^{\alpha_n} \cdot (p_n - 1) \qquad (21)$$

Оценим значение функции Эйлера для числа некоторой длины l_n. Минимальное ее значение будет соответствовать числу, составленному из простых чисел 2,3,5,7..., идущих подряд, начиная с 2. Это значение будет минимальным так-так в этом случае в расчете $\varphi(n)$ множитель (p_i-1) будет встречаться максимальное число раз, а все остальные множители будут равны p_i. Экспериментально найдем минимальную оценку для $\varphi(n)$. Результаты эксперимента приведены на рис. 2. По оси x показана длина числа n в битах, а по оси y значение разности длины чисел n от длины значения функции Эйлера $\varphi(n)$.

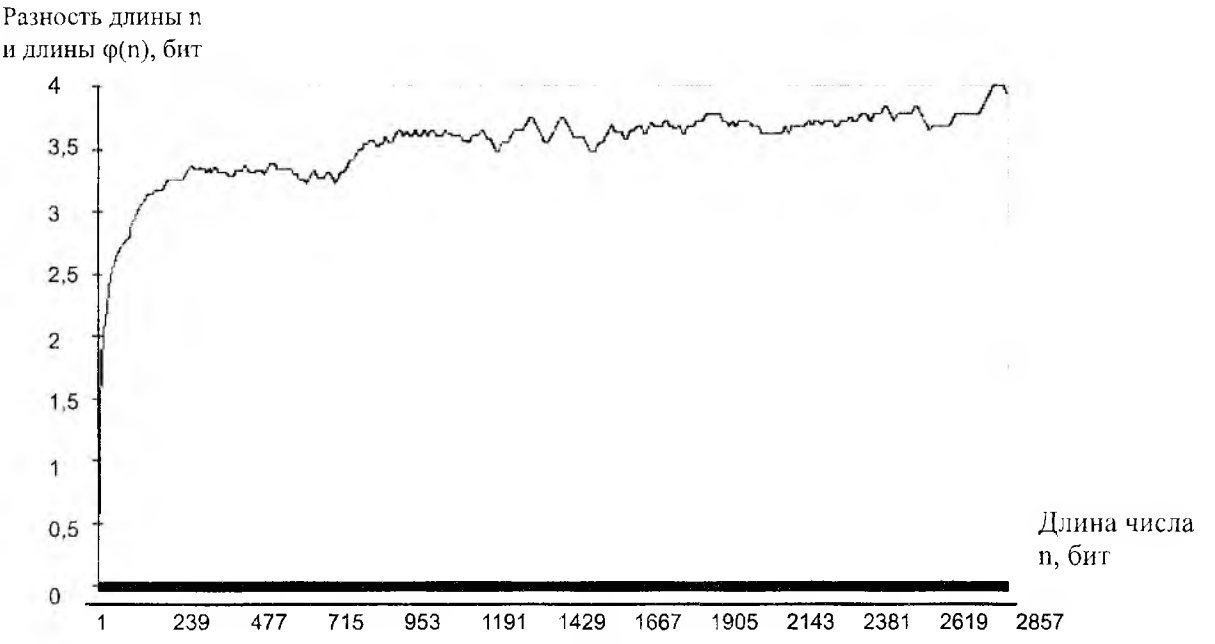


Рис. 2

Приведем пример оценки длины функции Эйлера для длины модуля преобразования N RSA-системы равному 256 бит. Из условия задачи следует, что число N состоит из двух простых множителей длиной $l_P=l_Q=128$ бит. Из $\varphi(N)= (P-1)*(Q-1)$ следует, что $l_{\varphi(N)}=256$ бит. Зная, что минимальное значение длины функции Эйлера можно вычислить по формуле

$$\varphi(\varphi(N)) = \varphi((P-1)*(Q-1)) \qquad (22)$$

можно оценить минимальную и максимальную границы для функции Эйлера в следующем виде: min $l_{\varphi(\varphi(N))}$= 250 бит, max $l_{\varphi(\varphi(N))}$= 256 бит.

Разница длин $l_{\varphi(\varphi(N))}$ и $l_{\varphi(N)}$ была определена в ходе эксперимента и составляет максимум 3 бита. Для получения результата предполагалось, что разложение чисел P-1, Q-1 может содержать сомножители любой длины. Используя результаты рис. 2 можно оценить степень уменьшения числа ключей E_k, если известна длина модуля N.

Заключение

Криптосистемы с открытыми ключами нашли применение при реализации модели взаимного недоверия и взаимной защиты. В этом случае один из ключей является личным, а второй открытым. Личный должен храниться как конфиденциальный и не выходить из-под влияния пользователя. Открытый ключ доступен всем пользователям сети. Основным требованием к открытому ключу является обеспечение его целостности и аутентичности на всех этапах жизненного цикла. Объявление одного из ключей открытым, существенно снижает криптостойкость. Если криптосистема с открытыми ключами применяется в информационных системах, в которых действует модель взаимного доверия, то в ней открытые ключи и/или параметры, например, значение модуля RSA-преобразования N, могут объявляться и использоваться на всем жизненном цикле как личные. В этом случае криптостойкость систем защиты существенно повышается, и она начинает приближаться по стойкости к симметричным системам. Наибольшая степень повышения стойкости достигается, если в RSA-системе и модуль N и открытый ключ объявляются и существуют в системе как личные. Если N или один из ключей являются открытыми, то в этом случае стойкость понижается, но остается на много выше, чем в стандартной RSA криптосистеме, в которой модуль N и ключ E являются открытыми.

В то же время, на наш взгляд, применение RSA систем в режиме защиты модуля или открытого ключа является оправданным, так как личный ключ остается конфиденциальным. Это означает, что при его использовании для цифровой подписи, лицо или объект, подписывающий информацию, несет за нее ответственность так же, как в системе взаимного недоверия. Именно в этом преимущества применения RSA системы в предлагаемых режимах обеспечения конфиденциальности открытых ключей и/или модулей преобразования.

Если RSA применяется в режиме конфиденциальности открытых ключей или модулей преобразования, то длины ключей или модулей можно уменьшить до 256 битов. Это позволит повысить скорость и уменьшить сложность прямых и обратных преобразований, причем, стойкость будет оставаться близкой к стойкости симметричных криптосистем.

Авторы будут признательны публичному обсуждению такого режима применения RSA системы и возможных областей применения. Считаем, что RSA система все-таки будет проигрывать большинству симметричных криптосистем, по крайней мере, с эквивалентной длиной ключа.

Список литературы: 1. *Диффи У.* Первые десять лет криптографии с открытым ключом // ТИИЭР. 1998. Т.76, №5. С.54-74. 2. *RSA: мифы и реальность* /Под. ред. И.Д. Горбенко, Е.Г. Качко, А.В. Свинарев и др. // Безопасность информации 1996. №2. С. 17-25. 3. *Виноградов И.М.* Основы теории чисел М.: Наука, 1981. 176 с. 4. *Menezes A., Oorschot P., Vanstone S.* Handbook of applied cryptography. CRC Press, 1996. 816 p. 5. *Buchmann J., Loho J., Zayer J.* An implementation of the general number field sieve. New York: Advances in cryptology: Proc. of Crypto'93. 1994. P.13-26.

Харьковский государственный технической университет радиоэлектроники

Поступила в редколлегию 27.03.2001

МЕТОДЫ АУТЕНТИФИКАЦИИ: ПРОБЛЕМЫ И ПРИНЦИПЫ РЕАЛИЗАЦИИ

УДК 681.3.06:519.248.681

О. А. ЗАМУЛА, канд. техн. наук, Г. М. ГУЛАК, Є. В. ПОПОВИЧ

МЕТОДИ АВТЕНТИФІКАЦІЇ В БЕЗУМОВНО СТІЙКИХ КРИПТОСИСТЕМАХ

Вступ

Згідно встановлених на сьогодні норм та вимог нормативних документів системи захисту інформації повинні надавати користувачу послуги спостереженості, доступності (управління доступом) та цілісності. Теоретичною основою побудови систем та засобів надання вказаних послуг є загальна теорія автентичності (справжності). На сьогодні ця теорія одержала значний розвиток, особливо в практичній площині. Але, на наш погляд, вона потребує подальшого узагальнення, класифікації, пояснення нових задач та проблемних питань, визначення методів та способів їх розв'язання. Метою статті є розгляд та аналіз методів та проблемних питань забезпечення автентичності в різноманітних інформаційних технологіях, включаючи телекомунікаційні системи. В даній статті ми розглянемо загальні питання теорії автентифікації та здійснення автентифікації в безумовно стійких системах. Справа в тому, що на сьогодні в більшості інформаційних технологій більш гостро стоять питання цілісності, автентичності та доступності. Вважаємо за необхідне запропонувати для обговорення та можливого використання наступні поняття та визначення.

Цілісність інформації – це властивість інформації, яка полягає в тому, що вона не може бути змінена випадково або навмисно неавторизованим користувачем і/або процесом і може бути використана за призначенням.

Доступність – властивість ресурсу системи або комп'ютерної системи (автоматизованої системи), яка полягає в тому, що авторизований користувач і/або процес, який володіє відповідними повноваженнями, може використати ресурс згідно з правилами і з визначеною якістю, в тому числі за рахунок використання криптографічних перетворень.

Спостереженість – властивість ресурсу системи (комп'ютерної системи, об'єкта комп'ютерної системи, інформації), що дозволяє реєструвати роботу користувачів та процесів, використання ресурсу системи, однозначно установлювати ідентифікатори (імена) причетних до певних подій користувачів та процесів, а також реагувати на ці події з метою мінімізації можливих втрат в системі, у тому числі за рахунок використання криптографічних перетворень.

Аналіз показує, що для забезпечення автентичності необхідно розглядати всі об'єкти та суб'єкти, що мають інформаційні співвідношення. В якості основи для побудови системи захисту повинна бути вибрана модель взаємної недовіри і взаємного захисту. Схематично така модель подана на рис. 1.

В даній моделі, з точки зору інформаційних співвідношень, приймають участь джерело та приймач інформації, арбітр, криптоаналітик. Для такої моделі джерело інформації, приймач інформації та арбітр є сторонами, що не довіряють один одному та повинні бути захищені від обману. Автентичність, як показав аналіз, може досягатися за рахунок розв'язання наступних задач:

1. Встановлення справжності користувача, що намагається вступити в інформаційні співвідношення (доступ до інформації, що захищається до ресурсу системи і т.п.).

2. Автентифікація системи - процедура встановлення справжності мережі, системи, до якої отримано доступ.

3. Автентифікація програмного забезпечення та даних - процедура встановлення цілісності програмного забезпечення та даних, які протягом деякого часу могли знаходитись за межами контролю володаря, а також підтвердження їх справжності (авторства).

4. Автентифікація повідомлень - процедура перевірки цілісності повідомлень та підтвердження авторства.

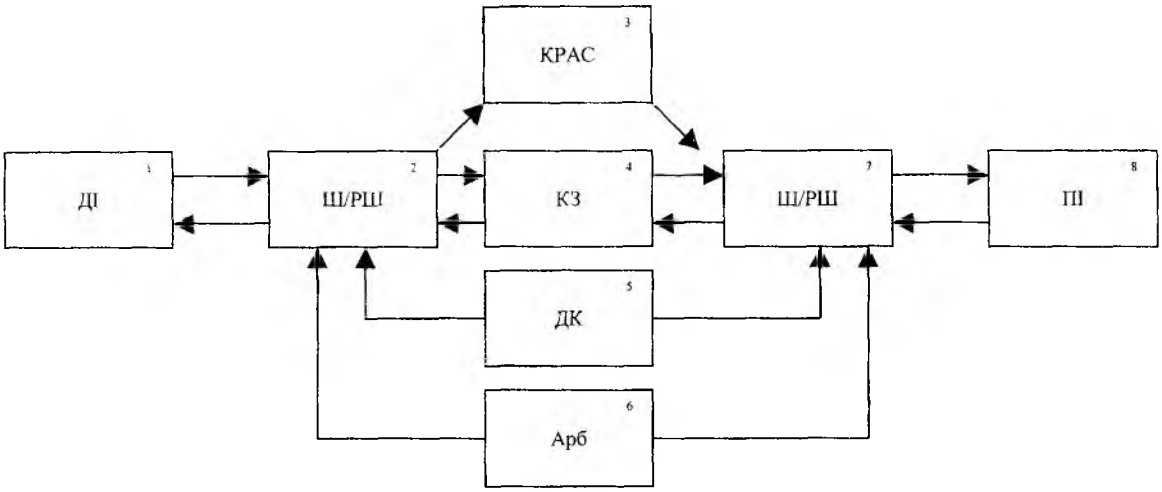


Рис. 1

Введені позначення:
1, 8 – джерело (приймач) інформації (ДІ, ПІ);
2, 7 – пристрої зашифрування/розшифрування (ЗШ/РШ);
3 – криптоаналітична система (КРАС);
4 – канал зв'язку та телекомунікаційна система, носій інформації;
5 – джерело ключа (ДК); 6 – арбітр (Арб);
8 – приймач (джерело) інформації.

2. Основні загрози порушення автентичності

Проведений аналіз показав, що в розглянутій моделі можуть бути реалізовано ряд загроз порушення автентичності [1-2].

Зі сторони об'єкта або суб'єкта А:

А1: абонент А формує повідомлення M_i та надсилає його абоненту В, а потім відмовляється від факту надсилання повідомлення M_i.

А2: абонент не формував і не передавав повідомлення M_i, але стверджує, що передавав.

А3: абонент А передав повідомлення M_j, а стверджує, що передав M_i.

А4: абонент передав повідомлення M_i у момент часу t_v, а стверджує, що передав повідомлення у $t_v \pm \Delta t$.

Під час реалізації загроз абонент А буде прагнути максимізувати обсяг втрати користувача В.

Зі сторони об'єкта або суб'єкта В:

В1: абонент В отримує від абонента А повідомлення M_i, а потім відмовляється від факту його отримання.

В2: абонент В отримує повідомлення M_i, а потім змінює його на M_i і стверджує, що отримав саме його.

В3: абонент А надсилає повідомлення у момент часу t_v, абонент В отримує це повідомлення у $t_v + \Delta t_1$, а стверджує, що отримав його у $t_v + \Delta t_2$.

В4: абонент В створює повідомлення M_i, а стверджує, що отримав його від абонента А.

Зі сторони арбітра:

АР1: арбітр може видати неправильне рішення відносно аналізу вищезазначених загроз Аі або Ві.

АР2: арбітр може прийняти неправильне рішення по відношенню до обох абонентів з метою збільшення свого виграшу.

АР3: арбітр приймає сторону зловмисника та компрометує систему (розголошує її)

КРАС1: імітація неправдивої криптограми з ймовірністю $P_i(C')$.

КРАС2: заміна істинної криптограми C неправдивою C' з ймовірністю $P_n(C')$.

КРАС3: повторна передача повідомлення, яке було передано раніше з ймовірністю $P_{pn}(C)$.

КРАС4: дезорганізація системи за рахунок передавання неправдивих команд управління.

КРАС5: модифікація повідомлення з ймовірністю $P_м(C)$.

У вірно спроектованій системі захисту, відповідно до політики безпеки, повинні бути перекриті основні загрози та мінімізовані втрати відповідно з вимогами до системи. Наприклад, під час вибору

стратегії подій криптоаналітик буде намагатися максимально збільшити ступінь втрат, тобто максимізувати ймовірність обману $P_{обм}$ [1]:

$$P_{обм}=max\{P_i,P_n,P_{рп},P_M,...\}, \quad (1)$$

де $P_i,P_n,P_{рп},P_M,...$ відповідні ймовірності, задані в наведеній вище моделі загроз. В якості загроз криптоаналітику необхідно вибирати одну або декілька, але таких, які мають найбільшу ймовірність успіху.

Покладемо, що на виході джерела повідомлень формується повідомлення M_i, де $i = \overline{l..n_M}$. Під час відображення повідомлення M_i, в криптограму C_i кількість станів джерела криптограм змінюється в межах $j = \overline{l..n_C}$. Якщо відомі розміри множини повідомлень n_M та множини криптограм n_C, то ймовірність обману в загальному випадку для моделі, що визначена на рисунку 1 $P_{обм}$, можна визначити як:

$$P_{обм} = \frac{n_M}{n_C}. \quad (2)$$

Аналіз співвідношення (2) показує, що:

1. Неможливо досягнути ймовірності $P_{обм}= 0$ тому, що для цього потрібно, щоб $n_C \to \infty$, а $n_M \to 0$.
2. З метою зменшення ймовірності $P_{обм}$ необхідно, щоб $n_M << n_C$.
3. Множина станів джерела криптограм може бути збільшена за рахунок внесення надлишковості, перш за все за допомогою збільшення довжини криптограм ($l_C >> l_M$), і, як наслідок, за рахунок збільшення кількості можливих криптограм.
4. Якщо $n_C = n_M$, то $P_{обм}= 1$.

3. Основні положення теорії автентичності Сімонсона

Аналіз джерел [1-3] показав, що на нинішній час з використанням інформаційного підходу розроблена тільки теорія оцінок ймовірності обману, яка одержала назву теорії Сімонсона [2].

В моделі, що розглядає Сімонсон, існує три учасника інформаційних співвідношень - джерело, одержувач та криптоаналітик. Водночас вважалося, що з метою автентифікації використовується, як і в схемі Шеннона, одноразовий ключ, а криптоаналітик знаходиться між джерелом та одержувачем (рис. 2).

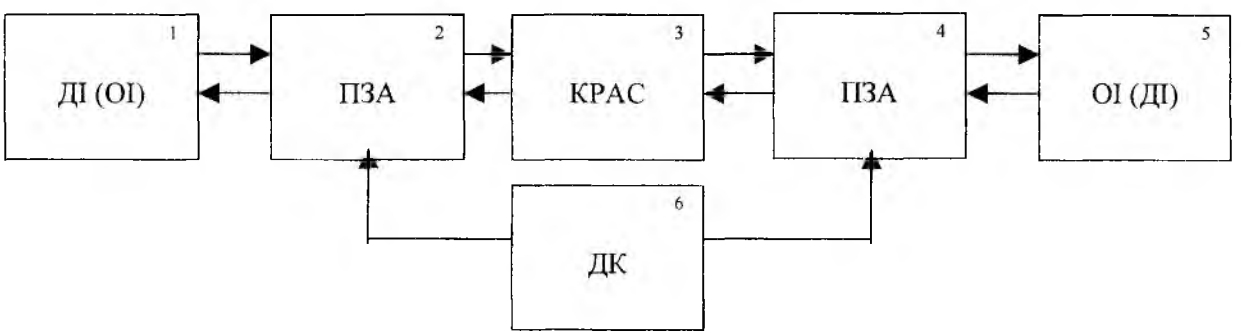


Рис. 2

Введені позначення:
1 – джерело інформації;
2, 4 – пристрої забезпечення автентифікації (ПЗА);
3 - криптоаналітична система (КРАС);
5 - одержувач (джерело) інформації; 6 - джерело ключів.

Сімонсон показав, що ймовірність обману в сенсі імітації P_i можна визначити, як [2]:

$$log_2 P_{обм} \geq -I(C,K), \quad (3)$$

де $I(C,K)$ - надлишковість (інформація), що вводиться з метою рішення задачі автентіфікації, наприклад ключа автентіфікації K_a.

В цьому випадку ключ автентифікації використовується з метою рішення завдання автентифікації, а ключ шифрування - для забезпечення конфіденційності. Тобто з теорії Сімонсона витікає, що з

метою забезпечення послуги конфіденційності повинен використовуватись ключ шифрування $K_ш$, а з метою забезпечення послуги автентифікації (цілісності та справжності) повинен використовуватися ключ автентифікації K_a.

Якщо $I(C,K)$ – взаємна інформація між джерелом криптограм та ключів, то [1, 2]

$$I(C,K)=H(C)-H(C/K). \tag{4}$$

$$H(C)=\sum_{i=1}^{n_C} P(C_i)\log P(C_i). \tag{5}$$

Розв'язуючи (3) отримаємо, що

$$P_{обм} \geq 2^{-I(C,K)}. \tag{6}$$

У випадку m - ічного алфавіту

$$P_{обм} \geq m^{-I(C,K)}. \tag{7}$$

Співвідношення (6) та (7) дозволяють отримати нижню оцінку для ймовірності обману, але наскільки реальна ймовірність буде близькою до неї, визначити неможливо. Сімонсон визначив ідеальну систему автентифікації, як криптографічну систему, в якій досягнута нижня межа ймовірності обману, тобто

$$P_{обм} = m^{-I(C,K)}. \tag{9}$$

Для випадку, коли в якості $I(C,K)$ можлива надлишковість, яка вводиться з метою забезпечення автентифікації у вигляді ключа автентифікації K_a довжиною $l_н$ коду автентіфікації повідомлень (КАП) та покласти

$$I(C,K) = l_н. \tag{10}$$

де індекс $н$ означає надлишковість. Тоді співвідношення (10) приймає вигляд:

$$P_{обм} \geq 2^{-l_н}. \tag{11}$$

З даного співвідношення випливає, що з метою зменшення ймовірності обману необхідно збільшувати надлишковість.

4. Аналіз методів забезпечення автентичності в безумовно стійких криптосистемах

Розглянемо загальні умови та методи забезпечення автентичності в безумовно стійких криптосистемах. Для них є вірними наступні ствердження про умови та стійкість в сенсі автентичності, яка може бути досягнена:

Ствердження 1. Нехай в системі Вернама здійснюється зашифрування інформації за правилом $C_i=(M_i+K_i) mod\ m$, а розшифрування – за правилом $M_i=(C_i-K_i) mod\ m$, тобто реалізується поточний метод шифрування. Тоді застосування поточного шифру Вернама [2] є необхідною, але недостатньою умовою забезпечення автентичності. Розглянемо простий випадок, коли $m=2$. В цьому випадку $C_i=(M_i \oplus K_i)$. Нехай в системі присутній криптоаналітик (відповідно до рисунку 2) та який перетворює C_i шляхом складання її з випадковою або спеціальною послідовністю символів R_i. Тоді на виході криптоаналітичної системи криптограма має вигляд:

$$C_i' = C_i \oplus R_i = M_i \oplus K_i \oplus R_i. \tag{11}$$

Одержувач здійснює розшифрування за правилом:

$$M_i' = C_i' \oplus K_i = M_i \oplus K_i \oplus R_i \oplus K_i = M_i \oplus R_i. \tag{12}$$

Аналіз (12) показує, що, якщо $R_i \neq 0$, то $M_i' \neq M_i$. За результатами криптоаналітичної атаки M_i відтворилося в M_i', де M_i (залежно від R_i) може бути як випадковим, так і повідомленням, яке має певний зміст та дозволене в даній системі, наприклад, «ТАК» може бути змінено на «НІ». Таким чином, в класичній системі Вернама завжди існує можливість заміни змісту повідомлень, тобто порушення автентичності.

Для довільної потужності алфавіту m:

$$C_i' = (C_i+R_i) mod\ m = (M_i+K_i+R_i) mod\ m, \tag{13}$$

а при розшифруванні:

$$M_i^{'}=(C_i^{'}-K_i)mod\ m=(M_i+K_i+R_i-K_i)mod\ m=(M_i+R_i)mod\ m. \quad (14)$$

Таким чином, для довільної потужності алфавіту в поточній системі Вернама застосування поточного шифрування є необхідною, але недостатньою умовою забезпечення автентичності. Твердження 1 доведено.

Розглянемо приклад, який підтверджує висновок загальної теорії Сімонсона про те, що з метою забезпечення автентичності необхідно внесення до криптограми надлишковості, а в класичній системі Вернама надлишковість відсутня.

Покладемо, що в системі Вернама, з метою виявлення втручань криптоаналітика, застосовуються групові лінійні коди, тобто використовується попереднє кодування інформації. Для цього випадку справедливе ствердження 2.

Ствердження 2. Якщо в системі Вернама повідомлення створюється за результатами кодування блоків M_{ij}, де j - кодова комбінація систематичного коду (групового коду визначення помилок), які створюються з повідомлення M_i додаванням $M_{ij}^{надл}$:

$$M_j = M_{ij} \mid M_{ij}^{надл}, \quad (15)$$

тобто перетворюємо повідомлення M_i з довжиною l_m в повідомлення M_j з довжиною $l_m + l_{надл}$:

$$M_i^{l_m} \Rightarrow M_i^{l_m} l_m^{надл} = M_j, \quad (16)$$

де $j = \overline{1, l_m + l_{надл}}$, тоді застосування системи Вернама є необхідною але недостатньою умовою аутентифікації.

Доказ.

Дійсно, нехай джерело повідомлень після здійснення групового системного кодування здійснює зашифрування відповідно з правилом:

$$C_j=(M_j+K_j)mod\ m \quad j = \overline{1, l_m + l_{надл}}. \quad (17)$$

За аналогією з (13) виконаємо перетворення C_j та отримаємо $C_j^{'}$:

$$C_j^{'} = (C_j+R_j)mod\ m = (M_j+K_j+R_j)mod\ m. \quad (18)$$

Після розшифрування отримаємо:

$$M_j^{'} = (C_j^{'}-K_j)mod\ m = (M_j+K_j+R_j-K_j)mod\ m = (M_j+R_j)mod\ m. \quad (19)$$

Якщо R_j випадкова або псевдовипадкова послідовність, то значення (19) буде послідовністю змінених кодових комбінацій, оскільки приймач не знає R_j послідовності, і декодер не зможе виправити помилки, і M_j+R_j буде являти собою повідомлення, яке не можливо прочитати (декодувати).

У цьому випадку декодер визначить помилки, а система може автоматично відмовитись від викривленого повідомлення яке нав'язується.

Таким чином, застосування групового систематичного коду є необхідною та достатньою умовою, коли R_j випадкова або псевдовипадкова.

В випадку, коли криптоаналітик формує R_j у вигляді послідовності комбінацій цього ж групового систематичного коду, причому йому відомі початок та кінець комбінацій коду, співвідношення (18) буде мати вигляд:

$$C_j^{'} = (C_j+R_j^k)mod\ m = (M_j+K_j+R_j^k)mod\ m = ((M_j+K_j)+R_j^k)mod\ m. \quad (20)$$

У випадку, коли R_j^k є послідовністю комбінацій групового систематичного коду, то враховуючи його замкнутість, $M_j+R_j^k$ являє собою кодову комбінацію з цього коду M_j^{ξ}, і тоді співвідношення (20) має вигляд:

$$C_j^{'}=(M_j^{\xi}+K_j)mod\ m. \quad (21)$$

На прийомному боці, відповідно до (19), розрахуємо $M_j^{'}$

$$M_j^{'} = (M_j^{\xi}+K_j-K_j)mod\ m = M_j^{\xi}. \quad (22)$$

Оскільки M_j^{ξ} є послідовністю комбінацій групового систематичного коду, то під час декодування факт модифікації повідомлення не буде визначений, і таким чином ствердження 2 підтверджується.

Розглянемо ствердження 3, в якому доводяться необхідні та достатні умови забезпечення автентичності та визначення навмисного впливу криптоаналітика.

Ствердження 3. Нехай в системі Вернама використовується поточне шифрування, тоді необхідними та достатніми умовами забезпечення аутентичності є:

1. Використання поточного шифрування.
2. Використання групового коду з визначенням помилок.
3. Формування ключової послідовності з використанням нелінійних алгоритмів перетворення відкритої інформації та криптограми.

Дійсно, нехай як і раніше

$$C_j = (M_j + K_j) mod\ m, \tag{23}$$

де M_j - послідовність комбінацій групового коду. На відміну від (23)

$$C_j = (M_j + K_j^{'}) mod\ m, \tag{24}$$

де:
$$K_j^{'} = \Psi[K_j, M_{j-\nu} \cup (\cap) C_{j-\nu}]. \tag{25}$$

Нехай функція Ψ в даному випадку реалізується схемою наведеною, на рис. 3.

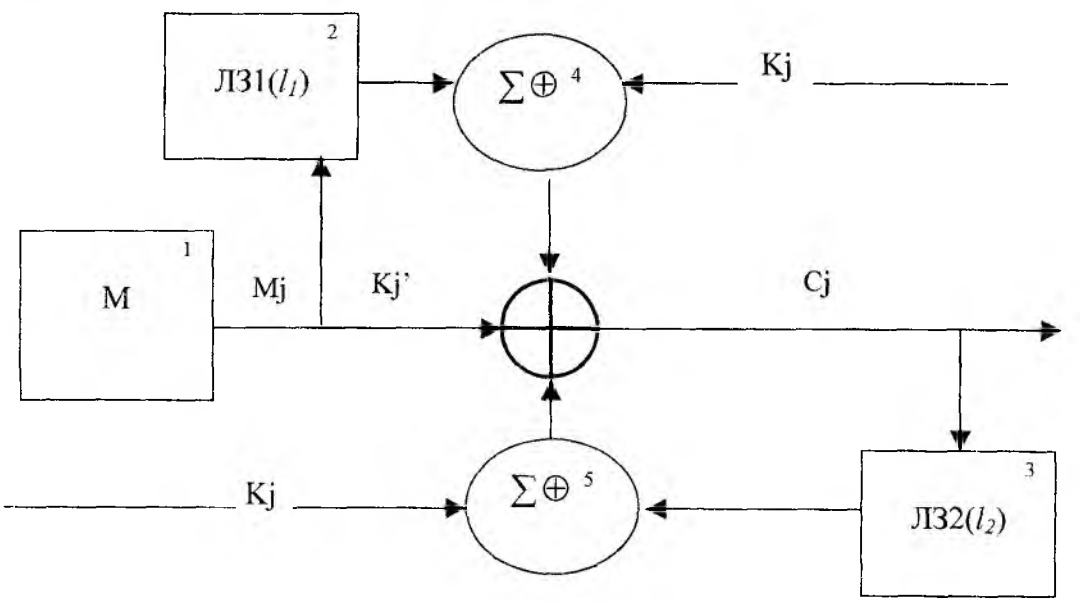


Рис. 3

Введені позначення: l_1 l_2 – довжина ліній затримки ЛЗ1 та ЛЗ2 відповідно;
$\sum\oplus$ - суматор за модулем два.

У вигляді співвідношення цю схему можна описати таким чином

$$K_j^{'} = \Psi[\ldots] = K_j \oplus \sum_{i=1}^{l_1} \oplus M_{j-i} \cup (\cap) \oplus \sum_{i=1}^{l_2} \oplus C_{j-i}. \tag{25}$$

Враховуючи співвідношення (25), маємо:

$$C_j = M_j \oplus K_j \oplus \sum_{i=1}^{l_1} \oplus M_{j-i} \cup (\cap) \oplus \sum_{i=1}^{l_2} \oplus C_{j-i},\ j = \overline{1, l_M}. \tag{26}$$

Нехай криптоаналітик, як і раніше здійснює криптоперетворення C_j криптограми

$$C_j^{'} = \varphi(C_j, R_j). \tag{27}$$

Тобто

$$C_j' = \varphi(C_j, R_j) = C_j \oplus R_j = M_j \oplus K_j \oplus \sum_{i=1}^{l_1} \oplus M_{j-i} \cup (\cap) \oplus \sum_{i=1}^{l_2} \oplus C_{j-i} \oplus R_j. \quad (28)$$

На прийомному боці реалізована наступна схема розшифрування (рис. 4).

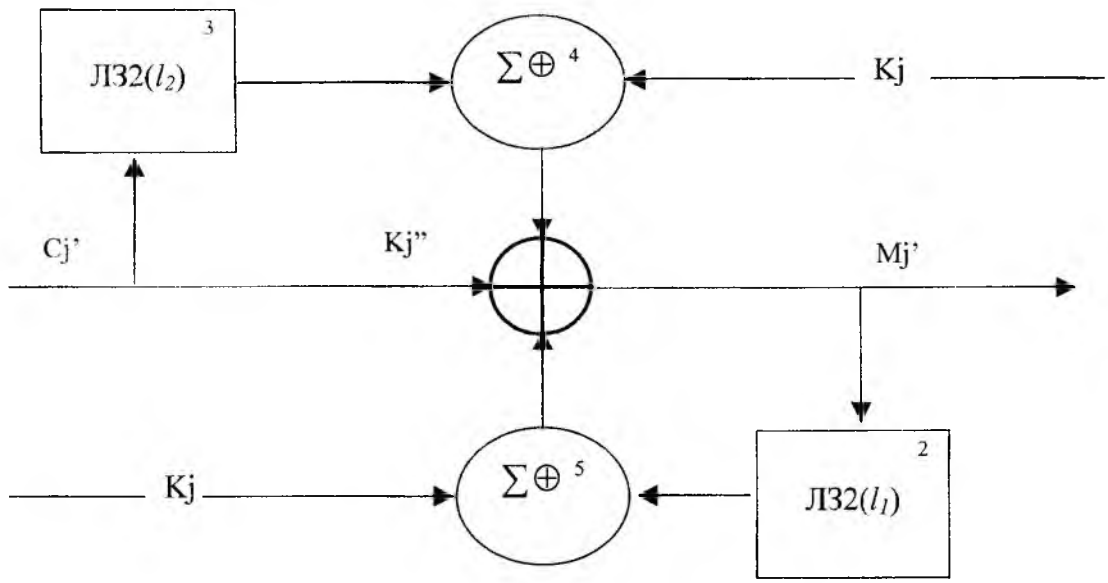


Рис. 4

У вигляді співвідношення цю схему можна описати наступним чином

$$M_j' = C_j' \oplus K_j''. \quad (29)$$

$$K_j'' = K_j \oplus \sum_{i=1}^{l_1} \oplus M_{j-i}' \cup (\cap) \oplus \sum_{i=1}^{l_2} \oplus C_{j-i}'. \quad (30)$$

Підставимо співвідношення 28 та 30 у співвідношення (29)

$$M_j' = M_j \oplus \underline{K_j} \oplus \sum_{i=1}^{l_1} \oplus M_{j-i} \cup (\cap) \oplus \sum_{i=1}^{l_2} \oplus C_{j-i} \oplus R_j \oplus$$

$$\oplus \underline{K_j} \oplus \sum_{i=1}^{l_1} \oplus M_{j-i}' \cup (\cap) \oplus \sum_{i=1}^{l_2} \oplus C_{j-i}' =$$

$$= M_j \oplus R_j \oplus (\sum_{i=1}^{l_1} \oplus M_{j-i} \oplus \sum_{i=1}^{l_1} \oplus M_{j-i}') \cup (\cap) \oplus (\sum_{i=1}^{l_2} \oplus C_{j-i} \oplus \sum_{i=1}^{l_2} \oplus C_{j-i}'). \quad (31)$$

Ми отримали співвідношення для розшифрування відповідно до схеми, зображеної на рис. 4. Спочатку для аналізу візьмемо лише частину із зворотним зв'язком за криптограмою.

$$M_j' = M_j \oplus (\sum_{i=1}^{l_2} \oplus C_{j-i} \oplus \sum_{i=1}^{l_2} \oplus C_{j-i}') \oplus R_j. \quad (32)$$

Аналіз співвідношення 32 показує, що у випадку, коли $R_j = 0$ (тобто R_j – відсутнє), а C_j' не містить помилок, тоді:

$$M_j' = M_j. \quad (33)$$

Іншими словами, реалізована безпомилкова передача. У випадку коли $R_j = 0$ (R_j – відсутнє), а C_j' містить помилки, тоді:

$$\sum_{i=1}^{l_2} \oplus C_{j-i} \oplus \sum_{i=1}^{l_2} \oplus C'_{j-i} \neq 0, \tag{34}$$

$$M'_j \neq M_j, \tag{35}$$

та декодер групового коду визначить помилку.

Таким чином, система визначає помилки природного походження. У випадку, коли $R_j \neq 0$, сума:

$$(\sum_{i=1}^{l_2} \oplus C_{j-i} \oplus \sum_{i=1}^{l_2} \oplus C'_{j-i}) \oplus R_j \neq 0, \tag{36}$$

незалежно від наявності помилки в каналі зв'язку. Якщо R_j випадкове, то схема працює, як і раніше, та визначає помилки, і декодер автоматично може відмовитись від повідомлення. Якщо R_j - груповий код, то він створить помилки в ключі K_j''. Тобто будь-який символ R_j викривить хоча б один символ криптограми C_j, та цей символ затримується в лінії затримки l_2 разів та в середньому скривдить $l_2/2$ символів, що призведе до появи пакету помилок, які визначать груповий код з визначенням помилок. Очевидно, що підібрати R_j криптоаналітик не може у зв'язку з тим, що він не знає ключа та повідомлення.

Розглянемо випадок, коли використовується лише зворотній зв'язок за повідомленням.

$$M_j' = M_j \oplus (\sum_{i=1}^{l_1} \oplus M_{j-i} \oplus \sum_{i=1}^{l_1} \oplus M'_{j-i}) \oplus R_j. \tag{37}$$

В цьому випадку криптоаналітик також не знає повідомлення та не в змозі підібрати відповідне R_j, тоді:

$$\oplus (\sum_{i=1}^{l_1} \oplus M_{j-i} \oplus \sum_{i=1}^{l_1} \oplus M'_{j-i}) \oplus R_j \neq 0. \tag{39}$$

Таким чином, ствердження 3 доведене.

Висновок

Основним засобом забезпечення послуг цілісності, доступності та спостереженності є застосування методів автентифікації. Методи автентифікації можуть застосовуватись для встановлення справжності користувача, системи, повідомлень, програмного забезпечення.

На сьогоднішній день немає навіть прикладної теорії автентифікації. Застосування елементів теорії Сімонсона дозволяє одержати граничні оцінки ймовірностей обману, які може досягти зловмисник.

Основним методом забезпечення автентичності є внесення в повідомлення надлишковості, яка може формуватись у вигляді контрольних сум, збиткових символів кодів, які визначають помилки, криптографічних контрольних сум (кодів автентифікації, імітовкладок) та хеш-функцій, а також цифрових підписів.

В потокових системах цілісність повідомлень може забезпечуватись за рахунок модифікації ключової послідовності (в безумовно стійких системах) або потокових гам (в обчислювально стійких системах) з використанням символів повідомлень та криптограм.

Список літератури: 1. *Boca Raton.* Authentication codes that parmit arbitration presented at the 18-th Southeastern conf., on Combinatories, Grapt Theory Computing, 1987. Feb. 23-27. 2. *Г.Дж. Симмонс.* Обзор методов аутентификации информации // ТИИЭР. 1988. 76 (5). С.105-125 (Малый тематический выпуск. Защита информации). 3. *Neuman B.C., Ts'o T.* Kerberos: An Authentication Service for Computer Networks//IEEE Comm. Magazine. 1994. Vol. 32. № 9. P. 33-38.

Харківський державний технічний університет радіоелектроніки

Надійшла до редколегії 29.03.2001

УДК 681.3.06:519.248.681

І. Д. ГОРБЕНКО, д-р техн. наук, О. А. ЗАМУЛА, канд. техн. наук,
Є. В. ПОПОВИЧ, Т. О. ГРІНЕНКО

МЕТОДИ ЗАБЕЗПЕЧЕННЯ АВТЕНИЧНОСТІ З ВВЕДЕННЯМ НАДЛИШКОВОСТІ

Згідно з основними положеннями міжнародних та державних стандартів системи захисту інформації повинні надавати користувачам такі послуги, як цілісність, доступність та спостережність. Одним із основних методів забезпечення цих послуг є застосування процедур та алгоритмів автентифікації. На сьогодні теорія та практика автентифікації одержали значний розвиток [1-3]. Але досі немає робіт, в яких методи автентифікації розглядались би в порівнянні. Метою даної статті є класифікація, характеристика та порівняння відомих методів автентифікації інформації, яка обробляється та передається з введенням надлишковості.

Аналіз низки джерел показав, що автентифікація повідомлення M_i на основі надлишковості може бути забезпечена за рахунок використання:

- однонаправленої хешфункції (h);
- ключових хешфункцій або кодів автентифікації ($КА$);
- цифрового підпису ($ЦП$);
- іншої надлишкової інформації, що формується відповідно з прийнятими правилами, які будемо називати контрольною сумою ($КС$).

Позначимо автентифіковане повідомлення, як M_i^A та розглянемо суть та відмінні риси методів та реалізованих на їх базі алгоритмів.

1. Автентифікація з використанням однонаправлених хешфункцій

Як відомо під однонаправленою хешфункцією розуміється відображення інформації M_i в її стислий образ h:

$$h=H(M_i), \tag{1}$$

де H – функція стиснення (розтиснення) M_i в h.

Особливістю такого відображення є те, що прямий розрахунок, відповідно до (1), повинен носити не більш, ніж поліноміальну складність. У випадку, коли довжина повідомлення менша за довжину хешфункції ($l_m<l_h$), функція H – розтискаюча, а коли довжина повідомлення більша за довжину хешфункції ($l_m>l_h$), функція H – стискаюча.

Довжина l_h, як правило, фіксована, або має декілька фіксованих значень (128, 160, 256 біт) та не залежить від довжини повідомлення.

Однонаправлена функція повинна бути стійкою проти криптоаналітичних атак, основними з яких є:

1. Стійкість проти знаходження прообразу повідомлення M_i:

$$M_i=H^{-1}(h), \tag{2}$$

при відомому значенні хеш-функції.

Складність таких розрахунків повинна бути не нижче за субекспоненціальну складність.

2. Стійкість за другим прообразом, яка заключається в тому, що при відомій $h=H(M_i)$ необхідно знайти таке повідомлення M_j, для якого

$$H(M_i)=H(M_j). \tag{3}$$

Задача пошуку такого повідомлення повинна носити не нижчу за субекспоненціальну складність.

3. Стійкість до колізій, суть якої заключається в складності знаходження двох повідомлень M_i та M_j, для яких

$$H(M_i)=H(M_j). \tag{4}$$

Задача пошуку таких повідомлень повинна бути експоненціально складною.

З вищезазначеного можна зробити такі висновки:

1. Застосування автентифікації на основі однонаправленої хеш-функції дозволяє забезпечити аутентичність (цілісність та достовірність) на підставі внесення та додавання до повідомлення M_i інформації h .

2. Автентифіковане повідомлення M_i^A є сукупність інформацій M_i та h ($M_i^A=\{M_i, h\}$).

3. При використанні однонаправленої хешфункції здійснюється відображення "багато в одне", тобто в конкретне значення хешфункції h може відобразитись велика кількість повідомлень.

4. Відповідно до парадоксу дня народження [1] для однонаправленої хешфункції існує атака типу колізій, розрахункова складність якої для повідомлення з основою алфавіту m має вигляд:

$$I \cong \sqrt{m^{l_h}}, \tag{5}$$

та з основою $m=2$:

$$I \cong \sqrt{2^{l_h}}. \tag{6}$$

Проведений аналіз показує, що розв'язання задач автентифікації та шифрування, наприклад, в безумовно стійкій системі, у випадку використання однонаправленої хешфункції, повинен складатись з наступних кроків:

1. Для повідомлення M_i розраховується хешфункція h та формується автентифіковане повідомлення M_i^A.
2. Далі повідомлення M_i^A зашифровується, наприклад, в системі Вернама [3].
3. При отриманні повідомлення розшифровується в системі Вернама.
4. Для розшифрованого повідомлення $M_i^{A'}=\{M_i', h\}$ розраховується хешфункція $h'=H(M_i')$.
5. Здійснюється порівняння розрахованої хешфункції h' та h, яка знаходиться при повідомленнях. В разі збігання даних хешфункцій повідомлення вважається достовірним та цілісним, в протилежному випадку повідомлення таким не вважається.

На цей час розроблена велика кількість однонаправлених хешфункцій основними з яких є: MD4, MD5, RIPEMD-128, SHA-1, RIPEMD-160, ГОСТ 34.311-95 [1-3].

Як показує аналіз даних хешфункцій, всі вони є середньоскладними, мають поліноміальну складність та забезпечують допустимі швидкості хешірування. Оцінка стійкості, в сенсі захищеності від обману під час знаходження колізій (6) за використанням оцінки Сіменсона [3], є:

$$P_{обм} \geq 2^{-l_h}. \tag{7}$$

Таблиця 1

Назва хешфункції	Довжина (біт)	Ймовірність обману
MD4	128 (256)	$3\cdot10^{-39}$
MD5	128	$3\cdot10^{-39}$
RIPEMD-128	128	$3\cdot10^{-39}$
SHA-1	160	$6{,}8\cdot10^{-49}$
RIPEMD-160	160	$6{,}8\cdot10^{-49}$
ГОСТ 34.311-95	256	$3\cdot10^{-77}$

В табл. 1 наведені значення ймовірності обману, які розраховані відповідно до (5) та (6) для вказаних хеш-функцій.

Даний метод є симетричним, оскільки не забезпечує реалізацію моделі взаємної недовіри. Він може використовуватися в разі, коли користувачі системи довіряють один одному.

2 Автентифікація з використанням ключових хешфункцій

У випадку, коли для забезпечення автентифікації надлишковість вноситься за рахунок використання ключових хешфункцій або кодів автентифікації (КА) [2], виникають деякі розходження з розглянутими раніше однонаправленими хешфункціями. Основною різницею є то, що:

$$KA=H(M_i, K_j). \tag{8}$$

Для ключової хешфункції основною перевагою є захищеність від атаки типу колізій. Для оцінки якості забезпечення автентифікації можна користуватися співвідношенням (8), але крім довжини хешфункції l_h, необхідно враховувати довжину ключа автентифікації l_{KA} та вибрати з них той показник, який приводить до найбільшої ймовірності обману. Оскільки для ключової хешфункції відсутня атака типу колізій, то довжина l_{KA} може бути менша за l_h при забезпечені тієї ж стійкості.

Ключова хешфункція є симетричною. Тому, відповідно, ключ автентифікації у всіх абонентів повинен бути однаковим і, як наслідок, за допомогою цієї функції не можливо забезпечити модель взаємної недовіри.

В якості ключових хешфункцій можна назвати наступні: MDC-2, MDC-4, DES в режимі роботи коду автентифікації, ГОСТ 28147 в режимі виробки імітовставки.

В табл. 2 наведені значення ймовірності обману для ключових хешфункцій, отриманих відповідно до співвідношення (7).

Таблиця 2

Назва алгоритму	Довжина ключа (біт)	Довжина імітовставки (біт)	Ймовірність обману
MDC-2	56	128	$1{,}4\cdot10^{-17}$
MDC-4	56	128	$1{,}4\cdot10^{-17}$
DES	56	64	$5{,}4\cdot10^{-20}$
ГОСТ 28147	256	64	$5{,}4\cdot10^{-20}$

Безпосередньо реалізація системи захисту з забезпеченням автентифікації на основі ключової хешфункції виконується аналогічно, як і для однонаправленої хешфункції. Різниця лише в тім, що отримувач повинен володіти ключем автентифікації *Kj*. В алгоритмах ГОСТ 28147, RIENDAEL, IDEA, DES, передбачено стандартний режим створення (виробітка) ключової хешфункції.

3. Автентифікація з використанням цифрового підпису

Суть цього метода автентифікації полягає в обчисленні для кожного повідомлення або інформації цифрового підпису [1]. Відмінною рисою цього методу є те, що цифровий підпис являє собою криптоперетворення від хешфункції *h*, особистого ключа $K_л$ та інших параметрів *Pr*:

$$ЦП=F(H(M),K_л,Pr,. \qquad (9)$$

де *H(M)* – як правило однонаправлена хешфункція; $K_л$ – особистий ключ відправника або автора повідомлення; *Pr* – додаткові параметри, в якості яких можуть використовуватися ідентифікатори відправника та одержувача, час створення та відправки повідомлення, час життя повідомлення, символи керування і т.і.

При використанні цифрового підпису може бути реалізована модель взаємної недовіри, взаємного захисту завдяки тому, що кожен користувач може генерувати сам собі необхідну пару ключів (особистий – секретний та загальний – відкритий). Особистий ключ зберігається в таємниці та не випускається з під контролю. Відкритий ключ розсилається безпосередньо або через відповідний центр у вигляді сертифікатів всім іншим користувачам.

Створити повідомлення може лише користувач, що володіє особистим ключем, перевірити цілісність та достовірність (автентифікувати) повідомлення може кожен, хто володіє відкритим ключем.

Аналіз показує, що, з точки зору складності розрахунків, цифровий підпис у порівнянні з розрахунками однонаправленої або ключової хешфункції має складність на декілька порядків більшу, а здійснення цифрового підпису вимагає більш потужних ресурсів, а також багатослівної арифметики.

При цьому, попередній аналіз показує, що:

1. Використання цифрового підпису в системах захисту забезпечує найбільшу стійкість автентифікації. Це пов'язано з тим, що цифровий підпис шифрується разом з повідомленням, і виділити його без знання ключа практично не можливо. В цьому випадку відсутні аналітичні атаки, крім атаки типу брутальна сила.

2. У випадку реалізації моделі взаємної недовіри та взаємного захисту складність криптоаналітичної атаки, метою якої є підробка цифрового підпису, зводиться до задачі, що стала вже класичною, субекспоненціального та експоненціального. Зокрема, при використанні цифрового підпису на базі RSA алгоритму та застосуванні методу загального решета числового поля [2], складність факторизації можна визначити формулою:

$$I=exp(\delta(\ln N)^{\nu}(\ln\ln N)^{1-\nu}), \qquad (10)$$

де *N* – модуль перетворення; ν – параметр, що залежить від методу вирішення задачі; δ – константа (залежить від використаного математичного методу).

При використанні протоколу Діфі-Хелмана основна частина процесу криптоаналізу складається з вирішення дискретно-логарифмічного порівняння виду

$$Y=(a^X) \bmod p, \qquad (11)$$

де *Y* – відкритий ключ; *a* та *p* – загальномережеві параметри; X – особистий ключ.

Задача криптоаналізу складається з находження особистого ключа

$$X=(\log_a Y) \bmod p. \qquad (12)$$

Складність вирішення такої задачі залежить від математичного апарату, який використовується, а також від потужності обчислювальної системи та програмної платформи.

Достатньо прийнятою апроксимацією складності є співвідношення (10) з відповідно вибраними параметрами δ та ν.

Очевидно, що найбільш перспективним є цифровий підпис, який реалізований в класі перетворень еліптичних кривих над розгорнутим полем [2]. У цьому випадку основною відзнакою від вищезазначених алгоритмів є те, що розмір модуля перетворень може бути значно зменшений до розмірів $\approx 2^{163}/2^{257}$ [2]. Крім того, до таких же розмірів може бути зменшена довжина загального та особистого ключів. Основною перевагою цифрового підпису на базі еліптичних кривих є те, що складність криптоаналізу для нього залишається експоненціальною та оцінюється для оптимального методу Поларда [2], як:

$$I=\sqrt{\frac{\pi n}{4}}, \qquad (13)$$

де n – порядок базової точки.

В нашому випадку n може приймати значення 2^{160}-2^{256}.

4. Автентифікація з використанням контрольних сум

Особливістю даного методу автентифікації є те, що для повідомлення розраховуються некриптографічні контрольні суми невеликого обсягу. Аналіз джерел [1-3] показує, що їх можна віднести до однонаправлених хешфункцій, оскільки контрольна сума розраховується аналогічно до співвідношення (1). Але такі контрольні суми не є розрахунково та колізійно стійкими, тобто для них може бути вирішена задача типу (2), завдяки чому в криптосистемах вони мають досить обмежене використання.

Висновок

Таким чином, проведений аналіз показав, що задача автентифікації є самостійною криптографічною задачею та потребує окремого розгляду.

На цей час немає єдиної теорії автетифікації, тобто забезпечення цілісності та достовірності інформації на всіх її життєвих етапах. В якості оцінки можуть використовуватися оцінки, які надає теорія Сімонсона.

Проведений аналіз показав, що існує декілька методів забезпечення автентичності повідомлень. В якості основної ознаки їх класифікації можна використати застосування або незастосування надлишковості. Проведений вище аналіз показав, що надійна автентифікація може бути забезпечена лишу під час застосування надлишковості.

За способом формування надлишковості методи можна поділити на однонаправлені та ключові. Відмінною ознакою є те, що ключові хешфункції потребують використання симетричних криптоперетворень та ключа.

Для моделі взаємної недовіри та взаємного захисту може використовуватися лише цифровий підпис.

Використання будь якого методу автентифікації необхідно узгоджувати з моделлю загроз, політикою інформаційної безпеки та політикою надання послуг користувачам.

Список література: 1. *Brown L.*, LOKI – a cryptographic primitive for authentication and secrecy application in Proceedings of AUSCRYPT 90, 1990. 2. *Krawczyk H.*IETF Draft: Keyed-MD-5 for Message Autentification, November, 1995. 3. *Симонс Г. ДЖ.* Обзор методов аутентификации // ТИИЭР. 1988. №5. С.105-125.

Харківський державний технічний університет радіоелектроніки

Надійшла до редколегії 5.04.2001

УДК 681.3.06

Г. З. ХАЛИМОВ, канд. техн. наук, А. А. КУЗНЕЦОВ

АУТЕНТИФИКАЦИЯ С ПРИМЕНЕНИЕМ АЛГЕБРОГЕОМЕТРИЧЕСКИХ КОДОВ

Введение

Перспективным направлением разработки схем аутентификации является применение строго универсального класса хеш-функций. Одним из известных способов построения такого класса является использование простых ортогональных массивов. Однако, схемы аутентификации на основе ортогональных массивов имеют недостаток, связанный с тем, что размер ключевых данных должен превышать объем передаваемых данных [1-4]. Это ограничение было снято в конструкции Д. Стинсона [5-7]. Стинсон рассмотрел композиционную конструкцию схемы аутентификации, составленную из универсального и строго универсального классов хеш-функций и показал, что такая схема аутентификации является строго универсальным классом хеш-функций. Полученные результаты практически одновременно натолкнули авторов статьи [8] на необходимость использования в композиционной конструкции алгебраико-геометрических кодов большой размерности. Как следует из [8], q-ичный код большой длины и с большим кодовым расстоянием эквивалентен универсальному классу хеш-функций, и их применение позволяет получить схему аутентификации с хорошими параметрами. Наиболее перспективными для этих целей являются коды Рида-Соломона, перекрученные БЧХ коды (twisted BCH codes), алгеброгеометрические коды, ассоциированные с кривыми Артин-Шрейера, Эрмита, Сузуки [9-11].

Задачей данной работы является изложение общетеоретических вопросов построения аутентификации с применением кодовых конструкций и исследование практических схем с оценкой их параметров. С этой целью в разделе 1 приводится композиционная конструкция Стинсона для построения строго универсального семейства хеш-функций. В разделе 2 рассматривается теория применения алгеброгеометрических кодов для целей универсального хеширования, а также – примеры и оценки для схем аутентификации с кодовыми конструкциями.

1. Композиционная конструкция Стинсона для построения строго универсального класса хеш-функций

В работах Стинсона [5-7] рассмотрена каскадная конструкция построения строго универсальных хеш-функций $\varepsilon - SU(N, n, m)$ с условием того, что объем ключевых данных меньше объема исходного сообщения $N < n$. Это снимает ограничение на схемы аутентификации с ортогональными массивами.

Композиционную конструкцию лучше всего пояснить с помощью языка отображений. Пусть X, Y, U являются множествами из n, m, u элементов, $n < m < u$. H_1 есть множество функций f_1, осуществляющих отображение $X \to U$, а H_2 – множество функций f_2, осуществляющих отображение $U \to Y$. Тогда $H = H_2 \circ H_1$ есть множество функций f, являющееся композицией $f = f_1 \circ f_2$.

Основной результат конструкции Стинсона представлен следующей теоремой [6].

Теорема 1. Композиция из универсального класса хеш-функций $\varepsilon_1 - U(N_1, n, u)$ и строго универсального класса хеш-функций $\varepsilon_2 - SU(N_2, n, m)$ является строго универсальным классом с параметрами $\varepsilon - SU(N_1 N_2, n, m)$, где $\varepsilon = \varepsilon_1 + \varepsilon_2 - \varepsilon_1 \varepsilon_2$.

Составляющая H_2 в композиционной конструкции $H = H_1 \circ H_2$ представляет класс строго универсальных хеш-функций, и его можно построить, используя простые ортогональные массивы [4]. Компонента H_1 является универсальным классом хеш-функций. Для его построения используем результат следующей теоремы [7].

Теорема 2. Код длины N с n – кодовыми словами, основанием m и относительным кодовым расстоянием $1 - \varepsilon$ эквивалентен универсальному классу хеш-функций с параметрами $\varepsilon - U(N, n, m)$.

Таким образом, действие функций f_1 из множества H_1 заключается в отображении m-ичного кодового слова, соответствующего передаваемому сообщению в i-й кодовый символ, соответствующий f_1 функций.

Очевидно, что необходимо использовать коды с большим основанием, большой длины и большим относительным кодовым расстоянием, так как в этом случае получим большой объем ключей аутентификации и малую вероятность коллизии кодов аутентификации.

Рассмотрим композиционную конструкцию с кодом Рида-Соломона. Справедлива следующая теорема.

Теорема 3. Пусть q – простое число, a, b, k – целые числа, $a > b$. Тогда композиционная конструкция с кодом РС, построенного по основанию q^a, является SU классом хеш-функций с параметрами

$$\varepsilon - SU\left(q^{2a+b}, q^{ak}, q^b\right),\ \varepsilon < \frac{k}{q^a} + \frac{1}{q^b}. \qquad (1)$$

Действительно, применение РС – кодов к композиции Стинсона задает множество H_1 универсальных хеш-функций с параметрами $\frac{k}{q^a} - U\left(q^a, q^{ka}, q^a\right)$, так как по теореме 2 имеем основание кода $m = q^a$, длину $N_1 = q^a$ и для k информационных символов – число кодовых символов $n = q^{ka}$.

Кодовое расстояние $d = n - k$, следовательно, $\varepsilon = \frac{k}{q^a}$. Используя простые ортогональные массивы, построим универсальное семейство хеш-функций H_2 с параметрами $\frac{1}{q^b} - SU\left(q^{a+b}, q^a, q^b\right)$ [4]. И, наконец, по теореме 1 имеем $\varepsilon = \frac{k}{q^a} + \frac{1}{q^b} - \frac{k}{q^{a+b}} < \frac{k}{q^a} + \frac{1}{q^b}$, $N_1 N_2 = q^{2a+b}$.

Конструкция Стинсона предполагает использование длинных m-ичных кодов. Перспективными для этих целей являются алгеброгеометрические коды.

2. Применение алгеброгеометрических кодов для универсального хеширования

Впервые алгеброгеометрический подход к построению кодов был предложен Гоппой В.Д. в 1981 году [12]. В работах [13] показано, что по алгебраическим кривым можно строить коды с очень хорошими асимптотическими свойствами. Доказано существование бесконечных серий q-ичных линейных кодов, параметры которых (при $q = 2^{2n} > 49$ и $N \to \infty$) лежат выше границы Варшамова – Гильберта. Результаты по алгеброгеометрическим кодам обобщим в следующей теореме.

Теорема 4. Пусть зафиксированы: гладкая алгебраическая кривая X рода g над F_q, наборы точек $P = \{P_1, P_2, ..., P_n\} \subseteq X(F_q)$, $Q = \{Q_1, Q_2, ..., Q_s\} \subseteq X(F_q)$, $Q \cap P = \varnothing$ (условие не существенно и снимается использованием пучковой конструкции), эффективный дивизор $D = \sum u_Q Q$ степени $\deg(D) = \sum u_Q$ и $\deg(D) > 2g - 2$. С дивизором D ассоциируем линейное пространство $L(D)$ из рациональных функций $f_0, f_1, ..., f_{k-1}$ на X как множество всех функций f таких, чтобы порядок f в каждой точке Q удовлетворял $div(f) \geq -u_Q$.

Тогда линейный код C над F_q определим как

$$C = \left\{ f(P_1), ..., f(P_n) \middle| f \in L(D) \right\}.$$

Причем, алгебраико-геометрический код имеет параметры

$$[n, \deg(D) - g + 1, d],\ d \geq n - \deg(D),$$

а двойственный к нему код также является алгебраико-геометрическим с параметрами

$$[n, n-\deg(d)+g-1, d^{\perp}],\ d \geq n-\deg(D)-2g+2.$$

Для универсального хеширования алгеброгеометрическая кодовая конструкция определяется следующей теоремой.

Теорема 5. Пусть задана алгеброгеометрическая кривая X над F_q с $N+1$ рациональными точками. Пусть P является одной из этих рациональных точек с порядками полюса u_i, где $u_0 = 0 < u_1 < u_2 \ldots$ и $u_i < N$. Тогда алгеброгеометрический код размерности i и минимальным расстоянием $d \geq N - u_i$ образует универсальный класс хеш-функций $\varepsilon - U(N, q^i, q)$, где $\varepsilon \leq \frac{u_i}{N}$.

Пример 7. Пусть F – алгебраическое покрытие F_q. Точки X определяются гомогенными координатами (x, y). Точки на X с координатами в F_q называются рациональными точками. Они имеют значения $P_i = (\alpha_i, 1)$, $0 \leq i \leq q-1$ и $Q = (1,0)$ – особая точка (точка неопределенности). Пусть Z – множество рациональных функций на X, которые определены в каждой P_i с коэффициентами в F_q и которые имеют полюс порядка меньше, чем m в точке Q и нет других полюсов. Рациональная функция f имеет вид $\frac{\alpha(x,y)}{\beta(x,y)}$, где $\alpha(x,y)$ и $\beta(x,y)$ гомогенные полиномы степени $< k$. Алгеброгеометрический код C определим как

$$C = \left\{ (f(P_0), f(P_1), \ldots, f(P_{q-1})) \middle| f \in Z \right\}.$$

Код C по результатам теоремы 4 имеет размерность пространства $Z(mQ), g = 0$, $k = \dim C = m - g + 1 = m + 1$ и минимальное расстояние $d \geq n - m$.

Таким образом, получим PC код в алгеброгеометрической интерпретации и универсальный класс хеш-функций $\frac{k}{q} - U(q, q^k, q)$.

Очевидно, что для построения универсального хеш-класса $\varepsilon - U(N, q^i, q)$, $\varepsilon \leq \frac{u_i}{N}$ с малой вероятностью коллизии ε и большим значением объема хешируемого сообщения q^i необходимо использовать алгебраические коды по кривым с как можно большим числом рациональных точек N, а при заданном значении i с порядком полюса u_i в особой точке как можно меньшим.

В [14] рассмотрен класс рациональных функций $K_q^{(r)}$, определенных на F_q при $r \geq 2$. Число рациональных точек для поля функций $K_q^{(r)}$ равно $N = q^r + 1$. Существует рациональная точка Вейерштрасса P_0, для которой порядки полюса образуют подгруппу

$$\sum_{i=1}^{r} q^{r-i}(q+1)^{i-1} N_0 .$$

Пусть порядок полюса P_0 равен $\sum_{i=1}^{r} \alpha_i q^r (q+1)^{i+1}, \alpha_i \geq 0$, через ω обозначим вес порядка $\omega = \sum_{i=1}^{r} \alpha_i$. Значение порядка полюса веса ω однозначно определяется как линейная комбинация слагаемых $q^{r-i}(q+1)^{i-1}$. Вес ω однозначно определен, если $\omega < \frac{q}{r-1}$.

Для задачи хеширования интерес представляют алгебраические кривые, которые допускают регулярное построение в произвольных конечных полях F_q. К таким кривым относятся: проективная прямая, кривые Эрмита, кривые Сузуки [15].

В примере 1 рассмотрена проективная прямая с кодами Рида-Соломона в конструкции хешируемой функции с параметрами (1).

Кривые Эрмита определены уравнением $x^{q+1}+y^{q+1}+z^{q+1}=0$ над полем F_{q^2} с числом рациональных точек $N=q^3+1$. Рациональные точки кривой являются точками Вейерштрасса. Подгруппа порядков полюса образуется значениями q и $q+1$.

Для заданного веса $\omega < g$ существует $\frac{(\omega+1)\omega}{2}+1$ рациональных функций, порядки полюса которых в P_0 не превышают ωq, что дает следующее [15]:

Теорема 6. Для каждого простого q и $\omega < q$, коды ассоциированные с Эрмитовой кривой над полем F_{q^2} образуют универсальный класс хешрующих функций

$$\frac{\omega}{q^2}-U(q^3,q^{\omega^2+\omega+2},q^2).$$

Кривые Сузуки определяются уравнением

$$x^{q_0}(z^q+zx^{q-1})=y^{q_0}(y^q+yx^{q-1})$$

над полем F_q, где $q=2^{2f+1}, q_0=2^f$, имеют q^2+1 рациональных точек, которые являются точками Вейерштрасса и образуют подгруппу для порядка полюса вида

$$qN_0+(q+q_0)N_0+(q+2q_0+1)N_0.$$

С помощью комбинаторного анализа можно показать, что для заданного веса $\omega < q$ существует ровно $1+\frac{(2\omega+1)(\omega+1)\omega}{2}$ рациональных функций, порядки полюса которых в P_0 не превышают $\omega \cdot q$, что дает следующее [15]:

Теорема 7. Пусть $q+2^{2f+1}, q_0=2^f$. Для заданного веса $\omega < q_0$ коды, ассоциированные с кривой Сузуки над полем F_q, образуют универсальный клас хешрующих функций

$$\frac{\omega}{q}-U(q^2,q^{\frac{1+(2\omega+1)(\omega+1)\omega}{6}},q).$$

Пример 2. Пусть $f=1, q_0=2, q=8$, тогда уравнение кривой Сузуки имеет вид

$$x^2(z^8+zx^7)=y^2(y^8+yx^7),$$

число точек кривой равно $N=65$. Зафиксируем особую точку $P_0(x,y,z)=(0,0,1)$, вес $\omega=2$ и значение порядка полюса $u_{P_0}=\omega q=16$.

Для веса $\omega=2$ существует ровно $\frac{(2\omega+1)(\omega+1)\omega}{2}+1=6$ рациональных функций $f_0, f_1 \ldots\ldots f_5$, степень дивизора которых $P_0 \geq -16$. Для построения алгеброгеометрического кода C зададим базис в линейном пространстве $L(u_{P_0}P_0)$. Первая базисная функция f_0 является единичным вектором с порядком $u_0=0$. Преобразуем уравнение к виду

$$\frac{1}{x^2} = \frac{z^8 + zx^7 + y^3x^5}{y^{10}}.$$

Второй базисной f_1 функцией является $f_1 = \frac{y^2}{x^2}$, которая имеет в точке $P_0 = (0,0,1)$ порядок полюса, равный $u_1 = -q = -8$, а в качестве третьей функции можно взять $f_2 = \frac{z^2}{x^2}$ с порядком $u_2 = -(q + q_0) = -10$. Остальные функции построим, как комбинации из f_1 и f_2. Получим $f_3 = \frac{y^4}{x^4}(u_3 = -16)$, $f_4 = \frac{y^2z^2}{x^4}(u_4 = -18)$, $f_5 = \frac{z^4}{x^4}(u_5 = -20)$. Порождающая матрица кода C будет иметь вид

$$G = \begin{vmatrix} f_0(P_i) \\ f_1(P_i) \\ f_2(P_i) \\ f_3(P_i) \\ f_4(P_i) \\ f_5(P_i) \end{vmatrix}, \; i = \overline{1,64}.$$

Алгеброгеометрический код C по теореме 4 имеет параметры (64,10,>44). Уточнение кодового расстояния Хемминга с помощью полного перебора всех кодовых слов показывает, что построен код (64,6,48). Этому коду соответствует универсальный хеширующий класс функций $\frac{1}{4} - U(64,8^6,8)$, что согласуется с результатами теоремы 7.

Кривые Эрмита и Сузуки определены для конечных полей F_q, с ограничением размерности для q.

Для произвольных значений q рассмотрим алгебраические кривые с уравнениями вида:

$$x(z^q + zx^{q-1}) = y(y^q + yx^{q-1}), \tag{2}$$

$$x(z^q + zx^{q-1}) = y(y^q + y^{q-1}x + yx^{q-1} + x^q), \tag{3}$$

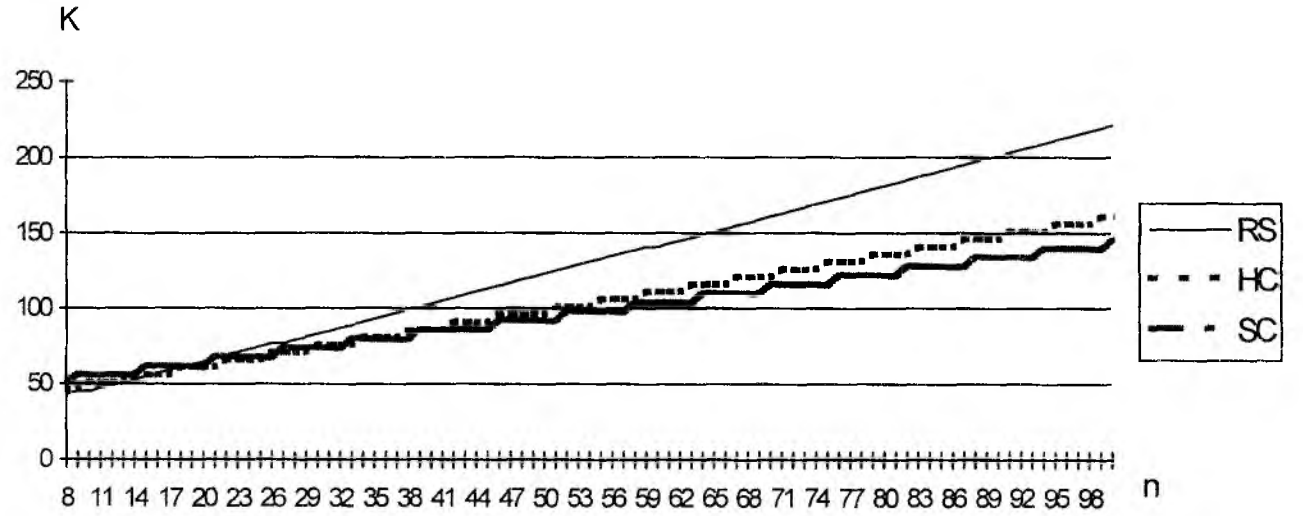
$$x(z^q + z^{q-1}x + zx^{q-1} + x^q) = y(y^q + y^{q-1}x + yz^{q-1} + x^q). \tag{4}$$

Число рациональных точек для этих кривых равно $q^2 + 1$. В конечных полях с $q = 2^{2f+1}$ многообразие точек кривых 2 и 3 совпадает с точками кривых Сузуки, что свидетельствует об эквивалентности этих многообразий. Для других значений q рациональные точки кривых 2-4 являются точками Виерштрассе, и подгруппа порядков полюса образуется значениями q и $q+1$. Для заданного веса $\omega < q$ существует $\frac{(\omega+1)\omega}{2} + 1$ рациональных функций, порядки полюса которых в P_0 не превосходят ωq, что позволяет сформулировать следующую теорему.

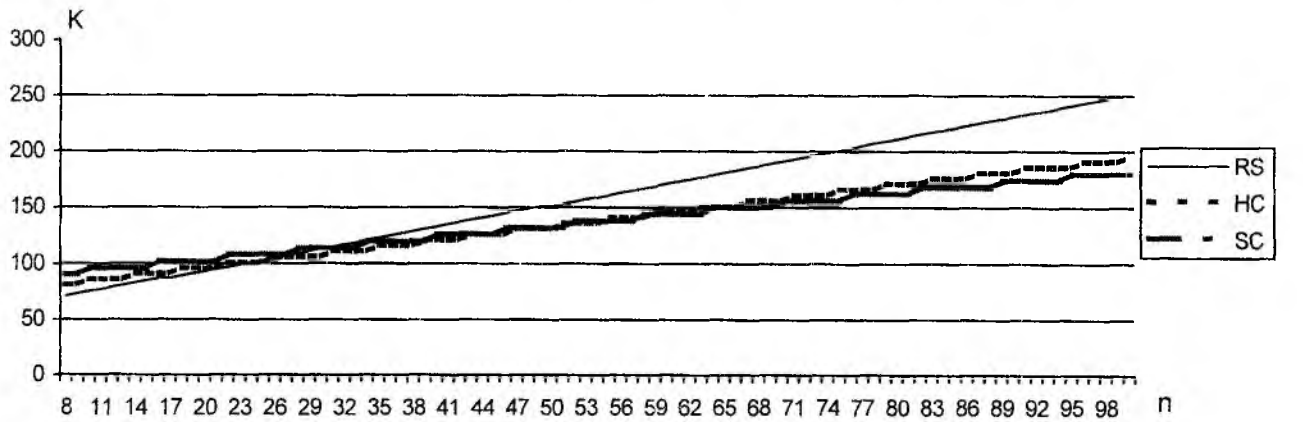
Теорема 8. Для каждого простого q и $\omega < q$ коды, ассоциированные с кривыми 2-4 над F_q, образуют универсальный класс хеширующих функций

$$\frac{\omega}{q} - U\left(q^2, q^{\frac{\omega^2+\omega+2}{2}}, q\right).$$

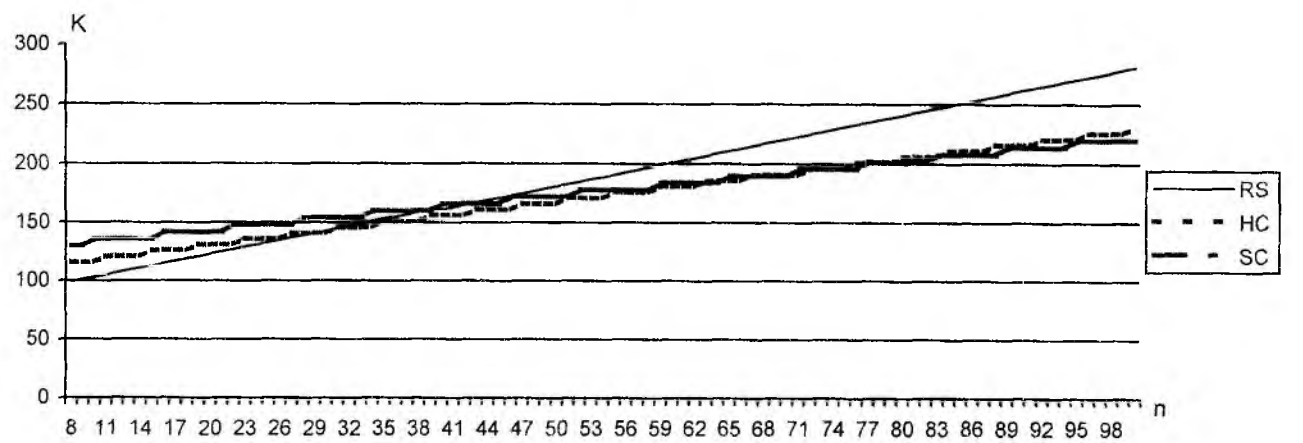
Результирующие соотношения для строго универсального класса хеширующих функций в конструкции Стинсона при использовании кодов Рида-Соломона, алгеброгеометрических кодов, ассоциированных с кривыми Эрмита (HC), Сузуки (SC), и кривыми 2-4, определяются следующей теоремой.



а



б



в

Рис. 1

Теорема 9. Пусть q – простое число, a, b – натуральные числа, причем $a > b$. Тогда строго универсальный класс хеш-функций в композиционной конструкции с алгебраическим кодом имеет следующие параметры:

$$\frac{2}{q^b} - SU\left(q^{2a+b}, q^{aq^{a-b}}, q^b\right) \text{ (RS-код)}, \tag{5}$$

$$\frac{2}{q^b} - SU\left(q^{\frac{5a}{2}+b}, q^{\frac{a\left(q^{2(a-b)}+q^{(a-b)}+2\right)}{2}}, q^b\right) \text{ (НС, } a\text{-четное)}, \quad (6)$$

$$\frac{2}{q^b} - SU\left(q^{3a+b}, q^{a\left(\frac{q^{3(a-b)}}{3}+\frac{q^{2(a-b)}}{2}+\frac{q^{a-b}}{6}\right)}, q^b\right) \text{ (SC-код, } a\text{-нечетное)}, \quad (7)$$

$$\frac{2}{q^b} - SU\left(q^{3a+b}, q^{a\left(q^{2(a-b)}+q^{a-b}+2\right)}, q^b\right) \text{ (C2-4-код)}. \quad (8)$$

Таким образом, схема аутентификации в композиционной конструкции имеет параметры: вероятность коллизий $P_K \le \frac{2}{q^b}$, длина аутентификатора равна b q-ичных символов, объем ключа аутентификации и объем хешируемых данных соответственно определяются первым и вторым параметрами в обозначении SU.

На рис. 1 приведены зависимости объема ключа аутентификации K от объема сообщений с использованием выражений (5-7). Рассмотрены случаи:

а) $b = 11$, $P_k = 2^{-10} \approx 10^{-3}$, $2^8 \le 2^n \le 2^{100}$;

б) $b = 21$, $P_k = 2^{-20} \approx 10^{-6}$, $2^8 \le 2^n \le 2^{100}$;

в) $b = 31$, $P_k = 2^{-30} \approx 10^{-9}$, $2^8 \le 2^n \le 2^{100}$.

Из анализа полученных результатов следует, что при небольшой длине хешируемого сообщения $< 2^{25}$ наиболее экономичным с точки зрения объема ключа аутентификации является схема с RS-кодами. При умеренной длине сообщений $2^{25} < 2^n < 2^{53}$ выигрыш в длине ключа аутентификации обеспечивается НС-кодами, а при $2^n > 2^{53}$ предпочтение имеют коды по кривым Сузуки.

Список литературы: 1. *Wegman M.N., Carter J.L.* New hash functions and their use in authentication and set equality // J. Computer and System Sci. 22 (1981), 265-279. 2. *Carter J.L., Wegman M.N.* Universal classes of hash functions // J. Computer and System Sci. 18 (1979), 143-154. 3. *D.R. Stinson.* Resilient functions and large sets of orthogonal arrays // University of Nebraska, Computer Science and Engineering Department, Report Series: UNL-CSE-93-005 (1993). 4. *Халимов Г.З., Кузнецов А.А.* Аутентификация и универсальное хеширование. //Радиотехника. 2001. Вып. 119. 5. *Stinson D.R.* Universal Hashing and Authentication Codes // Designes, Codes and Criptography 4 (1994), 369-380. A preliminary version appeared in the Procecdins of CRYPTO 91, Lecture Notes in Computer Science 576 (1992), 74-85. 6. *Stinson D.R.* Combinatorial techniques for universal hashing // Journal of Computer and Systems Science 48(1994), 337-346. 7. *Stinson D.R.* On the connections between universal hashing, combinatorial designes and error-correcting codes // Congressus Numerantium 114(1996), 7-27. 8. *Bierbrauer J., Johansson T., Kabatianskii G., Smeets B.* On families of hash function via geometric codes and concatenation // Advances in Cryptology – CRYPTO 93, Lecture Notes in Computer Science 773(1994), 331-342. 9. *Hansen J.P., Stichtenoth H.* // Group Codes on Certain algebraic curves with many rational point, AAECC 1(1990), 67-77. 10. *Hansen J.P.* Group Codes on Deligne-Lusztig, Varieties, Coding Theory and Algebraic Geometry, Proceedings Luminy 1991, Springer LNM 1518, 1992. 11. *Garsia A., Stichtenoth H.* A tower of Artin-Schreiber extensions of function fields attaining the Drinfeld-Vladut bound // Inventiones Mathematicae 121(1995), 211-222.12. *Гоппа В.Д.* Коды на алгебраических кривых // Докл. АН СССР. 1981. Т.259, № 6. С. 1289-1290. 13. *Цфасман М.А.* Коды Гоппы, лежащие выше границы Варшамова – Гилберта // Проблемы передачи информации. М. 1982, №3. С. 3-6. 14. *Pellikaan R., Shen B.Z., G.J.M. van Wee.* Which linear codes are algebraic-geometric? // IEEE Trans. Inform. Theory. IT-37 (1991), 583-602. 15. *Bierbrauer J.* Universal hashing and geometric codes, to appear in Designs // Codes and Cryptography, 1994.

Харьковский военный университет

Поступила в редколлегию 14.03.2001

УДК 681.3.06

Г. З. ХАЛИМОВ, канд. техн. наук, А. А. КУЗНЕЦОВ

АУТЕНТИФИКАЦИЯ И УНИВЕРСАЛЬНОЕ ХЕШИРОВАНИЕ

Введение

Теория аутентификации является математически сопряженной с теорией помехоустойчивого кодирования. В обеих схемах в последовательность передаваемых символов вводится избыточная информация, в результате чего только часть множества всех возможных последовательностей может быть передана в канал связи. Разница заключается в том, что в то время, как в теории кодирования имеется одно правило кодирования, соответствующее фиксированному коду, в кодах аутентификации имеется много правил кодирования, из которых передатчик или приемник может выбирать для использования одно конкретное (секретное) правило. Причем, теория кодирования занимается группированием наиболее вероятных изменений принимаемых сообщений вокруг исходного кодового слова как можно ближе к нему, а теория аутентификации – распределением этих изменений по возможности равномерно по всему множеству сообщений. Взаимосвязь кодирования и аутентификации впервые была отмечена Г. Д. Симмонсом в [1, 2]. Практическое применение теории кодирования для построения схем универсального кодирования и аутентификации предложено в работах [3-7]. Основная идея конструкции кодов аутентификации, изложенная в этих работах, заключается в применении строго универсального класса хеш-функций, построенного с использованием теории ортогональных массивов ОА (orthogonal arrays). Отметим, что в отечественной литературе по теории кодирования с понятием ортогональные массивы отождествляются t-схемы и ортогональные таблицы (см. [8]).

Предложенные в работах [3, 4] схемы аутентификации с использованием простых ортогональных массивов имеют преимущество заключающееся в том, что для них получены точные выражения, связывающие вероятность коллизии с объемом ключевых данных. Недостатком таких схем аутентификации. как показано в данной работе, является то, что объем ключевых данных должен превышать объем передаваемых сообщений.

Задачей данной статьи является изложение общетеоретических вопросов построения аутентификации с применением строго универсального хеширования на основе ортогональных массивов. С этой целью в разделе 1 приводятся определения универсальных и строго универсальных классов хеш-функций для построения кодов аутентификации. В разделе 2 рассматривается теория ортогональных массивов: определения и их свойства; приводятся практические схемы аутентификации с оценкой их параметров.

1. Определение схемы аутентификации строго универсальным классом хеш-функций

Определение универсальных классов хеш-функций для многочисленных применений в криптографии, аутентификации, теории сложности, в алгоритмах рандомизации было предложено Картером и Вергманом в [8]. Приведем основные определения в изложении Биербрауэра [11].

Определение 1 (универсальный класс хеш-функций). Пусть X , Y являются множествами из n и m элементов соответственно, и H есть множество из N функций, осуществляющих отображение

$$f : X \to Y .$$

Пусть $0 < \varepsilon < 1$. H является ε *– универсальным хеш-классом (сокращенно* $\varepsilon - U(N, n, m)$*),* если для двух различных элементов $x_1, x_2 \in X$ существует не больше, чем $N \cdot \varepsilon$ функций $f \in H$ таких, что $f(x_1) = f(x_2)$.

Универсальный класс хеш-функций можно представить матрицей, состоящей из N строк, n столбцов, элементы которой принимают одно из m значений. Каждая функция $f \in H$ связывается со строкой и определяет правило отображения элементов множества X (номеров столбцов матрицы) в элементы Y (собственные значения элементов матрицы). Если заданы два разных значения x_1 и x_2, тогда вероятность того, что $f(x_1) = f(x_2)$ (вероятность коллизии), не может быть больше ε.

Пример 1. Рассмотрим матрицу отображения $X \to Y$ вида

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
f_1	0	1	2	0	1	2	0	2	1
f_2	0	1	2	1	2	0	2	1	0
f_3	0	1	2	2	0	1	1	0	2

y_1	y_2	y_3
0	1	2

Эта матрица является $(1/3)-U(3,9,3)$ классом. Действительно, для любой пары столбцов x_i, x_j вероятность коллизии $(f(x_1)=f(x_2))$ по всему множеству $f \in H$ не будет превышать 1/3.

Пример 2. Пусть множество элементов Y принадлежит полю F_q, $k<q$, X является полем многочленов $X=\left\{p(X)\middle| p(X)=\sum_{i=0}^{k-1} a_i X^i \in F_q[X]\right\}$, и $H=\{f_u | u \in F_q\}$, где

$$f_u(p(X))=p(u).$$

Тогда H есть $\frac{k}{q}-U(q,q^k,q)$.

Действительно, множество элементов X состоит из всех q^k полиномов $p(X)$ степени $<k$ с коэффициентами $a_i \in F_q$. Элементы матрицы отображения $X \to Y$ задаются путем вычисления $p(X)$ в точках $u \in F_q$. Любые два полинома степени $<k$, определенные на множестве значений u, будут совпадать не больше, чем в k точках, тогда вероятность коллизии ε не будет превышать значение $\frac{k}{q}$.

Определение 2 (строго универсальный класс хеш-функций). Пусть X, Y являются множествами из n и m элементов, соответственно, и H есть множество из N функций, осуществляющих отображение

$$f: X \to Y.$$

Пусть $0<\varepsilon<1$. H является ε – *строго универсальным хеш-классом* (сокращенно $\varepsilon - SU(N,n,m)$), если выполняются следующие условия:

1) для каждого $x \in X$ и $y \in Y$ число функций $f \in H$ таких, что $f(x)=y$ равно $\frac{N}{m}$;

2) для любых различных элементов $x_1, x_2 \in X$ и не обязательно различных $y_1, y_2 \in Y$ число ν функций $f \in H$ таких, что $f(x_1)=y_1$, $f(x_2)=y_2$ не превышает $\nu \le \varepsilon \cdot \frac{N}{m}$.

Первое утверждение определяет, что число записей со значением y в каждом столбце матрицы отображения $X \to Y$ встречается одинаковое число раз.

Пример 3. Рассмотрим расширенный (4, 2, 2) код РС над $GF(4)$. Кодовые образуют матрицу отображения $X \to Y$ типа $\frac{1}{4}-SU(16,4,4)$.

	x_1	x_2	x_3	x_4
f_1	0	0	0	0
f_2	0	1	α	β
f_3	0	α	β	1
f_4	0	β	1	α
f_5	1	1	1	1

f_6	1	α	0	β
f_7	1	0	β	α
f_8	1	β	α	0
f_9	α	α	α	α
f_{10}	α	β	0	1
f_{11}	α	0	1	β

f_{12}	α	1	β	0
f_{13}	β	β	β	β
f_{14}	β	1	0	α
f_{15}	β	0	α	1
f_{16}	β	α	1	0

Действительно, для каждого $x \in X$ и $y \in Y$ число функций $f(x) = y$ равно $\frac{N}{m} = 4$, а число ν функций f таких, что $f(x_1) = y_1$, $f(x_2) = y_2$, равно $\nu = \varepsilon \cdot \frac{N}{m} = 1$.

Утверждение 1. Для схемы аутентификации с использованием универсального класса хеш-функций вероятность коллизии кодов аутентификации $P_{кол} \le \varepsilon$.

Пусть состояние источника сообщений определяется X. Секретный ключ является одной из функций $f \in H$. В передаваемое сообщение включается аутентификатор $y = f(x)$. Нарушитель может наблюдать передаваемое сообщение x и его аутентификатор $y = f(x)$. Для подмены сообщения он должен сформировать x' и соответствующий аутентификатор $y' = f(x')$. Вероятность подмены будет определяться условной вероятностью $P(f(x') = y' \mid f(x) = y)$, которая, по определению $\varepsilon - SU(N, n, m)$ функций, не может превышать ε.

Действительно, по формуле Байеса имеем

$$P(f(x'), f(x)) = P(f(x')) \cdot P(f(x') \mid f(x)). \qquad (1)$$

Так как число функций f для которых $f(x) = y$, равно $\frac{N}{m}$, тогда $P(y) = \frac{1}{m}$. Число функций f, для которых $f(x) = y$, $f(x') = y'$, $\nu \le \varepsilon \frac{N}{m}$, следовательно, $P(y, y') \le \varepsilon \cdot \frac{1}{m}$.

Подставляя оценки $P(y)$, $P(y, y')$ в (1), получим

$$P(f(x') \mid f(x) \le \varepsilon. \qquad (2)$$

Теорема 1. Пусть q – простое число, a, b, k целые числа, $a > b$. Тогда существует $\frac{k}{q^b} - SU(q^{a+b}, q^{ka}, q^b)$ семейство хеш-функций.

Фиксируем сюрективное F_q – линейное отображение

$$\varphi : F_{q^a} \to F_{q^b}.$$

X является множеством полиномов $p(X)$, определенным над F_{q^a} степени $\le k$, без постоянного члена.

Элементы матрицы отображения $X \to Y$ на пересечении (u, υ) строки, $u \in F_{q^a}$, $\upsilon \in F_{q^b}$, и $p(X)$ столбца определим как

$$\varphi(p(u)) + \upsilon = y.$$

Очевидно, что число элементов $y \in F_{q^b}$ в столбце равно $N = q^{a+b}$, и они повторяются точно $\frac{N}{m} = q^a$ раз. Зафиксируем $p_1(X)$, $p_2(X)$ и $y_1, y_2 \in F_{q^b}$, то есть

$$\varphi(p_1(u)) + \upsilon = y_1;$$
$$\varphi(p_2(u)) + \upsilon = y_2.$$

В силу линейности отображения φ справедливо выражение

$$\varphi((p_1 - p_2)(u)) = y_1 - y_2.$$

Каждое решение u этого уравнения даст нам уникальную пару (u, υ). Число решений (u, υ) будет $\le kq^{a-b}$, т. к. мощность обратного отображения $M = \varphi^{-1}(y_1 - y_2)$ $F_{q^b} \to F_{q^a}$ равна $|M| = q^{a-b}$, а $p_1(X) - p_2(X)$ есть полином степени $\le k$. Тогда $\varepsilon = kq^{a-b} / q^a = k / q^b$.

Как следует из приведенного доказательства, применение многочленов для отображения $X \to Y$ является эффективным средством для построения хорошего класса хеш-функций, что и будет рассмотрено ниже.

Легко заметить, что второе условие в определении 2 строго универсального семейства хеш-функций соответствует обобщенному определению ортогональных массивов. Действительно, предположим, что для множества функций $f: X \to Y$ это условие выполнено. Тогда число v равно точно числу функций f, для которых справедливо $f(x_1) = y_1$ и $f(x_2) = y_2$ для всех наборов (x_1, x_2) и (y_1, y_2). В этом случае очевидно, что $N = vm^2$. Тогда $\varepsilon = \frac{1}{m}$ и матрица отображения $X \to Y$ является ортогональным массивом силы 2 по определению.

1. Применение ортогональных массивов для построения схемы аутентификации

Ортогональные массивы изучаются давно, методы их построения и свойства достаточно полно представлены в [9, 10,11]. Изложим основные результаты.

Определение 3. Пусть X, Y являются множествами из k и υ элементов, соответственно, и H есть множество функций, осуществляющих отображение $f: X \to Y$. *Ортогональным массивом* $OA_{\lambda}(t,k,\upsilon)$ называется массив элементов $y_i \in Y$ со столбцами, соответствующими элементам множества X и строками, определяемыми элементами множества m, в котором для любой выборки из t элементов $y_1, y_2, ..., y_t$ из Y существует только λ функций $f \in m$, для которых справедливо $f(x_i) = y_i$, $i = 1,2,...,t$.

Определение 4. Массив $OA_{\lambda}(t,k,\upsilon)$ называется *простым*, если каждая строка повторяется только (точно) один раз.

Основная конструкция OA массивов определена следующей теоремой [11].

Теорема 2. Пусть q простое число, m, n, t – целые числа, $n \geq m$, $2 \leq t \leq q^n$. Зафиксируем сюръективное F_q – линейное отображение $\varphi: F_q^n \to F_q^m$. Для каждого t набора $(z, a_1, a_2, ..., a_{t-1})$, где $z \in F_{q^m}$, $a_j \in F_{q^n}$, $i = 1,2,...,t-1$, определим отображение $f = f(z, a_1, a_2, ..., a_{t-1}): F_q^n \to F_q^m$, вида

$$f(x) = \varphi\left(\sum_{j=1}^{t-1} a_j x^j\right) + z. \qquad (3)$$

Тогда массив, составленный из отображений вида (3), является ортогональным с параметрами $OA_{q^{(t-1)(n-m)}}\left(t, q^n, q^m\right)$.

Пусть $p(x) = \sum_{i=1}^{t-1} a_i x^i$. Для каждой пары $(p(x), z)$ справедливо

$$\varphi(p(x_j)) + z = \alpha_j,\ j = 1,2,...,t;\ z, \alpha_1, \alpha_2, ..., \alpha_t \in F_{q^m}.$$

По условию линейности отображения φ получим

$$\varphi(p(x_1) - p(x_j)) = \alpha_1 - \alpha_j,\ j = 2,3,...,t.$$

Пусть $U_j = \varphi^{-1}(\alpha_1 - \alpha_j)$, $j = 2,3,...,t$ обратное отображение $F_q^n \to F_q^m$. Так как отображение φ сюръективно получим $|U_j| = q^{n-m}$. Для каждого $(t-1)$ набора $(u_2, ..., u_t)$, $u_j \in U_j$ существует единственный многочлен $p(x) \in F_q^n[X]$ степени $< t$ удовлетворяющий условию

$$p(x_1) - p(x_j) = u_j\ \ (j = 2,3,...,t).$$

Число $(t-1)$ наборов $(u_2, ..., u_t)$ точно равно $q^{(t-1)(n-m)}$, что и требовалось доказать.

Рассмотрим вопрос, когда OA массив является простым. Пусть $(p_1(x), z_1)$ и $(p_2(x), z_2)$ опре-

деляют строки массива OA и

$$p_i(x) \in P = \left\{ p(x) \middle| p(x) \in F_q^n[X], p(0) = 0, \deg(p(x)) < t \right\}.$$

Строки в OA являются равными, если справедливо следующее

$$\varphi(p_1(u)) + z_1 = \varphi(p_2(u)) + z_2 \text{ для вех } u \in F_q^n .$$

Полагая $u = 0$, получим $z_1 = z_2$. Условие равенства определяется условием

$$\varphi((p_1 - p_2)(u)) = 0 , \ \forall u \in F_q^n .$$

Определим $P_0 = P_0(q, n, m, t, \varphi)$ как множество полиномов $p(x) \in P$, удовлетворяющих условию

$$\varphi(p(u)) = 0 , \ \forall u \in F_{q^n} .$$

Предложение 1. Две строки $(p_1(x), z_1)$ и $(p_2(x), z_2)$ ортогонального массива равны тогда и только тогда, когда

$$z_1 = z_2$$
$$(p_1(x) - p_2(x)) \in P_0$$

Очевидно, что строки в OA повторяются с периодом мультипликативного порядка q^ρ. Ниже приведем без доказательства несколько важных утверждений [11].

Утверждение 1. Пусть q, n, m, t, φ определены как в теореме 2. Тогда существует простой ортогональный массив с параметрами

$$OA_{q^{(t-1)(n-m)-\rho}}\left(t, q^n, q^m\right).$$

Лемма 1. Если $t \le q^m$, тогда $\rho = \dim(P_0) = 0$ и простой ортогональный массив будет иметь параметры

$$OA_{q^{(t-1)(n-m)}}\left(t, q^n, q^m\right).$$

Теорема 3. Если $m = 1$, $\varphi = trace : F_q^n \to F_q$, $t > q^i$, тогда P_0 включает F_q^n полиномы вида

$$(a_i x)^{q^i} + \ldots + (a_1 x)^q + a_0 x , \text{ где } \sum_j a_j = 0 .$$

Тогда $\rho = \dim(P_0) \ge i \cdot n$ и существует простой ортогональный массив с параметрами

$$OA_{q^{(t-1)(n-m)-in}}\left(t, q^n, q\right).$$

Теорема 4. Если $m = 1$, $\varphi = trace : F_q^n \to F_q$, $t = q + 1$, тогда

$$P_0 = \left\{ (ax)^q - ax \middle| a \in F_q^n \right\}.$$

Тогда $\rho = \dim(P_0) = n$ и существует простой ортогональный массив с параметрами

$$OA_{q^{(t-1)(n-m)-1}}\left(q+1, q^n, q\right).$$

Теорема 5. Пусть $m < n$. Тогда $\rho = \dim(P_0) = n$ и существует простой ортогональный массив с параметрами

$$OA_{q^{q^m(n-m)-n}}\left(q^{m+1}, q^n, q^m\right).$$

Теорема 6. Пусть $n \ge m + 2$, $t > q^{m+1}$. Тогда $\rho = \dim(P_0) \ge 2n - m$.

Рассмотрим следующие примеры.

Пример 4. Пусть $q = 2$, $n = 4$, $m = 2$. Построим поле F_2^4 с помощью многочлена $z^4 + z + 1$ и примитивного элемента $\alpha = z$. Линейное отображение $\varphi : F_2^4 \to F_2^2(0, \beta^0, \beta, \beta^2)$ определим как проекцию первых двух координат элементов поля F_2^4. В результате получим следующую матрицу преобразований

$0 = 0000 \leftrightarrow 0 = 00$ $\quad \alpha^4 = 0011 \leftrightarrow 0 = 00$ $\quad \alpha^8 = 0101 \leftrightarrow \beta^0 = 01$ $\quad \alpha^{12} = 1111 \leftrightarrow \beta^2 = 11$

$\alpha = 0010 \leftrightarrow 0 = 00$ $\quad \alpha^5 = 0110 \leftrightarrow \beta^0 = 01$ $\quad \alpha^9 = 1010 \leftrightarrow \beta = 10$ $\quad \alpha^{13} = 1101 \leftrightarrow \beta^2 = 11$

$\alpha^2 = 0100 \leftrightarrow \beta^0 = 01$ $\quad \alpha^6 = 1100 \leftrightarrow \beta^2 = 11$ $\quad \alpha^{10} = 0111 \leftrightarrow \beta^0 = 01$ $\quad \alpha^{14} = 1001 \leftrightarrow \beta = 10$

$\alpha^3 = 1000 \leftrightarrow \beta = 10$ $\quad \alpha^7 = 1011 \leftrightarrow \beta = 11$ $\quad \alpha^{11} = 1110 \leftrightarrow \beta^2 = 11$ $\quad \alpha^0 = 0001 \leftrightarrow 0 = 00$

Очевидно, что мощность обратного отображения φ^{-1} равна $2^{n-m} = 2^2$.

Пример 5. Пусть $q = 2$, $n = 3$, $m = 1$, $t = 3$. По теореме 4 существует простая ортогональная матрица с параметрами $OA_2(3,2^3,2)$, которая является кодом Рида – Малера первого порядка. Для $R(1,3)$ выпишем все кодовые слова.

Очевидно, что комбинация из $t = 3$ произвольных элементов x_i для любой строки массива $f_j(x)$ повторяется по всем строкам $\lambda = 2$ раз.

Пример 6. Пусть $q = 2$, $n = 4$, $t = 2$. По теореме 2 и лемме 1 существует простой ортогональный массив с параметрами $OA_4(2,2^4,2^2)$. Для построения $OA_\lambda(t,q^n,q^m)$ используем линейное отображение $\varphi : F_2^4 \to F_2^2$, как в примере 4 с функцией $f = f(z,a)$ вида

$$f(x) = \varphi(ax) + z.$$

Ортогональный массив будет иметь вид матрицы, в которой строки определяются функциями f_i с параметрами $a_i \in F_{2^4}$, $z_i \in F_{2^2}$, столбцы – значениями $x_i \in F_{2^4}$, а элементы – значениями $y_i \in F_{2^2}$. Для $z = 0$ получим матрицу следующего вида.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
f_0	0	0	0	0	0	0	0	0
f_1	0	0	0	0	1	1	1	1
f_2	0	0	1	1	0	0	1	1
f_3	0	1	0	1	0	1	0	1
f_4	0	0	1	1	1	1	0	0
f_5	0	1	0	1	1	0	1	0
f_6	0	1	1	0	0	1	1	0
f_7	0	1	1	0	1	0	0	1
f_8	1	1	1	1	1	1	1	1
f_9	1	1	1	1	0	0	0	0
f_{10}	1	1	0	0	1	1	0	0
f_{11}	1	0	1	0	1	0	1	0
f_{12}	1	1	0	0	0	0	1	1
f_{13}	1	0	1	0	0	1	0	1
f_{14}	1	0	0	1	1	0	0	1
f_{15}	1	0	0	1	0	1	1	0

$f_i(a_i x_j)$	x_j															
	0	1	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^9	α^{10}	α^{11}	α^{12}	α^{13}	α^{14}
$f_0(0x_j)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$f_1(\alpha^0 x_j)$	0	0	0	0	β	0	1	β^2	β	1	β	1	β^2	β^2	β^2	β
$f_2(\alpha^1 x_j)$	0	0	0	β	0	1	β^2	β	1	β	1	β^2	β^2	β^2	β	0
$f_3(\alpha^2 x_j)$	0	0	β	0	1	β^2	β	1	β	1	β^2	β^2	β^2	β	0	0
$f_4(\alpha^3 x_j)$	0	β	0	1	β^2	β	1	β	1	β^2	β^2	β^2	β	0	0	1
$f_5(\alpha^4 x_j)$	0	0	1	β^2	β	1	β	1	β^2	β^2	β^2	β	0	0	1	β
$f_6(\alpha^5 x_j)$	0	1	β^2	β	1	β	1	β^2	β^2	β^2	β	0	0	1	β	0
$f_7(\alpha^6 x_j)$	0	β^2	β	1	β	1	β^2	β^2	β^2	β	0	0	1	β	0	1
$f_8(\alpha^7 x_j)$	0	β	1	β	1	β^2	β^2	β^2	β	0	0	1	β	0	1	β^2
$f_9(\alpha^8 x_j)$	0	1	β	1	β^2	β^2	β^2	β	0	0	1	β	0	1	β^2	β
$f_{10}(\alpha^9 x_j)$	0	β	1	β^2	β^2	β^2	β	0	0	1	β	0	1	β^2	β	1
$f_{11}(\alpha^{10} x_j)$	0	1	β^2	β^2	β^2	β	0	0	1	β	0	1	β^2	β	1	β

$f_{12}(\alpha^{11}x_j)$	0	β^2	β^2	β^2	β	0	0	1	β	0	1	β^2	β	1	β	1
$f_{13}(\alpha^{12}x_j)$	0	β^2	β^2	β	0	0	1	β	0	1	β^2	β	1	β	1	β^2
$f_{14}(\alpha^{13}x_j)$	0	β^2	β	0	0	1	β	0	1	β^2	β	1	β	1	β^2	β^2
$f_{15}(\alpha^{14}x_j)$	0	β	0	0	1	β	0	1	β^2	β	1	β	1	β^2	β^2	β^2

Если дополнить эту матрицу еще тремя, при $z_1 = 1$, $z_2 = \beta$, $z_3 = \beta^2$, тогда получим $OA_4(2,16,4)$. Действительно, анализ приведенной матрицы показывает, что существует самое большее $\lambda = 4$ функций, для которых справедливо $f(x_1) = y_1$ и $f(x_2) = y_2$. Данный ортогональный массив является семейством строго универсальных хеш-функций. По определению 2 имеем следующие параметры. Общее число функций $N = 64$. Число записей со значением у в каждом столбце матрицы отображения $X \to Y$ встречается $\frac{N}{2^m} = 16$ раз. Число функций $f \in H$ таких, что $f(x_1) = y_1$, $f(x_2) = y_2$, не превышает $\nu \le 4$, так как $\lambda = 4$. Вероятность коллизии ε будет равна $\varepsilon \cdot \frac{N}{2^m} = \lambda$. $\varepsilon = \frac{1}{4}$ и имеем $\frac{1}{4} - SU(64,16,4)$ семейство хеш-функций. Результаты этого примера обобщим в следующем утверждении.

Утверждение 3. Пусть $t = 2$, q – простое число, n, m – целые числа. Тогда ортогональный массив с параметрами $OA_{q^{n+m}}(2, q^n, q^m)$ является строго универсальным классом хеш-функций $\frac{1}{q^m} - SU\left(q^{n+m}, q^n, q^m\right)$.

Действительно, при $t = 2$ сюрективное отображение $\varphi : F_{q^n} \to F_{q^m}$ в виде $\varphi(p(x)) + z = y$ образует ортогональный массив по теореме 2. Так как $p(x) = ax$ множество полиномов первой степени без постоянного члена, то это соответствует результату теоремы 1 для случая $k = 1$, что и завершает доказательство.

Преимуществом конструкции $\frac{1}{q^m} - SU\left(q^{n+m}, q^n, q^m\right)$ семейства хеш-функций, описанной в примере 6, является то, что для вычисления кода аутентификации необходимо использовать операции умножения элементов в F_{q^n}, проектирование m координат $F_{q^n} \to F_{q^m}$ и сложение в F_{q^m}. Практические схемы для таких вычислений легко реализуются. Очевидно, что в такой схеме аутентификации вероятность коллизии теоретически является наименьшей и определяется как величина, обратная мощности множества кодов аутентификации. При этом объем ключевых данных определяется как мощность множества функций отображения $X \to Y$ и равен произведению объема исходных сообщений на мощность множества кодов аутентификации. Практически это означает, что по закрытому каналу связи необходимо передавать ключевых данных больше, чем по открытому, – информационных данных.

Список литературы: 1. *Simmons G.I.* Authentication theory/coding theory, presented at Crypto'84, Santa Barbara, CA, P. 19-22, 1984. 2. *Симмонс Г.Д.* Обзор методов аутентификации информации // ТИИЭР 1988. Т.76, № 5. С.105-125. 3. *Wegman M.N., Carter J. L.* New hash functions and their use in authentication and set equality // J. Computer and System Sci. 22 (1981), 265-279. 4. *Carter J.L., Wegman M. N.* Universal classes of hash functions // J. Computer and System Sci. 18 (1979), 143-154. 5. *Stinson D.R.* Universal Hashing and Authentication Codes // Designes, Codes and Criptography 4 (1994), 369-380. A preliminary version appeared in the Procecdins of CRYPTO 91, Lecture Notes in Computer Science 576 (1992), 74-85. 6. *Kabatianskii G., Smeets B., Johansson T.* On the Cardinality of Systematic Authentication Codes Via Error-Correcting Codes. //IEEE Transactions on Information Theory. 1991. Vol. IT-42. 583-602. 7. *Bierbrauer J., Johansson T., Kabatianskii G., Smeets B.* On families of hash function via geometric codes and concatenation. //Advances in Cryptology – CRYPTO 93, Lecture Notes in Computer Science. 1994. 773. 331-342. 8. *Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А.* Теория кодов, исправляющих ошибки. М.: Связь, 1979. 744. С. 9. *Mukhopadhyay A.L.* Construction of some series of ortogonal array // Sankya B43 (1981), 81-92. 10. *Stinson D.R.* Resilient functions and large sets of orthogonal arrays // University of Nebraska, Computer Science and Engineering Department, Report Series: UNL-CSE-93-005 (1993). 11. *Bierbrauer J.* Construction of orthogonal arrays, to appear in Journal of Statistical Planning and Inference.

Харьковский военный университет

Поступила в редколлегию 14.03.2001

УДК 681.3.06:519.248.681

О. А. ЗАМУЛА, канд. техн. наук, Є. В. ПОПОВИЧ, Ю. І. ГОРБЕНКО

УМОВИ ТА МОЖЛИВОСТІ СТВОРЕННЯ БЕЗУМОВНО СТІЙКИХ КРИПТОСИСТЕМ

Вступ

На сьогодні використовуються, удосконалюються та розроблюються криптографічні системи (криптосистеми), які забезпечують різноманітний рівень криптостійкості. В ряді джерел [1-3] наведено умови створення криптосистем з різними рівнями стійкості. Проведений аналіз [1] показав, що, якщо в основу класифікації покласти рівень стійкості, то існуючі криптосистеми можна поділити на чотири класи:

1. Безумовно стійкі криптосистеми (БСК).
2. Розрахунково стійкі криптосистеми.
3. Доказуємо стійкі криптосистеми (імовірно стійкі).
4. Розрахунково нестійкі криптосистеми (тимчасової стійкості).

Умови та можливості реалізації таких криптосистем залежать від рівня розвитку математичних методів та систем криптоаналізу, тому створення умов і можливостей їх реалізації змінюються з часом. На сьогодні, на наш погляд, вже можна говорити та створювати криптосистеми та засоби, які забезпечують в різноманітних інформаційних технологіях вказані рівні стійкості. Особливо актуальними є задачі створення безумовно стійких криптосистем.

Метою статті є розгляд умов та можливостей створення на сучасному етапі розвитку безумовно стійких криптосистем.

1. Модель взаємодії користувачів

На рис. 1 наведена спрощена схема інформаційних співвідношень між двома абонентами А1 та А2.

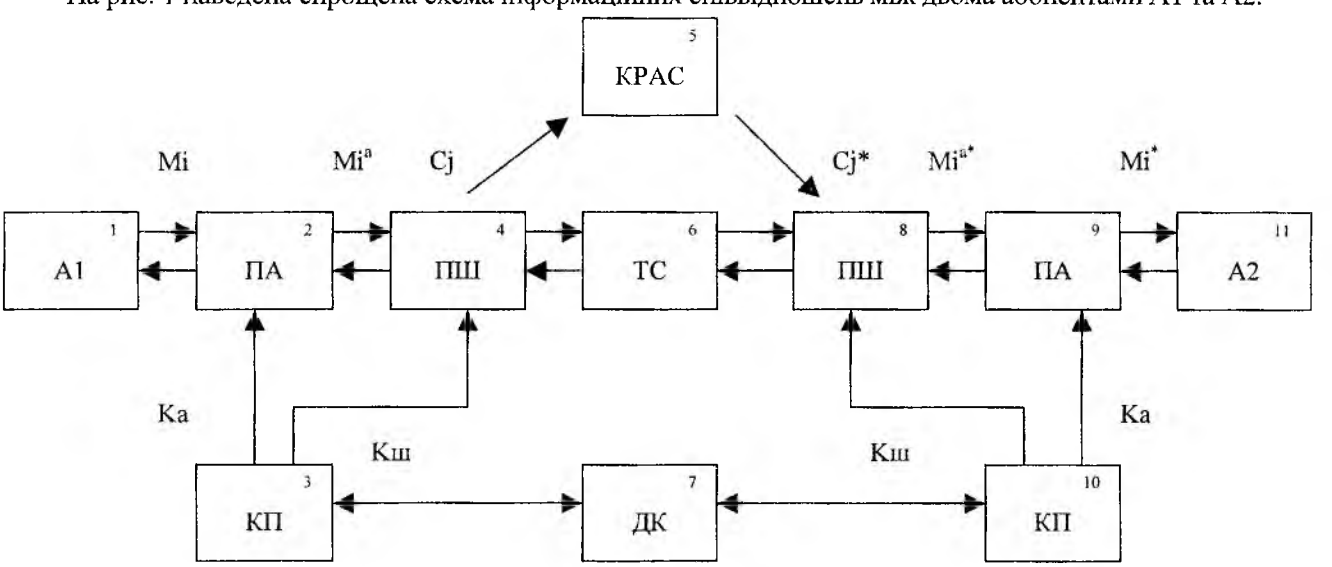


Рис. 1

На рис. 1 введені такі позначення: 2, 9 – пристрої автентифікації; 4, 8 – пристрої шифрування;3, 10 – ключові пристрої; 7 – джерело ключа; 5 – криптоаналітична система (криптоаналітик); 6 – комунікаційна система.

Будемо вважати, що А1 та А2 являють собою джерела інформації M_i з довільною потужністю алфавіту m та з відомою апріорною статистичною ймовірністю $P(M_i)$ для усіх повідомлень ($i = \overline{1, n_m}$) та ентропією джерела інформації $H(M_i)$. Аналогічно для А2 – m, $P(M_j)$ для усіх $j = \overline{1, n_m}$, *та* $H(M_j)$.

Оскільки повідомлення передається відкритою телекомунікаційною системою, то повинні бути забезпечені конфіденційність та автентичність відповідного рівня. Будемо вважати, що пристрій автентичності забезпечує надання користувачу послуг цілісності та справжності, а пристрій шифрування – послугу конфіденційності.

З метою забезпечення цих послуг використаємо криптоперетворення відповідних класів. Для реалізації криптоперетворень будемо використовувати ключі аутентіфікації K_a та ключі шифрування / розшифрування $K_ш$. Також будемо вважати, що ключі генеруються джерелом ключів та за слушними протоколами розповсюджуються абонентам А1 та А2.

Протидію системі здійснює криптоаналітична система (криптоаналітик). В подальшому будемо розглядати цю систему як зовнішню, так і внутрішню, наприклад, санкціонований користувач. Будемо вважати, що криптоаналітик перехоплює криптограми з необхідною йому імовірністю. Пристрій аутентифікації здійснює криптографічні перетворення за ключем K_a з метою забезпечення цілісності та достовірності повідомлення M_i. В подальшому будемо його подавати як M_i^a:

$$M_i^a=F_a(M_i,K_a,P_r), \tag{1}$$

де F_a – функція аутентіфікації; P_r – параметр перетворювання.

Пристрій шифрування здійснює зашифровування або розшифровування повідомлення M_i^a. Будемо вважати відомою статистику появи шифрограми $P(C_j)$ та статистику появи шифрограми M_i^a – $P(C_j/M_i^a)$ після їх зашифрування. При цьому:

$$C_j=F_з(M_i^a,K_з,P_r), \tag{2}$$

де $F_з$ – функція зашифрування.

Криптоаналітична система, що перехоплює криптограму C_j, має можливість розрахувати апостеріорну статистику $P(M_i^a/C_j)$ – ймовірність того, що C_j містить в собі M_i^a.

Абонент А2 приймає криптограму C_j^* – (знак "*" означає, що вона могла бути випадково або навмисно викривлена). Пристрій шифрування 8 здійснює розшифрування даної криптограми, при цьому M_i^{a*} утворюється як:

$$M_i^{a*}=F_p(C_j^*,K_p,P_r), \tag{3}$$

де F_p – функція розшифрування.

Пристрій аутентифікації здійснює криптоперетворення повідомлення M_i^{a*} з метою контролю цілісності та достовірності змісту повідомлення M_i^*. Якщо повідомлення M_i^* достовірне та цілісне, воно видається абоненту А2. При цьому на вхід А2 надходить повідомлення M_i^*, і, якщо $M_i^* = M_i$, то ми будемо вважати, що передача здійснена без порушення автентичності та цілісності, а якщо $M_i^* \neq M_i$, то порушена цілісність та автентичність при передачі повідомлення.

Будемо також вважати, що джерело ключів формує на своєму виході ключі K_i^a та $K_j^ш$ рівноймовірно, випадково та незалежно з ентропією джерела ключів $H(K^a)$ та $H(K^ш)$ з відповідними параметрами, і що інформаційна система здійснює періодичне передавання криптограм, і криптоаналітик їх перехоплює. Також будемо вважати, що на основі апріорної та апостеріорної статистики криптоаналітик розрахував $P(M_i/C_j)$ для $i = \overline{1, n_M}$ та $j = \overline{1, n_C}$. При цьому розмірність апостеріорного ряду

$$n_p = n_м \times n_c. \tag{4}$$

Якщо розмірність ряду, що визначена формулою (4) дуже велика, то практично побудувати або розрахувати ряд неможливо. В зв'язку з цим криптоаналітик повинен вести криптоаналіз з використанням апостеріорної ентропії H(M/C). При цьому апостеріорна ентропія розраховується згідно співвідношень:

$$H(M/C_j)=-\sum P(M_i/C_j)\ log_2P(M_i/C_j), \tag{5}$$

Та

$$H(M/C)=\sum_j P(C_j)H(M/C_j)=\sum_{j=1}^{n_c}\sum_{i=1}^{j n_m} P(C_j)\ P(M_i/C_j)\ log_2P(M_i/C_j). \tag{6}$$

До початку ведення криптоаналізу криптоаналітик має невизначенність (ентропію) відносно повідомлення:

$$H(M)=-\sum_i^{n_m} P(M_i)\ log_2P(M_i). \tag{7}$$

$H(M)$ – отримано при умові, що криптоаналітик знає апріорний ряд $P(M_i)$ – ймовірність появи повідомлення на виході джерела повідомлень. Після перехвату необхідної кількості криптограм невизначеність криптоаналітика відносно джерела повідомлень визначається за формулою (6).

2. Умови реалізації безумовно стійкої криптосистеми

Розглянемо умови реалізації безумовно стійкої криптосистеми, використовуючи модель, приведену на рис. 1.

Визначимо, перш за все, яку кількість інформації може отримати криптоаналітик. Умовна ентропія $H(M/C)$ характеризує невизначеність криптоаналітика після значної кількості отриманих криптограм, причому $H(M/C)$ характеризує середню невизначеність криптоаналітика відносно джерела повідомлень. Кількість інформації, яку він отримує про джерело повідомлень після проведення криптоаналізу, визначається формулою:

$$\Delta I=H(M)-H(M/C). \tag{8}$$

При умові, що криптоаналітик не отримує ніякої інформації про джерело повідомлень ($\Delta I=0$), маємо:

$$H(M)=H(M/C). \tag{9}$$

Умова (9) і є умовою реалізації безумовно стійкої криптосистеми. Причому, скільки б шифрограм не перехоплював криптоаналітик, він не збільшить своїх знань про джерело інформації. В цьому випадку криптоаналіз є безглуздим.

Коли криптоаналітик розкриває систему, тобто ($H(M/C)=0$), то $\Delta I=H(M)$. Це означає, що кількість інформації, яку він отримав, дорівнює ентропії джерела повідомлення Н(М). В більшості криптосистем:

$$0<H(M/C)<H(M). \tag{10}$$

Співвідношення, що наведено вище, торкається конфіденційності.

Наведемо теорему [1], яка визначає необхідні та достатні умови реалізації безумовно стійких криптосистем. При цьому зазначимо, що за сучасним поглядом співвідношення (9) є як необхідною, так і достатньою умовою, але воно не визначає практичних методів досягнення мети.

Теорема 1. Необхідною та достатньою умовою забезпечення безумовної стійкості у системі, схема якої наведена на рис. 1, є наступне:

$$P(C_j/M_i)=P(C_j), \tag{11}$$

тобто ймовірність появи криптограми не повинна залежати ні від того, яке повідомлення вибрано на виході джерела повідомлень, ні від того, який ключ з'явився на виході джерела ключів.

З (11) випливає, що в безумовно стійких криптосистемах (теоритично стійких системах) кожне повідомлення M_i повинно з однаковою ймовірністю відображатися в кожну криптограму. При цьому ми не накладали обмежень ні на потужність алфавіту повідомлення та ключа, ні на довжину повідомлень та криптограм.

Будемо вважати, що джерело повідомлень має алфавіт m_M, джерело криптограм – m_C, а довжина повідомлень та криптограм відповідно – l_M та l_C.

Доведення теореми 1. Для доведення теореми розглянемо апостеріорні ймовірності $P(M_i/C_i)$, уважаючи, що криптоаналітик перехвачує необхідну йому кількість криптограм. Тобто криптоаналіз відбувається в умовах вибору криптотексту (при відомому криптотексті). Використовуючи теорему Байєса, Р (M_i/C_i) може бути визначена як:

$$P(M_i/C_i)=\frac{P(M_i)P(C_i/M_i)}{P(C_j)}=\frac{P(M_i)P(K_{ij})}{\sum_{i=1}^{n_m}P(M_i)P(K_{ij})}. \tag{12}$$

Відповідно до розглянутого вище співвідношення (9) умовою безумовної стійкості є $H(M)=H(M/C)$. Стосовно до ймовірності появи повідомлення $P(M_i)$ та апостеріорної ймовірності $P(M_i/C_j)$ можливо записати, що в безумовно стійких системах ймовірність $P(M_i/C_j)$ під час криптоаналізу не повинна змінюватися відносно $P(M_i)$, тобто

$$P(M_i/C_j)=P(M_i). \tag{13}$$

Інакше, після криптоаналізу, криптоаналітик не отримує будь якої додаткової інформації і його апостеріорні знання не збільшуються відносно джерела інформації для $i=\overline{1,n_M}$ та $j=\overline{1,n_C}$. Розділимо ліву та праву частини співвідношення (12) на $P(M_i) \neq 0$. В результаті маємо:

$$\frac{P(M_i/C_i)}{P(M_i)}=\frac{P(C_j/M_i)}{P(C_j)}=1, \tag{14}$$

звідки $P(C_j/M_i) = P(C_j)$. Таким чином умова (11) є як необхідною, так і достатньою.

Таким чином в безумовно стійкій системі ймовірність появи криптограми на виході пристрою шифрування не повинна залежати ні від ймовірності появи повідомлень ні від ймовірності появи ключа. Крім того, кількість криптограм повинна бути не менша за кількість повідомлень M_i. Для однозначності дешифрування, це означає, що кількість ключів повинна бути не менш за кількість повідомлень. Тобто:

$$N_K \geq N_M. \tag{15}$$

З кута зору розглянутого вище інформаційного підходу це означає, що ентропія джерела ключа повинна бути більшою або рівною ентропії джерела повідомлень:

$$N(K) \geq N(M). \tag{16}$$

Для вихідних даних, наведених на рис. 1, кількість повідомлень N_M довжиною l_M при m_M алфавіті

$$N_M = m_M^{l_M}. \tag{17}$$

Для ключів з потужністю алфавіту m_K та довжиною l_M

$$N_K = m_K^{l_K}. \tag{18}$$

Якщо вважати, що всі повідомлення та ключі є равноіймовірні, маємо:

$$P(M_i)=\frac{1}{N_M}=m_M^{-l_M}, \tag{19}$$

$$P(K_j)=\frac{1}{N_K}=m_K^{-l_K}. \tag{20}$$

Після підстановки (19) та (20) в (16) отримаємо:

$$H(K)=-\sum_{i=1}^{n_K}P(K_j)\log_2 P(K_j)=\left(N_K\frac{1}{N_K}\log_2\frac{1}{N_K}\right)=\log_2 N_K=\log_2 m_K^{l_K}. \tag{21}$$

За аналогією –

$$H(M)=\log_2 m_M^{l_M}. \tag{22}$$

Після підстановки (21) та (22) в (16), отримаємо:

$$\log_2 m_K^{l_K} \geq \log_2 m_M^{l_M},$$

або

$$l_K \log_2 m_K \geq l_M \log_2 m_M$$

Якщо потужність алфавіту джерела повідомлень та джерела ключів однакова ($m_K=m_M$), а це майже завжди так, то:

$$l_K \geq l_M. \tag{23}$$

Інакше:

$$l_K=\frac{\log_2 m_M}{\log_2 m_K}l_M \tag{24}$$

3. Методи реалізації безумовно стійкої криптосистеми

Проведений аналіз показав, що висунутим вимогам (вибір ключів здійснюється рівноймовірно, випадково та незалежно, а також виконується умова (23)) задовольняє криптосистема відома під назвою «Система Вернама» [3]. В ній зашифрування здійснюється методом потокового криптографічного перетворення за правилом:

$$C_i=(M_i+K_i^1)mod\ m. \tag{25}$$

де m – потужність алфавіту C_i.

Принципова вимога до цього перетворення – це $l_{K_i} \geq l_{M_i}$. Розшифрування в такій системі здійснюється за правилом:

$$M_i=(C_i-K_i^P)mod\ m. \tag{26}$$

Безпосередній аналіз співвідношень (2.25) та (2.26) означає, що для розшифрування повідомлення M_i необхідно забезпечити синхронізацію K_i^3 та K_i^P.

Оцінимо стійкість такої системи проти різноманітних криптоаналітичних атак. Оскільки така система безумовно стійка, то при додержанні усіх вищезазначених вимог найкраща атака – це атака типу "брудна сила".

З метою оцінки складності реалізації такої атаки, можна використати показник безпечного часу системи

$$t_б = P_p \frac{N_K}{\gamma K}. \tag{27}$$

де P_p – ймовірність успішного рішення задачі; N_K – кількість ключів; γ – продуктивність аналітичної системи (кількість переборів за секунду); K – коефіцієнт перерахунку, який дорівнює 3.1×10^7 с/рік для отримання значення $t_б$ в роках.

При умові (23) кількість ключів визначається $N_K = m^{l_k}$, а $t_б$ визначається:

$$t_б = P_p \frac{m^{l_k}}{\gamma K}. \tag{28}$$

Для m=2 (двійковий алфавіт)

$$t_б = P_p \frac{2^{l_k}}{\gamma K}. \tag{29}$$

Для m=256=2^8

$$t_б = P_p \frac{256^{l_k}}{\gamma K}. \tag{30}$$

В табл. 1 наведено значення $t_б$ для безумовно стійкої криптосистеми, в якій зашифрування та розшифрування здійсняються згідно з правилами (25) та (26). Розрахунки виконано при P_p=1 та $\gamma=10^{12}$ операцій за сек.

Таблиця 1

Значення $t_б$ для безумовно стійкої системи

довжина *байт*	безпечний час системи $t_б$ *років*
8	$1,34*10^{-1}$
16	$4,17*10^{18}$
32	$2,59*10^{57}$
64	$6,35*10^{134}$
128	$1,63*10^{289}$
256	$5,74*10^{597}$
512	$3,7*10^{1214}$
1024	$7,47*10^{2447}$

Розрахуємо також відстань рівнозначності l_0 для безумовно стійкої криптосистеми.

Відомо, що [1]:

$$H(M/C)=H(K)-l_C r log_2 m, \tag{31}$$

де $H(K)$ – ентропія джерела ключів; l_C – довжина криптограми; r – збитковість мови; m – потужність алфавіту.

Враховуючи, що криптоаналіз можливий лише за умов $H(M/C) = 0$, з (31) отримаємо:

$$H(K)-l_0 r log_2 m=0, \tag{32}$$

де l_0 – відстань рівнозначності.

Звідки:

$$l_0 = \frac{H(K)}{r \log_2 m}. \quad (33)$$

Для безумовно стійкого шифру (правила (25), (26)):

$$H(K) = \log_2 N_k = \log_2 2^{l_K} = l_K. \quad (34)$$

Таким чином:

$$l_0 = \frac{l_K}{r \log_2 m}. \quad (35)$$

Оскільки, як правило, $r < 1$ (тобто будь яка мова має збитковість) то:

$$l_0 > l_K. \quad (36)$$

Із (36) випливає, що для розкриття шифрограми необхідно, щоб криптоаналітик отримав криптограму довжиною, більшою за довжину ключа.

Таким чином, для реалізації безумовно стійкої системи шифрування необхідно, щоб довжина ключа була не менш за довжину повідомлення і ключі в системі вибирались би джерелом ключів рівноймовірно, випадково та незалежно.

4. Проблемні питання реалізації та області застосування безумовно стійких криптосистем

Вище показано, що безумовна стійкість може бути досягнута при умові, що довжина ключа не менша довжини повідомлення, а ключі формуються випадково з рівномірним законом розподілу та є незалежними. Тому першою проблемою, складність якої стримує впровадження безумовно стійких криптосистем, є проблема генерування, розповсюдження, установки та використання ключів. Сутність її полягає у виконанні вимоги появлення символів "1" та "0", на різних довжинах ключів l_k ймовірності повинні бути близькими до 0,5. Навіть невеликі відхилення ймовірностей від 0,5 не дозволяють реалізувати безумовну стійкість. Далі, якщо необхідно забезпечити шифрування значних об'ємів інформації, то необхідно розповсюджувати великі об'єми ключів з великою захищеністю від можливою компрометації. При використанні ключів, з однієї сторони необхідно здійснити узгоджене їх використання, в змісті побітової синхронізації, а з другої – їх узгодженого знищення після використання.

Разом з тим, сучасні досягнення у галузі створення та використання носіїв інформації, які можуть бути використані в якості носіїв ключів, роблять можливим розповсюдження та використання ключів. Наприклад, навіть звичайна дискета може бути використана в якості носія ключів для сотень коротких повідомлень. Тому безумовно стійкі засоби криптографічного захисту інформації можуть бути реалізовані з використанням навіть звичайних персональних комп'ютерів. Це підтверджено на практиці, що буде розглянуто в подальшому.

Важливим є забезпечення також цілісності та автентичності ключів, які використовуються. Сутність цієї задачі полягає в тому, що ключі та їх носії повинні бути захищені від порушення цілісності та викривлення. Крім того, навіть в безумовно стійкій системі необхідно забезпечити потенціально досягаєму автентичність захисту інформації. На наш погляд ця проблема потребує окремого розгляду.

Щодо застосування безумовно стійких систем, то вони можуть бути використані для захисту конфіденційної інформації, різного призначення ключів, наприклад, головних ключів (ключів сертифікації та транспортних ключів). При цьому реалізація процедур зашифрування та розшифрування є простою (правила (25) та (26)) і може виконуватись з великою швидкістю.

Метою цієї статі є бажання авторів звернути увагу на можливість реалізації та застосування криптосистем, які забезпечують безумовну стійкість. Наведені в табл. 1 значення безпечного часу показують, що повідомлення з довжиною 32 байти і більше можуть бути захищені з великою стійкістю.

Список літератури: 1. *Шеннон К.* Теория связи в секретных системах. Работы по теории информации и кибернетике. М.: Изд. иностр. лит., 1963. С. 333-402. 2. *Диффи У., Хеллмэн М.Э.* Новые направления в криптографии // ТИИЭР. 1976. 22. С. 644-654. 3. *Брикелл Э.Ф., Одлижко Э.М.* Криптоанализ: обзор новейших результатов // ТИИЭР: Малый тематический выпуск. Защита информации. 1988. 76(5). С. 75-94.

Харківський державний технічний університет радіоелектроніки

Надійшла до редколегії 27.03.2001

УДК 681.3.06: 519.248.681

А. В. СВИНАРЕВ канд. тех. наук, А. Н. ЛЕПЕХА, Р. В. ОЛЕЙНИКОВ,
А. И. ШУМОВ, А. А. ГАЙОВИЧ, А. А. КАЗЬМИН

ПРИНЦИПЫ ФУНКЦИОНИРОВАНИЯ ПРОТОКОЛА IPSec

Введение

Бурное развитие сети Internet привело к проникновению новых информационных технологий практически во все сферы человеческой деятельности. Благодаря этому процессу стал возможным и обыденным ежедневный обмен сообщениями для пользователей с разных континентов, для работы и развлечения доступны гигантские массивы информации, множество организаций и ведут единый электронный документооборот в территориально разнесённых филиалах, множество фирм получили возможность вести бизнес из любой точки мира, где есть доступ к Сети. Дистанционное образование, дешевые международные телефонные переговоры, возможность найти интересующую информацию практически по любому вопросу, консультации множества специалистов были бы невозможны без происходящей информационной революции.

1. Основные угрозы, возникающие в IP-сетях

Своим рождением сеть Internet во многом обязана протоколу IP, на основании которого построены большинство национальных, корпоративных и академических сетей. Такая популярность вызвана его гибкостью, надёжностью и удобством в обслуживании. Неоспоримыми преимуществами IP является гибкость и простота маршрутизации пакетов, передаваемых по сети. Благодаря поддержке на подавляющем большинстве современных платформ, он является идеальным средствам построения негетерогенных сетей и является неотъемлемой частью широко распространённых операционных систем, таких как Unix и Windows.

Поскольку протокол IP появился более 30 лет назад, его разработчики не могли предвидеть столь широкое распространение, а также проблемы, возникающие в связи с этим. Сети, базирующиеся на протоколе IPv4 (наиболее распространённые в настоящее время), не имеют встроенных средств обеспечения безопасности взаимодействующих узлов, что связано со структурой самого протокола. Вместе со всеми преимуществами, гибкость IP протокола позволяет реализовать ряд опасных угроз:

- Spoofing – подмена IP-адресов, в результате чего невозможно гарантировать подлинность взаимодействующих сторон;
- Sniffing – приём всех IP-пакетов, пересылаемых по определённому сегменту сети, что даёт злоумышленнику возможность получить всю переданную по сети информацию;
- Highjacking – это комбинация первых двух методов, когда после установления соединения злоумышленник отключает легального абонента и вступает в информационный обмен вместо него.

Существуют различные решения проблемы обеспечения безопасности взаимодействующих сторон. Одним из вариантов является использованием современных криптографических алгоритмов на прикладном уровне (в соответствии с семиуровневой моделью OSI). Наиболее яркий пример – использование пакета PGP для обеспечения целостности, подлинности, аутентичности и конфиденциальности файлов, передаваемых по сети. Другим вариантом может стать применение семейства протоколов SSL/TLS для защиты на транспортном уровне.

Недостатком первого решения является невозможность обеспечить прозрачное для пользователя взаимодействие в реальном масштабе времени. Второе решение требует наличия специальных сетевых приложений с поддержкой SSL/TLS и кроме того, использование защиты на транспортном уровне (и выше) не позволяет скрыть топологию взаимодействия, что даже при использовании криптографических методов позволяет злоумышленнику получать информацию о взаимодействующих узлах, а также проводить атаки типа DoS (Denial of Service – отказ в обслуживании). Защита на прикладном уровне эффективна для электронной почты, на транспортном – для ведения электронной коммерции, однако для других задач, например при построении защищенных корпоративных сетей, этого явно недостаточно.

Наиболее универсальным и эффективным решением является обеспечение защиты на сетевом уровне. Поскольку на разных уровнях взаимодействуют различные протоколы, в зависимости от архитектуры сети и типа коммуникации, но, в конце концов, вся информация, подлежащая передаче по сети, поступает на сетевой уровень, где для информационного обмена используется только один протокол – IP. Таким образом, обеспечение безопасности сетевого уровня обеспечивает защиту любых соединений в сети для всех приложений. Более того, эти услуги совершенно прозрачны для пользователя и приложений.

2. Архитектура стандарта IPSec

Для защиты сетевого протокола следующей версии (IPv6) был разработан IP Security Protocol (IPSec). Его основные функции – обеспечение конфиденциальности и целостности передаваемых пакетов, сокрытие топологии взаимодействия в сети, аутентификации источника сообщений, а также защита против атак типа replay (повторного приёма пакетов) и DoS. Однако благодаря гибкости и масштабируемости IPSec совместим и с используемым IPv4.

К настоящему моменту на Украине нет литературы, которая бы описывала и проводила анализ IPSec протокола. Цель данной статьи восполнить этот пробел и описать принципы работы IPSec и способы, с помощью которых он предоставляет безопасность функционирования IP протокола. Другой задачей будет предоставить описание управления ключевыми структурами в открытых сетях между взаимодействующими сторонами.

IPSec предоставляет все необходимые средства для построения виртуальных защищённых сетей (VPN — Virtual Private Network) на основании открытых каналов связи. С помощью IPSec можно обеспечить защищенные каналы между произвольными сетями, используя для этого любые незащищенные линии коммуникации, например, Internet.

Существует два вида реализации протокола IPSec, предлагаемых в [1]:

- Программное решение, в виде драйвера для операционной системы – Bump-in-the-Stack (BITS);
- Аппаратное решение, с встроенной специализированной операционной системой – Bump-in-the-Wire (BITW).

BITS добавляет дополнительный заголовок в стек протокола TCP/IP (рис.1).

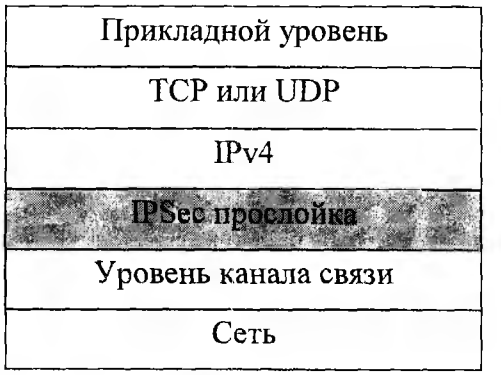


Рис. 1

Аппаратная реализация (BITW), представляет собой некоторое криптографическое устройство со специализированной операционной системой (шлюз), находящееся между защищенной локальной сетью и любой открытой сетью, через которую происходит взаимодействие с другой защищённой сетью.

IPSec функционирует в двух режимах:

- Транспортном;
- Туннельном.

В транспортном режиме IPSec обеспечивает защиту для транспортного уровня и выше (для TCP/UDP протоколов). В этом случае используется схема встраивания заголовка IPSec между заголовком IP и TCP (рис. 2). Заголовок IP не модифицируется и передается в открытом виде, защищенном только кодом аутентификации (HMAC).

Недостаток этого режима в том, что он не обеспечивает сокрытия топологии сети. К преимуществам можно отнести то, что этот режим полностью прозрачен для приложений, в отличие от TLS/SSL протоколов.

IP заголовок	IPSec заголовок	TCP/UDP заголовок	Данные

Рис. 2

Туннельный режим обеспечивает конфиденциальность всего пакета, включая и IP заголовок, как показано на рис.3.

Новый IP заголовок	IPSec заголовок	IP заголовок	TCP/UDP заголовок	Данные

Рис. 3

Этот режим также позволяет разрешить проблему использования зарезервированных адресов, которые недопустимо использовать в Internet. Как правило, при реализации этого варианта между открытой сетью и защищенной локальной сетью ставиться специальное устройство – шлюз, которое выполняет все необходимые функции.

Существует две реализации туннельного режима:

- Сеть-в-сеть (network-to-network);

- Узел-в-сеть (host-to-network).

В первом случае (сеть-в-сеть) используется связь между двумя виртуальными частными сетями, имеющих свои внутренние адреса, через специализированные устройства.

Во втором случае происходит связь между закрытой виртуальной сетью со шлюзом и некоторым узлом, с допустимым для Internet адресом, в открытой сети.

Стандарт IPSec – набор протоколов и компонентов. Базовый комплект состоит из следующих протоколов:

1. ISAKMP – Internet Security Association and Key Management Protocol – определяет правила инициализации SA (Security Association) – защищенного соединения. Задаются методы начальной аутентификации сторон, выбора метода установки защищенного канала и выработка общего секрета. Базируется на алгоритме Диффи-Хеллмана.
2. IKE – Internet Key Exchange – Задает правила установки безопасного соединения. Могут создаваться несколько защищенных соединений для разных категорий трафика. Использует общий секрет, полученный с помощью ISAKMP для генерации ключей. Кроме IP, может использоваться и с другими протоколами.
3. AH – Authentification Header – Обеспечивает целостность, аутентификацию и защиту от повторения пакетов. Это протокол, имеющий стандартный номер 51;
4. ESP – Encapsulating Security Payload – Обеспечивает конфиденциальность, целостность, аутентификацию и защиту от повторения. Протокол имеет стандартный номер 50.

Рассмотрим каждый из протоколов более подробно.

3. Управление ключами в IPSec

ISAKMP – это первая фаза работы IPSec протокола. По умолчанию протокол ISAKMP использует IKE протокол, который функционирует в 3-х режимах, используя первые два режима протокола IKE (Основной или Активный) для установления безопасного канала соединения [2].

Активный и Основной режимы практически одинаковы за исключением того, что А4ктивный режим использует меньшее число обменов, но не предоставляет аутентификационную защиту взаимодействующих узлов, так как стороны передают свои идентификаторы до того, как будет установлен безопасный канал. Оба режима еще будут рассмотрены далее в статье.

По сути, протокол ISAKMP должен установить безопасный канал для будущих обменов ключами. В результате его работы в одном из приведенных выше режимов, стороны должны выбрать поддерживаемый обеими сторонами алгоритм шифрования, хеширования, псевдослучайную функцию, метод аутентификации, выработать общий секрет и тип защиты – ESP, AH, либо оба совместно.

Основной режим выполняется в три этапа с использованием двухсторонних обменов (рис. 4). Сначала стороны обмениваются базовыми алгоритмами и хешами (пункт 1 и 2 на рис. 4). Далее (пункт 3 и 4) они обмениваются открытыми ключами по Диффи-Хеллману и обмениваются случайными числами, которые другая сторона обязана подписать и вернуть для своей идентификации. И, наконец, (пункт 5 и 6) они проверяют полученные идентификаторы.

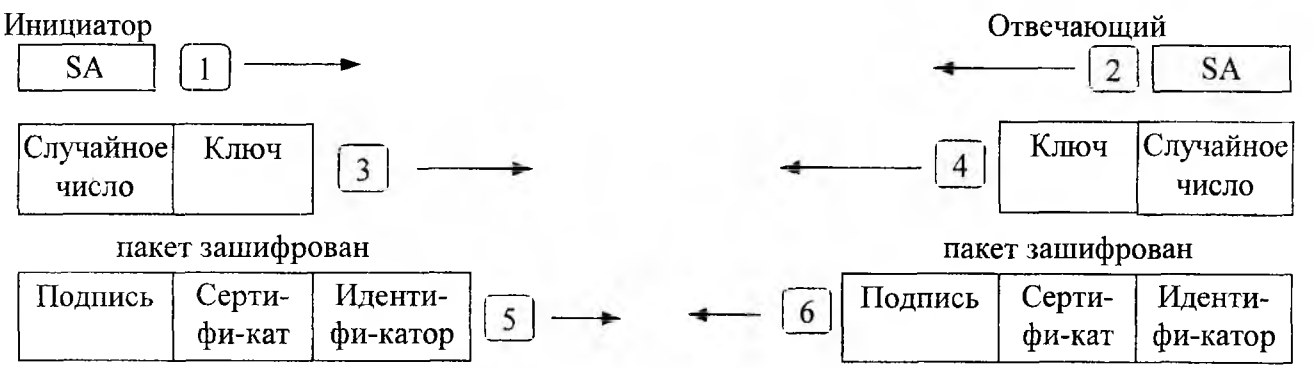


Рис. 4

Активный режим (рис. 5), являющийся упрощенной альтернативой Основному режиму, выполняется за три этапа, но в данном случае нарушается идентификационная защита, что открывает путь для атаки типа man-in-the-middle. Таким образом, эта схема будет непригодна для соединений, требующих надёжной защиты.

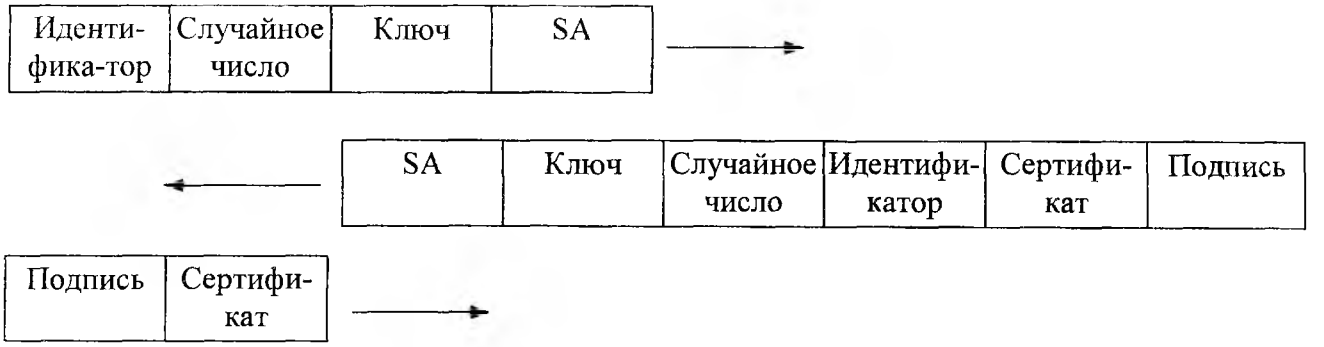


Рис. 5

При двухстороннем обмене стороны обмениваются псевдослучайными числами. Данное число можно получить как результат работы хеш-алгоритма. Для повышения безопасности лучшим вариантом будет использование криптографических генераторов случайных чисел. Существуют и другие *варианты работы ISAKMP протокола, описываемые в источнике [3].* Схема их работы похожа на приведенные выше схемы (рис. 4, 5). И более того, в них имеется много недостатков, которые будут рассмотрены ниже.

Перейдем к рассмотрению работы IKE протокола в Быстром режиме. Его задача, используя установленный предыдущими обменами защищенный канал, задать список общих IPSec сервисов и обновить или сгенерировать ключи.

Работа этого режима выглядит аналогично Активному режиму (рис. 5) при работе ISAKMP (IKE) протокола в Основном режиме. Однако при работе протокола IKE в Быстром режиме все пакеты шифруются и всегда передаются с хеш-значением, поэтому здесь гарантируется конфиденциальность и целостность передаваемых данных. Общие ключи могут получаться либо хешированием уже существующих ключей, полученных при работе ISAKMP протокола, либо стороны обмениваются случайными числами через установленный безопасный канал и используют хеш этих чисел как сеансовый ключ для работы.

При работе этого протокола вырабатывается также общий набор поддерживаемых алгоритмов шифрования, аутентификации, хеширования и генерации ключей. Весь этот набор заноситься в базу данных. Далее стороны вырабатывают так называемый SPI (Security Parameter Index) – число, с помощью которого обе стороны смогут в последствии делать предложения по использованию желательного набора правил для данного соединения. SPI вырабатывается совместно с IP адресом, протокольными данными и тем самым уникально идентифицирует этот набор IPSec SA, который должен быть идентичен для обеих сторон.

Возможно, лучшим решением было бы создание этих IPSec SA для каждого нового соединения, что будет включать и генерацию новых ключей. Это позволит производить частую смену ключей и воспрепятствует статистическому накоплению данных атакующим.

4. Криптографические преобразования в IPSec

Следующим компонентом, входящего в состав IPSec, является протокол AH, описываемый в документе [4]. Его цель состоит в предоставлении услуг аутентификации для IP пакета, но он не решает проблему конфиденциальности. Этот протокол используется либо сам по себе, либо совместно с протоколом ESP, который будет рассмотрен ниже.

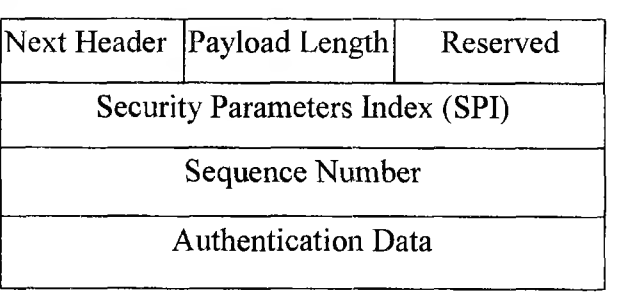


Рис. 6

На рис. 6 представлен заголовок протокола AH. Опишем его поля:

- Next Header – задает следующий (вложенный) протокол (либо TCP, либо UDP);
- Payload length – длина заголовка AH (может изменяться в зависимости от используемых алгоритмов аутентификации;
- Sequence Number – номер пакета. Служит для предотвращения повтора пакетов;

- SPI – служит для сопоставления пакета конкретному SA (защищенному каналу);
- Authentication Data – содержит ICV (Integrity Check Value), который формируется на основе кодов аутентификации заголовка IP пакета и его данных, используя 16-байтный хеш. Обязательной является поддержка алгоритмов HMAC-MD5-96 и HMAC-SHA-96.

АН протокол можно использовать как в транспортном режиме, чтобы защитить протоколы и данные верхнего уровня, так и в туннельном режиме, защищая весь пакет в целом. На рис. 7 показано встраивание заголовка АН в IP пакет.

IPv4 заголовок PROTOCOL=51	АН заголовок NEXT HEADER=6	TCP/UDP заголовок	Данные

Рис. 7

В транспортном режиме некоторые из полей IP заголовка не могут быть включены в процесс получения ICV из-за того, что они могут изменяться во время передачи. Поэтому вариантом усиления защиты может быть использование туннельного режима работы.

Оставшийся протокол, который необходимо рассмотреть – ESP, обеспечивает конфиденциальность передаваемых данных с помощью шифрования [5]. Подобно АН протоколу он также предоставляет услуги целостности, аутентификации и обнаруживает повторно принятые пакеты. Однако, в отличие от протокола АН, аутентифицируется не весь пакет, а только инкапсулированные в него данные, как показано на рис. 8.

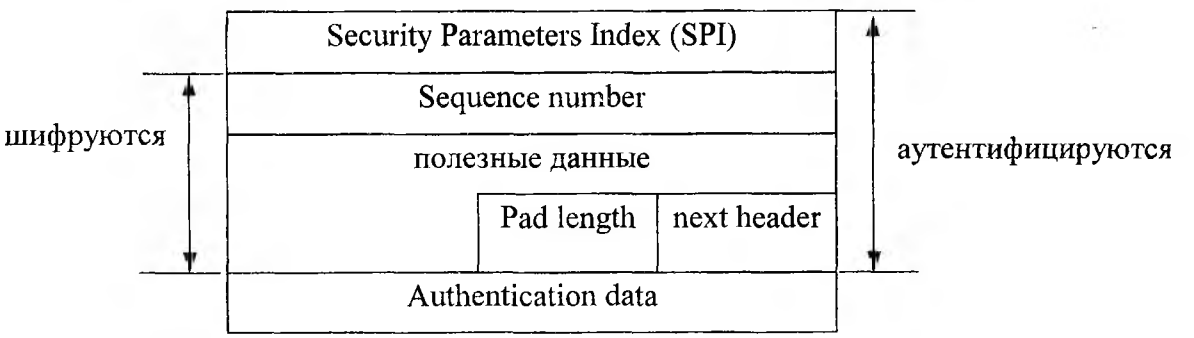


Рис. 8

Опишем представленные в ESP протоколе поля:

- SPI – служит для сопоставления пакета конкретному SA (защищенному каналу);
- Sequence Number – номер пакета, служит для предотвращения повторного приёма пакетов;
- Initialization Vector – 64-битовый случайный вектор, используемый для шифрования (необходим для режима CBC алгоритма DES, используемого для шифрования данных);
- Protected Data – зашифрованное содержимое IP пакета (в транспортном режиме), либо весь пакет целиком (в туннельном режиме). Для шифрования используется алгоритм шифрования DES (длина ключа 56 бит), а также возможно использование TDES (длина ключа 168 бит) и IDEA (128 бит);
- Pad Length – длина дополнения данных для выравнивания по 8-байтовой границе (требование режима CBC алгоритма шифрования DES);
- Authentication Data – данные аутентификации.

Протокол ESP встраивается в IP пакет сразу после заголовка, как показано на рис. 9. Он может использоваться как сам по себе, так и совместно с протоколом АН.

IPv4 заголовок PROTOCOL=50	ESP заголовок NEXT HEADER=6	АН заголовок NEXT HEADER=6	TCP/UDP заголовок	данные

Рис. 9

5. Недостатки IPSec

Анализируя набор IPSec протоколов, можно сказать, что в целом выбор IPSec будет одним из лучших вариантов решения при создании VPN, как с точки зрения безопасности, так и с точки зрения удобства для пользователей и экономической выгоды. Это позволяет использовать открытые сети без привлечения специализированных линии связи.

Однако IPSec имеет некоторые недостатки, которые должны учитываться при разработке протокола.

Как отмечается в [6], стандарт IPSec является слишком сложным, чтобы провести полный анализ его безопасности, что делает весьма вероятным выявление в будущем каналов уязвимости. Другими следствиями высокой сложности стандарта являются:

1. В документации по IPSec не определяются задачи защиты, функциональность не продумана: некоторые протоколы являются функционально избыточными (например, AH), и вместе с тем некоторые протоколы неполно обеспечивают некоторые услуги, (обмен сертификатами).
2. Администраторы безопасности могут допускать ошибки при конфигурировании системы защиты. Например, ассоциации безопасности в IPSec являются однонаправленными, следовательно, систему можно сконфигурировать так, чтобы защита выполнялась для исходящего трафика и не выполнялась для входящего.
3. Документация трудна для чтения и содержит ошибки.

Кроме этого, в [6] приведены и другие замечания по стандарту.

Во-первых, AH протокол защищает не все поля IP заголовка, поэтому при использовании в критичных, с точки зрения безопасности, системах, особенно в банковских, следует использовать туннельный режим работы. Лучше всего для этого использовать специализированное аппаратное криптографическое устройство. При этом можно исключить протокол AH из работы, а протокол ESP должен всегда включать аутентификацию.

Другим недостатком можно считать предлагаемые к использованию криптографические алгоритмы. В частности, в протоколе ESP базовым является DES (ключ длиной 56 бит). Сегодня он уже не обеспечивает требуемый уровень безопасности. В будущих вариантах протокола IPSec предлагается его надо исключить даже для применения в не требующих особой защиты системах. Альтернативные алгоритмы TDES и IDEA также со временем потребуют замены. К настоящему времени как стандарт XXI века принят криптографический симметричный алгоритм RIJNDAEL, главным достоинством которого является отсутствие каких-либо эффективных криптоаналитических атак, кроме прямого перебора. Для аутентификации можно применять современные алгоритмы на эллиптических кривых, обладающих высокой стойкостью. Следует заметить и о возможной замене обычного алгоритма Диффи-Хеллмана стандартом X.9-63 – Диффи-Хеллман на эллиптических кривых для использования в протоколах ISAKMP и IKE. Несомненным достоинством IPSec протокола является возможность подключения новых криптографических алгоритмов. Таким образом, устранить перечисленные недостатки не составляет особого труда.

Стоит отметить ошибку, которая скрыта в функционировании протокола ESP. Как уже отмечалось ранее, данный протокол сначала осуществляет зашифрование и лишь потом аутентификацию. Такой режим работы открывает дорогу ряду атак. Решением здесь может стать либо изменение порядка операций – сначала аутентификация, потом зашифрование – либо подписывать ключ шифрования вместе с зашифрованным текстом.

В документации по протоколу IPSec, как уже отмечалось ранее, рекомендуется использовать криптографические алгоритмы шифрования в режиме CBC. Однако в данном режиме высока вероятность коллизий, а также усложняется аппаратная реализация протокола. Этот факт следует учитывать в будущих разработках.

Помимо перечисленных выше недостатков спецификация протокола допускает использование переменного числа циклов в шифре, что ослабляет криптографическую стойкость. Следует убрать эту возможность.

Как уже упоминалось, протокол ISAKMP имеет различные схемы реализации и является в действительности набором протоколов, которые обязаны установить безопасный канал. В статье было рассмотрено использование протокола ISAKMP как работа первых двух режимов протокола IKE. Предоставленные в источнике [1] схемы обменов имеют существенные недостатки. Так в некоторых типах обменов, предоставляемых этим протоколом, имеется вероятность осуществления известных атак. Базовый режим в [1] аналогичен Активному режиму в Основном режиме работы протокола IKE. Недостатком здесь является отсутствие идентификационной защиты, что открывает путь для

атаки man-in-the-middle. Аналогичная атака возможна при использовании режима Аутентификация (Authentication Only Exchange), где вообще не осуществляется шифрование. Анализ показал, что предлагаемые схемы не предоставляют полной защиты как от replay (повтор пакета), так и DoS (отказ в обслуживании) атак. Это вызвано плохой продуманностью предлагаемой архитектуры протокола ISAKMP, о чем говорят ошибки, встречающиеся в документации и слишком общее описание его работы.

В рассматриваемом варианте, с использованием протокола IKE, имеется возможность осуществления атаки типа proposal, которая подразумевает под собой, что злоумышленник может навязать использование сторонами ослабленных вариантов криптографических алгоритмов (с малой длиной ключа, либо использование устаревших алгоритмов и т.п.). Это можно увидеть, анализируя Основной режим, представленный на рис. 4.

Еще одно замечание относительно протокола IKE, которое рассматривалось на конференции посвященной протоколу IPSec и проходившей во Франции в октябре 2000 года, заключалось в том, что протокол IKE при работе идентифицирует только удаленный компьютер, но не самого пользователя. Таким образом, при его работе используются сертификаты компьютера, но не самого пользователя. Как показывает практика, наибольшее количество нарушений безопасности вызывается самими служащими, а не внешними злоумышленниками. Поэтому данный момент требует обратить на себя внимание. На данный момент некоторые зарубежные корпорации предлагают использование дополнительных протоколов, которые позволяют осуществлять аутентификацию пользователя. Так, например, фирма Microsoft в собственных разработках протокола IPSec применяет протокол L2TP, который функционирует совместно с IPSec, работающим в туннельном режиме.

Тем не менее, несмотря на приведенные замечания, IPSec является на сегодняшний день лучшим стандартом защиты на сетевом уровне.

Заключение

На основании вышеизложенного материала можно говорить о том, что в целом, учитывая замечания, указанные в статье, применение протокола IPSec будет наиболее универсальным средством для решения задач обеспечения безопасности при взаимодействии в открытых сетях. Данная тема является одной из самых популярных за рубежом. Об свидетельствуют многочисленные публикации в журналах и Интернете, ежегодные конференции посвященные протоколу IPSec, на которых обсуждаются пути улучшения его функциональности. Многие компании предлагают свои услуги по тестированию продуктов реализующих протокол IPSec, предъявляя конкретные требования к конкурсантам. Многие из этих требований уже фактически стали стандартами для этого протокола. Предпосылками для внедрения IPSec протокола на отечественном рынке может служить то, что сегодня на рынке господствуют операционные системы зарубежных стран, которые не отвечают требованиям необходимой степени защиты и не дают гарантии, что в них не имеется закладок, оставленных разработчиками. Использование таких операционных систем в требующих повышенной безопасности организациях слишком большой риск. Не меньший риск и покупка средств защиты, предлагаемых на иностранных рынках. Поэтому внедрение IPSec, разработанного отечественными специалистами, позволит решить эти проблемы наиболее универсальным способом.

Список литературы: 1. *Ken Masica.* Understanding the IP Security Protocol // Internet Security. Oct. 2000. Vol.3. No. 5. pp. 38-42. 2. *RFC 2409.* The Internet Key Exchange (IKE), Request for Comments: 2409 / Network Working Group, 1998. 3. *RFC 2408.* Internet Security Association and Key Management Protocol (ISAKMP), Request for Comment 2408. / Network Working Group, 1998. 4. *RFC 2402.* IP Authentication Header (AH), Request for Comments: 2402 / Network Working Group, 1998. 5. *RFC 2406.* IP Encapsulating Security Payload (ESP), Request for Comments: 2406 / Network Working Group, 1998. 6. *Niels Ferguson, Bruce Schneier.* A Cryptographic Evaluation of IPsec. /Counterpane Internet Security, Inc., 1999.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 29.03.2001

МЕТОДЫ И СРЕДСТВА ФОРМИРОВАНИЯ И ИССЛЕДОВАНИЯ СЛУЧАЙНЫХ И ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

УДК 681.324.067

А. А. ТОРБА, канд. техн. наук, С. Г. ЕЛАКОВ, А. З. СТЕПЧЕНКО

ГЕНЕРАЦИЯ РАВНОВЕРОЯТНЫХ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НА ОСНОВЕ ФИЗИЧЕСКИХ ДАТЧИКОВ

Применение физических датчиков позволяет генерировать случайные последовательности, которые не будут коррелированны на сколь угодно длинном расстоянии. Такие последовательности действительно являются случайными, т.к. они не могут быть воспроизведены в заданном порядке, не могут быть повторены в следующем опыте и являются полностью непредсказуемыми.

Для генерации случайных последовательностей достаточно внесение одного какого-нибудь случайного (непредсказуемого) параметра в детерминированный процесс. Простейшим примером является считывание состояний детерминированного счетчика в случайные моменты времени. Обязательным условием для генерации случайных последовательностей является многократное переполнение счетчика между считываниями. Известен пример генерации случайных чисел при считывании состояний счетчика в таймере IBM PC в моменты нажатий произвольных клавиш. Для набора случайного числа длиной 512 бит необходимо нажать 32 клавиши. Это занимает много времени (до одной минуты), однако такой метод может быть реализован на программном уровне и не требует дополнительного оборудования.

Физические датчики шума (резисторы, полупроводниковые и вакуумные электронные приборы) генерируют случайные последовательности импульсов различной амплитуды и с широким частотным спектром. Наиболее удобно применять в вычислительных устройствах физические датчики шума на основе полупроводниковых приборов с Зенеровским пробоем (стабилитронов).

Генераторы шума КГ401А при токе 50...100 мкА формируют случайные импульсы амплитудой 0,1...1 В с максимальной частотой до 3-х МГц (график Uшд на рис. 1, а). Преобразование этих импульсов в логические уровни цифровых микросхем (график Uтш на рис. 1, б) реализуется усилителями-ограничителями (компараторами) с небольшим гистерезисом на входе – триггерами Шмитта (TS).

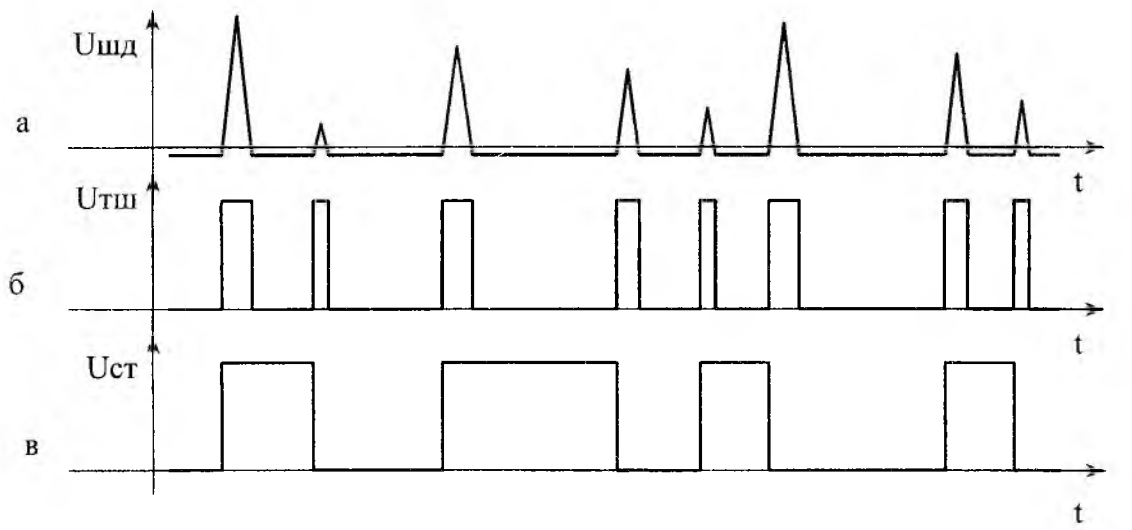


Рис. 1

В выходном сигнале триггера Шмитта (график Uтш на рис. 1, б) преобладает уровень логического нуля. При считывании этого сигнала в произвольные моменты времени формируемая случайная последовательность будет содержать значительно больше нулевых битов, чем единичных.

Для выравнивания вероятностей "0" и "1" выходной сигнал триггера Шмитта подается на счетный триггер (см. рис. 2). Выходной сигнал счетного триггера (график Uст на рис. 1, в) с равной веро-

ятностью принимает значения «логического нуля» и «логической единицы» в произвольные моменты времени. Считывание случайных битов можно производить в детерминированные моменты времени.

Достоинством данного метода формирования случайных битов является малая зависимость параметров формируемых последовательностей от режимов первичного генератора шумовых импульсов и закона распределения во времени случайных импульсов. Обязательным условием независимости элементов генерируемых случайных последовательностей является многократное срабатывание счетного триггера в течение интервала времени между считываниями.

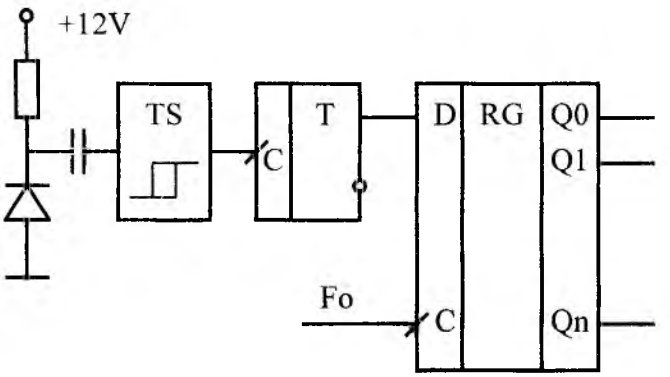


Рис. 2

Состояния счетного триггера считываются с частотой Fo сдвигающим регистром RG (рис. 2). Частота Fo выбирается в 5...10 раз меньше, чем средняя частота шумовых импульсов на выходе триггера Шмитта. Это необходимо для многократного срабатывания счетного триггера между соседними считываниями случайных битов. При этом исключается взаимное влияние вероятности появления очередного бита от состояния предыдущего бита.

Реальная средняя частота шумовых случайных импульсов диода КГ401А составляет 2...3 МГц, поэтому детерминированная частота считывания Fo выбирается около 200 кГц (период формирования случайных битов – 5 мкс). При этом время формирования случайного слова (длиной 16 бит) равно 80 мкс. Время формирования ключа длиной 512 бит составит – 2,56 мс.

Одноканальная схема формирования случайных битов, включающая шумовой диод, усилитель-ограничитель (триггер Шмитта) и счетный триггер (см. рис. 2), не обеспечивает необходимую надежность генерации равновероятных битов в случае изменения параметров источника шума или усилителя-ограничителя на основе триггера Шмитта. Повышение надежности канала формирования случайных битов достигается горячим резервированием, то есть параллельной работой нескольких каналов [1]. На рис. 3 приведена схема генератора случайных последовательностей с двумя каналами формирования случайных битов (первый канал выделен на рис. 3 пунктиром). Возможно применение трех и более аналогичных каналов формирования случайных битов.

Выходные равновероятные случайные логические сигналы всех каналов объединяются схемой ИСКЛЮЧАЮЩЕЕ ИЛИ (схемой суммирования по модулю 2) и считываются в сдвигающий регистр с частотой Fo (см. рис. 3).

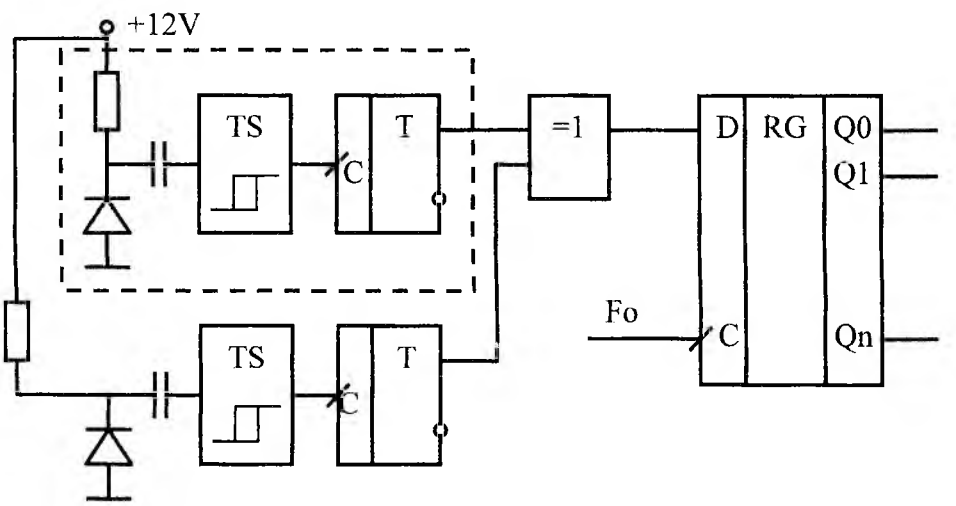


Рис. 3

Функционирование генератора случайных последовательностей (см. рис. 2 или 3) возможно только в составе программно-аппаратного комплекса, включающего в себя:

- собственно генератор равномерно распределенных случайных чисел,
- программный драйвер считывания случайных чисел,
- программы тестирования случайных последовательностей.

При экспериментальных исследованиях генераторов случайных последовательностей (см. рис. 2 или 3), реализованных на микросхемах ТТЛШ, было обнаружено, что при частоте входных импульсов шумового генератора около 2 МГц вероятность «нулевых» битов превышает вероятность «единичных» битов на величину примерно:

$$\Delta P = P(0) - P(1) = 0{,}001.$$

При увеличении входной частоты шумовых импульсов – разность вероятностей ΔР также увеличивается.

Это объясняется особенностями схемотехники выходного каскада микросхем ТТЛШ. Выходное сопротивление каскада в состоянии «логическая единица» значительно больше выходного сопротивления каскада в режиме «логический нуль». Поэтому время перезарядки «паразитных емкостей» нагрузки элемента ТТЛШ через выходное сопротивление каскада будет различным. В результате: счетный триггер дольше переходит из состояния "0□1", чем "1□0". Поэтому в выходной последовательности генераторов случайных чисел в среднем на 1000 «нулей» формируется примерно 999 «единиц».

Существует несколько алгоритмов выравнивания вероятностей случайных битовых последовательностей.

Первый алгоритм позволяет значительно уменьшить разность вероятностей генерируемых случайных битов. Для этого из двух последовательных случайных битов формируется их логическая функция ИСКЛЮЧАЮЩЕЕ ИЛИ. Промежуточный регистр RG1 запоминает два последних генерируемых случайных бита (рис. 4). В выходной регистр RG2 записывается логическая функция ИСКЛЮЧАЮЩЕЕ ИЛИ этих битов, но с частотой в два раза меньше, чем Fo (счетный триггер Т2 делит частоту считывания Fo на два). Вероятность единичного формируемого бита на *входе регистра RG1* обозначим P(1), а вероятность нулевого – P(0) = P(1) + Δ.

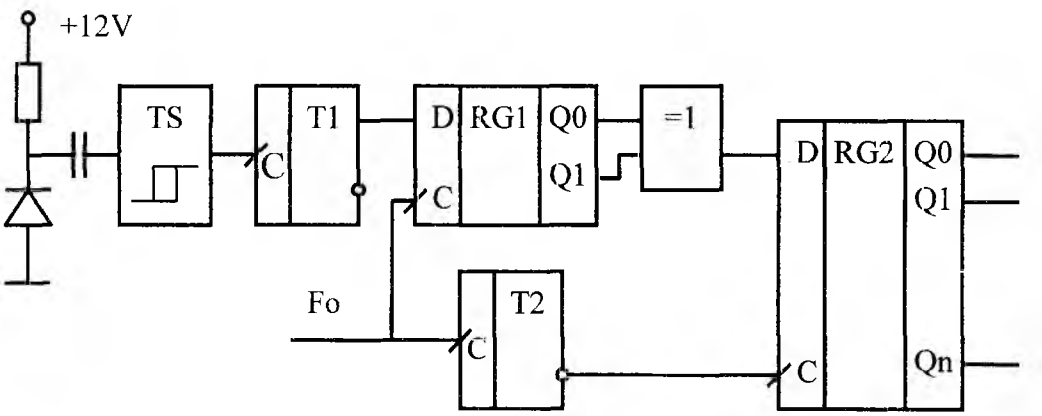


Рис. 4

Сумма вероятностей P(0) + P(1) всегда равна единице.

Запишем все комбинации битов на *выходе промежуточного регистра RG1* и вероятности этих комбинаций (с учетом полной статистической независимости генерируемых соседних случайных битов) (см. табл. 1).

Таблица 1

Q1	Q2	Вероятности
0	0	[P(1) + Δ] * [P(1) + Δ]
0	1	[P(1) + Δ] * P(1)
1	0	P(1) * [P(1) + Δ]
1	1	P(1) * P(1)

На выходе схемы ИСКЛЮЧАЮЩЕЕ ИЛИ (см. рис. 4) формируется логический нуль при комбинациях, соответствующих первой и последней строкам табл. 1. Поэтому вероятность "нулей" P(0)' на выходе схемы ИСКЛЮЧАЮЩЕЕ ИЛИ равна:

$$P(0)' = [\,P(1) + \Delta\,] * [\,P(1) + \Delta\,] + P(1) * P(1).$$

Логической единице на выходе схемы ИСКЛЮЧАЮЩЕЕ ИЛИ будут соответствовать две средние строки в таблице 1, поэтому вероятность "единиц" P(1)' будет равна:

$$P(1)' = [\,P(1) + \Delta\,] * P(1) + P(1) * [\,P(1) + \Delta\,].$$

Разность вероятностей на выходе схемы ИСКЛЮЧАЮЩЕЕ ИЛИ равна:

$$\Delta' = P(0)' - P(1)' = \Delta^2.$$

Учитывая малую величину разности вероятностей Δ (примерно 0,001), можно утверждать, что ее квадрат будет значительно меньше.

К недостаткам этого метода (метода «Дельта-квадрат») можно отнести в два раза меньшую скорость формирования случайных битов и, хотя и маленькую, но не нулевую, разность вероятностей "0" и "1".

Второй метод еще в два раза уменьшает скорость формирования случайных последовательностей, но позволяет выровнять вероятности "0" и "1". Идея этого метода понятна из анализа табл. 1. Вероятности второй и третьей строк равны. Поэтому при комбинации сигналов на выходах промежуточного регистра RG1, соответствующей второй строке, в выходной регистр RG2 записывается нулевой бит, а при комбинации, соответствующей третьей строке, – единичный бит. Комбинации сигналов, соответствующие первой и последней строкам, не используются.

Для этого нулевой логический сигнал с выхода логической схемы ИСКЛЮЧАЮЩЕЕ ИЛИ (контролирующей выходы промежуточного регистра RG1) запрещает запись в выходной регистр RG2 (см. рис. 5) при комбинациях, соответствующих первой и последней строкам табл. 1.

Описанные схемы формирователей случайных последовательностей (рис. 4 и 5) обладают относительно небольшой скоростью генерации случайных бит, которая при частоте шумовых импульсов кремниевого диода КГ401А около 2МГц – не превышает десятков килогерц.

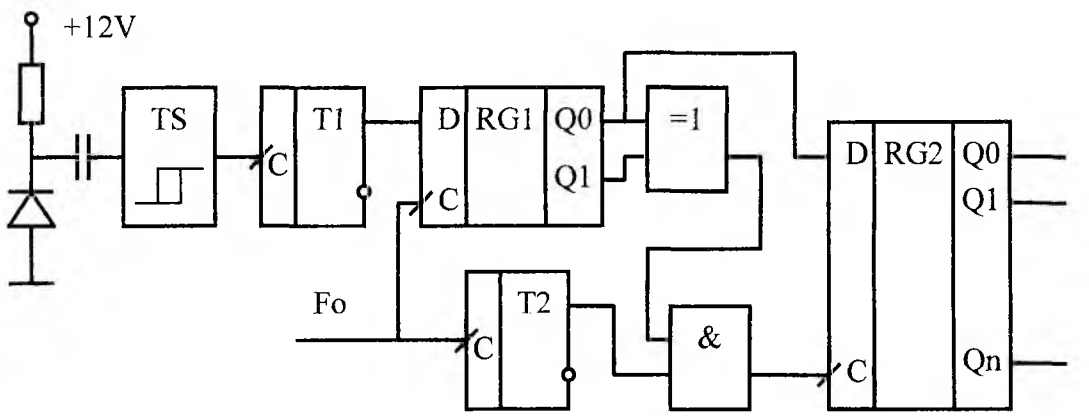


Рис. 5

Значительно повысить скорость формирования случайных битов позволяет схема генератора псевдослучайных последовательностей на основе сдвигающего регистра с обратными связями и случайным инвертированием входного сигнала регистра при помощи элемента ИСКЛЮЧАЮЩЕЕ ИЛИ (рис. 6) [2].

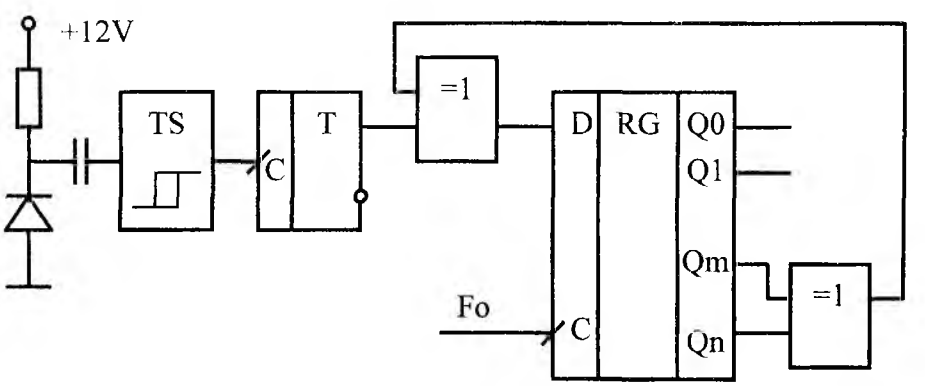


Рис. 6

Частота генерации случайных бит Fo может многократно превышать среднюю частоту датчика шума. Реально максимальное значение частоты Fo ограничено быстродействием элементов сдвигающего регистра и может превышать 100 МГц (для современных КМОП микросхем).

Для увеличения надежности генераторов случайных последовательностей методом "горячего резервирования" в схемы на рис. 4, 5 и 6 необходимо ввести многоканальные формирователи случайных битов [1].

Применение современных интегральных микросхем в значительной степени определяет массогабаритные параметры устройства. Генераторы случайных последовательностей реализованы на логических микросхемах средней степени интеграции микромощных серий ТТЛШ и сопряжением с каналом персонального компьютера через ISA-слот, а также на программируемых логических интегральных схемах (ПЛИС) и сопряжением через слот PCI.

Разработан программный драйвер для сопряжения формирователей с вычислительной системой и программы тестирования случайных последовательностей.

При тестировании генератора проверялось соответствие генерируемой случайной последовательности условиям:

- равновероятности с использованием критерия Пирсона;
- случайности по критерию серий;
- некоррелированности по коэффициентам корреляции разрядов байтов случайной последовательности;
- независимости по методу сопряженности признаков;
- однородности по методу проверки гипотезы о совпадении распределений.

Результаты экспериментальных исследований подтвердили равновероятный закон распределения генерируемых случайных последовательностей.

Экспериментально проверена эффективность горячего резервирования для двух каналов генерации случайных битов. На вход элемента ИСКЛЮЧАЮЩЕЕ ИЛИ (см. рис. 3) вместо одного из каналов генерации случайных битов подавались:

- постоянные логические уровни "0" или "1";
- прямоугольный сигнал детерминированного генератора.

Тестирование выходных случайных последовательностей не выявили отклонений от равновероятного закона распределения во всех экспериментах.

Недостатком схем генераторов случайных последовательностей с сопряжением через ISA-слот или PCI-слот является необходимость разборки корпуса компьютера при инсталляции генератора (разборка компьютера в некоторых условиях эксплуатации не допускается).

Этот недостаток был устранен в генераторе случайных последовательностей с сопряжением через внешний COM-порт компьютера.

Значительно уменьшить габариты схемы удалось реализацией логических схем обработки случайных сигналов на однокристальном микроконтроллере.

Разработан программно-аппаратный комплекс генерации случайных последовательностей, который включает (рис. 7):

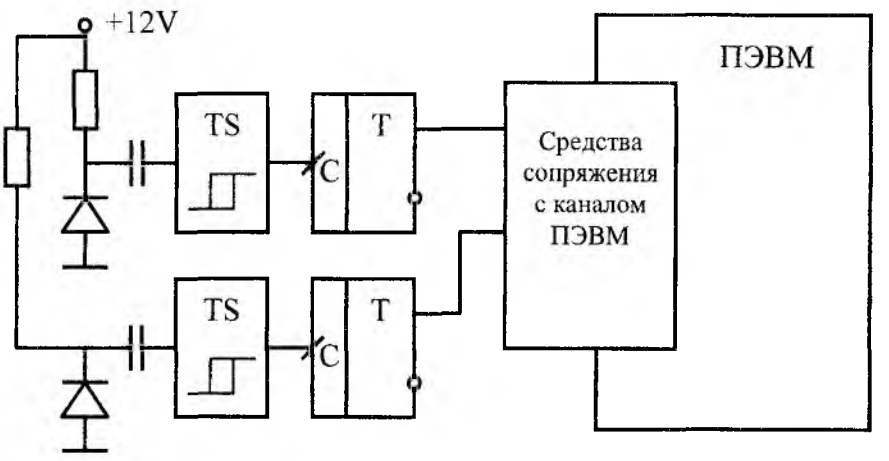


Рис.7

- два физических датчика случайного сигнала на основе шумовых диодов, работающих в режиме Зенеровского пробоя;
- два компаратора напряжения для преобразования аналоговых сигналов в логические уровни цифровых микросхем;
- два счетных триггера, преобразующих импульсы с выходов компараторов в сигналы, с равной вероятностью принимающие нулевые и единичные логические уровни;
- программно-аппаратные средства на основе однокристального микроконтроллера для сопряжения каналов формирования равновероятных битовых последовательностей с COM-портом ПЭВМ;
- программные средства для управления датчиком случайных последовательностей, его диагностики и тестирования.

Алгоритм функционирования такого генератора случайных последовательностей реализован на программном уровне в однокристальном микроконтроллере и полностью совпадает с логикой работы описанных генераторов. В этой схеме (см. рис. 7) также применяется "горячее резервирование" каналов формирования случайных битов.

Скорость генерации случайных последовательностей ограничена скоростью передачи COM-порта и составляет 9,6 Кбит/сек.

Тестирование случайных последовательностей на соответствие указанным выше условиям подтвердили равновероятный закон распределения. Экспериментально подтверждена эффективность "горячего резервирования" на случай отказа одного из генераторов шума.

Применение в формирователе случайных последовательностей микроконтроллеров позволяет реализовать на программно-аппаратном уровне средства защиты от несанкционированного доступа к аппаратным и программным ресурсам ПЭВМ.

Список литературы: **1**. *Патент Украины № 33361.* МКИ^{6} G 06F7/58, G 07C15/00. Генератор равномерно распределенных случайных чисел / *И.Д. Горбенко, А.А. Торба, С.Г. Елаков и др.* Опубл. Бюл. №1 от 15.02.2001. **2**. Заявка № 99116006 от 02.11.1999 (Решение на выдачу патента Украины от 26.05.2000). Способ генерации случайных чисел и устройство для его осуществления / *А.А. Торба.*

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 27.03.2001

УДК 681.3.06

М. А. КРИВОШЛЫК, Ю. И. ГОРБЕНКО, В. Н. ВЕРВЕЙКО

АЛГОРИТМЫ И СРЕДСТВА ТЕСТИРОВАНИЯ СЛУЧАЙНЫХ И ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Введение

Безопасность большинства криптографических систем зависит от способов формирования и использования ключей и параметров. Предельные характеристики стойкости достигаются в случае, если ключи и параметры выбираются случайно, равновероятно и независимо из полного пространства. Для выполнения этих условий используются генераторы случайных и псевдослучайных последовательностей. Случайный генератор – это устройство или алгоритм, который выдает последовательность статистически независимых символов с основанием алфавита m. Псевдослучайным генератором (ПСГ) называется детерминированный алгоритм (совокупность алгоритмов, устройство, совокупность алгоритмов и устройств), который, используя действительно случайную последовательность (СП) длиной k, формирует на ее основе псевдослучайную последовательность длины $l>>l_0$, где l_0 – допустимая длина. Вход в ПСГ называется начальным состоянием или зерном *(seed)*, на выходе ПСГ формируется псевдослучайная последовательность. Основным применением генераторов ПСП и СП в криптографии является формирование ключей, параметров и синхромаркеров.

Чтобы убедиться, что генератор безопасен, он должны быть подвергнут ряду статистических испытаний с целью подтверждения у сформированной им последовательности таких характеристик, которые ожидаются у случайных последовательностей. Псевдослучайные последовательности оцениваются с использованием ряда количественных показателей. Основными показателями являются[1,2]:

- случайность, равновероятность, независимость, однородность;
- период построения l_n ПСП;
- основание алфавита m;
- вероятность перекрытия в пространстве или во времени двух сегментов Y_r и Y_μ;
- структурная скрытность (эквивалентная сложность) S_l последовательности Y;
- энтропия источника начальных значений (seed);
- расстояние равнозначности l_0 конкретной последовательности Y_v;
- безопасное время генератора ПСЧ $t_Б$
- сложность I_Y формирования последовательности Y;
- длина параметров обратной связи (выход-вход) генератора;

Тесты, которые мы рассмотрим, помогают обнаружить уязвимые места, которые может иметь генератор. Анализ свойств ПСП и СП может быть выполнен путем исследования выходной последовательности генератора с использованием статистических тестов. Каждый статистический тест позволяет определить, обладает ли последовательность такими свойствами, какими обладает истинно случайная последовательность. Если последовательность не проходит хотя бы один из тестов, генератор либо может быть отвергнут как не обеспечивающий свойств случайности, либо может быть подвергнут дальнейшему тестированию. Если тестируемая последовательность проходит все статистические тесты, с определенной вероятностью генератор признается случайным. Более точно термин “признается” следует понимать как “не отвергается”.

При построении ключей одной из основных задач является получение случайных или псевдослучайных последовательностей, которые неотличимы от случайных, обладают большим периодом и другими свойствами, перечисленными выше. Перед использованием сгенерированной последовательности чисел $X = \{x_1, x_2, \ldots, x_n\}$ необходимо убедиться, что случайная величина X обладает равномерным законом распределения, её реализации случайны и независимы.

Статистическая гипотеза H представляет собой предположение относительно распределения случайной величины. Проверка статистической гипотезы представляет собой процедуру, позволяющую на основании значений случайной переменной сделать вывод о справедливости или ошибочности с определенной вероятностью выдвинутой гипотезы. Важным при гипотетическом тестировании является понятие уровня значимости α.

Уровень значимости α проверки статистической гипотезы H представляет собой вероятность того, что мы отвергнем гипотезу H, являющуюся на самом деле истинной.

Правильный выбор уровня значимости α для проверки является очень важной задачей. Если мы возьмем α слишком большим, то велика вероятность того, что мы отвергнем гипотезу, являющуюся на самом деле истинной. С другой стороны, если мы возьмем α слишком маленьким, то велика вероятность того, что мы примем гипотезу, являющуюся на самом деле ложной. На практике обычно используются значения уровня значимости от 0,001 до 0,05.

Математическая статистика дает нам возможность построить статистические тесты для проверки гипотез о равномерности, случайности и независимости случайных величин. Для этих целей можно использовать критерий χ^2 Пирсона. В США принята методика тестирования ПСП и СП, базирующаяся на критерии χ^2 Пирсона. Она зарегистрирована как FIPS 140-1. Недостатком этой методики является то, что с ее использованием можно протестировать ПСП или СП строго определенной длины – 20000 битов. Целью настоящей статьи является разработка методики проверки на случайность ПСП и СП произвольной длины (ограниченной только снизу), а также рассмотрение как частного случая методики, определенной FIPS 140-1.

Базовые статистические вероятностные тесты

В качестве базовых рекомендуется использовать следующие тесты[1,2]:

- частотный (монобитный) тест;
- тест двухбитных серий;
- тест Поккера;
- общий тест серий;
- автокорреляционный тест.

1.1 Монобитный тест

Цель этого теста состоит в том, чтобы определить, является ли количество «0» и «1» в последовательности s приблизительно таким, каким оно ожидается для случайной последовательности. Пусть n_0 и n_1 обозначают количество нулей и единиц в s, соответственно. Тогда параметр ПСП

$$X_1 = \frac{(n_0 - n_1)^2}{n}$$

подчиняется χ^2 - распределению с одной степенью свободы (если $n \geq 10$).

1.2 Тест двухбитных серий

Цель этого теста состоит в том, чтобы определить, является ли число появлений 00, 01, 10 и 11 как последовательностей s приблизительно таким же, как ожидается для случайной последовательности. Пусть n_0 и n_1 обозначают количество нулей и единиц в s, соответственно, и пусть n_{00}, n_{01}, n_{10}, n_{11} обозначают число появлений 00, 01, 10, 11 в s, соответственно. Заметим, что $n_{00} + n_{01} + n_{10} + n_{11} = (n-1)$, так как подпоследовательности могут перекрываться. Используемый статистический параметр равен

$$X_2 = \frac{4}{n-1}\left(n_{00}^2 + n_{01}^2 + n_{10}^2 + n_{11}^2\right) - \frac{2}{n}\left(n_0^2 + n_1^2\right) + 1$$

и подчиняется χ^2 - распределению с двумя степенями свободы (если n≥21).

1.3 Тест Покера

Пусть m будет положительным целым числом, таким, что $\left\lfloor \frac{n}{m} \right\rfloor \geq 5 \cdot 2^m$, и пусть $k = \left\lfloor \frac{n}{m} \right\rfloor$. Разделим последовательность s на k неперекрывающихся частей, каждая длиной m, и пусть n_i будет числом появлений i-го типа последовательности длины m, $1 \leq i \leq 2^m$. Тест Покера определяет, действительно ли каждая последовательность длиной m появляется приблизительно столько же раз в s, сколько ожидается для случайной последовательности. Используемый статистический параметр равен

$$X_3 = \frac{2^m}{k}\left(\sum_{i=1}^{2^m} n_i^2\right) - k,$$

что приблизительно соответствует χ^2 - распределению с 2^m -1 степенями свободы. Заметим, что тест Покера является обобщением частотного теста: установка $m = 1$ в тесте Покера дает частотный тест.

1.4 Тест серий

Тест серий позволяет определить, действительно ли число серий либо нулей, либо единиц различных длин в последовательности s такое же, как ожидается для случайной последовательности. Ожидаемое число интервалов (или блоков) длиной i в случайной последовательности длиной n равно $e_i = (n - i + 3)/2^{i+2}$. Пусть k будет равным наибольшему целому числу i, для которого $e_i \geq 5$. Пусть Bi, Gi будут числом блоков нулей и единиц, соответственно, длиной i в s для каждого i, $1 \leq i \leq k$.

Используемый статистический параметр равен

$$X_4 = \sum_{i=1}^{k} \frac{(B_i - e_i)^2}{e_i} + \sum_{i=1}^{k} \frac{(G_i - e_i)^2}{e_i}$$

и приблизительно соответствует χ^2 - распределению с $2k$ -2 степенями свободы.

1.4 Автокорреляционный тест

Цель автокорреляционного теста состоит в том, чтобы проверить степень связи между последовательностью s и апериодическим ее сдвигом. Пусть d будет фиксированным целым числом, $1 \leq d \leq \lfloor n/2 \rfloor$. Число битов в последовательности s не совпадает с их числом в d-сдвигах, и равно $A(d) = \sum_{i=0}^{n-d-1} s_i \oplus s_{i+d}$. Используемый статистический параметр равен

$$X_5 = 2\left(A(d) - \frac{n-d}{2}\right)/\sqrt{n-d}$$

и приблизительно соответствует $N(0,1)$ нормальному распределению, если n-$d \geq 10$. Так как малые значения $A(d)$ являются столь же маловероятными, как и большие, то должен использоваться двусторонний тест.

2. Пример решения задачи тестирования последовательности

Рассмотрим последовательность s длины $n = 128$, полученную путем четырехкратного копирования следующей последовательности:

0110 0100 0111 1010 1100 1000 1111 0101.

Проверим эту последовательность по статистическим критериям:

2.1 Частотный (монобитный) тест

$n = 128$, $n_0 = 60$, $n_1 = 68$, тогда $X_1 = \frac{(60-68)^2}{128} = 0{,}5$. Поскольку $X_1 < X_{1пр} = \chi^2(1;0.05) = 3{,}84$ [2] и $X_1 = 0{,}5 < 3{,}84$, то тест пройден, и гипотеза не отвергается.

2.2 Двухбитный тест серий

$n_{00} = 24$, $n_{01} = 36$, $n_{10} = 35$, $n_{11} = 32$ и значение статистического параметра

$X_2 = \frac{4}{127}\left(24^2 + 36^2 + 35^2 + 32^2\right) - \frac{2}{127}\left(60^2 + 68^2\right) + 1 = 2{,}295$. Поскольку

$X_2 < X_{2пр} = \chi^2(2;0{,}05) = 5{,}99$ [2] и $X_2 = 2{,}295 < 5{,}99$, то тест пройден, и гипотеза не отвергается.

2.3 Тест Покера

Пусть $m = 3$ и $k = 42$. Блоки 000, 001, 010, 011, 100, 101, 110, 111 появляются 4, 6, 5, 6, 5, 6, 4, 6 раз

соответственно, а значение статистического параметра

$$X_3 = \frac{2^3}{42}\left(\sum_{i=1}^{8} 4^2 + 6^2 + 5^2 + 6^2 + 5^2 + 6^2 + 4^2 + 6^2\right) - 42 = 1{,}048.$$

Поскольку $X_3 < X_{3пр} = \chi^2(7;0{,}05) = 14{,}076$ и $X_3 = 1{,}048 < 14{,}076$, то тест пройден, и гипотеза не отвергается.

2.4 Тест серий

Здесь $e_1 = 16{,}25$, $e_2 = 8{,}0625$ и $k = 2$. Имеется 20, 8 блоков единиц длиной 1, 2 соответственно, и 20, 8 интервалов нулей длиной 1, 2 соответственно. Значение статистического параметра X_4 равно 1,73.

Поскольку $X_4 < X_{4пр} = \chi^2(2;0{,}05) = 5{,}99$ и $X_4 = 1{,}73 < 5{,}99$, то тест пройден, и гипотеза не отвергается.

2.5 Автокорреляционный тест

Если $d = 8$, то $A(8) = 61$. Значение статистического параметра

$$X_5 = 2\left(61 - \frac{128-8}{2}\right)/\sqrt{128-2} = 0{,}18$$

Так как $X_{5пр} \le X_5 < X_{5пр} = N(0;1)$ и X_5 находится в допустимом интервале, $-2{,}17 \le 0{,}18 \le 2{,}17$, то тест пройден и гипотеза о нормальном законе распределения не отвергается.

Таким образом, рассмотренная ПСП проходит все использованные тесты по критерию χ^2, и гипотеза о том, что она обладает свойствами случайной последовательности не отвергается. Дело в том, что выбранная 32-разрядная последовательность является линейной рекуррентной последовательностью максимального периода с добавленным одним битом, и она должна успешно проходить тесты на псевдослучайность.

3. Тестирование источников случайных и псевдослучайных последовательностей на основе методики американского федерального стандарта FIPS 140-1

В американском федеральном стандарте FIPS 140-1 используется четыре статистических теста на случайность: монобитный тест, блочный тест, тест серий и тест длин серий. В этих тестах для удовлетворительных значений статистических параметров задаются границы. Отдельная битовая строка длиной 20000 битов, получаемая из генератора, подвергается проверке по каждому из четырех названных тестов. Если какой-нибудь из тестов не пройден, то считается, что генератор не прошел тестирование. FIPS 140-1 рекомендуется применять для технологического тестирования аппаратных генераторов случайных чисел.

Периодом периодической последовательности s называется наименьшее положительное число n, для которого s периодическая. Если s — периодическая последовательность периода n, то циклом s является подпоследовательность s.

Пусть s — последовательность. Серией s называется подпоследовательность s, состоящая из последовательных 0 или 1. Серия, состоящая из 0, называется интервалом, а серия, состоящая из 1, называется блоком.

3.1. Монобитный тест

Цель этого теста состоит в том, чтобы определить, является ли количество нулей и единиц в последовательности s приблизительно таким, каким оно ожидается для случайной последовательности. Пусть n_1 обозначает количество нулей (или единиц) в s. Число должно удовлетворять условию $9654 < n_1 < 10346$.

3.2. Блочный тест

Пусть m положительное целое число, такое, что $\left\lfloor \frac{n}{m} \right\rfloor \ge 5 \cdot (2^m)$, и пусть $k = \left\lfloor \frac{n}{m} \right\rfloor$. Разобьем последовательность s на k непересекающихся подпоследовательностей, каждая длиной m, и пусть n_i бу-

дет числом появлений *i-го* типа последовательности длиной *m*. Блочный тест определяет, действительно ли последовательности длиной *m* появляются в *s* приблизительно столько же раз, сколько ожидается для случайной последовательности. Для применения критерия используется расчет параметра

$$X_3 = \frac{2^m}{k}\left(\sum_{i=1}^{2^m} n_i^2\right) - k \ ,$$

что приблизительно соответствует χ^2 распределению с 2^m-1 степенями свободы. Статистический параметр, задаваемый уравнением, вычисляется для $m = 4$. Статистика должна удовлетворять условию $1{,}03 < X_3 < 57{,}4$.

3.3 Тест серий

Цель теста серий состоит в том, чтобы определить, действительно ли число серий различных длин в последовательности *s* такое же, как ожидается для случайной последовательности. Пусть *k* равно наибольшему количеству битов в последовательности. Для каждого *i* от 1 до k подсчитывается число интервалов и блоков длиной *i* (в целях упрощения теста серии длиной, больше 6, рассматриваются как серии длиной 6). Тест серий пройден, если количество серий нулей и единиц последовательности находится в пределах соответствующего интервала, заданного табл. 1.

Таблица 1

Длина серии	Требуемый интервал
1	2267-2733
2	1079-1421
3	502-748
4	223-402
5	90-223
6	90-223

3.4 Тест длин серий

Цель теста длин серий состоит в том, чтобы определить, действительно ли максимальная длина серии в последовательности *s* такая же, как ожидается для случайной последовательности. Тест длины серий пройден, если длина любой серии анализируемой последовательности не превышает 34.

Заключение

Использование критерия χ^2 позволяет проверить ПСП и СП на степень их "похожести" на случайную последовательность.

При создании стандарта статистического тестирования Украины в качестве базовых можно использовать частотный тест, двухбитный тест, тест Покера, общий тест серий и автокорреляционный тест. Эти тесты могут применяться при разработке новых программных и аппаратных генераторов ПСП и СП, а также в качестве технологического теста.

Предлагаемый набор тестов является более расширенным по сравнению со стандартом FIPS 140-1 и может применяться для ПСП и СП неограниченных длин.

Список литературы: 1. *FIPS PUB 140-1.* Cryptographic modules security requirements // NIST, 1993. **2.** *Менезис А., Ван Оршот П., Ватсон С.* Прикладная криптография Гл. 5 // CRC Press, 1996. **3.** *ANSI X9.17.* "American National Standard for Financial Institution Key Management (Wholesales)" American Bankers Association, 1985.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 04.04.2001

УДК 681.3.06

Т. А. ГРИНЕНКО, Ю. И. ГОРБЕНКО, С. Ю. ОРЛОВА

МЕТОД ФОРМИРОВАНИЯ И СВОЙСТВА ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ

Введение

Известно, что требуемый уровень криптографической защиты информации может быть обеспечен при условии, что ключи, параметры и пароли порождаются в системе случайно, равновероятно и независимо [1,2]. Для решения этих задач используются генераторы случайных и псевдослучайных последовательностей (ПСП). При программной и программно-аппаратной реализации криптографических средств в большинстве случаев практически невозможно или весьма сложно использовать чисто случайные процессы. Весьма сложно, а в ряде случаев и невозможно, использовать случайные реализации, когда последовательности большого периода должны однозначно восстанавливаться в пространстве и времени на различных объектах. В этом случае используют псевдослучайные последовательности с требуемыми свойствами. Так, для применения в криптографии ПСП должны обладать рядом свойств "случайности" их появления, иметь требуемый период повторения, высокую (требуемую) структурную скрытность законов формирования и др. [2]. Важной оперативно-технической характеристикой является вычислительная сложность формирования ПСП, которая может оцениваться, например, скоростью формирования символов в ПСП. В литературе эти требования интегрированы в понятии криптографически сильного генератора (КСГ). По определению Яо [4] КСГ можно считать такой генератор, в котором выходную последовательность генератора невозможно отличить от чисто случайной. Более точно криптографически сильным называют такой генератор, у которого никаким полиномиальным тестом невозможно определить, порождена ли его выходная последовательность псевдослучайным алгоритмом или выбрана случайно и равновероятно из множества всех последовательностей длины l.

Второе определение КГС основано на понятии непредсказуемости закона формирования символов и, в общем, звучит как требование невозможности предсказать никаким полиномиальным алгоритмом следующий бит последовательности на выходе генератора, если известны все предыдущие биты. Не останавливаясь на подробном анализе известных результатов в области КСГ, отметим, что общим подходом к созданию КСГ должно быть требование экспоненциальной сложности криптоанализа ПСП (в смысле определения закона её формирования). Поэтому поиск КСГ нужно вести с использованием преобразований, которые обладают экспоненциальной сложностью задач криптоанализа. К классу таких преобразований относятся преобразования на эллиптических кривых.

В 90-е годы разработан математический аппарат и созданы криптографические системы, преобразования в которых осуществляются в группах на эллиптических кривых [6]. В то же время в доступной литературе только указывается, буквально в виде термина, на возможность построения генератора ПСП на эллиптических кривых, но нет практических предложений и исследования свойств, формируемых таким образом ПСП. Целью настоящей статьи и является разработка и исследование свойств ПСП, формируемых с использованием генератора на эллиптических кривых.

1. Метод формирования псевдослучайных чисел на эллиптических кривых

Метод формирования базируется на рекуррентном вычислении и преобразовании псевдослучайных последовательностей и чисел на эллиптических кривых в полях Галуа [6].

Эллиптическая кривая (ЭК) над простым полем $GF(p)$, где p – простое число, определяется множеством точек $P_i = (x_i, y_i)$.

Будем полагать, что координаты x_i и y_i принимают значения над простым полем $GF(p)$, т.е. в интервале $[1, p\text{-}1]$. Причём, каждая точка этой кривой удовлетворяет уравнению [6]:

$$y_q^2 = (x_q^3 + ax_q + b) \bmod p \qquad (1)$$

Кроме того, за счёт выбора параметров кривой выполняется условие:

$$4a^3 + 27b^3 \neq 0 (\bmod\ p) \qquad (2)$$

Суть метода формирования псевдослучайных чисел заключается в следующем. В качестве начального значения принимается случайное или псевдослучайное число $a_0 \in [1, P\text{-}1]$. Назовём его раз-

мерностью пространства внутренних состояний l_0. Правило формирования ПСП на эллиптической кривой представим в виде:

$$a_i = a_{i-1} * G(\text{mod } p), \qquad i = 1,2,\ldots,n \tag{3}$$

где G - базовая точка эллиптической кривой (2) порядка n; p – простое число; a_0 – случайное начальное состояние генератора (число); знак "*" означает операцию скалярного умножения, соответственно.

Если точка G имеет порядок n, то максимальный период повторения будет равным n. Значение G выбирается случайно из полного множества $\{G\}$.

При правильном выборе параметров a и b эллиптической кривой порядок эллиптической кривой $u_{ЭК}$ может изменяться для простого поля $GF(p)$ в интервале $p-2\sqrt{p+1} \le u_{ЭК} < p+ +2\sqrt{p+1}$ и зависит от значений a и b.

Для конкретно выбранной базовой точки порядок эллиптической кривой моет быть связан с порядком базовой точки соотношением:

$$u_{эк} = h \cdot n, \tag{4}$$

причём, h предпочтительно выбирать равным 2, 4, 8.

Соотношение (4) реализует скалярное умножение, его можно записать в виде:

$$a_i = (\underbrace{G + G + G + \ldots + G}_{a_{i-1}}) \bmod p \tag{5}$$

Поскольку G - это точка на эллиптической кривой, то она имеет две координаты – $\{x_G, y_G\}$. Значение a_i, в общем случае, также принимает значение на эллиптической кривой, то есть имеет две координаты (x_{ai}, y_{ai}). В связи с тем, что a_{i-1} должно быть числом, его будем формировать как:

$$a_{i-1} := \psi(a_{i-1}). \tag{6}$$

В простом случае:

$$a_{i-1} := x(a_{i-1}) \qquad \text{или} \qquad a_{i-1} := y(a_{i-1}). \tag{7}$$

В более сложном случае:

$$a_{i-1} := \psi(x_{ai-1}, y_{ai-1}), \tag{8}$$

где ψ - функция нелинейного отображения, например, хеш-функция.

Таким образом, a_{i-1} всегда будет целым числом.

Анализ (5) показывает, что оно имеет большую сложность, если его рассчитать выполнением операции сложения. Расчет в (5) можно свести к удвоению точки G и сложению двух различных точек.

Пусть точка имеет вид $G(x_1, y_1)$. Тогда удвоение:

$$G + G = 2(x_1, y_1) = (x_3, y_3). \tag{9}$$

Здесь координаты (x_3, y_3) результирующей точки вычисляются следующим образом:

$$x_3 = \lambda^2 - 2x_1;\ y_3 = \lambda(x_1 - x_3) - y_1;\ \lambda = \frac{3x_1^2 + a}{2y_1}. \tag{10}$$

При сложении точек G_1 и G_2 имеем:

$$G_1 + G_2 = (x_1, y_1) + (x_2, y_2) = (x_3, y_3), \tag{11}$$

в этом случае:

$$x_3 = \lambda^2 - x_1 - x_2;\ y_3 = \lambda(x_1 - x_3) - y_1;\ \lambda = \frac{y_2 - y_1}{x_2 - x_1}. \tag{12}$$

Все операции в (10) и (12) выполняются по модулю p. Комбинируя операции удвоения и сложения, найдем скалярное произведение (5), то есть, $a_i=(x_{ai}, y_{ai})$.

В случае, когда нужно обеспечить меньшую вычислительную сложность формирования ПСП, а она для правила (3) хотя и носит полиномиальный характер, но остаётся ещё значительной, правило (3) можно реализовать в виде операций рекуррентного сложения точек на эллиптической кривой. Тогда:

$$G_i = (G_{i-1} + G_k) \bmod p, \tag{13}$$

где G_k – есть ключевая точка (ключ) на эллиптической кривой, или

$$G_i = (2G_{i-1}) \bmod p. \tag{14}$$

где G_0 – базовая точка на эллиптической кривой.

2. Разработка алгоритма формирования псевдослучайных чисел на эллиптической кривой

Анализ показывает, что для реализации генератора ПСП на эллиптических кривых необходимо сформировать общесистемные параметры – простое число p требуемой величины, и базовую точку G порядка n. Это вполне разрешимая задача и требует отдельного рассмотрения. Здесь мы выберем из [7] стандартные общесистемные параметры.

Входные данные: параметры эллиптической кривой, т.е. простое число p, длиной 192 бита, параметры кривой a и b, а также базовая точка G порядка n с коэффициентом связи h.

Выходные данные: последовательность a_n, полученная посредством применения функции выделения координаты x точки G: $\psi=\psi(x)$, и последовательность a_j, полученная посредством применения хеш-функции MD5: $\psi=\psi(x,y)$.

Алгоритм.

1. Согласно [8] параметры области случайной эллиптической кривой над F_p определены вектором T=(p,a,b,G,n,h), где конечное поле F_p определено как:

p= FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF = 2^{192}-2^{64}-1

Кривая $E: y^2=x^3+ax+b$ над F_p определяется как:

a= FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFC
b= 64210519 E59C80E7 0FA7E9AB 72243049 FEB8DEEC C146B9B1

Базовая точка:

G= 04 188DA0E B03090F6 7CBF20EB 43A18800 F4FF0AFD 82FF1012
07192B95 FFC8DA78 631011ED 6B24CDD5 73F977A1 1E794811

Порядок n точки G и кофактор (коэффициент связи) h:

n= FFFFFFFF FFFFFFFF FFFFFFFF 99DEF836 146BC9B1 B4D22831
h= 01

2. Начальное значение a_0, целое 192-битное число, выбираем случайным образом.
3. Определяем длину последовательности *count*.
4. Открываем файл *gen.out* для записи выходной последовательности a_i и файл *genhash.out* для записи выходной последовательности a_j.
5. Расчёт выходной последовательности:

Для k от 0 до *count* выполнять:

5.1. $a_k=a_0*G(mod\ p)=(x_{a_0}, y_{a_0})$

5.2. $a_i := x_{a_0}$;

5.3. Запись a_i в *gen.out*.

5.4. $a_i := Hash(x,y)$;

5.5. Запись a_j в *genhash.out*.

Результатом данного алгоритма будут две выходные последовательности, записанные в соответствующие файлы. Длина каждого отдельного числа последовательности, полученной в результате

выделения координаты x полученной точки, равна 192 битам. Длина чисел второй последовательности равна 128 битам.

В процессе разработки программного модуля была использована библиотека многократной точности, разработанная на кафедре БИТ ХТУРЭ.

3. Анализ генератора псевдослучайных чисел на эллиптической кривой

Результаты исследования генератора псевдослучайных чисел на эллиптической кривой показали, что выходные последовательности, сформированные с помощью такого генератора, обладают гарантированной длиной периода равной порядку n базовой точки G. В данной работе рассматривались последовательности, длина периода которых равна

FFFFFFFF FFFFFFFF FFFFFFFF 99DEF836 146BC9B1 B4D22831.

При этом модуль преобразований остаётся равным 192 битам.

Для проверки статистических свойств последовательностей использовались критерии χ^2-Пирсона, Колмогорова и Мизеса, а также программный комплекс, реализующий проверку последовательностей при помощи теста Маурэра [9]. Результаты исследований приведены в табл. 1-4.

В табл. 1 приведены результаты тестирования выходной последовательности a_i ГПСЧ на ЭК, сформированной с использованием функции выделения координаты x полученных точек, по критериям χ^2-Пирсона, Колмогорова и Мизеса. Причём, в указанной таблице приведены средние значения вероятности для критериев χ^2-Пирсона и Колмогорова и средние значения критерия Мизеса, вычисленные по некоторому количеству выборок для каждого указанного в таблице объёма. Количество выборок определённого объёма зависит от длины файла, например, если файл длиной 12288 байт, то количество выборок объёмом 5120 байт будет равной двум и объёмом 10240 байт - одной выборке. В процессе тестирования использовался файл длиной 48432 байта.

Таблица 1

Последовательность	Объём выборки	$P(\chi^2)$				$P(D_n\sqrt{n} < \lambda_a)$	ω^2
		1-мерный	2-мерный	3-мерный	4-мерный		
a_i	5120	0,182	0,182	0,111	0,182	0,1	0,043
	10240	0,077	0,077	0,054	0,077	0,514	0,1
	20480	0,14	0,14	0,103	0,14	0,814	0,114
	40960	0,33	0,33	0,212	0,33	0,886	0,171

Все значения, приведённые в табл. 1, не должны превышать 0.95. Из таблицы следует, что все последовательности a_i, сформированные генератором на эллиптической кривой, удовлетворяют требованиям независимости, равновероятности и однородности.

Таблица 2

Последовательность	Номер выборки	X_u	Z_u
a_i	1	3.317838	1.088926
	2	3.302538	-1.430245
	3	3.311281	0.009319
	4	3.309282	-0.319836
	5	3.309543	-0.276801
	6	3.314171	0.485174

В табл. 2 приведены значения X_u и Z_u вычисленные в результате проверки последовательности a_i при помощи универсального теста Маурэра. При этом значение параметра L было равно 4.

Значения Z_u, приведённые в табл. 2, должны попадать в интервал от –2,32638 до 2,32638. Это означает, что последовательность a_i практически не сжимаемая, а значит, удовлетворяет свойствам случайности и принадлежит равномерному закону распределения.

Аналогично, в табл. 3-4 приведены результаты тестирования выходной последовательности a_i, сформированной ГПСЧ на ЭК с применением хэш-функции, по критериям χ^2-Пирсона, Колмогорова и Мизеса (табл. 3), и универсального теста Маурэра (L=4) (табл. 4). В процессе тестирования использовался файл длиной 32768 байт.

Таблица 3

Последовательность	Объём выборки	$P(\chi^2)$				$P(D_n\sqrt{n} < \lambda_a)$	ω^2
		1-мерный	2-мерный	3-мерный	4-мерный		
a_j	5120	0,173	0,173	0,289	0,173	0,743	0,2
	10240	0,489	0,489	0,112	0,489	0,743	0,171
	20480	0,438	0,438	0,173	0,438	0,743	0,1

Все значения, приведённые в табл. 3, не должны превышать 0.95.

Так же как и данные, приведённые в табл. 2, значения Z_u, приведённые в табл. 4, должны попадать в интервал от –2,32638 до 2,32638. Анализ представленных результатов показывает, что последовательности, сформированные по рекуррентному правилу (3) и последующим хешированием, также удовлетворяют необходимым требованиям.

Таблица 4

Последовательность	Номер выборки	X_u	Z_u
a_j	1	3,302240	-1,479258
	2	3,312920	0,279148
	3	3,306862	-0,718212
	4	3,319351	1,338085

Кроме того, исследование предложенных последовательностей при помощи статистических тестов показало, что применение хеш-функции в виде правила (8) при одинаковом количестве испытаний в большинстве случаев даёт лучшие результаты, чем при применении правила выделения (7).

Заключение

С учётом того, что рассматриваемые ПСП сформированы с помощью ГПСЧ на ЭК, и при преобразованиях используются такие сложно обратимые операции как сложение или скалярное умножение на эллиптических кривых, можно высказать предположение, что формируемые последовательности относятся к классу криптографически сильных.

Применение функции нелинейного отображения $\psi = \psi(x, y)$ значительно увеличивает структурную скрытность ПСП и, как показано ниже, улучшает значения статистических критериев, однако, вычислительная сложность при этом несколько увеличивается, но остаётся приемлемой.

С учётом того, что порядок точек G на ЭК заведомо большой (2^{192}), можно утверждать, что вероятность перекрытия выходной последовательности есть достаточно малая величина.

Выборки из последовательностей, сформированных по правилам (6) и (7), являются случайными, равновероятными и независимыми.

С учётом сказанного можно утверждать, что генераторы ПСП на ЭК могут найти применение в криптографических и других приложениях. Кроме того, авторы понимают необходимость проведения более глубоких исследований по рассматриваемой проблеме.

Список литературы: 1. *Кнут Д.* Искусство программирования для ЭВМ. Т. 2. Получисленные алгоритмы. М.: Мир, 1977. 387 с. 2. *Шеннон К.Э.* Теория связи в секретных системах // Работы по теории информации. М.: ИЛ, 1963. 3. *Blum M., Micali S.* How to generate cryptographically strong sequences of pseudo-random bits // SIAM Journal of Computing, 13(4):850-864, 1984. 4. *Andrew C. Yao.* Theory and applications of trapdoor functions. In Proceedings of the 23rd Annual Symposium on Foundation of Computer Science, pages 80-91, IEEE Computer Society, 1982. 5. *Завадская Л.А., Фраль А.И.* Криптографически сильные генераторы псевдослучайных последовательностей // Безопасность информации: №1. С.7-11, 1997. 6. *Бондаренко М.Ф., Горбенко И.Д., Качко Е.Г.* Сущность и результаты исследований свойств перспективных стандартов цифровой подписи X9.62 –1998 и распределения ключей X9.63 – 1999 на эллиптических кривых // Радиотехника. 2000. Вып. 114. С.15-24. 7. ANSI X9.62 –1998: Certificate Managment. 8. *Simon Blake-Wilson, Minghula Qu.* Guidelines for Efficient Cryptography. Recommended Elliptic Curve Domain Parameters. Version 0.4. 1999. 9. *Menezes A., P. van Oorschot, and Vanstone S.* Handbook of Applied Cryptography. Chapter 5. Pseudo-random Bits and Sequences, CRC Press, 1997.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 14.02.2001

УДК 681.32

В. А. ГОРБАЧЕВ, канд. техн. наук, М. А. ВОЛК, канд. техн. наук,
С. Н. САРАНЧА, В. В. СТЕПАНЕНКО

ИССЛЕДОВАНИЕ МЕТОДОВ ГЕНЕРАЦИИ МНОГОСВЯЗНЫХ МАРКОВСКИХ ЦЕПЕЙ В ЗАДАЧЕ СЕРТИФИКАЦИИ МИКРОПРОЦЕССОРНЫХ КОМПОНЕНТОВ

Введение

Сложность современных микропроцессорных устройств, более широкий класс возможных неисправностей сверхбольших интегральных схем (СБИС) резко обостряют проблему их сертификации [4,5] и контроля [1,2]. Во-первых, значительная часть неисправностей микропроцессорных СБИС проявляется лишь на высоких частотах, в то время как на низких частотах микропроцессор функционирует правильно. Во-вторых, кроме функциональных неисправностей механизмов обработки, управления, хранения для микропроцессоров характерны так называемые "нелюбимые коды", т.е. неисправности типа "чувствительность к определенным последовательностям команд". Указанные особенности обусловливают необходимость широкого применения методов функционального контроля, при котором отвлекаются от конкретной вентильной реализации микропроцессорной СБИС и рассматривают лишь выполняемые схемой функции.

К настоящему времени сложилось два подхода к функциональному тестированию микропроцессорных СБИС. Первый основан на построении регулярных (детерминированных) тестов для заданного класса неисправностей СБИС. Основным его недостатком является сложность построения тестов, так как число неисправностей для современных СБИС огромно. Второй подход связан с использованием псевдослучайных тестов, для построения которых не требуется явного перечисления неисправностей. Однако было обнаружено, что формируемые псевдослучайные тесты сильно избыточны и содержат запрещенные коды и их комбинации, которые могут привести к непредсказуемому поведению объекта контроля. Устранить указанные недостатки можно путем управления процессом генерации, т.е. изменением вероятности порождения тестовых воздействий в зависимости от особенностей конкретного объекта контроля.

1. Псевдослучайный контроль МП

Процесс контроля микропроцессора (МП) заключается в следующем. На МП подается детерминированная последовательность команд, которая устанавливает МП в исходное состояние: загружает регистры, сбрасывает признаки результата, устанавливает счетчик команд и указатель отека и т.д. Далее на входы МП подается псевдослучайная последовательность команд. Отклики проверяемого МП анализируются путем сравнения либо с откликами эталонного МП, либо с заранее вычисленными характеристиками (сигнатуры, контрольные суммы, число переходов 0-1 и др.). Следует подчеркнуть, что рассматриваемая методика контроля ориентирована на проверку функций, реализуемых микропроцессором, независимо от его внутренней структуры.

Очевидно, что с целью обнаружения неисправностей, характерных для микропроцессорных СБИС, метод псевдослучайной генерации должен удовлетворять следующим требованиям:

- генерация тест-программ должна осуществляться на частоте не ниже рабочей частоты контролируемого микропроцессора;
- получаемые псевдослучайные тест-программы должны обладать заданной полнотой контроля.

Как известно, полнота контроля обеспечивается соблюдением условий загрузки, проявлением и транспортировкой неисправностей, что возможно лишь при управлении стохастическими характеристиками генерируемых тест-программ. Под термином "управление" подразумевается возможность задания условных либо безусловных вероятностей порождения отдельных команд соответствующей тест-программы.

В простейшем случае псевдослучайные последовательности могут состоять из команд, вероятности появления которых фиксированы и не зависят от того, какие команды поступали на МП в предыдущих тактах. К сожалению, получаемые тест-программы характеризуются низкой полнотой контроля [1].

Повысить полноту контроля можно за счет изменения вероятностей соответствующих команд в зависимости от того, какие команды поступали на МП в предыдущих тактах. В качестве математического аппарата управления процессом формирования тестовых программ используются многосвязные марковские цепи. Стохастические свойства порождаемых тест-программ задаются с помощью

соответствующей матрицы переходных вероятностей, что обеспечивает требуемое качество и эффективность контроля.

2. Генератор многосвязных Марковских цепей

В задачах технической диагностики и функционального контроля наиболее часто используются марковские цепи с дискретным временем и конечным числом состояний[3]. Рассмотрим общий случай многосвязной однородной Марковской цепи. Будем рассматривать единственно возможные и несовместимые события

$$A_1, A_2, \ldots, A_n$$

в неограниченном ряде испытаний с номерами 0, 1, 2,..., которые соединяются в γ-членные звенья с номерами 0,1,2,..., придавая номер h звену, объединяющему испытания с номерами

$$h, h+1, h+2, \ldots, h+\gamma-1$$

Однородной γ-связной Марковской цепью будем называть такую последовательность испытаний, при которых вероятность появления события A_β в испытании с номером $k > \gamma$ равна

$$p_{\alpha_1\alpha_2\ldots\alpha_\gamma\beta},$$

когда установлено появление событий

$$A_{\alpha_1}, A_{\alpha_2}, \ldots, A_{\alpha_\gamma} \qquad (1)$$

в γ предшествующих испытаниях, т.е. в испытаниях с номерами k-γ, k-γ+1, k-γ+2, k-1 (или в звене k-γ), при этом испытания с номерами k+1, k+2 и т. д. предполагаются еще не осуществленными и результаты их неопределенными и таким образом не влияют на результат.

Такое определение однородной многосвязной марковской цепи с конечным числом событий и дискретным временем (ММЦ) позволяет обобщить понятие МЦ. Например, «обычные» марковские цепи можно рассматривать как многосвязные со связностью 1. Независимые испытания, вероятность событий которых не зависят от предыдущих состояний системы, рассматриваются как ММЦ связности 0.

Законом цепи называется матрица порядка n^γ, составленная из переходных вероятностей $p_{\alpha_1\alpha_2\ldots\alpha_\gamma\beta}$. Данные переходные вероятности полностью определяют характер поведения ММЦ.

Задание переходных вероятностей возможно несколькими способами. Все примеры приведены для ММЦ связности n=2 с числом состояний γ=3.

1 способ. Задание переходных вероятностей определяется квадратной матрицей размерности n^γ, горизонтальными и вертикальными координатами в которой выступают звенья. Однако как видно из примера на рис. 1 матрица получается сильно разреженной ввиду наличия большого числа несовместных событий, и размерность матрицы значительно возрастает с ростом n или γ, что ведет к значительным непродуктивным затратам памяти при машинной реализации.

		Звено *t-1*							
		11	12	13	21	22	23	31	3
Звено *t-2*	11	p_{111}	p_{112}	p_{113}	0	0	0	0	
	12	0	0	0	p_{121}	p_{122}	p_{123}	0	
	13	0	0	0	0	0	0	p_{131}	p
	21	p_{211}	p_{212}	p_{213}	0	0	0	0	
	22	0	0	0	p_{221}	p_{222}	p_{223}	0	
	23	0	0	0	0	0	0	p_{231}	p
	31	p_{311}	p_{312}	p_{313}	0	0	0	0	
	32	0	0	0	p_{321}	p_{322}	p_{323}	0	
	33	0	0	0	0	0	0	p_{331}	p

Рис. 1

2 способ. Задание переходных вероятностей определяется прямоугольной матрицей размерности $n \times n^{\gamma}$, где по горизонтали располагаются состояния, а по вертикали звенья. Приведенный на рис.2 пример данной матрицы показывает неизбыточность данного способа представления данных.

		Состояние *t*		
		1	2	3
Звено *t-2*	11	p_{111}	p_{112}	p_{11}
	12	p_{121}	p_{122}	p_{12}
	13	p_{131}	p_{132}	p_{13}
	21	p_{211}	p_{212}	p_{21}
	22	p_{221}	p_{222}	p_{22}
	23	p_{231}	p_{232}	p_{23}
	31	p_{311}	p_{312}	p_{31}
	32	p_{321}	p_{322}	p_{32}
	33	p_{331}	p_{332}	p_{33}

Рис. 2

Аналогичным является представление матрицы переходных вероятностей в виде многомерного массива вида var P:array[1..3,1..3,1..3] в Pascal – нотификации.

3 способ. Все вышеперечисленные способы в машинной реализации являются применимыми при заранее известной размерности матрицы (то есть определены связность n и число состояний γ). В том случае, если данные параметры заранее неизвестны, или матрица переходных вероятностей является настолько разреженной, что не представляет смысла хранить ее в виде массива, предлагается представление в виде связного дерева, как показано на рис. 3.

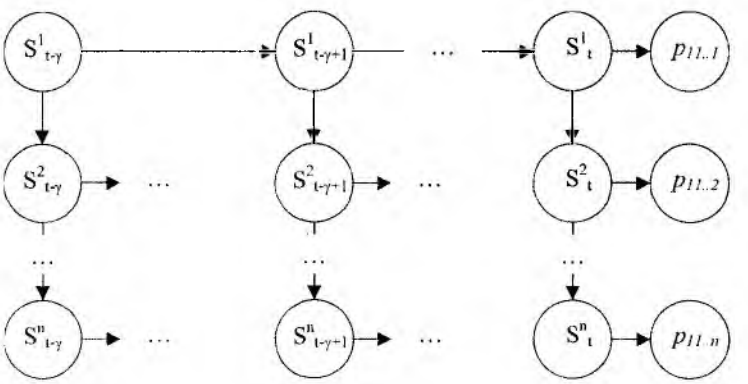


Рис. 3

Данная структура является универсальной с точки зрения программной реализации, позволяя строить матрицы переходных вероятностей для ММЦ с любыми n и γ.

Доказано[2], что наиболее благоприятным условием для генерации псевдослучайных тестовых программ является случай $H(P) \to 0$, где $H(P)$-энтропия матрицы переходных вероятностей. В то же время данное условие является противоречивым по отношению к сущности метода псевдослучайного тестирования, так как при $H(P) = 0$ каждая строка матрицы переходных вероятностей содержит только одну единицу и все остальные нули. В данном случае можно утверждать, что генерируемая тестовая программа будет *детерминированной* и постоянной. Разрешить данное противоречие можно с помощью рационального формирования матрицы переходных вероятностей. В то же время из условия $H(P) \to 0$ можно сделать вывод, что матрица переходных вероятностей будет разреженной. В таком случае наиболее рациональным с точки зрения затрат памяти при больших n и γ является именно способ представления матрицы переходных вероятностей в виде связного дерева.

3. Аппаратные и программные способы генерации ММЦ

Как известно, в общем случае для генерации случайной цепи с дискретными событиями необходимо произвести соответствие равномерно распределенной в диапазоне [0,1] случайной величины $x=R[0,1]$ будущему состоянию S_i с помощью соотношений:

$$S = \begin{cases} S_0 \ если \ x \le F(S_0) \\ S_1 \ если \ F(p_0) < x \le F(S_1) \\ \dots \\ S_n \ если \ x > F(S_{n-1}) \end{cases},$$

где $F(S)$ – функция распределения.

Представим теперь алгоритм генерации марковской последовательности произвольной связности как состящий из 2-х этапов. На первом этапе по сформированной случайной величине x и функции распределения $F(S)$ определяется будущее состояние цепи. На втором этапе производится модификация массива предыстории, в соответствии с ним формируется новая $F(S)$ и изменяется модельное время. Таким образом, для программной генерации многосвязных марковских последовательностей предлагается обобщенный алгоритм, представленный на рис. 4.

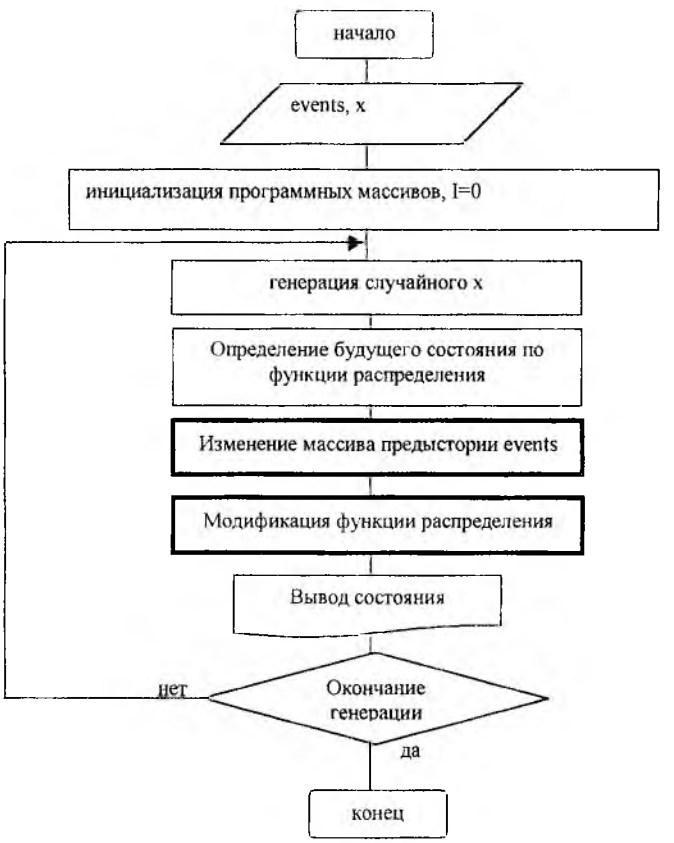


Рис. 4

В данном алгоритме выделенные вершины являются вариантными по отношению к размерности генерируемой цепи. Рассмотрим более подробно функции каждой вершины. В вершине «изменение массива предыстории events» происходит модификация занесением будущего состояния, представленная на рис. 5.

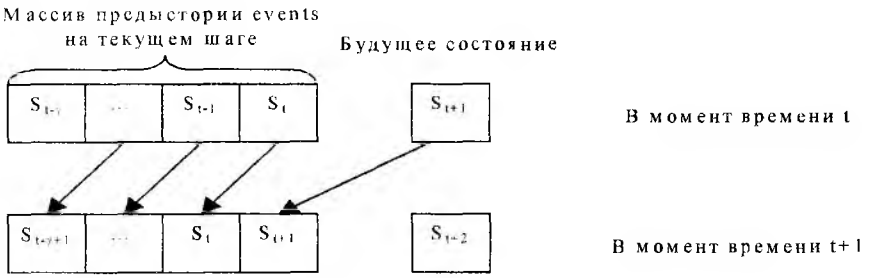


Рис. 5

В вершине «модификация функции распределения» происходит формирование новой функции распределения в соответствии с измененным массивом предыстории. Более детальные функции данной вершины сильно зависят от формы представления матрицы переходных вероятностей. Наиболее удобно данный алгоритм заложить в рамки объекта, реализующего функции загрузки матрицы переходных вероятностей из файла или базы данных, начальной инициализации, формирования случайной величины *x*, определения будущего состояния, изменения массива предыстории и получения соответствующей функции распределения. В целях исследования эффективности данного алгоритма составлена программа на языке C++ с применением библиотеки STL, содержащая объявление описанного выше класса, реализующего все 3 способа хранения матрицы переходных вероятностей и связанной с этим функциональности.

Для ускорения процесса формирования тестовых воздействий и разгрузки процессора возможно применение аппаратно реализованного генератора многосвязных марковских последовательностей. Структура данного формирователя представлена на рис. 6.

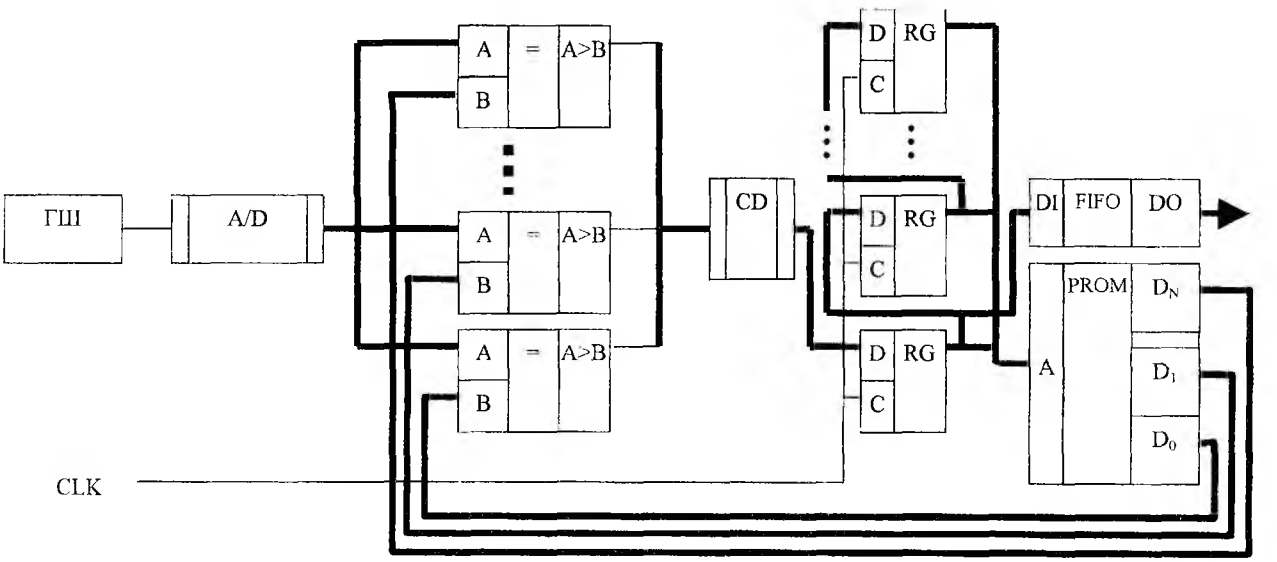


Рис.6

В структуре данного формирователя можно выделить следующие основные блоки:

- ГШ (генератор шума) и A/D (аналого-цифровой преобразователь), выполняющие функцию генерации случайного числа *x*.
- Блок цифровых компараторов, на вход *A* поступает случайная величина *x*, а на вход *B* – код $F(S_i)$ функции распределения, хранящейся в PROM.
- Приоритетный шифратор CD, формирующий код состояния на основе информации с выходов цифровых компараторов.
- Блок из γ регистров, реализующий функции хранения предыстории процесса. Выходы блока регистров подаются на адресный вход PROM, формируя таким образом соответствующую данной предыстории процесса функцию распределения $F(p)$.
- Буфер FIFO для хранения и воспроизведения формируемой последовательности.

Как видно из структуры данного формирователя, в нем используется второй способ представления матрицы переходных вероятностей, что предопределяет невозможность динамического изменения размерности матрицы и параметров цепи n и γ.

Код с выхода аналого-цифрового преобразователя, представляющий собой случайную величину x, параллельно поступает на входы блока цифровых компараторов, на вторые входы которых подается коды соответствующей предыстории функции распределения. Выходы компараторов подаются на вход приоритетного шифратора, формирующего код будущего состояния по позиции старшей единицы. Код будущего состояния с выхода приоритетного шифратора по переднему фронту тактового импульса сохраняется в младшем регистре блока регистров, реализующего функции массива предыстории. Одновременно с этим происходит сдвиг значений в остальных регистрах по направлению к старшему регистру, реализуя приведенное на рис. 5 преобразование. Код с младшего регистра сохраняется в буфере FIFO, что дает возможность чтения сформированной последовательности блоками в управляющий компьютер и воспроизведения последовательности. Так как выходы блока регистров составляют адресный вход для PROM, изменения в блоке регистров приводят к изменениям на выходах данных $D_0 \ldots D_n$, представляющих значения функции распределения для состояний $S_0 \ldots S_n$. Измененная функция распределения подается на вторые входы блока компараторов, завершая цикл формирования одного состояния. Данная структура реализована на базе программируемых ИМС ALTERA FLEX 8000 в САПР MAX + PLUS II и показала высокое быстродействие. В качестве элементов памяти использовались высокоскоростные микросхемы Motorola MCM 6806 BRJ-6 с временем доступа 6 нс. К преимуществам данной структуры можно отнести высокое быстродействие, реализуемое за счет параллельного сравнения и сдвига массива предыстории. К недостаткам следует отнести специфические требования к организации памяти и связанную с этим невысокую точность представления матрицы переходных вероятностей. Как видно из структуры формирователя, матрица переходных вероятностей хранится в параллельном виде. Это определяет, например, при связности n=2 с числом состояний γ=3 и восьмибитовом представлении вероятностей следующие требования к организации памяти: разрядность 16 бит, глубина 9. Можно сделать вывод, что данная структура более применима в случаях с небольшим числом состояний γ.

Для исследования алгоритмов формирования многосвязных марковских последовательностей была написана исследовательская программа с использованием пакета Inprise C Builder 4. Входными данными для программы являются количество состояний, связность МЦ, матрица переходных вероятностей и начальное состояние. Все исходные данные хранятся в базе данных типа DBF, доступ к которой осуществляется через библиотеку BDE. Программа генерирует γ-связные ММЦ и ведет статистики появления заданного состояния или цепочки состояний. Управление программным обеспечением комплекса диагностики осуществляется с помощью технологии OLE, для чего необходимо изначально зарегистрировать сервер OLE в системе. Тестирование программы показало высокое быстродействие предлагаемого метода генерации МЦ.

Заключение

В данной статье рассмотрены вопросы сертификации микропроцессорных компонентов с точки зрения псевдослучайного тестирования. Особое внимание было уделено проблеме высокочастотной генерации тестовых воздействий на основе математического аппарата многосвязных марковских последовательностей. В качестве результатов исследования выступает разработанный способ представления матрицы переходных вероятностей в виде связного дерева, обобщенный алгоритм и структура аппаратного формирователя многосвязных марковских последовательностей. Тестирование с помощью разработанных программных модулей и платы прототипирования показало высокую эффективность предлагаемых методов.

Список литературы: 1. *Хаханов В.И.* Техническая диагностика элементов и узлов персональных компьютеров. ХТУРЭ.-К.:ИСМО, 1997. 308 с. 2. *Кистории И.Ф., Гремальский А.А.* Функциональный контроль микропроцессорных устройств. Минск: Знание, 1990. 90 с. 3. *Дынкин Е.Б.* Марковские процессы. М.: Физматгиз, 1963. 860 с. 4. *Анин Б.Б.* Защита компьютерной информации. СПб.: BHV, 2000. VIII. 368с.:ил. 5. *Программно аппаратные средства обеспечения информационной безопасности* / В.Г. Проскурин и др. // Защита в ОС. М.: Радио и связь, 2000. 166с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 5.04.2001

УДК 681.04

В. А. КРАСНОБАЕВ, д-р. техн. наук

МЕТОДЫ РЕАЛИЗАЦИИ МОДУЛЬНЫХ ОПЕРАЦИЙ В СИСТЕМАХ ЦИФРОВОЙ ОБРАБОТКИ ИНФОРМАЦИИ

В ряде источников показана высокая эффективность применения системы счисления в остаточных классах (СОК) при решении отдельных задач обработки цифровой информации (решение задач фильтрации, БПФ, ДПФ и др.) [1-4]. В этом аспекте целесообразны теоретические и практические исследования путей дальнейшего повышения эффективности решения задач цифровой обработки (в частности, осуществление операции преобразования Фурье [5]) информации на основе использования свойств СОК.

В [6] детально рассмотрено влияние основных свойств СОК на структуру и принципы функционирования ЭВМ. В частности, показано, что малоразрядность остатков α_i дает возможность реализации арифметических операций в СОК либо на базе малоразрядных двоичных сумматоров, либо в табличном варианте. При первом методе реализации арифметических операций проявляется (хотя и в значительно меньшей степени) тот же недостаток, что и в позиционных системах счисления (ПСС): наличие межразрядных связей в пределах данного остатка m_i СОК. При табличном варианте реализации арифметических операций отсутствуют межразрядные связи между обрабатываемыми операндами вообще, однако для достаточно большой разрядной сетки ЭВМ (для больших по величине модулей СОК) резко увеличивается количество и сложность оборудования операционного устройства (ОУ). Важно и актуально рассмотреть промежуточный вариант реализации арифметических операций в СОК, основанный на применении кольцевого сдвига путем использования кольцевых сдвигающих регистров (КСР).

В [7] сформулирован новый принцип реализации арифметических операций в СОК принцип кольцевого сдвига (ПКС), особенность которого заключается в том, что результат арифметической операции $(\alpha_i \pm \beta_i) \bmod m_i$ по произвольному модулю СОК, заданной совокупностью $\{m_j\}$ $(j = \overline{1,n})$ оснований, определяется только за счет циклических сдвигов заданной цифровой структуры. Действительно, известная теорема Кэли устанавливает изоморфизм между элементами конечной абелевой группы и элементами группы перестановок. В этом случае матрица сложения для произвольного m_i модуля СОК будет задана табл. 1 (для $m_i = 5$ – табл. 2).

Таблица 1

β_i	α_i				
	0	1	2	...	$m_i - 1$
0	0	1	2	...	$m_i - 1$
1	1	2	3	...	0
2	2	3	4	...	1
...	...	...	...	...	...
$m_i - 1$	$m_i - 1$	0	1	...	$m_i - 2$

Таблица 2

β_i	α_i				
	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Одним из следствий теоремы Кэли является вывод о том, что отображение элементов абелевой группы на группу всех целых чисел является гомоморфным. Это обстоятельство позволяет организовать процесс определения результата арифметических операций в СОК посредством использования ПКС.

Операнд в СОК представляется набором из n остатков $\{\alpha_i\}$, образованных путем последовательного деления исходного числа A на n взаимно попарно простых чисел $\{m_i\}$ для $(i = \overline{1,n})$. В этом случае совокупность остатков $\{m_i\}$ непосредственно отождествляется с суммой n простых полей Галуа вида $\sum_{i=1}^{n} GF(m_i)$.

Известно, что преобразование Фурье связано с вычисление полинома вида $P(x) = \sum_{i=1}^{n-1} \alpha_i x^i$. Одно из приложений преобразования Фурье – вычисление свертки $\sum_{i=1}^{n} \alpha_i \beta_i$ двух n - мерных векторов $A = (\alpha_1, \alpha_2, ..., \alpha_n)$ и $B = (\beta_1, \beta_2, ..., \beta_n)$. Таким образом, эта операция является полным аналогом реализации арифметических операций умножения двух чисел A и B в СОК с последующим сложением компонент типа $\alpha_i \beta_i (\mathrm{mod}\ m_i) + \alpha_j \beta_j (\mathrm{mod}\ m_j)$.

Для рассмотрения метода реализации арифметических операций в СОК достаточно рассмотреть вариант для произвольного конечного поля Галуа $GF(m_i)$ при i = const , т. е. для конкретной приведенной системы вычетов по модулю m_i.

Пусть для заданной операции модульного сложения $(\alpha_i + \beta_i) \mathrm{mod}\ m_i$ в поле $GF(m_i)$ составлена таблица Кэли (табл. 1). Из существования нейтрального элемента в поле $GF(m_i)$ следует, что в табл. 1 есть строка (столбец), в которой элементы данного поля стоят в порядке возрастания, а из того факта, что в поле вычетов $GF(m_i)$ эти элементы различны (порядок группы равен m_i), следует, что в каждой строке (столбце) табл. 1 содержатся все элементы поля ровно по одному разу. Использование перечисленных свойств позволяет реализовать операции модульного сложения и вычитания в СОК путем применения ПКС посредством n кольцевых $M = m_i([\log_2(m_i - 1)] + 1)$ - разрядных сдвигающих регистров (КСР).

Пусть произвольная алгебраическая система представлена в виде $S = \langle G, \otimes \rangle$, где G - непустое множество; $\otimes$ - тип операции, определенной для любых двух элементов α_i, $\beta_i \in G$. Операция $\oplus$ сложения в множестве классов вычетов R, порожденных идеалом J, образует новое кольцо, называемое кольцом классов вычетов R/J. Его можно представить в виде Z/m_i, где Z – множество целых чисел 0, ±1, ±2,.... Если основание СОК m_i – простое число, то Z/m_i - поле. Данное обстоятельство, как указывалось выше, и обусловливает возможность реализации арифметической операции сложения в СОК без межразрядных переносов путем кольцевого сдвига (посредством применения КСР).

На основе предложенного в [7] принципа разработан метод реализации арифметических операций в СОК (метод двоичного кодирования). Суть разработанного метода состоит в том, что исходная цифровая структура для каждого модуля (основания) СОК представляется в виде содержимого первой строки (столбца) таблицы модульного сложения (вычитания) $(\alpha_i \pm \beta_i) \mathrm{mod}\ m_i$ вида:

$$P_{\text{исх}}^{(m_i)} = \left[P_0(\alpha_0) \| P_1(\alpha_1) \| ... \| P_{m_i-1}(\alpha_{m_i-1}) \right], \qquad (1)$$

где $\|$ - операция конкатенации; $P_v(\alpha_v)$ - k-разрядный двоичный код, соответствующий значению α_v-го остатка $(\alpha_v = \overline{0, m_i - 1})$ числа по модулю m_i ; $k = \left[\log_2(m_i - 1) + 1\right]$. Для заданного модуля m_i=5 исходная цифровая структура содержимого КСР имеет вид:

$P_{U-}^{(5)} = \left[000 \| 001 \| 010 \| 011 \| 100\right]$.

Таким образом, посредством используемых в ПСС кольцевых регистров сдвига легко реализовать арифметические операции в СОК. При этом степени циклических перестановок, исходя из (1), определяются следующими выражениями:

$$\left[P_0(\alpha_0) \| P_1(\alpha_1) \| ... \| P_{m_i-1}(\alpha_{m_i-1}) \right]^z = \left[P_z(\alpha_z) \| P_{z+1}(\alpha_{z+1}) \| ... \| P_0(\alpha_0) \| ... \| P_{m_i-1}(\alpha_{m_i-1}) \right] \qquad (2)$$

$$\left[P_0(\alpha_0) \| P_1(\alpha_1) \| ... \| P_{m_i-1}(\alpha_{m_i-1}) \right]^{-z} = \left[P_{m_i-1-z}(\alpha_{m_i-1-z}) \| .. \| P_{m_i-z}(\alpha_{m_i-z1}) \| \right.$$
$$\left. \| ... \| P_0(\alpha_0) \| P_1(\alpha_1) \| ... \| P_{m_i-z-2}(\alpha_{m_i-z-2}) \right] \qquad (3)$$

Отметим, что $\left[P_0(\alpha_0)\|P_1(\alpha_1)\|...\|P_{m_i-1}(\alpha_{m_i-1})\right]^{m_i} = \varepsilon$, т.е. при $z = m_i$ все элементы упорядоченного множества $\{P_j(\alpha_i)\}$ $(j = \overline{0, m_i - 1})$ остаются на исходном месте. При технической реализации данного метода первый операнд α_i определяет номер α_{α_i} разряда $P_{\alpha_i}(\alpha_{\alpha_i})$ с содержимым результата модульной операции по модулю m_i, а второй операнд β_i - число разрядов КРС ($\beta_i k$ -двоичных разрядов), на которые необходимо провести сдвиги исходного (1) содержимого КРС в соответствии с алгоритмами (2), (3). Основными недостатками предложенного в работе [7] метода реализации арифметических операций в СОК является сравнительно большое время его реализации, что снижает эффективность использования ПКС. Этот недостаток обусловлен тем, что структура $P_{исх}^{(m_i)}$ (1) представлена набором исходных остатков первой строки матрицы $(\alpha_i + \beta_i) \bmod m_i$, отображаемых двоичным кодом. В этом случае время реализации модульного сложения двух операндов $A = (\alpha_1, \alpha_2, ..., \alpha_{n-1}, \alpha_n)$ и $B = (\beta_1, \beta_2, ..., \beta_{n-1}, \beta_n)$ в СОК определяется выражением [8]:

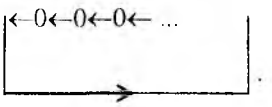
$$t_{сл} = k\beta_{max\ i}\tau, \tag{4}$$

где τ - время сдвига одного бита информации (одного двоичного разряда).

Рассмотрим метод реализации арифметических операций в СОК. Для метода унитарного кодирования, информационная структура $P_{исх}^{(m_i)}$ произвольного модуля m_i СОК, представляется в виде унитарного (m_i-1)-разрядного кода:

$$P_{исх}^{(m_i)} = \left[P(\alpha_{i-1})\|P(\alpha_{i-2})\|...\|P(1)\|P(0)\right], \tag{5}$$

где $P(\alpha_j)$ - двоичный разряд цифровой структуры (5), единичное состояние которого соответствует значению операнда $\alpha_i = \alpha_j$, представленного унитарным кодом $(\alpha_j = \overline{0, m_i - 1})$. В этом случае исходное состояние КРС состоит из m_i-1 двоичных разрядов и схематически может быть представлено в виде



При этом первый операнд $\alpha_i = \alpha_j$, отображаемый унитарным кодом по произвольному модулю m_i СОК. заносится в j-й разряд КРС, т.е переводит j-й двоичный разряд в единичное состояние. Второй операнд β_i указывает на число сдвигов z содержимого КРС, определяя время реализации арифметических операций по модулю m_i СОК, т. е.

$$t_{сл} = \beta_i\tau. \tag{6}$$

Отметим, что время реализации арифметической операции A + B в СОК будет определяться временем выполнения операции для максимального значения $(\beta_{max\ i}$ $(i = \overline{1, n})$ остатка из совокупности $\{\beta_i\}$ для данного операнда $B = (\beta_1, \beta_2, ..., \beta_n)$:

$t_{сл} = \beta_{max\ i}\tau$.(7)

Анализ выражений (4 и 7) показывает, что разработанный метод унитарного представления сокращает в $k = \left[\log_2(m_i - 1) + 1\right]$ раз время выполнения арифметических операций по сравнению с методом двоичного кодирования.

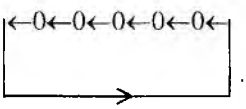
Алгоритм реализации арифметических операций в СОК посредством разработанного метода проиллюстрируем на примере операции сложения A + B для СОК, заданной основанием m_1=2, m_2=3, m_3=5. Пусть A = (0, 10, 100) и B = (1, 01, 010) (см. табл. 3). Так как $\beta_{max\ i} = \beta_3 = 010$, в соответст-

вии с выражением (7) $t_{сл} = \beta_3 \tau$ и алгоритм реализации операции сложения полностью определяется алгоритмом реализации модульного сложения $(\alpha_3 + \beta_3) \bmod m_3$.

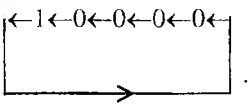
Таблица 3

A в ПСС	A в СОК при			A в ПСС	A в СОК при		
	m_1=2	m_2=3	m_3=5		m_1=2	m_2=3	m_3=5
0	0	00	000	15	1	00	000
1	1	01	001	16	0	01	001
2	0	10	010	17	1	10	010
3	1	00	011	18	0	00	011
4	0	01	100	19	1	01	100
5	1	10	000	20	0	10	000
6	0	00	001	21	1	00	001
7	1	01	010	22	0	01	010
8	0	10	011	23	1	10	011
9	1	00	100	24	0	00	100
10	0	01	000	25	1	01	000
11	1	10	001	26	0	10	001
12	0	00	010	27	1	00	010
13	1	01	011	28	0	01	011
14	0	10	100	29	1	10	100

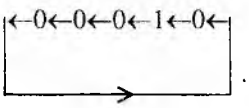
Исходное содержание КРС определяется в виде:



Первый операнд $\alpha_3 = 100$ дешифруется, и значение $\alpha_3 = 4$ в унитарном коде заносится в четвертый разряд КРС, содержание которого принимает вид:



Второй операнд $\beta_3 = 010$ также дешифруется, и полученное значение $\beta_3 = 2$ определяет число z сдвигов в положительном (против часовой стрелки) направлении содержимого КРС. В результате содержимое КРС представим следующим образом:



В соответствии с данными значениями кодов (табл. 4), посредством шифратора, по значению 00010 однозначно определяется результат операции $\alpha_3 + \beta_3$. Аналогично проводятся операции модульного сложения остатков по основаниям m_1 и m_2.

Таблица 4

Код	
входа шифратора	выхода шифратора
00001	000
00010	001
00100	010
01000	011

Проведем сравнительную оценку времени реализации арифметических операций в СОК и ПСС. Известно [9], что время реализации арифметических операций сложения и умножения в ПСС для l-байтовых $l = \overline{(1,4)}$ машинных слов определяется следующими выражениями:

$$t_{сл}^{(2)} = \tau(2\rho + 1), \qquad (8)$$

$$t_{умн}^{(2)} = \tau\rho^2, \quad (9)$$

где $\rho = 8 \cdot l$. А возможное и максимальное время реализации соответствующих арифметических операций для ПКС при применении метода двоичного представления – выражениями:

$$t_{сл}^{(2)} = (m_n - 1)k\tau, \quad (10)$$

$$t_{умн}^{(2)} = (m_n - 1)m_n k\tau/2. \quad (11)$$

Из выражений (6) и (7) видно, что максимально возможное время при использовании метода унитарного кодирования равно:

$$t_{сл}^{(2)} = (m_n - 1)\tau, \quad (12)$$

а для умножения в СОК:

$$t_{умн}^{(2)} = (m_n - 1)(m_n - 2)\tau, \quad (13)$$

так как $t_{умн}^{(3)} = \alpha_i(\beta_i - 1)\tau$, т. е. операнд α_i в унитарном коде заносится в КРС, а затем последовательно проводится сложение по схеме $\underbrace{\alpha_i + \alpha_i + \alpha_i + ... + \alpha_i}_{\beta_i}$.

Расчет (табл. 5), проведенный в соответствии с выражениями (8) - (13), показал высокую эффективность применения метода унитарного кодирования с точки зрения времени реализации арифметических операций в СОК по сравнению с методом двоичного кодирования и временем реализации таких же операций в ПСС.

Таблица 5

Разрядная сетка ЭВМ $l(m_n)$	Основания СОК $m_i\ (i=\overline{1,n})$	t[τ]					
		ПСС		СОК			
		Сложение	Умножение	Двоичное представление		Унитарное представление	
				Сложение	Умножение	Сложение	Умножение
l=1(m_n=7)	3,4,5,7	17	128	18	63	6	30
l=2(m_n=13)	2,5,7,9,11,13	33	512	48	312	12	132
l=3(m_n=19)	3,4,5,7,11,13,`7,19	49	1152	90	855	18	306
l=4(m_n=29)	2,3,5,7,11,13,17,19,23,29	65	2048	140	2030	28	756

Таким образом, полученные результаты могут быть использованы при оценке вычислительной сложности алгоритмов цифровой обработки информации, основанных на использовании преобразования Фурье.

Список литературы: 1. *Акушский И.Я., Юдицкий Д.И.* Машинная арифметика в остаточных классах. М.: 1968. 440 с. **2.** *Блейхут Р.* Быстрые алгоритмы цифровой обработки сигналов: Пер. с англ. М.: Мир, 1989. 448 с. **3.** *Кравченко В.Ф., Крот А.М.* Методы и микроэлектронные средства цифровой фильтрации сигналов и изображений на основе теоретико-числовых преобразований // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники. 1997. №6. С.3-31. **4.** *Червяков Н.И., Тынчеров К.Т., Велигоша А.В.* Высокоскоростная цифровая обработка сигналов с использованием непозиционной арифметики // Радиотехника. 1997. №10. С.23-27. **5.** *Лавриненко Д.И.* Применение быстрого преобразования Фурье в криптографических преобразователях// Радиотехника. 2000. Вып. 114. С.75-79. **6.** *Краснобаев В.А.* Основы создания вычислителей на основе остаточных классов // Системи обробки інформації. Харків:НАНУ, ПАНМ, ХВУ. 2001. Вип. 1(11). С.3-7. **7.** *Краснобаев В.А.* Принцип реализации арифметических операций в системе остаточных классов// АСУ и приборы автоматики. 1988. Вып.86. С. 82-85. **8.** *Долгов В.И., Краснобаев В.А., Кононова И.В.* Метод и алгоритмы реализации арифметических операций в системе остаточных классов // Электрон. моделирование. 1990. №5. С.70-72. **9.** *Краснобаев В.А., Ирхин В.П.* Алгоритмы реализации операции модульного умножения в системе остаточных классов // Электрон. моделирование. 1993. №5. С.20-26.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 17.04.2001

БЛОЧНЫЕ СИММЕТРИЧНЫЕ КРИПТОАЛГОРИТМЫ: ОСОБЕННОСТИ ПОСТРОЕНИЯ И КРИПТОАНАЛИЗ

УДК 681.3.06

С. А. ГОЛОВАШИЧ

БЕЗОПАСНОСТЬ РЕЖИМОВ БЛОЧНОГО ШИФРОВАНИЯ

Введение

Неотъемлемым компонентом современных систем криптографической защиты информации являются симметричные шифры. Основную область их применения составляют приложения, обрабатывающие большие объёмы конфиденциальной информации и предъявляющие высокие требования к производительности системы. Симметричные шифры принято разделять на два класса: поточные и блочные. Поточные шифры представляют класс симметричных криптоалгоритмов, которые обрабатывают по одному символу текста на каждом шаге, используя шифрующее преобразование, изменяющееся во времени. В отличии от них блочные шифры одновременно обрабатывают группы символов текста, используя фиксированное шифрующее преобразование [1]. Большинство современных шифров являются бинарными, т.е. символ представляет собой двоичную цифру (бит). Поэтому блочный шифр может рассматриваться как ключезависимая перестановка на множестве битовых векторов, соответствующих отдельным блокам.

В открытых компьютерных системах большее распространение получили блочные шифры, в то время как классические поточные шифры обычно ориентированы на аппаратную реализацию, являются секретными и используются преимущественно в специализированных системах связи. С целью устранения недостатков, свойственных шифрам подстановки, для блочных шифров был разработан ряд режимов, предназначенных для обработки больших объёмов информации. Эти режимы фактически определяют поточные схемы шифрования, построенные на базе блочного шифра. Далее, применительно к подобным схемам, преобразование, определённое алгоритмом блочного шифрования, будем называть (базовой) функцией шифрования, а всю схему — шифратором.

Целью данной статьи является анализ степени защищённости шифраторов, определённых стандартными режимами, от атак, используемых для нападения на блочные шифры. А также определение способов повышения безопасности этих режимов.

В общем случае, основной задачей криптоаналитика, атакующего систему шифрования, является определение секретного ключа либо синтез схемы, выполняющей преобразование данных эквивалентное расшифрованию (зашифрованию) на неизвестном ключе. Наиболее эффективные атаки криптоанализа блочных симметричных шифров строятся на основе знания криптоаналитиком как шифрованного, так и соответствующего ему открытого текста, т.е. известных пар «открытый – шифрованный текст». Эффективность многих из этих криптоатак повышается, если криптоаналитик может выбирать открытые тексты, поступающие на вход функции шифрования, т.е. атаки на основе «подобранных открытых текстов». Считается, что если функция шифрования устойчива к атакам на основе известных (подобранных) текстов, то она устойчива и к другим видам атак (атаки на основе связанных ключей учитывать не будем, т.к. они используют слабости процесса генерации ключей и для их реализации требуются «специфические» условия).

Стандартные режимы применения блочных шифров

В 1980 г. американским стандартом FIPS Pub 81, «DES Modes of Operation» [2] и затем в 1983 г. стандартом ANSI X3.106 [3] специально для алгоритма DES было определено четыре режима применения функции шифрования DEA. В общем виде для произвольных 64–битных и n–битных блочных шифров аналогичные режимы были определены стандартами ISO 8732 [4] и ISO/IEC 10116 [5] соответственно. Стандарт ГОСТ 28147–89 [6] также определяет четыре режима криптографического преобразования данных. По назначению и структуре эти режимы аналогичны режимам, определённым в указанных выше стандартах. Исключение составляет второй режим ГОСТа (режим гаммирования). По назначению он соответствует режиму OFB, но вместо обратной связи по гамме шифрующей, использует принцип «счётчика состояний». Подобная схема ранее была предложена Диффи и Хеллманом [7]. Рассмотрим все режимы, определённые указанными стандартами:

- Electronic Codebook (ECB) — режим электронной кодовой книги (ГОСТ: режим простой замены);
- Output Feedback (OFB) — режим обратной связи по выходу;
- Cipher Feedback (CFB) — режим обратной связи по шифртексту (ГОСТ: режим гаммирования с обратной связью);
- Cipher Block Chaining (CBC) — режим связки шифроблоков (ГОСТ: режим выработки имитовставки);
- «Counter» mode — режим «счётчика» (ГОСТ: режим гаммирования).

Так как большинство выше перечисленных режимов определяют поточные шифраторы, отметим, что они делятся на два класса: синхронные и самосинхронизирующиеся. В синхронных шифрах гамма шифрующая формируется независимо от обрабатываемого текста и определяется только ключом шифрования и внутренним состоянием шифратора. В самосинхронизирующихся шифрах гамма формируется как функция от ключа шифрования и некоторого фиксированного количества ранее сформированных символов криптограммы [1]. Для рассматриваемых схем отдельному символу будет соответствовать блок текста, обрабатываемый за одну итерацию шифрования.

Далее воспользуемся следующими обозначениями:
K – ключ шифрования;
n – размер блока базовой функции шифрования;
r – размер блока текста / гаммы (для поточных режимов);
L – длина текста (открытого либо зашифрованного) в блоках;
M_i, C_i – блоки соответственно открытого текста и криптограммы;
S_i – слово состояния шифратора (для режимов поточного шифрования);
$IV\ (SYN)$ – вектор начальной инициализации поточных режимов (синхропосылка);
$O_i = E_K(I_i)$ – функция зашифрования n–битного блока I_i на ключе K;
$I_i = D_K(O_i) = E_K^{-1}(O_i)$ – функция расшифрования n–битного блока O_i на ключе K;
I_i, O_i – соответственно входное и выходное значения функции шифрования E_K;
B_i – блок обратной связи;
G – рекуррентный генератор ПСЧ («счётчик»).
В используемых дальше соотношениях операция деления является целочисленной, т.е. деление на 2^k соответствует логическому сдвигу вправо на k разрядов.

Режим ECB

В режиме электронной кодовой книги (ECB) базовая функция шифрования применяется без дополнительных преобразований, т.е. отдельные блоки открытого текста независимо зашифровываются в блоки криптограммы. Прямое и обратное шифрующие преобразования для этого режима могут быть описаны следующими соотношениями:

$$C_i = E_K(M_i), \quad M_i = D_K(C_i), \quad i = 1 \ldots L.$$

Для этого режима свойственны все недостатки шифров подстановки: одинаковые блоки открытого текста отображаются на одинаковые блоки криптограммы; перестановка блоков криптограммы приводит к соответствующей перестановке блоков открытого текста и наоборот; модификация любого блока открытого текста после зашифрования сказывается только на соответствующем блоке криптограммы и наоборот. Независимость обработки отдельных блоков часто приводит к невозможности скрытия структуры защищаемой информации. Это, при определённых обстоятельствах, позволяет криптоаналитику получить интересующую информацию о зашифрованном сообщении только на основе криптограммы, без поиска ключа шифрования. Поэтому применение данного режима для обработки сообщений, превышающих один блок, не рекомендуется.

В этом режиме возможен криптоанализ базовой функции шифрования на основе «известных» и «отобранных» открытых текстов, т.к. значения входа и выхода функции шифрования соответствуют блокам открытого текста и криптограммы.

Режим OFB

Шифратор, определяемый режимом обратной связи по выходу (OFB), фактически соответствует поточному синхронному шифру. При этом размерность пространства состояний определяется длиной блока (входа) функции шифрования, а размерность выходного алфавита определяется длиной блока гаммы. В этом режиме блоки гаммы шифрующей Γ_i и блоки обратной связи B_i формируются на осно-

ве выхода базовой функции шифрования. В общем случае блоки Γ_i и B_i могут иметь некоторые длины r и t меньшие, чем длина блока n базовой функции шифрования, и формироваться как произвольные подмножества выходных битов функции E_K. Стандарт FIPS Pub 81 [2] предполагает возможность реализации OFB режима с идентичными блоками гаммы и обратной связи произвольной длины, т.е. $r = t$, $1 \le r \le n$, $B_i = \Gamma_i$.

В общем виде прямое и обратное шифрующие преобразования для случая r–битной гаммы и t–битной обратной связи могут быть описаны следующими соотношениями:

$$S_1 = SYN, \quad i = 1 \ldots L,$$
$$S_{i+1} = (S_i \times 2^t + B_i) \bmod 2^n, \quad O_i = E_K(S_i),$$
$$\Gamma_i = O_i / 2^{n-r}, \quad B_i = O_i / 2^{n-t},$$
$$C_i = M_i \oplus \Gamma_i, \quad M_i = C_i \oplus \Gamma_i.$$

Отметим, что для случая полной обратной связи имеем $r = t = n$.

Рассмотрим принципы выбора параметров r и t. Учитывая, что функция шифрования E_K определяет перестановку на множестве n–битных блоков и предполагая, что для случайного ключа K перестановка E_K выбирается действительно случайно из пространства всех возможных $(2^n)!$ перестановок, можно показать, что для случайно выбранных ключа и начального состояния ожидаемая длина цикла, до повторения состояния, будет равна приблизительно 2^{n-1}. С другой стороны, если $r < n$, то выходная последовательность формируется в соответствии с некоторой итеративной функцией, не являющейся перестановкой, и при допущении, что она ведёт себя как случайная функция, ожидаемая длина цикла будет порядка $2^{n/2}$ [1]. В связи с этим при реализации OFB режима рекомендуется использовать полную обратную связь (n бит). Такое требование предъявляется стандартом ISO/IEC 10116 [5]. Однако следует отметить, что для большинства блочных шифров задача доказательства (определения) минимально-возможного цикла, а следовательно, и минимального периода гаммы в этом режиме, на полном пространстве допустимых ключей и начальных состояний, является трудно разрешимой. В связи с этим возможна ситуация когда для некоторых значений ключа шифрования и начальных состояний длины циклов будут очень короткими, что может привести к повторению последовательности гаммы шифрующей при обработке больших потоков информации. С другой стороны, использование блоков гаммы длиной меньше n бит ограничивает возможности атакующего изучать свойства функции шифрования на основе известных пар «открытый–шифрованный» текст и, следовательно, усложняет задачу поиска ключа шифрования.

Данный режим обеспечивает большую надёжность, чем ECB и предлагается в качестве одного из основных режимов шифрования больших потоков данных. Но этот режим также не «скрывает» функцию шифрования от атак на основе известных пар текстов. Так, если криптоаналитику известен открытый текст для некоторой цепочки последовательных блоков криптограммы, тогда первый блок цепочки пропускается, т.к. для этого блока состояние шифратора полностью не определено, а начиная со второго блока цепочки, входное значение может быть вычислено как XOR предыдущего блока криптограммы с соответствующим ему блоком открытого текста. При этом следует учитывать, что значение синхромаркера SYN, используемое для инициализации шифратора на первом цикле, является открытой величиной и подаётся на вход функции шифрования в исходном виде, поэтому, если известная криптоаналитику цепочка начинается с первого блока, то для построения атаки могут использоваться все известные блоки, начиная с первого. Для случая «полной» (n–битной) обратной связи имеем:

$$I_i = C_{i-1} \oplus M_{i-1}, \quad O_i = C_i \oplus M_i.$$

Режим CFB

Шифратор, определяемый режимом обратной связи по шифртексту (CFB) фактически соответствует поточному самосинхронизирующемуся шифру. При этом, как и для OFB режима, размерность пространства состояний шифратора определяется длиной блока (входа) функции шифрования, а размерность выходного алфавита определяется длиной блока гаммы. Символы гаммы шифрующей формируются как некоторое подмножество выходных битов функции шифрования, т.е. длина блока гаммы r может быть меньше длины блока n базовой функции шифрования. В качестве обратной связи используется блок криптограммы, поэтому её разрядность также равна r. Стандарт ГОСТ 28147-89 (режим гаммирования с обратной связью) предусматривает использование только полной n-битной обратной связи.

В общем виде прямое и обратное шифрующие преобразования для случая r–битных гаммы и обратной связи могут быть описаны следующими соотношениями:

$$S_0 = SYN, \quad i = 1 \ldots L,$$
$$S_{i+1} = (S_i \times 2^r + C_i) \bmod 2^n,$$
$$\Gamma_i = E_K(S_{i-1}) / 2^{n-r},$$
$$C_i = M_i \oplus \Gamma_i, \quad M_i = C_i \oplus \Gamma_i.$$

Режим CFB, как и OFB, позволяет атаковать базовую функцию шифрования на основе известного открытого текста: входное значение функции шифрования соответствует ранее полученному шифртексту либо синхромаркеру для первого блока криптограммы, а выход функции шифрования может быть получен как XOR блоков криптограммы и соответствующего известного открытого текста. Для случая «полной» (n–битной) обратной связи имеем:

$$I_i = C_{i-1}, \quad O_i = C_i \oplus M_i.$$

Кроме того, если полученная криптограмма содержит две одинаковые последовательности блоков длиной $\lceil n / r \rceil$, значит открытый блок, следующий за каждой из этих последовательностей, будет зашифрован идентичным значением гаммы, однако при ограниченной длине текста ($L^2 << 2^n$) вероятность такого совпадения, в соответствии с парадоксом «дня рождения», является довольно низкой.

Режим CBC

Режим связки шифроблоков (CBC) по своим возможностям подобен режиму с обратной связью по шифртексту (CFB). Однако в отличие от CFB-режима, в режиме CBC операция блочного шифрования выполняется после «связывания» текущего блока с предшествующим, а при расшифровании используется обратная схема. Вследствие использования для обратного преобразования базовой функции блочного расшифрования, режим может манипулировать только блоками текста полной длины n.

Прямое и обратное шифрующие преобразования для этого режима могут быть описаны следующими соотношениями:

$$C_0 = SYN, \quad i = 1 \ldots L,$$
$$C_i = E_K(M_i \oplus C_{i-1}), \quad M_i = C_{i-1} \oplus D_K(C_i).$$

Основное достоинство данного режима перед CFB заключается в том, что последний блок криптограммы является ключезависимой нелинейной функцией от всех блоков криптограммы (в CFB-режиме последняя пара блоков открытого и шифрованного текстов связана функцией XOR). Это свойство позволяет использовать последний блок криптограммы (или его часть) в качестве кода аутентификации сообщения, т.е. получать код аутентификации L–блочного текста за L шагов. В соответствии с американскими и международными стандартами [2–5] этот режим может применяться как для шифрования, так и для аутентификации сообщений, в то время как аналогичный режим стандарта ГОСТ 28147–89 (режим выработки имитовставки) применяет функцию шифрования с сокращённым вдвое числом циклов и предписывает использование этого режима только для целей аутентификации, т.е. промежуточные блоки криптограммы не сохраняются.

Режим CBC, как и предыдущие режимы, позволяет атаковать непосредственно базовую функцию шифрования на основе пар известного (подобранного) открытого текста: входное значение функции шифрования может быть вычислено как XOR предыдущего блока шифртекста и текущего блока открытого текста, а выход функции шифрования соответствует текущему блоку криптограммы:

$$I_i = M_i \oplus C_{i-1}, \quad O_i = C_i.$$

Кроме того, если два соседних блока шифртекста появляются в криптограмме более одного раза, то значит второй блок в этой паре соответствует одинаковым блокам открытого текста, хотя, при ограниченной длине текста ($L^2 << 2^{2n}$), вероятность такого совпадения пренебрежительно мала.

Режим «счётчика»

Режим гаммирования в соответствии с ГОСТ 28147–89, как и режим OFB (в соответствии с FIPS), определяет шифратор, соответствующий поточному синхронному шифру, однако использует принцип «счётчика». В отличие от OFB, данный режим ГОСТа предполагает использование блоков гаммы полной длины n.

Прямое и обратное шифрующие преобразования могут быть описаны следующими соотношениями:

$$S_0 = E_K(SYN), \quad i = 1 \ldots L,$$
$$S_i = G(S_{i-1}) = G^*(S_0, i),$$
$$\Gamma_i = E_K(S_i),$$
$$C_i = M_i \oplus \Gamma_i, \quad M_i = C_i \oplus \Gamma_i.$$

Данный режим ГОСТа использует значение зашифрованного базовой функцией синхромаркера (открытый параметр) в качестве начального состояния n – разрядного линейного конгруэнтного генератора с известным большим периодом (близким к 2^n). Блоки гаммы получаются путём шифрования базовой функцией «задающей» последовательности, формируемой указанным генератором. Использование подобной схемы позволяет обеспечить доказуемый фиксированный период выходной гаммы при любом ключе шифрования и любом начальном состоянии (синхромаркере). Кроме того, при такой схеме невозможен криптоанализ функции шифрования на основе известных пар текстов, т.к. вход функции шифрования не может быть получен из открытого и зашифрованного текста путём применения тривиальных (не зависящих от ключа) преобразований.

Данная схема также обладает рядом недостатков. Так, использование в качестве генератора «счётчика» с известным коэффициентом приращения приводит к тому что атакующему известна дифференциальная разность любой пары элементов «задающей» последовательности. При «слабой» функции шифрования это свойство может быть использовано для построения атаки дифференциального криптоанализа. Кроме того, формируемая последовательность блоков гаммы шифрующей является периодичной, т.е. появление идентичных значений гаммы невозможно для соседних $\pm T$ блоков, где T – период гаммы, обеспечиваемый генератором. Это даёт криптоаналитику дополнительную информацию, особенно при наличии большого объёма известного открытого текста.

Рассматриваемый режим, в отличии от OFB, обладает свойством «произвольного доступа», т.е. возможностью выполнять зашифрование / расшифрование фрагментов потока с произвольным смещением. При этом, в отличие от CFB-режима, не требуется знания дополнительных блоков криптограммы. Потребность в указанном свойстве возникает при реализации функций «прозрачного» шифрования в устройствах хранения информации с произвольным доступом. Данное свойство обусловлено наличием вычислительно простого не итеративного соотношения:

$$S_i = G^*(S_0, i).$$

Наиболее простым генератором, удовлетворяющим указанному свойству, является «счётчик» — накапливающий сумматор с фиксированной величиной приращения.

«Счётчик» с фиксированным периодом

Рассмотрим способы построения генераторов «задающей» последовательности типа «счётчик» и их свойства.

Наиболее простым (классическим) «счётчиком» является схема, удовлетворяющая соотношению $G_1(X) = (X + 1) \bmod M$. Очевидным является факт, что период рекуррентной последовательности вида $X_{i+1} = G_1(X_i)$ будет равен M при любом начальном значении X_0. Однако существенным недостатком данной схемы является то, что при $L << M$ элементы последовательности $(X_i, \ldots, X_{i+L})$ будут отличаться друг от друга только в младших разрядах. Поэтому использование подобной «задающей» последовательности приводит к ограниченной «активизации» входов функции шифрования, что может быть использовано для повышения эффективности криптоанализа базовой функции.

В связи с этим более предпочтительной является следующая обобщённая схема:

$$G_C(X) = (X + C) \bmod M. \tag{1}$$

Утверждение 1. Если конгруэнтный генератор вида:

$$X_{i+1} = G_C(X_i) \tag{2}$$

удовлетворяет требованию $(C, M) = 1$, то формируемая этим генератором последовательность $\{X_i\}$ для любой начальной точки X_0 будет пробегать все значения в диапазоне $0, \ldots, M-1$ и соответственно иметь максимальный период повторения, равный M.

Указанное свойство следует из теории вычетов [8], учитывая, что любой элемент последовательности (2) может быть записан в следующем виде:

$$X_i = (X_0 + C \times i) \bmod M. \qquad (3)$$

Использованный в соотношении (1) коэффициент C определяет величину приращения (шаг) генератора вида (2). Его выбор определяет свойства формируемой последовательности, поэтому для обеспечения максимального периода, необходимо проверять условие $(C, M) = 1$, а для улучшения статистических свойств формируемых последовательностей — накладывать ограничение на длины 1- и 0-вых битовых серий.

Основным недостатком генераторов вида (3), как было отмечено выше, является фиксированная дифференциальная разность для любой пары элементов последовательности:

$$\Delta X_{ij} = C \times (j - i) \bmod M.$$

При этом указанная зависимость между элементами «задающей» последовательности, даже в случае секретного (зависящего от ключа и синхромаркера) коэффициента C, может быть использована для криптоанализа функции шифрования.

Для дальнейшего изложения воспользуемся следующими обозначениями для определения операций над n - разрядными целыми числами x и y:

- $x\,[+]_n\,y = (x + y) \bmod 2^n$;
- $x\,\{+\}_n\,y = ((x + y) + (x + y) / 2^n) \bmod 2^n$;
- $\{++\}_n\,x = x\,\{+\}_n\,1$.

Первая операция соответствует обычному n–разрядному суммированию. Вторая и третья операции, соответственно, определяют функции сложения и инкрементирования по модулю 2^n-1, при этом в области значений результата выполняется замена $0 \rightarrow 2^n$-1, т.е. «запрещённым» является значение 0. Преимуществом второй операции перед обычным сложением является влияние каждого разряда аргумента на все разряды результата. Аппаратно эти функции реализуются на основе обычного сумматора путём дополнительного прибавления бита внешнего переноса (переполнения) к младшему разряду результата.

Рассмотрим свойства генератора G, применяемого в режиме гаммирования ГОСТа. Он представляет собой два параллельно работающих «счётчика» сравнительно малого периода. Периоды этих «счётчиков» взаимно просты, поэтому общий период комбинированного генератора равен произведению периодов «счётчиков» его составляющих. Генератор ГОСТа может быть описан следующими соотношениями:

$$S_{i+1} = G(S_i), \quad S_i = \{Y_i, X_i\},$$
$$X_{i+1} = X_i\,[+]_r\,C_1, \quad Y_{i+1} = Y_i\,\{+\}_r\,C_2,$$

где X_i, Y_i – значения r - разрядных регистров на шаге i ($r = 32$); C_1, C_2 – константы.
Учитывая свойства операций $[+]_r$ и $\{+\}_r$ получаем:

$C_1 = 01010101h \Rightarrow (C_1, 2^{32}) = 1 \Rightarrow T_1 = 2^{32}$,
$C_2 = 01010104h \Rightarrow (C_2, 2^{32}-1) = 1 \Rightarrow T_2 = 2^{32}-1$,
$(T_1, T_2) = 1 \Rightarrow T = T_1 \times T_2 = 2^{32} \times (2^{32}-1) = 2^{64} - 2^{32}$.

Пути повышения безопасности стандартных режимов

Как было показано выше, все рассмотренные режимы обладают определёнными недостатками. Определим принципы построения поточных схем шифрования на базе блочных криптоалгоритмов, устраняющие обнаруженные слабости:

1. Период гаммы, формируемой шифратором, должен удовлетворять некоторой нижней границе T_{min} при любом ключе шифрования и векторе инициализации.

2. Функция смены состояний (формирования «задающей» последовательности) должна быть нелинейной и ключезависимой.

3. Шифратор должен скрывать своё текущее состояние, а его выход (блок гаммы Γ_i) должен оставлять неопределённость относительно текущего состояния шифратора S_i, т.е. пространство состояний должно превышать пространство выходов.

Применение приведенных выше принципов построения поточных режимов шифрования позволяет минимизировать, по сравнению со стандартными режимами, количество дополнительной ин-

формации, которую может извлечь криптоаналитик из известных пар текстов.

Для реализации первого принципа при построении шифратора в качестве основы может использоваться схема на базе «счётчика», гарантирующая необходимый период.

Для реализации второго принципа критичный параметр генератора (величина приращения «счётчика») может быть сделан динамически изменяемой величиной и вычисляться как функция от ключа шифрования и текущего состояния, т.е. указанный параметр может формироваться по принципу OFB- (CFB-) режима. Однако такое решение вступает в конфликт с реализацией первого принципа. Для решения этого противоречия можно предложить использовать схему на базе «счётчика» с «плавающим» периодом, рассмотренную далее.

Для реализации третьего принципа необходимо, чтобы каждый блок гаммы формировался путём криптографического «сжатия» текущего внутреннего состояния шифратора. Наиболее простым решением этой задачи, при идентичной разрядности генератора и базовой функции шифрования, является формирование блоков гаммы как некоторого подмножества выходных битов функции шифрования. Тогда, если используется m–разрядный генератор и число его состояний $M \approx 2^m$, а выход шифратора составляет n разрядов ($m > n$), то, при условии равновероятного распределения выходных значений шифратора по состояниям генератора, вероятность «угадывания» состояния шифратора по известному блоку гаммы (при фиксированном ключе) составит $2^n / M \approx 2^{n-m}$.

В случае выполнения последнего требования становится возможным повторение отдельных блоков гаммы в пределах периода, что позволяет сократить информацию о состоянии криптосистемы и открытом тексте, которые атакующий может извлечь из перехваченной криптограммы.

«Счётчик» с «плавающим» периодом

Рассмотрим принципы построения и свойства криптографически стойких генераторов типа «счётчик» с динамически изменяемым коэффициентом приращения.

Определение 1. Под *генератором псевдослучайных чисел с «плавающим» периодом* (*либо «плавающим» ГПСЧ*) будем понимать ГПСЧ, формирующий последовательность чисел с переменным периодом T, зависящим от начального состояния S_0 и управляющего параметра K генератора, фактическое значение которого всегда находится в диапазоне предельных значений: $T_{min} < T = \lambda (S_0, K) < T_{max}$.

Определение 2. *Счетчиком с «плавающим» периодом*, будем называть «плавающий» ГПСЧ, удовлетворяющий следующим рекуррентным соотношениям:

$$S_{i+1} = S_i\,[+]_m\,(C \times N_i), \qquad N_i = \{{+}{+}\}_n\,\mathrm{H}_K(S_i),$$

где m – разрядность «счётчика»; n – разрядность нелинейной обратной связи; C – «коэффициент подъёма» (нечётная константа, т.к. $(C, 2^m) = 1$); S_i – состояние генератора на шаге i (разрядность m бит); N_i – «шаг подъёма» на шаге i (разрядность n бит); K – управляющий параметр генератора (ключ шифрования); H_K – криптографическая функция нелинейного «сжатия» $S_i \rightarrow N_i$ ($m > n$), параметризованная ключом K.

«Шаг подъёма» N_i в последнем соотношении отражает количество «шагов» базового «счётчика» с фиксированным периодом отделяющих состояния S_i и S_{i+1}. При этом необходимым требованием является условие $N_i \geq 1$, для его обеспечения используется операция $\{{+}{+}\}_n$. Величину $C \times N_i$ назовём динамически изменяемым приращением счётчика.

Отметим, что «счётчик» с фиксированным периодом может рассматриваться как частный случай «счётчика» с «плавающим» периодом, когда $\forall i$ выполняется $N_i = 1$. Кроме того, вместо операции $[+]_m$ возможно применение операции $\{+\}_m$, тогда $(C, 2^m\text{-}1) = 1$

Утверждение 2. Период m–разрядного «плавающего» счётчика с n–разрядным «шагом подъема» *при любых начальном состоянии* S_0 *и управляющем ключе* K *будет находиться в интервале:* $2^{m-n} < T \leq 2^m$.

Указанное свойство является следствием двух предельных случаев: учитывая, что нелинейная функция H_K является «сжимающей» ($m > n$), можно предположить, что при некоторых значениях K и S_0 для всех элементов последовательности S_i выполняется одно из следующих условий:

1) если $N = \mathrm{H}_K(S_i) = 1, \forall i$, то $T = T_{max} = M / N = 2^m$;
2) если $N = \mathrm{H}_K(S_i) = 2^n\text{–}1, \forall i$, то $T = T_{min} = M / N = 2^m / (2^n\text{-}1) \approx 2^{m-n}$.

Величина неопределенности криптоаналитика относительно «расстояния» (дифференциальной разности) между любыми двумя соседними состояниями S_i и S_{i+1} определяется ключезависимой нелинейной обратной связью — «шагом подъёма» N_i. Увеличение разрядности этой обратной связи n будет повышать неопределённость указанного «расстояния», но в то же время снижать нижнюю границу возможного периода генератора T_{min}. Поэтому компромиссным решением является выбор разрядности обратной связи $n = m / 2$.

Следует отметить, что для «счётчиков» большой разрядности ($m > 2r$, где r – разрядность регистров базового процессора) операция $C \times X$ в общем случае является сравнительно трудоёмкой. Однако учитывая, что величина приращения состояния $\Delta S_{i,\,i+1}$ зависит от трудно предсказуемого «шага подъёма» N_i, использование открытого значения C в шифраторах на базе «плавающего счётчика» не приводит к существенному снижению общей стойкости. Поэтому в таких схемах значение C может быть константным и выбираться с учётом оптимизации производительности (например, умножение на константу $C = 2^{m-r} + ... + 2^r + 1$ может быть реализовано только командами сложения определённых регистров, содержащих X).

"Усиленные" режимы поточного шифрования

В качестве примера рассмотрим несколько схем поточных синхронных шифров на базе «плавающего счётчика» и функции блочного шифрования с длиной блока n, равной разрядности «счётчика». В предлагаемых схемах выход функции шифрования разделён на две равные составляющие (по $n/2$ бит):

1) блок гаммы текста Γ^T_i — используется для зашифрования / расшифрования i–го блока текста;

2) блок гаммы обратной связи Γ^{FB}_i — формирует обратную связь, определяющую величину «шага подъёма» N_i.

При условии применения стойкой функции блочного шифрования каждая из указанных половинок может рассматриваться как результат однонаправленного криптографического «сжатия» текущего состояния шифратора S_i двумя различными функциями. Так как размер выходного слова шифратора (гаммы текста) и соответственно единицы обработки текста равен $n/2$, то под блоком далее будем понимать вектор длиной $n/2$ бит.

На рис. 1 приведена структурная схема усиленного режима поточного шифрования «с последовательным доступом». Период гаммы шифрующей для этой схемы будет «плавающим» в диапазоне $2^{n/2} < T \le 2^n$, а его фактическое значение будет зависеть от ключа шифрования K и вектора инициализации (синхропосылки SYN).

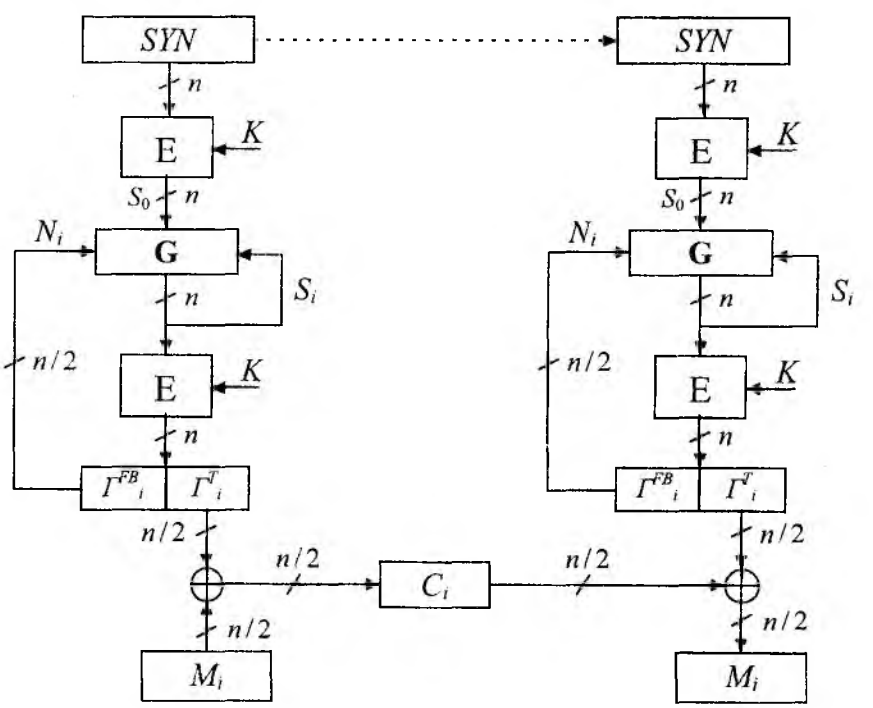


Рис. 1

Процедуры зашифрования и расшифрования для этого режима могут быть описаны следующими соотношениями:

$$S_0 = E_K(SYN), \quad N_0 = 1, \quad i = 1 \ldots L,$$
$$S_i = \mathrm{G}(S_{i-1}, N_{i-1}) = S_{i-1}[+]_m (C \times N_{i-1}), \quad N_i = \{++\}_{n/2} \Gamma^{FB}{}_i,$$
$$\{\Gamma^{FB}{}_i, \Gamma^{T}{}_i\} = E_K(S_i),$$
$$\Gamma^{T}{}_i = E_K(S_i) \pmod{2^{n/2}}, \quad \Gamma^{FB}{}_i = E_K(S_i) / 2^{n/2},$$
$$C_i = M_i \oplus \Gamma^{T}{}_i, \quad M_i = C_i \oplus \Gamma^{T}{}_i,$$

где $\Gamma^{FB}{}_i$ – значение гаммы обратной связи для i-го шага (блока); $\Gamma^{T}{}_i$ – значение гаммы текста для i-го шага (блока).

Остальные обозначения соответствуют введенным ранее.

Данная схема объединяет достоинства OFB-режима (нелинейная ключезависимая обратная связь) и режима «счётчика» (гарантированный период). Однако в отличие от OFB-режима, входные значения функции шифрования недоступны криптоаналитику, а, в отличие от режима гаммирования ГОСТа, разность между элементами «задающей» последовательности неизвестна и изменяется на каждом шаге. Также возможно появление идентичных блоков гаммы текста в пределах одного периода.

Рассмотренная выше схема может быть адаптирована для выполнения одновременно с процессом шифрования функций аутентификации исходного сообщения, т.е. вычисления имитовставки. Такая модификация исходной схемы не приводит к снижению безопасности криптосистемы. Соответствующая структурная схема режима поточного шифрования с аутентификацией приведена на рис. 2.

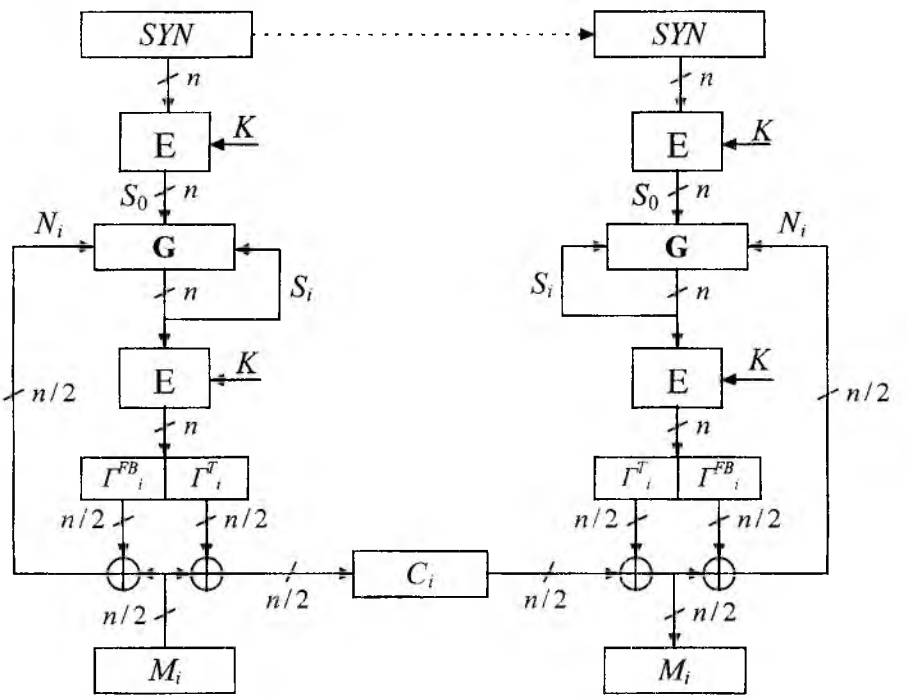


Рис. 2

Соотношения, описывающие процесс прямого и обратного преобразования для этой схемы, аналогичны соотношениям базовой схемы. Исключение составляет выражение для «шага подъёма»:

$$N_i = \{++\}_{n/2} (M_i \oplus \Gamma^{FB}{}_i).$$

В соответствии с последним соотношением «шаг подъёма» N_i является функцией от блока открытого текста M_i и криптографически «сжатого» образа текущего состояния $\Gamma^{FB}{}_i$, т.е. можно сказать, что обратная связь, аналогично CFB-режиму, представляет собой блок «внутренней» криптограммы, значение которой не выходит за пределы шифратора. Благодаря использованию не пересекающихся векторов в качестве гаммы обратной связи и гаммы текста, наличие известной пары «открытый–шифрованный» текст не позволяет определить значение обратной связи N_i. На основе указанной пары криптоаналитик может определить только поток гаммы текста ($\Gamma^{T}{}_1, \ldots, \Gamma^{T}{}_L, \Gamma^{T}{}_{L+1}$), каждый из блоков которой, при фиксированном (неизвестном) ключе K и «хорошей» функции E_K, с вероятностью $2^{-n/2}$

мог быть получен в одном из $2^{n/2}$ состояний (при $L << T_{min}$). Кроме того, в отличие от CFB-режима, данная схема свободна от недостатка «выявления коллизий», т.е. обнаружение в потоке криптограммы совпадающих блоков не увеличивает информацию криптоаналитика о неизвестных блоках гаммы шифрующей и открытого текста.

Предлагаемая схема объединяет достоинства CFB-режима и схем на базе «счетчиков». В этой схеме состояние шифратора перед обработкой очередного блока текста зависит от предыдущего состояния и обработанного на прошлом шаге блока сообщения. Иначе говоря, текущее состояние шифратора зависит от исходного вектора инициализации (синхропосылки), всех обработанных блоков сообщения и ключа шифрования. При этом влияние последнего блока текста на вектор очередного состояния носит линейный предсказуемый характер. Поэтому для получения кода аутентификации сообщения, после обработки всех блоков сообщения, необходимо выполнить ещё один дополнительный шаг «шифрования», после чего выход функции шифрования $\{\Gamma^{FB}_{L+1}, \Gamma^{T}_{L+1}\}$ либо его часть может использоваться в качестве криптографической контрольной суммы – имитовставки.

Для последней схемы понятие периода отсутствует, так как очередное состояние шифратора определяется не только текущим состоянием, но и значением обрабатываемого блока данных. Однако использование n-разрядного «плавающего счётчика» с $n/2$-разрядным «шагом подъема» позволяет гарантировать, что шифратор может оказаться в состоянии, равном данному не ранее, чем через $2^{n/2}$ шагов (блоков текста).

Безопасность обеих «усиленных» схем поточного шифрования зависит от длины блока базовой функции шифрования, определяющего минимальный период шифратора. Для современных коммерческих приложений, предъявляющих повышенные требования к уровню безопасности, можно рекомендовать использование в качестве базового алгоритма блочные шифры с длиной блока не менее 256 бит, т.е. длины блоков гаммы текста и обратной связи составляют по 128 бит, а минимальный период — $T_{min} = 2^{128}$.

Выводы

Проведенный анализ стандартных режимов применения блочных шифров показал существование у криптоаналитика потенциальной возможности атаковать непосредственно базовую функцию блочного шифрования на основе известных пар «открытый–шифрованный текст» в любом из режимов, предусмотренных международным стандартом ISO/IEC 10116 и соответствующих режимах ГОСТ 28147-89. Кроме того, в режиме ECB возможно построение атак на основе «подобранных» открытых текстов. Режим CBC также позволяет выполнять атаки такого типа, но для этого атакующему необходима возможность динамически формировать очередной блок открытого текста по значению последнего блока криптограммы. Реализация подобной атаки для CFB-режима теоретически возможна, но требует от криптоаналитика наличия «специального» доступа к аппаратуре шифрования, позволяющего определить значение блока гаммы до момента формирования обратной связи. Другим недостатком режимов ECB, CFB и CBC является возможность обнаружения коллизий (повторение блоков открытого текста либо гаммы) по шифрованному тексту, хотя вероятность возникновения таких коллизий в режимах CFB и CBC, при длине блока 128 и более бит, является сравнительно низкой. Основным недостатком режима OFB, кроме отмеченных выше, является зависимость периода гаммы шифрующей от свойств базовой функции шифрования и сложность доказательства его нижней границы. Режим «счётчика» (режим гаммирования ГОСТа) также обладает рядом недостатков: использование на входе функции шифрования «задающей» последовательности с известными дифференциальными свойствами, а также однозначное соответствие блока гаммы текущему состоянию шифратора и, как следствие, невозможность появления идентичных блоков гаммы в пределах одного периода.

Для устранения обнаруженных слабостей стандартных режимов, при построении перспективных схем поточного шифрования на базе блочных симметричных криптоалгоритмов, можно рекомендовать придерживаться приведенных выше принципов. Учитывая требование второго принципа (использовать криптографически стойкую функцию смены состояний), можно заключить, что в приложениях, предъявляющих повышенные требования к безопасности конфиденциальной информации, не следует использовать шифраторы «с произвольным доступом», т.к. принцип их построения противоречит этому требованию.

Обе приведенные в статье схемы шифраторов «с последовательным доступом» удовлетворяют указанным выше требованиям и могут рекомендоваться в качестве альтернативных режимов применения блочных шифров. Стойкость приведенных схем может быть повышена путём сокращения дли-

ны блока гаммы текста, однако «ценой» за такое «усиление» будет снижение общей производительности шифратора, что, для ряда приложений, может быть вполне приемлемо.

Список литературы: 1. *Menezes A. , P. van Oorschot, Vanstone S.* Handbook of Applied Cryptography, CRC Press, 1996. 2. *FIPS 81. «DES modes of operation».* Federal Information Processing Standards Publication 81. U.S. Department of Commerce / National Bureau of Standards, National Technical Information Service, Springfield, Virginia. 1980. 3. *ANSI X3.106, «American National Standard for Information Systems – Data Encryption Algorithm – Modes of Operation».* American National Standards Institute. 1983. 4. *ISO 8732. «Banking – Key management (wholesale)», International Organization for Standardization.* Geneva, Switzerland, 1988 (first edition). 5. *ISO/IEC 10116. «Information processing – Modes of operation for an n–bit block cipher algorithm».* International Organization for Standardization, Geneva, Switzerland, 1991 (first edition). 6. *ГОСТ 28147-89.* Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования. М.: Госстандарт СССР. 7. *Diffie W. Hellman M.E.* «Privacy and authentication: An introduction to cryptography». Proceedings of the IEEE, 67 (1979), pp. 397–427. 8. *Михелович Ш.Х.* . Теория чисел. М.: «Высшая школа», 1962. 259 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 10.04.2001

УДК 681.3.06: 519.248.681

Р. В. ОЛЕЙНИКОВ

ДИФФЕРЕНЦИАЛЬНЫЙ КРИПТОАНАЛИЗ АЛГОРИТМА ШИФРОВАНИЯ ГОСТ 28147-89

Симметричные блочные алгоритмы шифрования широко применяются в современных системах криптографической защиты информации. Первоначально они использовались исключительно для шифрования в блочном режиме, однако в настоящее время сфера их применения значительно расширилась. Блочные алгоритмы в качестве базовых примитивов используются при построении поточных шифров, криптографических генераторов псевдослучайных последовательностей и хеш-функций, а также при выработке кодов аутентификации сообщений. Учитывая высокую производительность симметричных шифров, их часто применяют вместе с несимметричными алгоритмами, получая эффективную реализацию модели взаимного недоверия и взаимной защиты. Стойкость, надежность и производительность современных систем криптографической защиты информации в значительной степени определяется характеристиками применяемых симметричных алгоритмов шифрования.

С принятием шифра DES [1] в качестве национального стандарта США в 1977г. наметился подход использовать в системах защиты исключительно стандартизированные алгоритмы. В пользу этого подхода свидетельствуют множество фактов взлома различных алгоритмов, которые не являлись стандартными и использовались без надлежащего анализа. Одним из последних примеров стал алгоритм шифрования A5/1, применяемый для обеспечения конфиденциальности в системе мобильной связи GSM. После детального анализа была обнаружена атака, позволяющая восстанавливать содержание зашифрованных сообщений на достаточно мощном персональном компьютере, хотя сам алгоритм A5/1 устойчив к силовым атакам.

Национальный институт стандартов США при проведении международного конкурса на алгоритм шифрования AES одним из главных требований считал обеспечение стойкости нового алгоритма к различным атакам. Это же условие является одним из основных в европейском проекте создания криптографических стандартов Nessie.

В нашей стране в качестве симметричного блочного шифра используется ГОСТ 28147-89. Кроме обеспечения конфиденциальности, он применяется для выработки кода аутентификации в ГОСТ 34.311-95, в свою очередь, используемом в стандарте цифровой подписи ГОСТ 34.310-95. Принципы проектирования алгоритма его разработчики оставили закрытыми, как и правила генерации долговременных ключей–одного из основных элементов, влияющих на стойкость шифра. Несмотря на более чем десятилетнюю историю, в открытой печати, как отечественной, так и зарубежной, практически отсутствуют сведения об анализе стойкости алгоритма, и нет ни одной публикации об успешной атаке всего алгоритма. В настоящей работе производится исследование стойкости ГОСТ 28147-89 к дифференциальному криптоанализу, и впервые описываются условия, при которых существует атака на алгоритм, более эффективная, чем полный перебор ключей.

Дифференциальный криптоанализ относится к классу атак с выбранными открытыми текстами. Предполагается, что криптоаналитик имеет возможность выбирать значения открытых блоков, подаваемые на вход шифратора и получать соответствующие им зашифрованные значения. При выполнении криптоанализа изучается прохождение разностей между обрабатываемыми текстами (блоками) через циклы шифрования алгоритма. Обычно для вычисления разности выбирается операция, обратная к операции введения ключа, что позволяет при расчете вероятности прохождения разности исключить влияние ключа. В классическом варианте атаки на алгоритм DES [2] используют операцию побитового сложения по модулю 2. Основным объектом атаки является нелинейное преобразование, выполняемое цикловой функцией. Шифр уязвим для дифференциального криптоанализа, если вероятность прохождения разностей через нелинейное преобразование будет достаточно высокой. Как правило, при изучении прохождения разностей, делается предположение о независимости используемых на каждом цикле шифрования подключей, что значительно упрощает анализ. Результатом атаки является вычисление подключей шифрования на последних циклах, по которым восстанавливается полный ключ шифрования.

Для дальнейшего изложения потребуется несколько определений [2]:

Разность (дифференциал) – результат некоторой операции над входными, выходными или промежуточными значениями двух параллельно шифруемых блоков. В ГОСТ 28147-89 ключ вводится с помощью операции сложения по модулю 2^{32}. Для вычисления разности возможно использование двух операций: вычитания по модулю 2^{32} или сложения по модулю 2. Атака с

использованием последней будет более простой, поэтому для вычисления разностей выбрана операцию сложения по модулю 2.

Совокупность разностей открытых и зашифрованных на одном ключе блоков (в дальнейшем обозначаемых как Ω_P и Ω_T), а также соответствующие им промежуточные значения на выходе каждого из n циклов шифрования называется ***n- цикловой характеристикой***. Каждая характеристика имеет определённую *вероятность*, в соответствии с которой при шифровании случайно выбранных открытых блоков с разностью, соответствующей характеристике, все разности при дальнейших преобразованиях, вплоть до зашифрованных блоков, будут также соответствовать характеристике. Если входная разность одной характеристики соответствует выходной разности другой, то их можно объединить. Итоговая вероятность полученной характеристики будет соответствовать произведению вероятностей исходных [2]. С точки зрения криптоанализа лучшая характеристика имеет наибольшую вероятность (но в то же время меньшую 1).

Верной парой называется совокупность открытых блоков X и X' и результат их шифрования Y и Y', разность которых, а также все промежуточные значения при шифровании соответствуют заданным характеристикой разностям.

Подстановкой (или S-блоком) в ГОСТ 28147-89 является перестановка чисел от 0 до 15 в произвольном порядке. Восемь подстановок формируют заполнение узлов замены (таблицу подстановок), используемую как долговременный ключ. Подстановка называется *активной* на i-м цикле, если в используемой характеристике на этом цикле перестановке соответствует ненулевая разность.

Таблицей распределения разностей называется математическая конструкция, описывающая дифференциальные свойства подстановки и позволяющая определить вероятность преобразования входной разности в выходную для заданного ***S***-блока. Для подстановок ГОСТа входные и выходные разности могут изменяться в диапазоне от 0h до 0Fh, а вероятность – принимать значения от 0/16 до 16/16.

Таблица восстановления входных значений – элемент, позволяющий по заданной входной и выходной разности определить все возможные значения на входе подстановки, которые могут дать заданное преобразование разностей.

Цикловая функция – основной элемент криптографического преобразования, итеративно повторяемый при шифровании. В ГОСТ 28147-89 цикловую функцию составляет сложение с ключом, подстановка и циклический сдвиг влево на 11 разрядов.

Сложностью атаки на алгоритм шифрования является количество необходимых операций шифрования для восстановления ключа при условии, что криптоаналитик обладает всей остальной необходимой для него информацией (описание алгоритма шифрования, определённое количество пар открытых и зашифрованных на искомом ключе блоков и т.д.). Сложность дифференциальной атаки зависит от вероятности лучшей характеристики. В свою очередь, вероятность характеристики зависит от числа циклов, количества активных подстановок на каждом цикле и вероятностей нужного преобразования разностей в активных ***S***-блоках. Поскольку количество циклов в ГОСТ 28147-89 равно 32, то критерием выбора лучшей характеристики будет минимальное количество активных подстановок и набольшая вероятность заданного преобразования в каждом ***S***-блоке.

Кратко остановимся на свойствах подстановок с точки зрения дифференциального криптоанализа. Наиболее вероятную характеристику будут формировать подстановки с максимальными вероятностями преобразования разностей. Поскольку долговременный ключ стандартом не определяется, то имеется возможность выбора таблицы подстановок с любыми свойствами, в том числе уязвимыми для дифференциальной атаки – с вероятностью преобразования разностей, равной 1. Пример перестановки, для которой входная разность, равная 01h, всегда преобразуется в выходную 02h, приведен в табл. 1. Приблизительную оценку количества перестановок ГОСТа, обладающих таким свойством, можно найти в [3].

Таблица 1

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
6	4	1	3	0	2	7	5	14	12	13	15	8	10	9	11

Более детальную зависимость преобразований входных разностей в выходные для приведенной подстановки можно увидеть из распределения разностей в таблице 2. В левой колонке выбирается входная разность $\Delta_{вх}$, в верхней строке – выходная $\Delta_{вых}$, а число на пересечении выбранной колонки и столбца даёт количество входных пар, соответствующих заданному преобразованию (пустая ячейка означает, что выбранное преобразование невозможно). Разделив это число на 16 (общее количество пар, формирующих разность), можно получить вероятность преобразования $p_{пр}$ выбранных разностей в ***S***-блоке.

Таблица 2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	16															
1			16													
2		4		4					8							
3		4		4		8										
4			4		12											
5			12		4											
6		8				4		4								
7				8		4		4								
8									8				4		4	
9											8		4		4	
A										4		4				8
B										4		4		8		
C									4		4				8	
D									4		4		8			
E										8				4		4
F												8		4		4

Имея таблицы распределения разностей всех 8 перестановок, можно определить вероятность преобразования произвольных входных разностей для цикловой функции ГОСТа в произвольные выходные. Рассмотрим подробнее прохождение разностей через цикл шифрования (рис. 1).

Сначала разность попадает на сумматор, где к обрабатываемому полублоку добавляется подключ. Операция введения ключа (сложение по модулю 2^{32}) является нелинейной по отношению к операции вычисления разности (сложению по модулю 2), поэтому при преобразовании возможно искажение разностей. В [2] для решения этой проблемы (при рассмотрении модификаций алгоритма DES) предлагается строить таблицу распределения разностей, при этом полагая, что значение ключа является случайной равномерно распределённой величиной. Однако это не решает проблемы, поскольку при выполнении дифференциальной атаки все блоки шифруются на одном ключе, который является постоянным.

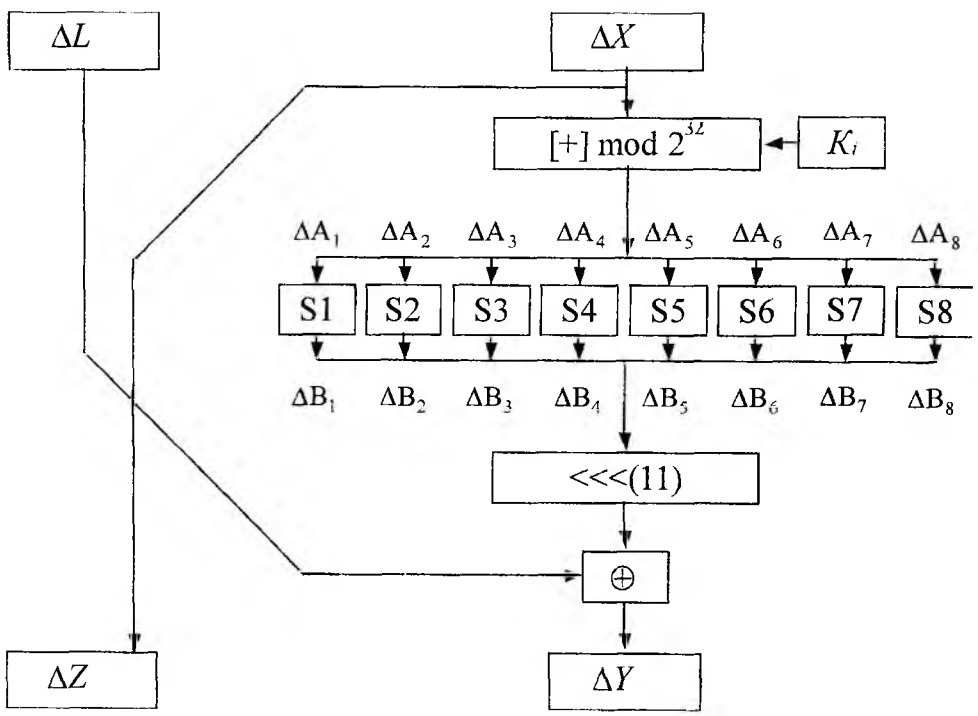


Рис.1

Более предпочтительным вариантом является построение отдельных таблиц распределения разностей для каждого из значений ключа, поскольку разные ключи дают разные вероятности преобразования. Вычислить, проанализировать и сохранить таблицу по размеру сумматора (2^{32}x2^{32}) представляется достаточно ресурсоемкой задачей, поэтому имеет смысл разбить входную 32-разрядную разность на 8 блоков по 4 бита (по размеру ***S***-блока). При таком размере вектора разности нужно построить всего 16 таблиц (для всех 4-битовых значений ключа от 0 до 0Fh размера 16x16). Пример преобразования разностей для ключа, равного 0Eh, приведен в табл. 3 (для полного анализа преобразования на сумматоре требуется, соответственно, 16 таких таблиц).

Можно отметить, что для выбранного ключа входная разность 01h всегда преобразовывается в выходную 01h (вероятность преобразования p_s=16/16=1). Эта особенность сохраняется для всех четных 4-битовых блоков подключа. Для нечетных блоков подключа это преобразование является невозможным (вероятность преобразования p_s=0/16=0). Следующей важной особенностью таблиц распределения разностей при ключевом преобразовании так же, как и при преобразовании на ***S***-блоках, является прохождение нулевой разности через сумматор без искажения. Одним из наиболее оптимальных вариантов построения характеристики будет комбинирование максимального количества нулевых 4-битовых разностей с минимальным количеством единичных с прохождением разностей через сумматор без искажения.

Таблица 3

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	16															
1		16														
2							8								8	
3								8								8
4					8								8			
5						8								8		
6			8								8					
7				8								8				
8									16							
9										16						
A							8								8	
B								8								8
C					8								8			
D						8								8		
E			8								8					
F				8								8				

В случае активной младшей подстановки на ключ шифрования накладывается дополнительно ограничение: для прохождения разностей через сумматор без искажения младший бит подключа должен быть равным нулю. Вероятность прохождения разности 01h через сумматор без искажения во всех остальных тетрадах (4-битовых блоках) равна 1/2 из-за возникновения переносов из младших разрядов в старшие при сложении с ключом по модулю 2^{32}. При этом предполагается, что нулевые и единичные биты ключа имеют равную вероятность появления, а в используемых характеристиках разности 01h всегда предшествует нулевая, что определяет возникновение синхронного переноса (одинакового в двух шифруемых блоках). Перенос в активную тетраду из предыдущей возникает с вероятностью 1/2, что эквивалентно прибавлению единицы к активным четырём битам ключа. При использовании четного ключа получается искажение разностей. Для нечетного ключа это можно представить как преобразование в четный и соответствующее прохождение разностей через сумматор без искажения. Отсюда получается вероятность преобразования, равная 1/2, для всех активных тетрад, в которые возможен перенос из младших разрядов вне зависимости от конкретных значений битов ключа.

В среднем при сложении перенос распространяется не более, чем на 2 разряда, что меньше длины анализируемых блоков. Поэтому преобразования 4-битовых разностей при добавлении 32-битового подключа можно считать независимыми событиями, и при вычислении вероятности

преобразования разности для всего сумматора следует перемножить вероятности для каждого 4-битового элемента.

Исходя из вышеизложенного, можно определить вероятность прохождения разности заданного вида (и вычисленной по модулю 2) через ключевой сумматор без искажения:

$$p_{np} = \left(\frac{1}{2}\right)^{k_a}, \tag{1}$$

где k_a – количество активных (ненулевых) тетрад в разности при условии, что младшая тетрада неактивна или в текущем цикле шифрования добавляется чётный подключ (с младшим битом, равным нулю). Если это условие не выполняется, разность обязательно будет искажена (вероятность прохождения равна нулю).

После сумматора разность попадает на ***S***-блоки. Поскольку они не определены стандартом, их можно выбрать со следующим свойством: как и для подстановки, приведенной в таблице 1, входная разность, равная 01h, всегда преобразуется в выходную 02h. В этом случае разность проходит через подстановку без изменений.

Затем в цикловой функции следует циклический сдвиг влево на 11 разрядов. Фактически это является переименованием двоичных переменных, представляющих каждый разряд разности. После сдвига активные тетрады займут другие позиции, а активная разность 02h будет преобразована в 01h. Пример прохождения разности $\Omega_{[}$ = 00101000h через один цикл шифрования приведен в таблице 4, при этом вероятность преобразования при сложении с ключом вычислена по формуле (1).

Таблица 4

Операция	Входное значение	Выходное значение	Вероятность преобразования
Сложение с подключом	00101000h	00101000h	2^{-2}
Подстановка	00101000h	00202000h	1
Сдвиг влево	00202000h	01000001h	1

Итоговая вероятность прохождения разностей через один цикл шифрования (вероятность одноцикловой характеристики) будет равна произведению вероятностей преобразования на каждом из элементов:

$$p_u = p_{sm} \cdot p_s \cdot p_r = 2^{-2}. \tag{2}$$

Отсюда следует, что, в среднем, в одном из четырёх шифрований на одном цикле будет выполняться нужное преобразование разностей.

Характеристику можно распространить на все 32 цикла шифрования путём комбинирования одноцикловых. Опять же, итоговая вероятность построенной характеристики будет равна произведению вероятностей характеристик, её составляющих:

$$p_\Omega = \prod_{i=1}^{32} p_i \,. \tag{3}$$

Размер блока в ГОСТ 28147-89 составляет 64 бита, что даёт 2^{64} возможных входных значений. Поэтому вероятность характеристики, пригодной для дифференциальной атаки, ограничена снизу значением $p_\Omega = 2^{-64}$. Если для заданного набора подстановок не существует характеристик, удовлетворяющих выбранному условию, тогда алгоритм неуязвим для дифференциального криптоанализа. Для подстановок с дифференциальными свойствами, аналогичными свойствам подстановки, приведенной в таблице 1, вероятность 32-цикловой характеристики (с младшими активными битами в тетрадах) будет находиться в пределах от $p_{\Omega_{\min}} = 2^{-154}$ (Ω_P = 0000 0000 1111 1111h и Ω_T = 1111 1111 0000 0000h) до $p_{\Omega_{\max}} = 2^{-33}$ (Ω_P=0000 0010 0100

0001h и Ω_T= 0000 0010 0101 0001h). Пример прохождения характеристики (Ω_P=00000000 00000001h, Ω_T=00100000 00010100h, $p_\Omega = 2^{-38}$) приведен в [3].

При выполнении атаки выбирается случайный открытый блок X, вычисляется соответствующий характеристике блок $X' = X \oplus \Omega_P$, после чего оба блока зашифровываются на секретном (искомом) ключе. В среднем требуется выполнить $(p_\Omega)^{-1}$ шифрований для того, чтобы полученные значения Y и Y' удовлетворяли соотношению $Y \oplus Y' = \Omega_T$. Найдя верную пару, можно получить значения входных и выходных разностей в активном ***S***-блоке на последнем цикле.

Левая половина зашифрованного блока соответствует входному значению цикловой функции на последнем цикле, поэтому известны значения X и X' для 32-го цикла (см. рис.1). $\Delta X = X \oplus X'$ можно вычислить непосредственно из зашифрованных блоков, но это значение уже известно, поскольку для верной пары характеристики известны все промежуточные разности. Предположив, что при сложении с ключом разность не исказилась, можно разделить ΔX на 8 значений по четыре бита, получив при этом входные разности ΔA_i для каждого ***S***-блока. Значение ΔL и ΔY также известно из характеристики, и из них вычисляется $\Delta B = \Delta L \oplus \Delta Y$. Разбив ΔB по ***S***-блокам, получим значения выходных разностей для каждого из них.

Как следует из табл. 2, только часть входных значений может вызвать заданное преобразование разностей. Например, входная разность $\Delta_{вх}$= 0Ah преобразуется в выходную $\Delta_{вых}$= 09h только четырьмя входными значениями. Соответственно, по известным входным и выходным разностям можно определить подмножество возможных входных значений. Удобнее всего для этого использовать таблицы восстановления входных значений. Пример для выходной разности $\Delta_{вых}$= 09h приведен в табл. 5. В левой колонке выбирается входная разность $\Delta_{вх}$, в верхней строке входное значение A_i, а число на пересечении выбранной колонки и столбца показывает вероятность заданного преобразования разностей. Пропущенные строки и пустые ячейки означают, что для выбранных значений преобразование невозможно. Для восстановления входных значений подстановки при любых выходных разностях требуется, соответственно, 16 таких таблиц.

Таблица 5

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
A					1	1									1	1
B	1	1									1	1				
E			1	1			1	1	1	1			1	1		

Получив верную пару, в соответствии с характеристикой можно определить ΔA_i и ΔB_i для активного на последнем цикле ***S***-блока (см. рис. 1). По разностям можно определить все входные значения, причем, чем меньше вероятность преобразования $\Delta A_i \rightarrow \Delta B_i$, тем меньше остаётся вариантов входных значений A_i. Поскольку $\Delta X = \Delta Z$, а также $X = Z$ и $X' = Z'$, можно вычислить варианты нескольких битов подключа шифрования на последнем цикле, соответствующие активному ***S***-блоку: $K_8^{S_i} = A_i - Z_i (\mathrm{mod}16)$. В верной паре присутствуют два шифртекста Y и Y' (соответственно два значения Z и Z'), для которых выполняется определение вероятных битов ключа шифрования. Найдя пересечение возможных значений, можно дополнительно уменьшить их количество. Для вероятности преобразования разности на ***S***-блоке, равной 2/16, остаётся одно возможное (верное) значение четырёх битов ключа для одной найденной верной пары.

Заметим, что при использовании преобразования разностей с вероятностью 1 (см. таблицу 2) все входные значения вызовут заданное преобразование, что не позволит ограничить количество возможных входных значений. Поэтому, хотя для построения характеристики используются переход с вероятностью 1, для восстановления ключей необходимо использовать искажение разностей при преобразовании на сумматоре в ходе последнего цикла шифрования. В этом случае удаётся скомбинировать характеристику с высокой вероятностью и переход с низкой вероятностью на последнем цикле, что в результате даст возможность определить несколько битов ключа со сравнительно низкими вычислительными затратами. Для переходов с вероятностью, меньшей 1,

вероятность 32-цикловых характеристик в ГОСТе будет меньше 2^{-64}, из чего следует, что с высокой степенью вероятности верной пары для таких характеристик вообще не существует.

При поиске верной пары для характеристики задаётся разность для открытых блоков, а после выполнения шифрования проверяется разность между зашифрованными блоками. В случае, когда нужно найти пару, соответствующую характеристике до предпоследнего цикла и последующим искажении на сумматоре, выполняются несколько дополнительных проверок. Значение левой половины зашифрованного блока (соответственно и разницы между ними) останется без изменения ($X = Z$, см. рис. 1). Отсюда же следует, что значение ΔL с высокой вероятностью соответствует характеристике. После сумматора изменится активная разность ΔA_i. Также возможно изменение разностей в блоках, старших по отношению к активному. Значение разности на выходе S блоков можно получить по формуле $\Delta B = (\Delta Y \oplus \Delta L) >>> 11$ (результат сложения по модулю 2 сдвигается вправо на 11 разрядов). Здесь ΔL известна из характеристики, а ΔY вычисляется по правой половине зашифрованных блоков. Из ΔB находится искомая разность ΔB_i. Разность на входе S-блока определяется искажением на ключевом сумматоре. Рассматривая различные варианты искажения разности на сумматоре, можно найти преобразование на S-блоке, дающее однозначное определение ключевых битов, причём вероятность нахождения верной пары для такой характеристики будет лишь немного меньше (порядок разности вероятностей примерно равен 2^{-3}).

Используя приведенную технику, можно восстановить 28 из 32 битов на последнем цикле шифрования. Для получения оставшихся четырёх битов необходимо использовать искажение разностей на предпоследнем цикле шифрования. Контроль соответствия характеристике обрабатываемых пар выполняется путём сравнения уже известных значений на выходе предпоследнего цикла. Они получаются после расшифрования с использованием вычисленных битов ключа шифрования.

После получения всех 32 битов подключа на последнем цикле шифрования можно выполнить расшифрование на 1 цикле и затем применить описанную методику к ГОСТу с 31 циклом для получения следующих 32 битов ключа. Продолжая атаку, возможно вычислить все 256 битов сеансового ключа. Сложность вычисления составит менее 2^{50} эквивалентных операций шифрования с использованием алгоритма ГОСТ 28147-89. Дополнительно повысить эффективность атаки позволит квартет-метод из [2].

Описанное нападение применимо лишь к "слабым" сеансовым ключам, описанным в [3]. Дополнительным условием является применение "слабых" долговременных ключей, оптимизированных для выполнения дифференциальной атаки.

Рассмотренная атака была реализована для ГОСТа с восемью циклами шифрования. Поиск подключа на последнем цикле шифрования занимал несколько минут на ЭВМ класса 486SX-33. Криптоанализ полного 32-циклового алгоритма, по нашим оценкам, займёт время порядка одного месяца при использовании вычислительной сети из 500 компьютеров класса Celeron.

Для защиты ГОСТ 28147-89 от атак дифференциального криптоанализа достаточно использование долговременных ключей, не позволяющих строить характеристики с вероятностью, превышающей 2^{-64}. Методы формирования таких подстановок предполагается рассмотреть в одной из будущих публикаций.

Список литературы: 1. *FIPS PUB 46-3. Federal Information Processing Standards Publication.* U.S. Department Of Commerce/National Institute of Standards and Technology. Data Encryption Standard (DES). 1999. 2. *Biham E., Shamir A.* Differential Cryptanalysis of the Data Encryption Standard. Springer-Verlag. New York. 1993. 3. *Долгов В.И., Лисицкая И.В., Олейников Р.В.*. "Слабые" ключи в алгоритме шифрования ГОСТ 28147-89 // Радиотехника. 2000. Вып 114. С. 63–68.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 19.03.2001

УДК 681.3.06: 519.248.681

В. И. ДОЛГОВ, д-р техн. наук, И. В. ЛИСИЦКАЯ, канд. техн. наук,
Р. В. ОЛЕЙНИКОВ, С. А. ГОЛОВАШИЧ, А. С. КОРЯК.

ДОПОЛНИТЕЛЬНЫЕ ТРЕБОВАНИЯ К ОТБОРУ ТАБЛИЦ ПОДСТАНОВОК ДЛЯ ГОСТ 28147-89

В существующих работах обосновывается методика отбора случайных подстановок и случайных таблиц подстановок в интересах построения долговременных ключей для алгоритма ГОСТ 28147-89 (в дальнейшем просто ГОСТ)[1-5]. В последующих публикациях [6-7] делается, однако, вывод, что хотя используемые в шифрах DES и ГОСТ таблицы подстановок (*S* блоков) и удовлетворяют критериям случайности [5], однако выполнение одних только показателей случайности не дает уверенности в безопасности шифра. Необходимо еще выполнить проверку защищенности шифра от известных криптоаналитических атак. Для шифров DES и ГОСТ речь, прежде всего, идет об атаках дифференциального и линейного криптоанализа [8,9]. В работе [8] показано, что случайные *S*-блоки делают шифр DES более слабым к этим атакам по сравнению с *S*-блоками разработчиков стандарта. А в работе [10] прямо указывается на то, что для окончательной уверенности в пригодности случайных таблиц подстановок в качестве долговременных ключей для шифра ГОСТ также необходима проверка устойчивости этого шифра и к отмеченным криптоаналитическим атакам.

Следует сразу подчеркнуть, что анализу стойкости шифра DES уделено очень большое внимание в печати, и с момента его появления он находится под постоянным и пристальным вниманием специалистов. Именно для этого шифра и были впервые разработаны эффективные принципы дифференциального, а в дальнейшем и линейного криптоанализа. Наибольшее различие между шифрами DES и ГОСТ автор работы [11] видит в том, что в цикловой функции ГОСТа используется циклический сдвиг вместо перестановки. Он отмечает, что в DES перестановка приводит к существенному увеличению лавинного эффекта. Если в шифре DES для того, чтобы изменение одного входного бита повлияло на каждый выходной бит, достаточно 5 циклов, то в ГОСТе для этого необходимо уже 8-9 циклов. Однако разработчикам ГОСТа удалось добиться некоторого повышения его надежности или повышения трудности его криптоанализа, о чем по всей вероятности и свидетельствует весьма ограниченное число посвященных этому шифру публикаций. Этому, конечно, в значительной степени способствует также то, что таблица подстановок в шифре ГОСТ является конфиденциальным параметром и может меняться по желанию пользователя, что создает дополнительные трудности для криптоанализа.

В этой работе мы все же хотим высказать некоторые соображения, связанные с анализом устойчивости шифра ГОСТ к атакам дифференциального и линейного криптоанализа, которые позволяют обосновать дополнительные ограничения к отбору *S*-блоков и придать методике отбора случайных таблиц подстановок для шифра ГОСТ 28147-89 более завершенный вид.

Как известно, для шифра DES операцией, позволяющей исключить при вычислении дифференциалов влияние ключевых бит, является операция суммирования по модулю 2 (XOR). Именно с помощью операции XOR в этом алгоритме в цикловую функцию вводятся биты подключа. В ГОСТе ключевые биты вводятся в цикловую функцию путем суммирования по модулю 2^{32}, что при вычислении разностей вызывает определенные проблемы, связанные с необходимостью учета переносов разрядов. Поэтому первые наши предложения по построению атак дифференциального криптоанализа ГОСТ [12] были предприняты при использовании разностей, вычисляемых опять-таки с помощью операции XOR.

Переходя к построению дифференциальной характеристики, полезно напомнить о выводе работы [12] о том, что при использовании для формирования разностей операции суммирования по модулю 2 для ГОСТа существуют "слабые" подстановки, которые позволяют построить характеристики для этого шифра, имеющие высокие вероятности. Основу этих "слабых" *S*-блоков составляют специфические подстановки. Они обладают тем свойством, что для них вероятность перехода фиксированной входной разности в фиксированную выходную разность равна 1.

И в наших исследованиях мы остановились на использовании подстановок, обладающих отмеченным свойством. Однако мы здесь приведем пример построения дифференциальной характеристики, отличающейся от работы [12], которая примечательна тем, что имеет более высокую вероятность. Она включает в себя две шестицикловые характеристики, представленные на рис. 1 и 2.

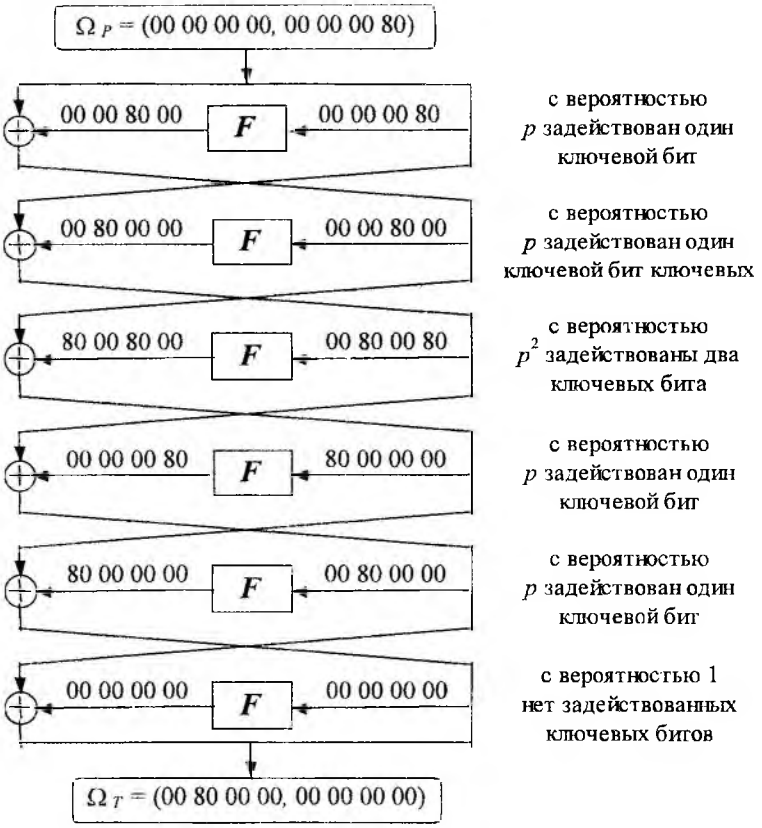

Ω_P = (00 00 00 00, 00 00 00 80)
00 00 80 00
F
00 00 00 80
с вероятностью p задействован один ключевой бит
00 80 00 00
F
00 00 80 00
с вероятностью p задействован один ключевой бит ключевых
80 00 80 00
F
00 80 00 80
с вероятностью p^2 задействованы два ключевых бита
00 00 00 80
F
80 00 00 00
с вероятностью p задействован один ключевой бит
80 00 00 00
F
00 80 00 00
с вероятностью p задействован один ключевой бит
00 00 00 00
F
00 00 00 00
с вероятностью 1 нет задействованных ключевых битов
Ω_T = (00 80 00 00, 00 00 00 00)

Рис.1

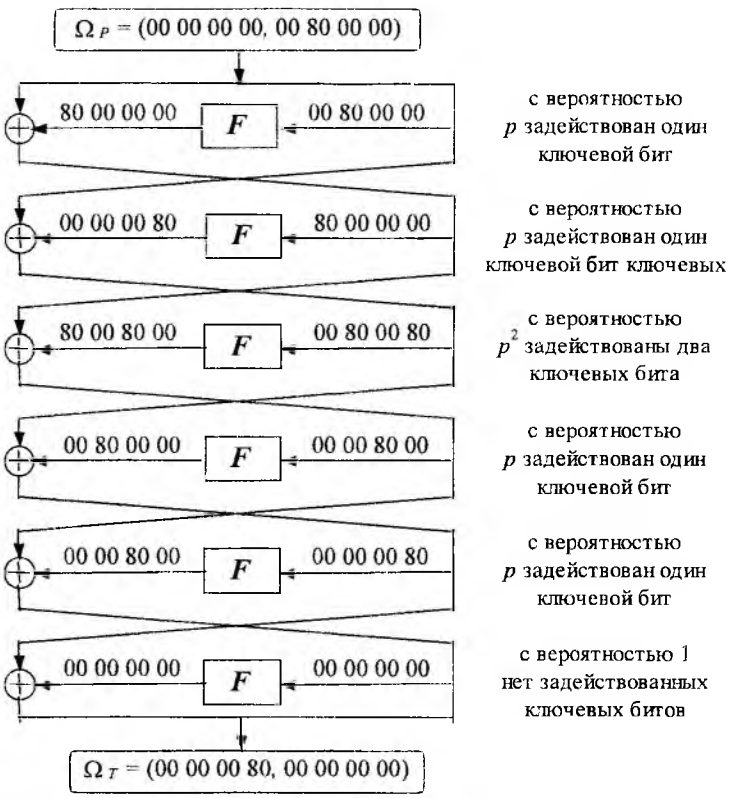

Ω_P = (00 00 00 00, 00 80 00 00)
80 00 00 00
F
00 80 00 00
с вероятностью p задействован один ключевой бит
00 00 00 80
F
80 00 00 00
с вероятностью p задействован один ключевой бит ключевых
80 00 80 00
F
00 80 00 80
с вероятностью p^2 задействованы два ключевых бита
00 80 00 00
F
00 00 80 00
с вероятностью p задействован один ключевой бит
00 00 80 00
F
00 00 00 80
с вероятностью p задействован один ключевой бит
00 00 00 00
F
00 00 00 00
с вероятностью 1 нет задействованных ключевых битов
Ω_T = (00 00 00 80, 00 00 00 00)

Рис.2

При построении этой дифференциальной характеристики использованы таблицы подстановок с переходами фиксированной разности 8h (1000_2) в 1h (0001_2). Всего в 32-цикловой характеристике, получающейся при итеративном продолжении шестицикловых характеристик (рис.1 и 2), используется 31 активный *S*-блок. Важной особенностью этой характеристики следует считать то, что из восьми *S*-блоков в ее построении участвуют только 4, а именно 2-й, 4-й, 6-й и 8-й (четные *S*-блоки), т.е. отмеченное выше ограничение на переходные характеристики касается только четырех из восьми *S*-блоков. На нечетные *S* блоки никаких ограничений не накладывается. Однако при этом возможно восстановить лишь биты ключа, соответствующие активным подстановкам на последних циклах, а не весь 256-битовый сеансовый ключ.

Приведем соображения по оценке возможностей реализации атак с применением указанной характеристики. Нас будет интересовать вероятность получения на входе *S*-блока разностей, не зависящих от битов ключа.

Рассмотрим тетрады полублоков, формирующие фиксированные входные разности $\Delta = 8\text{h}$. Всего возможно 16 значений таких разностей: 0-8, 1-9, 2-10, 3-11, 4-12, 5-13, 6-14, 7-15, 8-0, 9-1, 10-2, 11-3, 12-4, 13-5, 14-6, 15-7 (8 значений в прямом и 8 этих же значений в обратном сочетании). При суммировании тетрад, формирующих фиксированные разности, с соответствующими ключевыми битами возможны три ситуации:

1) переносов разрядов нет ни для одного из слагаемых;
2) имеется перенос только для одного из слагаемых;
3) оба четырехбитных слагаемых имеют переносы разрядов в следующую тетраду.

Заметим, что здесь, как и ранее, идентичные переносы разрядов в каждом из четырехбитных слагаемых не мешают построению интересующей нас характеристики, ей будет препятствовать только ситуация 2. Для того, чтобы оценить вероятность прохождения разности через сумматор по модулю 2^{32} без искажения, рассмотрим ситуации, которые могут возникнуть для различных пар входов. Они представлены вместе с показателями прохода через сумматор в табл. 1.

С учетом того, что значения ключевых тетрад равновероятны, для среднего значения числа входных разностей, которые проходят через сумматор по модулю 2^{32} без искажения, можно получить результат

$$N_{пр} = 1 + \frac{1}{16}(2 + 4 + 6 + 8 + 10 + 12 + 14) = 7.$$

В этой формуле дробь 1/16 перед скобкой есть вероятность выбора одного из равновероятных значений ключевого слова. Поэтому результирующая вероятность прохода через сумматор входной разности $\Delta = 8h$ без изменений равна $p_{пр} = 1/2$.

Заметим, что переносы, возникающие при прохождении через сумматор по модулю 2^{32} в предшествующей паре четырехбитных слов (а они могут быть только двойными), легко учитываются увеличением значения четырехбитного ключа для текущей пары слов на единицу. Это значит, что полученный выше результат не меняется. Кроме того, для полублоков, отличающихся 32-м битом, единица переноса уходит за границы 32-битного блока, и, следовательно, вероятности прохода соответствующих разностей ($\Delta = 8h$) через сумматор по модулю 2^{32} без переносов равны 1.

Таблица 1

Значение ключевых бит	Благоприятные сочетания пар выходов (от → до)	Число благоприятных исходов
0	0 - 8 → 7 - 15	8 (переносов нет)
1	0 - 8 → 6 - 14	7 (переносов нет)
2	0 - 8 → 5 - 13	6 (переносов нет)
3	0 - 8 →4 - 12	5 (переносов нет)
4	0 - 8 → 3 - 11	4 (переносов нет)
5	0 - 8 → 2 - 10	3 (переносов нет)
6	0 - 8 → 1 - 9	2 (переносов нет)
7	0 - 8 → 0 - 8	1 (переносов нет)

Продолжение табл. 1

Значение ключевых бит	Благоприятные сочетания пар выходов (от → до)	Число благоприятных исходов
8	0	0
9	7 - 15 → 7 - 15	1 (двойной перенос)
10	6 - 13 → 7 - 15	2 (двойной перенос)
11	5 - 12 → 7 - 15	3 (двойной перенос)
12	4 - 11 → 7 - 15	4 (двойной перенос)
13	3 - 10 → 7 - 15	5 (двойной перенос)
14	2 - 9 → 7 - 15	6 (двойной перенос)
15	1 - 8 → 7 - 15	7 (двойной перенос)

Все представленные выше результаты получены при использовании случайного набора пар шифруемых текстов и полного перебора по всему множеству возможных ключей. Для вероятности 32-цикловой характеристики, использующей переходы рис. 1 и 2, с учетом того, что из 32-х активных *S*-блоков, задействованных в ней, шесть имеют вероятности прохода входной разности через сумматор по модулю 2^{32} без изменения, равные 1, а 26 – равные $p_{пр} = 1/2$, приходим к результату:

$$P_{32} = \left(p_{пр}\right)^{26} \cdot p^{32} = 2^{-26} \cdot p^{32}. \qquad (1)$$

Заметим, что при получении этого и предыдущих результатов использовано предположение о взаимной независимости ключей на различных циклах шифрования, в то время как в шифре ГОСТ используется восемь сеансовых 32-битных подключей. При шифровании они циклически повторяются, причем, на последних восьми циклах порядок использования подключей меняется на обратный. Распределение подключей по циклам шифрования в привязке к дифференциальной характеристике, строящейся с использованием переходов рис. 1 и 2, представлено в табл. 2. В ячейках таблицы указаны ненулевые биты разностей пар правых полублоков, участвующих в построении этой характеристики.

Из таблицы следует, что преобразование выполняется с вполне определенным набором цикловых побитных разностей. Для определения вероятности благоприятного прохода соответствующих разностей через k активных *S*-блоков с переносами разрядов через циклы с одним и тем же ключом можно записать выражение:

$$P_k = \frac{1}{16} \cdot \left(1 + 2 \cdot \left(\left(\frac{14}{16}\right)^k + \left(\frac{12}{16}\right)^k + \left(\frac{10}{16}\right)^k + \left(\frac{8}{16}\right)^k + \left(\frac{6}{16}\right)^k + \left(\frac{4}{16}\right)^k + \left(\frac{2}{16}\right)^k\right)\right).$$

Отметим здесь, что переносы из 32-го бита "уходят" за пределы полублока и их можно игнорировать.

В итоге, с учетом распределения активных *S*-блоков с переносами по подключам, представленного в табл. 2, для вероятности 32-цикловой дифференциальной характеристики получаем результат

$$p_{пр} = \left(P_4\right)^4 \cdot P_5 \cdot P_3 \cdot \left(P_1\right)^2 = 6{,}28 \cdot 10^{-7} \approx 2^{-20}, \qquad (2)$$

что подтверждается экспериментально полученным $p_{пр} = 2 \cdot 10^{-7}$.

Таблица 2

Подключи K_i	K_1	K_2	K_3	K_4	K_5	K_6	K_7	K_8
Биты 1 - 8 циклов	8	16	8,24	32	24	–	24	32
Биты 9 - 16 циклов	8,24	16	8	–	8	16	8,24	32
Биты 17 - 24 циклов	24	–	24	32	8,24	16	8	–
Биты 32 - 25 циклов	32	24	–	24	32	8,24	16	8
Число активных *S* блоков с переносами	4	3	4	1	4	4	5	1

Как видно из результатов (1) и (2), наличие детерминированной связи в используемых ключах приводит к заметному увеличению вероятности получения приемлемой для атаки дифференциальной характеристики.

Таким образом, и для шифра ГОСТ 28147-89 при выполнении описанных условий открываются возможности для выполнения криптоанализа на основе атаки с выбранными открытыми текстами. Все рассмотренные атаки строятся на том, что существуют подстановки, для которых фиксированная входная разность с большой вероятностью переходит в фиксированную выходную (более того, как мы увидим из дальнейшего, существует значительное число подстановок, для которых этот переход осуществляется с вероятностью 1). Будем, как и ранее [12], такие подстановки называть "слабыми".

Очевидно, что в этих условиях значительный практический интерес приобретают ответы на два принципиальных вопроса:

1) обеспечивают ли предлагаемые в [5] критерии отбора подстановок и таблиц подстановок фильтрацию "слабых" подстановок и таблиц подстановок?

2) если же это не так, то какие нужно ввести дополнительные ограничения на отбор подстановок и таблиц подстановок, чтобы исключить "слабые" ключи?

Для ответа на эти вопросы и были предприняты исследования, направленные на изучение особенностей формирования и оценки характеристик множества подстановок и таблиц подстановок, являющиеся "слабыми" в рассматриваемом смысле.

Отметим в связи с этим результаты еще одной из редких работ [13], посвященных анализу шифра ГОСТ 28147-89. В этой работе при оценке устойчивости шифра ГОСТ к атакам дифференциального криптоанализа авторы обнаружили подстановки, для которых ненулевой входной дифференциал ΔX приводит к ненулевому выходному значению ΔY с вероятностью 1. К примеру, в подстановке $\pi = (15,13,9,11,7,14,10,3,2,12,8,6,0,5,1,4)$, все входы с дифференциалом $\Delta X = 3 = 0011$ приводят к выходному дифференциалу $\Delta Y = 4 = 0100$. Всего, как отмечают авторы [13], было найдено 1096 перестановок с таким свойством.

Выполненные нами теоретические расчеты позволили получить выражение для общего числа подстановок, для которых существует переход фиксированной входной разности в фиксированную выходную разность с вероятностью 1, в виде:

$$N_1 = 15^2 \cdot 8!2^8 = 2{,}32 \cdot 10^9.$$

Для вероятности P_{N_1} случайного выбора подстановки, удовлетворяющей рассматриваемому ограничению, соответственно получаем результат

$$P_{N_1} = \left(\frac{2 \cdot 10^{13}}{2{,}33 \cdot 10^9} \right)^{-1} = (9009)^{-1}.$$

Это означает, что на каждые 9000 случайных подстановок встречается одна подстановка с переходом фиксированной входной разности в фиксированную выходную разность с вероятностью 1. Полученный результат, очевидно, хорошо согласуется с данными статистического эксперимента. Более того, как показывает анализ и эксперименты, существуют подстановки, которые обладают отмеченным свойством одновременно для нескольких вариантов сочетаний входных и выходных разностей.

В результате существует вполне реальная возможность построения таблиц подстановок с одним и тем же фиксированным переходом входной разности в выходную. Так, если считать, что всего существует $N_1' = 8!2^8 = \cdot 10^7$ подстановок с одним и тем же фиксированным переходом входной разности в выходную, то из этих подстановок можно сформировать $\sim \left(10^7\right)^8 = 10^{56}$ различных таблиц подстановок полностью специфического типа.

Заметим также, что общее число подстановок, имеющих переход фиксированной разности $\Delta X = 8h$ в $\Delta Y = 1h$, определяется формулой

$$N_1 = 8! \cdot 2^8 = 10^7.$$

Наконец, отметим дополнительно, что для получения "слабой" таблицы подстановок в рассмотренном примере из восьми подстановок только четыре должны иметь специфический вид (четные *S* блоки).

Наш анализ показал, что из всех случайных подстановок, проходящих критерии, приведенные в [5], примерно половина оказались слабыми.

Изложенные соображения и результаты позволяют сформулировать дополнительные требования к отбору *S*-блоков, выполнение которых, на наш взгляд, позволит гарантировать устойчивость шифра ГОСТ 28147-89 к атакам дифференциального криптоанализа.

Требование 1. Для защиты от атак дифференциального криптоанализа максимально допустимое значение таблиц распределения разностей *S*-блоков для ненулевых входов не должно превышать значения 8, при этом не менее, чем 99% ячеек таблицы дифференциальных разностей *S* блоков не должны превышать значения, равного 4.

При этих ограничениях даже в том случае, если встречается таблица, которая имеет идентичные для всех подстановок переходы фиксированных входных разностей в выходные, равные максимально допустимым (8), для вероятности ранее рассмотренной дифференциальной характеристики, составленной из 31 активного *S* блока, получим оценочное значение:

$$2^{-21}\cdot\left(\frac{8}{16}\right)^{31}=2^{-53}.$$

Более высокий уровень защищенности можно обеспечить, если предъявить дополнительное ограничение на подстановки, из которых составляется таблица, по степени подобия соответствующих им таблиц распределения разностей. Такое ограничение может быть, например, представлено в виде дополнительного требования.

Требование 2. Подстановки, попавшие в таблицу, при выполнении требования 1, не должны иметь сколько-нибудь существенных идентичных переходов (переходов одного типа) входной разности в выходную, или в более конкретном выражении – *S*-блоки не должны иметь более двух максимально возможных значений переходов одного типа.

Тогда, полагая, что два из четырёх активных *S*-блоков, участвующих в формировании рассматриваемой характеристики, дают максимальное значение вероятности прохода *S*-блока с заданными (фиксированными) разностями (например, если считать, что максимальное значение в табл. 3 дают второй и шестой *S*-блоки), получим: $2^{-21}\cdot\left(\frac{8}{16}\right)^{8}\cdot\left(\frac{4}{16}\right)^{23}=2^{-73}$.

Этот результат уже представляется существенно более надежным.

Во всех представленных выше оценках определяются значения вероятностей дифференциальных характеристик путем статистического усреднения по всему возможному множеству значений ключей. Естественно, что эффективность атаки зависит от значения неизвестного ключа. Пределы, в которых меняется успех атаки с характеристикой, составленной из шестицикловых характеристик рис.1, рис.2, определяются границами изменения первого множителя в (1). Он может быть равным 1 при ключевых битах $K_h = 0000$ (с нулевыми значениями позиций) четырехбитных входов четных *S*-блоков и равным 0 при $K_h = 8h = 1000_2$ (K_h – четырехбитный сегмент ключа, взаимодействующий с ненулевой разностью $\Delta = 8h$).

Что касается линейного криптоанализа, то, как уже указывалось выше, в ГОСТе биты подключа вводятся на каждом цикле с помощью операции суммирования по модулю 2^{32}. Эта операция выполняется с переносом разрядов, что не позволяет построить линейные соотношения, зависящие только от битов подключей (не удается избавиться от зависимости результирующего линейного соотношения от битов шифруемого текста), т.е. биты ключа не могут быть объединены непосредственно. Эта же мысль проводится в [13].

Отмеченное позволяет утверждать, что атаки линейного криптоанализа (по крайней мере, в том варианте, который предлагается автором линейного криптоанализа М. Мацуи) для шифра ГОСТ неприменимы.

В результате приведенные выше дополнительные ограничения к отбору случайных *S*-блоков позволят, на наш взгляд, сформировать таблицы подстановок (долговременные ключи) для шифра

ГОСТ 28147-89, которые обеспечат стойкость этого алгоритма от рассмотренных криптоаналитических атак.

Остается заметить, что известные нам два варианта таблиц подстановок, построенных разработчиками алгоритма ГОСТ 28147-89, полностью удовлетворяют предложенным выше критериям отбора.

Список литературы: 1. *Горбенко И.Д., Лисицкая И.В.* Критерии отбора случайных таблиц подстановок для алгоритма шифрования по ГОСТ 28147-89 // Радиотехника. 1997. Вып 103. С. 121–130. 2. *Лисицкая И.В.* К вопросу построения долговременных ключей для алгоритма ГОСТ 28147-89 // Информационно-управляющие системы на железнодорожном транспорте. 1997. № 3. С. 54–57. 3. *Бильчук В.М., Лисицкая И.В.* К вопросу построения долговременных ключей для алгоритма ГОСТ 28147-89. Критерии отбора второго уровня // Информационно-управляющие системы на железнодорожном транспорте. 1998. № 1. С. 10–17. 4. *Горбенко И.Д., Лисицкая И.В., Коряк А.С.* Анализ стойкости алгоритма ГОСТ 28147-89 при использовании подстановок случайного типа // Радиоэлектроника и информатика. 1998. №1 (02). С. 39–43. 5. *Лисицкая И.В., Коряк А.С.* Уточненные критерии отбора таблиц подстановок с заданными критериями случайности. Радиотехника. 2000. Вып 114. С. 47–56. 6. *Лисицкая И.В., Головашич С.А., Олейников Р.В.* Построение таблиц подстановок для стандарта шифрования данных // Проблемы бионики. 1999. Вып 50. С. 185-194. 7. *Лисицкая, И.В., Олейников Р.В., Головашич С.А.* Анализ стойкости DES - подобных алгоритмов шифрования при использовании таблиц подстановок случайного типа. // Радиоэлектроника и информатика. 1999. № 1. С. 109-115. 8. *Biham E., Shamir A.* Differential Cryptanalysis of the full 16-round DES. Technical Report - Computer Science Department, Technion, Israiel, 1993. 9. *Matsui M.* Linear Cryptanalysis Method for the DES Cipher. // Proc. of Eurocrypt'93, Norway, 1993. 10. *Фаль А.М.* Алгоритм шифрования по ГОСТу 28147-89 и способы применения блочных шифров // Безопасность информации. 1995. №3. С. 8-11. 11. *Schneier B.* Applied Cryptography. Second Edition: protocols, algorithms, and Source code in C. Published by John Wiley & SonS. Inc, New York: ChicheSter BriSbane Toronto Singapore, 1996 – 758 p. 12. *Долгов В.И., Лисицкая И.В., Олейников Р.В.* "Слабые" ключи в алгоритме шифрования ГОСТ 28147-89 // Радиотехника. 2000. Вып 114. С. 63–68. 13. *C. Charnes, L. O'Connor, J. Pieprzyk, R. Safavi-Naini, Y. Zheng.* Further Comments on the Soviet Encryption Algorithm // Wollongong, NSW 2500, AUSTRALIA 1994. pp. 1-10.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 19.03.2001

УДК 681.3.06: 519.248.681

И. В. ЛИСИЦКАЯ, Т. В. ЦЕПУРИТ, В. В. ЛЕСНЯК, М. В. ПИНЧУК, А. П. МЕЛЕЦКИЙ

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ МОДЕРНИЗАЦИИ ШИФРА ГОСТ 28147-89 С ЦЕЛЬЮ ДАЛЬНЕЙШЕГО ПОВЫШЕНИЯ ЕГО БЕЗОПАСНОСТИ

Современный этап развития криптографии характеризуется появлением новых требований к системам шифрования. К наиболее актуальным можно отнести требование к увеличению длины шифруемого блока и требование обеспечения стойкости алгоритма к известным криптоаналитическим атакам. Это приводит к необходимости разработки новых алгоритмов, что само по себе является очень сложным и дорогостоящим. В данной работе была сделана попытка не разработки нового, а модернизации существующего и хорошо зарекомендовавшего себя алгоритма с целью приведения его в соответствие к современным требованиям.

Одной из наиболее универсальных и мощных криптоаналитических атак на симметричные системы шифрования в настоящее время является дифференциальный криптоанализ. Он был первым успешным криптонападением на американский стандарт DES, который до этого более 15 лет считался неуязвимым. Поэтому эта атака обязательно учитывается при оценке стойкости любой современной симметричной системы шифрования.

Напомним основные положения дифференциального криптоанализа [1]. Атакующий имеет возможность управлять разностями пар открытых (незашифрованных) блоков на входе шифратора и имеет доступ к его выходу. Для уязвимых алгоритмов существуют разности между парами открытых текстов, которые проходят через все циклы алгоритма шифрования с вероятностью, превышающей пороговую. Далее, зная входные и выходные значения открытых и зашифрованных текстов, криптоаналитик имеет возможность получить наиболее вероятные значения ключа шифрования. Успех атаки зависит от вероятности нахождения пары открытых текстов, разность которых приводит к специфической разности шифртекстов.

Для DES-подобных шифров (к числу которых относится и ГОСТ 28147-89) устойчивость к дифференциальному криптоанализу в значительной мере определяется свойствами применяемых при их построении нелинейных преобразований. В шифрах DES и ГОСТ 28147-89 нелинейными преобразованиями являются таблицы подстановок (так называемые S-блоки). Именно на основе анализа свойств S-блоков была предложена методика определения ключей для нескольких DES-подобных шифров со сложностью, меньшей, чем прямой перебор. Отечественный стандарт ГОСТ 28147-89 введен в действие гораздо позже DES. Несмотря на то, что и в ГОСТе нелинейным преобразованием, как и в DES, является подстановка, тем не менее, в открытой литературе практически нет публикаций, посвященных изучению его стойкости к различным атакам. Предполагается, что за счет использования вдвое большого числа циклов, чем DES, ГОСТ обладает более высокой защищенностью от многих известных криптоаналитических атак. Однако, в последнее время появились публикации, в которых идеи дифференциального криптоанализа успешно применены к шифру ГОСТ 28147-89 и доказано существование в этом алгоритме определённых слабостей. [2]

В нашей работе сначала предпринимаются шаги, направленные на дальнейшее повышение стойкости ГОСТ 28147-89 к атакам дифференциального криптоанализа, а затем рассматривается возможность модернизации этого алгоритма путем увеличения вдвое длины шифруемого блока. Для решения первой задачи предлагается заменить в каждом цикле стандартной процедуры криптопреобразований детерминированный сдвиг на 11 разрядов на нелинейную операцию параметрического (управляемого) циклического сдвига.

В рассматриваемой в работе реализации для определения параметра сдвига задействуются пять определенных битов шифруемого полублока, взятые после прохождения им таблицы подстановок. Эти пять битов задают текущее значение сдвига полублока на выходе цикловой функции в каждом цикле преобразований (в пределах от нуля до тридцати одного разрядов). Так как значение сдвига зависит от ключа и шифруемых данных, то он является случайной величиной.

На рис. 1 приведены результаты экспериментов по определению закона распределения значений случайного сдвига, задаваемого для обеспечения максимального быстродействия пятью младшими разрядами.

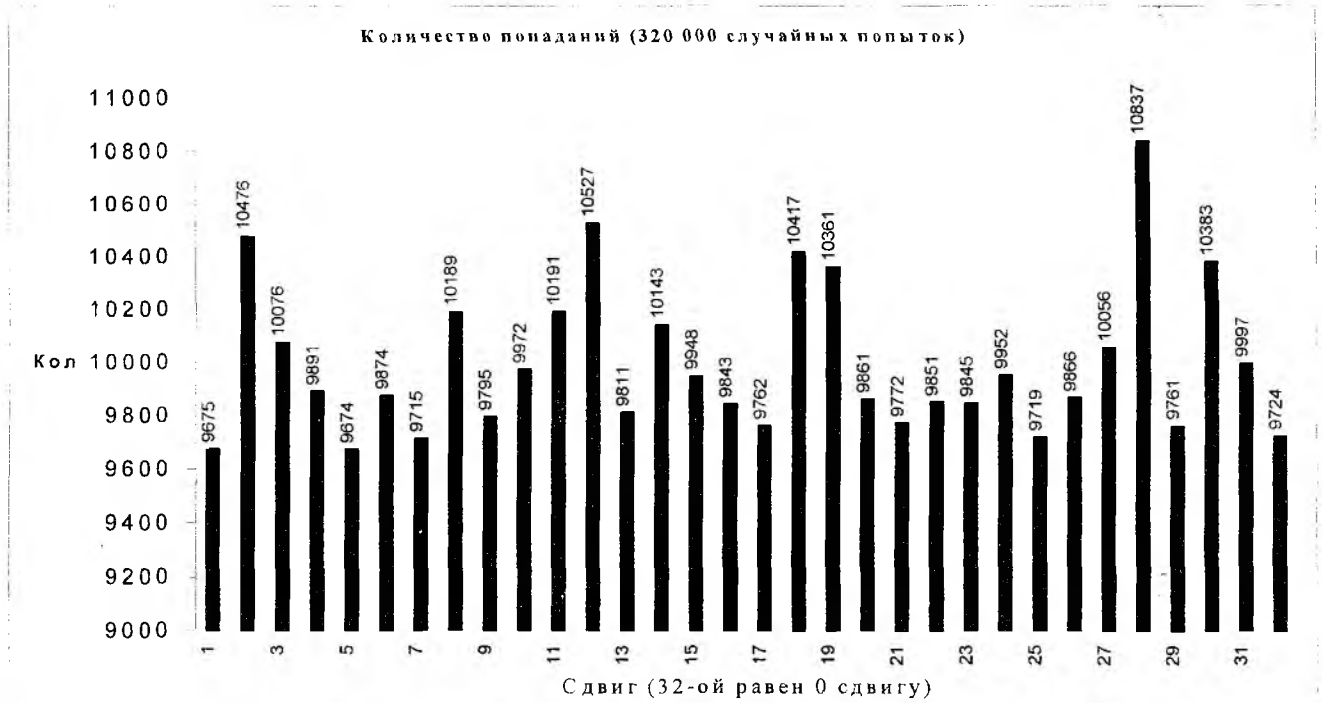


Рис. 1

На основе анализа приведенных данных сделан вывод, что закон распределения можно действительно считать достаточно близким к равномерному. Тогда, если считать, что каждое из значений сдвигов на отдельном цикле процедуры шифрования появляется равновероятно, для вероятности совпадения идентичных значений сдвигов на всех тридцати двух циклах можно получить оценку $(2^{-5})^{-32} = 2^{-160}$. Это позволяет сделать атаку дифференциального криптоанализа на модернизированный ГОСТ неэффективной даже при "благоприятных" ситуациях [2].

Проведение статистических испытаний являются, как известно, единственной стратегией испытаний больших криптографических систем с секретными ключами, построенных в виде чередующихся слоёв блоков замен и перестановок. Поэтому на следующем этапе была выполнена оценка показателей статистической безопасности стандартного и модернизированного алгоритмов.

За основу взяты три показателя стойкости (статистической безопасности [3]), которые принято сейчас использовать при проверке многоцикловых процедур современных блочных систем шифрования:

1. Число циклов алгоритма, начиная с которого две криптограммы, полученные шифрованием двух, отличающихся на один бит блоков данных (открытых текстов), становятся устойчиво независимыми (в том смысле, что при большем числе циклов они остаются независимыми). Другими словами, необходимо определить число циклов шифрования алгоритма, начиная с которого обеспечивается влияние любого (одного) входного бита, на каждый выходной бит - это, так называемый лавинный эффект .

2. Число циклов шифрования, при котором один и тот же открытый текст, зашифрованный на ключах, отличающихся одним битом, порождает устойчиво независимые (некоррелированные) криптограммы.

3. Коэффициент сжатия шифрованного текста при применении процедуры архивирования Лемпела-Зива, характеризующий степень его случайности.

На рис. 2-3 приведены результаты оценки глубины входа в модифицированный алгоритм при изменении одного бита сообщения для случаев выполнения процедуры шифрования на одном из долговременных ключей, сформированных разработчиками стандарта, когда сеансовый ключ случайный и когда сеансовый ключ нулевой.

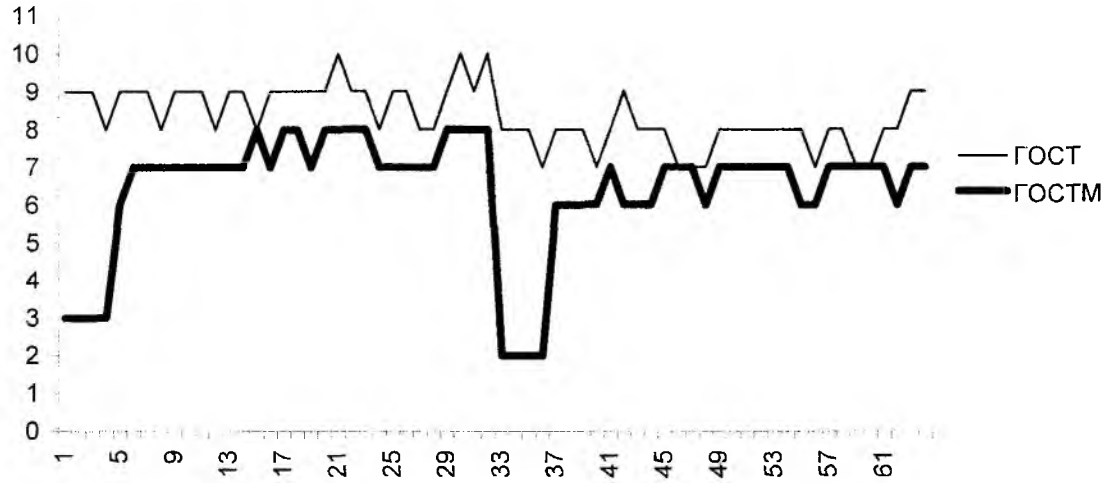


Рис. 2

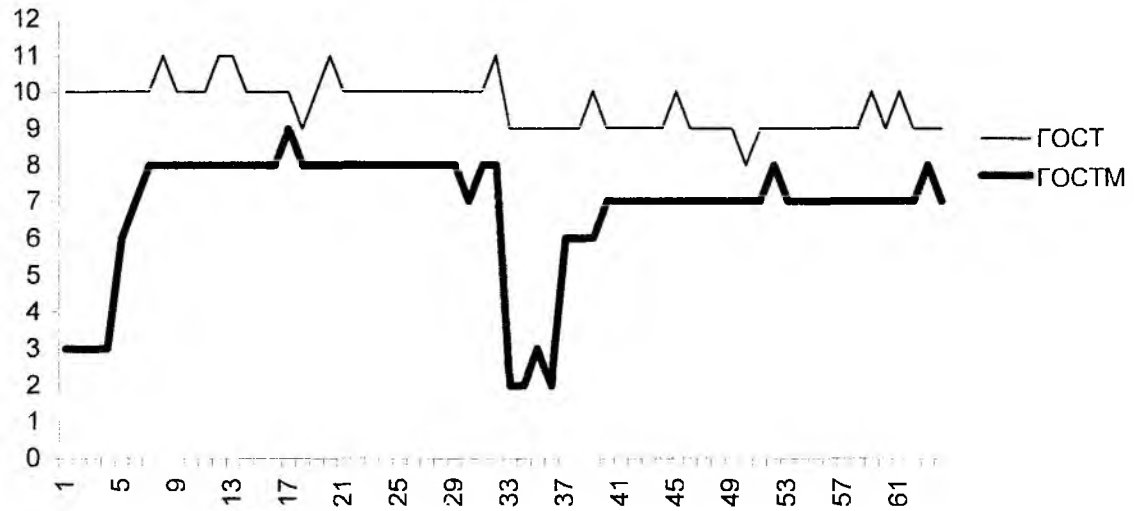


Рис. 3

Глубина входа в алгоритм при экспериментах определялась номером цикла, начиная с которого математическое ожидание числа единичных (ненулевых) бит m_w зашифрованного блока данных удовлетворяло условию $32-0{,}4 \le m_w \le 32+0{,}4$ [3].

На основе полученных данных сделан вывод, что модернизированная процедура построения шифра при использовании случайных таблиц подстановок обладает показателями лавинного эффекта, не уступающими стандартной. Глубина вхождения в алгоритм при использовании ненулевого сеансового ключа $(\overline{K} \neq 0)$, обеспечивающая зависимость всех выходных бит от любого входного бита, в стандартном ГОСТе равна 8-9 циклам. При нулевом сеансовом ключе $(\overline{K} = 0)$ появляется ещё один дополнительный "этаж". Для модернизированного ГОСТа, построенного с применением процедуры параметрического сдвига, лавинный эффект наступает на 2-8 циклах при ненулевом сеансовом ключе $(\overline{K} \neq 0)$ и на 2-9 при нулевом сеансовом ключе $(\overline{K} = 0)$.

Результаты оценки числа циклов алгоритма, при котором наступает статистическая независимость выходных (шифрованных) текстов при шифровании сообщений с помощью ключей $\vec{K}$, отличающихся одним битом, проиллюстрированы на рис 4.

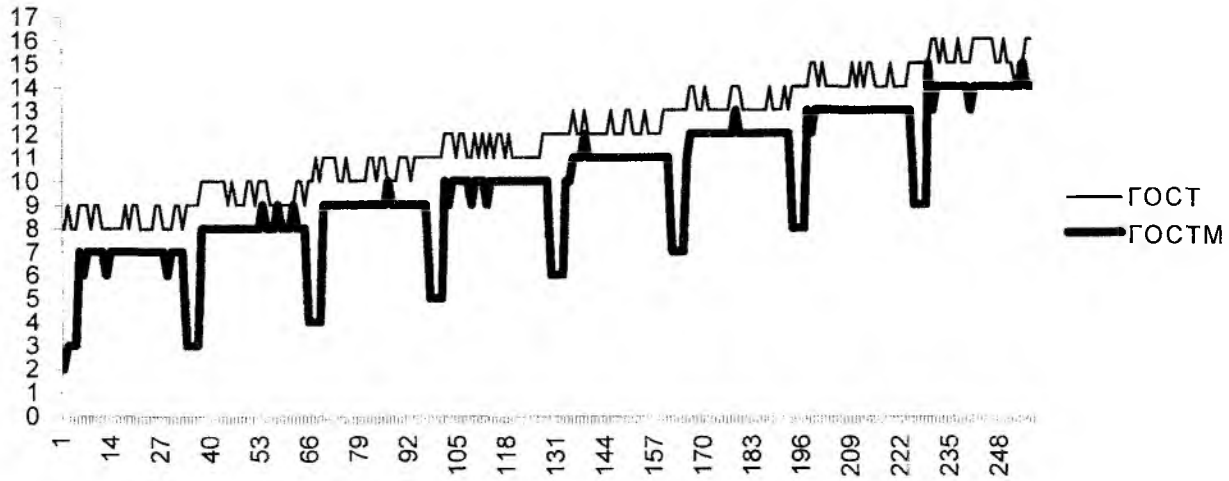


Рис. 4

Как и в предыдущем случае, экспериментальные данные показали улучшение показателей статистической безопасности при применении в алгоритме операции параметрического сдвига (8-16 циклов для ГОСТ 28147-89 и 2 –15 циклов для модернизированного алгоритма).

Таблица 1

	Исходный текст	ГОСТ	ГОСТМ
EXE	49,78	10,09	10,09
WORD	82,42	60,06	60,06
TXT	53,68	1,32	1,32

Проверка степени сжатия текстов проводилась с помощью процедуры Лемпела-Зива на ключах и подстановках случайного типа. Испытывались три варианта текстов: exe-файл, как текст усредненного типа, word-овский файл, как пример очень избыточного (с повторения-ми) текста, и обыкновенный текстовый файл. Результаты этой проверки иллюстрирует табл. 1.

Как видно из представленных результатов, во всех случаях, кроме ситуации с word-файлом, обеспечивается сжатие шифрованного текста менее, чем на 11% [3]. Результат с word-файлом свидетельствует лишь о том, что в этом случае имеется значительная часть повторяющихся (одинаковых) сообщений" при шифровании word-файла на разных сеансовых ключах он уже не сжимается.

На втором этапе была решена задача увеличения вдвое длины шифруемого блока. При построении усовершенствованного симметричного шифра сохранены все основные идеи, использованные в самом стандарте ГОСТ 28147-89, но они переработаны теперь для длины блока, равной 128 битам. Кроме того, как и в предыдущем случае, в цикловую

функцию шифра введена дополнительная операция параметрического циклического сдвига. Результаты проведенных статистических испытаний модернизированного таким образом алгоритма приведены на рис 5-6.

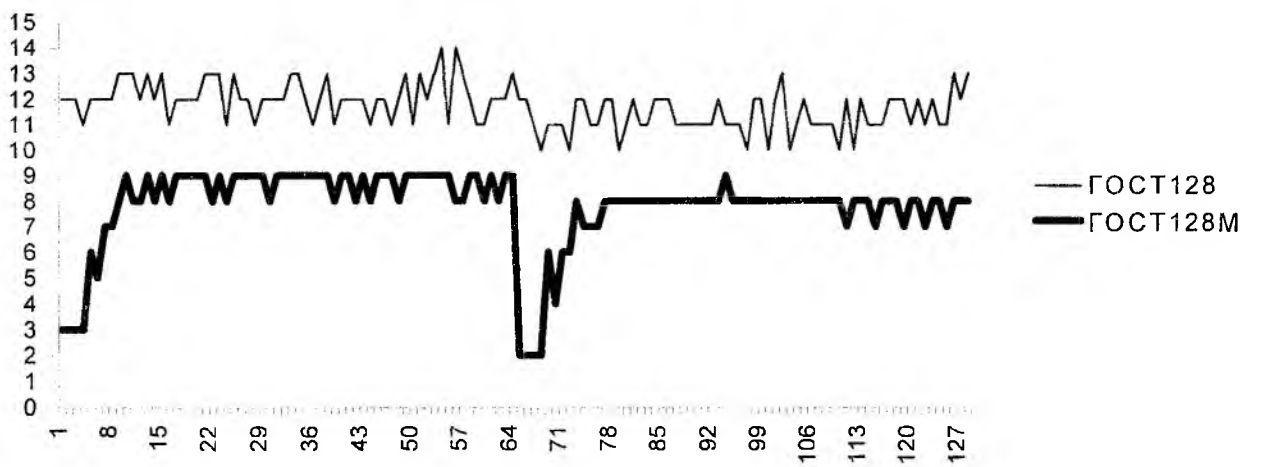


Рис. 5

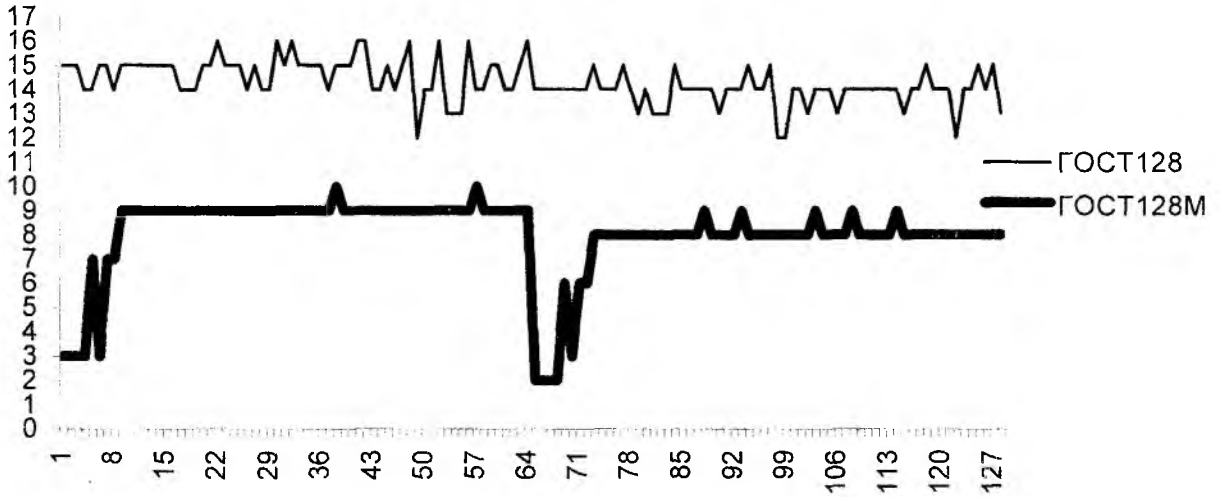


Рис. 6

При использовании в алгоритме ГОСТ–128 ненулевого сеансового ключа $(\overline{K} \neq 0)$ для обеспечения зависимости всех выходных бит от любого входного бита требуется вхождения в алгоритм на глубину до 10-14 циклов. При нулевом сеансовом ключе $(\overline{K} = 0)$ – на 12-16 циклов. Применение в алгоритме параметрического сдвига, как и в предыдущем случае, улучшает результаты.

При ненулевом сеансовом ключе $(\overline{K} \neq 0)$ лавинный эффект наступает на 2-9 циклах, при нулевом $(\overline{K} = 0)$ – на 3-10. Результаты оценки числа циклов алгоритма, при котором наступает статистическая независимость выходных (шифрованных) текстов при шифровании сообщений с помощью ключей $\vec{K}$, отличающихся одним битом, приведены на рис. 7.

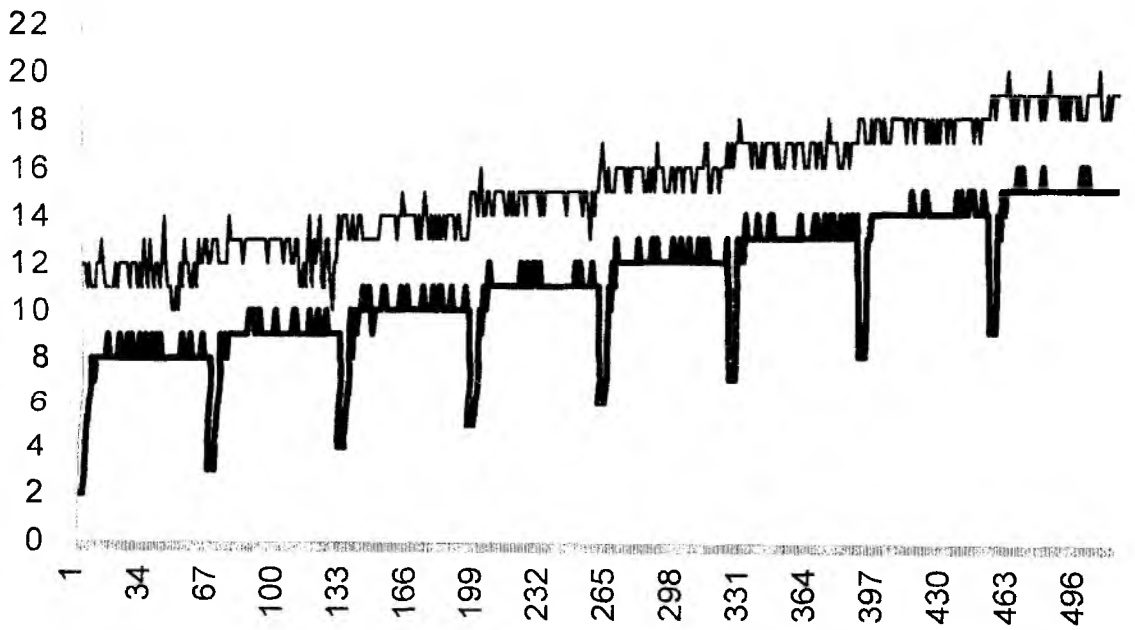


Рис. 7

Как и в предыдущем случае, экспериментальные данные показали улучшение показателей статистической безопасности при применении в алгоритме операции параметрического сдвига (10-20 циклов для ГОСТ -128 и 2–16 циклов для алгоритма с параметрическим сдвигом).

Глубина входа в алгоритм в этом случае определялась номером цикла, начиная с которого математическое ожидание числа единичных (ненулевых) бит m_w зашифрованного блока данных удовлетворяет условию $64 - 0{,}6 \le m_w \le 64 + 0{,}6$.

Результаты проверки степени сжатия текстов, зашифрованных на ключах - подстановках случайного типа, с помощью процедуры Лемпела-Зива приведены в табл. 2

Таблица 2

ТИП	Исходный текст	ГОСТ-128	ГОСТ-128М
EXE	49,78	5,752	5,744
TXT	53,68	0,145	-0,0071
WORD	77,73	39,15	39,16

Как видно, применение 128-битного алгоритма и параметрического сдвига и в данном случае улучшает характеристики случайности шифрованных блоков данных.

Несколько слов относительно быстродействия рассмотренных модернизированных алгоритмов. Наши эксперименты показывают, что введение операции параметрического сдвига не приводит к сколько-нибудь существенным потерям в скорости по отношению к стандартной процедуре шифрования.

В итоге можно сделать вывод, что существует реальная возможность повышения безопасности алгоритма шифрования по ГОСТ 28147-89. Мы продолжаем исследования в рассмотренных направлениях и считаем, что представленные в работе результаты могут заинтересовать специалистов.

Список литературы: 1. *Biham E., Shamir A.* Differential Cryptanalysis of DES-like cryptosystems. The Weizmann Institute of Science. Department of Applied Mathematics. Technical Report CS90-16. 1990. 2. *Долгов В.И., Лисицкая И.В., Олейников Р.В, Шумов А.И.* "Слабые" ключи в алгоритме шифрования ГОСТ28147-89// Радиотехника: Всеукр. межвед. науч.-техн. сб. 2000. Вып 114. С. 63-68. 3. *Горбенко И.Д., Лисицкая И.В., Коряк А.С.* Анализ стойкости алгоритма ГОСТ 28147-89 при использовании подстановок случайного типа // Радиотехника и информатика. 1998. №1(02). С.63-68.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 6.03.2001

УДК 681.3.06: 519.248.681

В. И. ДОЛГОВ, д-р техн. наук, В. И. РУЖЕНЦЕВ, М. А. ФЕДОТОВ,
А. П. МЕЛЕЦКИЙ, М. В. ПИНЧУК

MDES-128 С ТАБЛИЦАМИ ПОДСТАНОВОК СЛУЧАЙНОГО ТИПА

Американский стандарт шифрования DES является одним из наиболее популярных и известных симметричных шифров. Вот уже более 20 лет этот алгоритм находится в центре внимания специалистов. Его можно считать наиболее глубоко и досконально исследованным. Несмотря на многочисленные критические замечания, он выдержал проверку временем и сохранил авторитет одного из надежных для прошедшего этапа инструментов защиты информации во многих государственных и коммерческих структурах (автоматизированных системах управления). Сегодня, однако, прогресс в области вычислительной техники ставит на повестку дня вопрос замены этого стандарта новым. Размер ключа DES стал слишком мал для современных приложений. Не так давно в работах [1,2] показано, что, затратив около 1 млн. $ (против 20 млн. $, названных в работе 1977 года [3]), возможно создание специализированного аппаратного устройства, которое за время менее 1 часа выполнит атаку на ключ DES методом полного перебора. Можно напомнить и свежий пример, когда ключ DES был найден методом простого перебора с использованием инструментария, распространенного в Internet. Таким образом, и при использовании возможностей программных средств 56-битный ключ также может уже не обеспечить надежной защиты.

Здесь следует заметить, что на проблему слишком маленького размера ключа было обращено внимание почти сразу же после публикации DES [3]. Поэтому, для случаев, когда есть основания опасаться очень серьезного криптоанализа, уже давно обсуждаются возможности увеличения длины используемого ключа путем композиции DES шифрований. Простейшей из таких возможностей является последовательное использование дважды одного и того же шифра, однако для алгоритма DES с фиксированными таблицами *S*-блоков двойное шифрование с независимым выбором ключей не приводит в теоретическом плане к повышению стойкости процедуры шифрования. В этом случае существует способ оптимизации переборной процедуры, называемый встречной атакой [4], который требует для криптоанализа число шагов, пропорциональное 2^n, где n – длина ключа, т.е. совпадающее с числом возможных ключей для однократного DES (правда, взамен требуется машинная память такого же порядка).

Чаще DES используют в режиме тройного шифрования, при котором открытый текст шифруется 3 раза с тремя независимыми ключами. Данный метод получил название тройной DES. Ряд недостатков различных вариантов этого подхода детально обсуждаются в [5] и других работах. Тем не менее, Комитет X9.F.1 Американского Национального института Стандартов (ANSI) работал над принятием наборов режимов для трехкратного (тройного) шифрования с использованием DES [6.7].

Отметим также еще раз работу [5], в которой предлагается 6-ти цикловый шифр, который использует DES в качестве цикловой функции. Результатом является 128-битный блочный шифр, который назван DEAL–A, с 64-битными циклическими подключами, которые получаются из выбранного пользователем ключа, согласно алгоритму генерации ключей. Система так же быстра, как тройной DES, в котором используется 6 шифрований для того, чтобы зашифровать два 64-битных блока открытого текста. Заметим, наконец, что 128-битный блочный шифр DEAL–A, строящийся на основе DES, предлагался кандидатом на NIST AES стандарт.

Заключая приведенный небольшой анализ подходов к композиции шифров DES, следует констатировать, что во всех представленных выше случаях приходится мириться с понижением быстродействия криптографических преобразований – характеристикой, которая представляется одной из наиболее ценных для симметричных шифров.

Мы хотим открыть новую страницу в изучении и исследовании возможностей использования ставшей уже классической схемы преобразований, примененной в стандарте шифрования данных DES. Основное внимание настоящей работы сосредотачивается на развитии подхода, позволяющего преодолеть одновременно все ограничения стандарта DES – малую длину шифруемого блока и ключа, а также уязвимость стандарта к атакам дифференциального и линейного криптоанализов.

Для построения усовершенствованного симметричного шифра предлагается взять за основу идею, использованную при построении самого стандарта DES, но реализовать теперь ее для длины блока, равной 128 битам. Кроме того, в цикловую функцию шифра DES предлагается ввести дополнительную операцию – параметрический циклический сдвиг. Эта операция позволяет по-новому решить задачу обеспечения безопасности шифра.

Будем в дальнейшем называть рассматриваемый вариант построения новой процедуры шифрования DEA-128. В шифре DEA-128 в качестве левого и правого полублоков схемы Фестеля будут выступать уже 64-битные полублоки. Конечно, потребуется теперь использовать таблицу из 16 различных S-блоков и осуществить модификацию всех остальных преобразований для увеличенной в два раза длины блока: начальной и конечной перестановок IP и IP^{-1}, расширяющего преобразования E. таблицы перестановки P и алгоритма получения подключей K_i, $i = 1,2,\ldots,16$ [8]. Сосредоточим сначала внимание на детальном описании параметров алгоритма DEA-128.

Итак, в алгоритме DEA-128 входной блок A, состоящий из 128 двоичных символов, разбивается на две равные части по 64 бита: левый полублок L_0 и правый полублок R_0. Затем, как и в стандарте, осуществляется 16 циклов преобразования сообщения $A = (L_0R_0)$, так что в i-м цикле слово $(L_{i-1}R_{i-1})$ преобразуется в (L_iR_i) по правилам:

$$L_i = R_{i-1},\ R_i = L_{i-1} \oplus f(R_{i-1},K_i,K'_i),$$

где символ $\oplus$ обозначает операцию побитного суммирования по модулю 2; $f(R_{i-1},K_i,K'_i)$ – цикловая функция; (L_iR_i) – операция конкатенации (объединения) блоков L_i и R_i, $i = 0,1,\ldots,16$.

В цикловой функции $f(R_{i-1},K_i,K'_i)$ вычисление начинается с так называемого расширения, преобразующего 64 битный полублок R_i в 96-битный в соответствии с расширяющей таблицей E (фиксированного вида), которая, фактически, дополнительно вставляет копии 32 позиций. Полученные 96 бит гаммируются (побитно суммируются) с 96 битами подключа (K_i,K'_i), представляющего собой объединение двух 48-битных подключей, которые формируются в соответствии с алгоритмом развертывания подключей, используемым в самом стандарте DES. Результат гаммирования разбивается на 16 блоков по 6 бит, поступающих на S-блоки $S_1,S_2,\ldots,S_{16}$. В каждом из S-блоков, также как и в стандарте DES, входные 6 бит заменяются 4-мя выходными.

Шестнадцать 4-битных двоичных блоков, поступающих с выходов S-блоков, образуют 64-битный полублок. Двоичные элементы этого полублока подвергаются P-перестановке, задаваемой фиксированной таблицей.

Завершая описание основных операций DEA-128, остается заметить, что по аналогии со стандартом процедура шифрования начинается с начальной фиксированной перестановки IP и завершается применением к полученному результату обратной перестановки IP^{-1}.

Напоминаем также, что в самом конце выполнения всех 16 циклов алгоритма DES левый и правый полублоки меняются местами, так что результатом работы алгоритма DEA-128 является $E_{K,K}(A) = (R_{16}L_{16})$.

Приведем теперь конкретные параметры и характеристики модифицированных таблиц перестановок и подстановок, использованных в DEA-128.

Начальная перестановка IP, построенная по аналогии с идеей построения подстановки стандарта DES, представлена в виде табл. 1, где i_A обозначает номер позиции входного блока, а $IP(i_A)$ – номер позиции, в которой i_A-й бит входного блока окажется в результате перестановки.

Таблица 1

$IP(i_B)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
I_B	114	98	82	66	50	34	18	2	116	100	84	68	52	36	20	4
$IP(i_B)$	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
I_B	118	102	86	70	54	38	22	6	120	104	88	72	56	40	24	8
$IP(i_B)$	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
i_B	122	106	90	74	58	42	26	10	124	108	92	76	60	44	28	12
$IP(i_B)$	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
i_B	126	110	94	78	62	46	30	14	128	112	96	80	64	48	32	16
$IP(i_B)$	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
I_B	113	97	81	65	49	33	17	1	115	99	83	67	51	35	19	3
$IP(i_B)$	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
I_B	117	101	85	69	53	37	21	5	119	103	87	71	55	39	23	7
$IP(i_B)$	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
I_B	121	105	89	73	57	41	25	9	123	107	91	75	59	43	27	11
$IP(i_B)$	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
I_B	125	109	93	77	61	45	29	13	127	111	95	79	63	47	31	15

Математически эти преобразования можно представить в виде следующих двух соотношений:

$IP(i_A) = [2\cdot(64+1)-16\cdot i_A](\mathrm{mod}\ 130)$ для $IP(i_A) \le 64$ и

$IP(i_A) = [2\cdot(64+1)-16\cdot(i_A-64)-1](\mathrm{mod}\ 130)$ для $IP(i_A) > 64$

против $IP(i_A) = [2\cdot(32+1)-8\cdot i_A](\mathrm{mod}\ 66)$ $IP(i_A) \le 32$ и

$IP(i_A) = [2\cdot(32+1)-8\cdot(i_A-32)-1](\mathrm{mod}\ 66)$ $IP(i_A) > 32$ для стандартного DES.

Завершающая обратная перестановка представлена в табл. 2, где i_B обозначает номер позиции в блоке шифртекста, полученного в результате выполнения 16 циклов, а $IP^{-1}(i_B)$ – номер позиции, на которую поступает i_B-й бит указанного блока в результате перестановки.

Расширяющая перестановка E построена по принципу, используемому в алгоритме DES, и преобразует 64-битный полублок в 96-битный в соответствии с табл. 3, в которой j обозначает номер позиции бита в R_{i-1}, а j_E – номер позиции, куда поступает j-й бит в результате выполнения E-перестановки. В случае, когда j-й бит поступает в две позиции, номера позиций перечислены через запятую.

Таблица 2

$IP^{-1}(i_B)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
i_B	72	8	80	16	88	24	96	32	104	40	112	48	120	56	128	64
$IP^{-1}(i_B)$	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
i_B	71	7	79	15	87	23	95	31	103	39	111	47	119	55	127	63
$IP^{-1}(i_B)$	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
i_B	70	6	78	14	86	22	94	30	102	38	110	46	118	54	126	62
$IP^{-1}(i_B)$	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
i_B	69	5	77	13	85	21	93	29	101	37	109	45	117	53	125	61
$IP^{-1}(i_B)$	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
i_B	68	4	76	12	84	20	92	28	100	36	108	44	116	52	124	60
$IP^{-1}(i_B)$	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
i_B	67	3	75	11	83	19	91	27	99	35	107	43	115	51	123	59
$IP^{-1}(i_B)$	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
I_B	66	2	74	10	82	18	90	26	98	34	106	42	114	50	122	58
$IP^{-1}(i_B)$	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
i_B	65	1	73	9	81	17	89	25	97	33	105	41	113	49	121	57

Таблица 3

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
j_E	2,96	3	4	5,7	6,8	9	10	11,13	12,14	15	16	17,19	18,20	21	22	23,25
J	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
j_E	24,26	27	28	29,31	30,32	33	34	35,37	36,38	39	40	41,43	42,44	45	46	47,49
j	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
j_E	48,50	51	52	53,55	54,56	57	58	59,61	60,62	63	64	65,67	66,68	69	70	71,73
j	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
j_E	72,74	75	76	77,79	78,80	81	82	83,85	84,86	87	88	89,91	90,92	93	94	1,95

Для задания нелинейных S подстановок использовались 16 таблиц размером 4×16 каждая, отобранные в соответствии с методикой, предложенной в [9], т.е. при их формировании учитывалось выполнение критериев случайности.

Завершающим преобразованием в шифрующей функции алгоритма DES является P-перестановка. Она относится к перестановкам случайного типа (отвечает требованиям случайности [10]), а также удовлетворяет дополнительным ограничениям [11], представленным специалистами фирмы IBM – разработчиками стандарта: каждый из четырех выходов любого S-блока распределяется по позициям входов различных S-блоков на следующем цикле.

Для более полного соответствия DEA-128 с DES-64 нами применена P-перестановка, представляющая собой сдвоенную стандартную перестановку: биты с 1 по 33 преобразуются в соответствии

со стандартной перестановкой, а перестановка для битов с 33 по 64 рассчитывается по простой формуле:

$$P_{128}(i) = P_{64}(i) + 32.$$

Приведенный вариант перестановки был изучен, протестирован и при проверке лавинного эффекта показал лучшие результаты по сравнению с другими вариантами. Ниже приведена табл. 4 с *P*-перестановкой, используемой в обсуждаемом шифре.

Таблица 4

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
P(i)	16	7	20	21	29	12	28	17	1	15	23	26	5	18	31	10
i	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
P(i)	2	8	24	14	32	27	3	9	19	13	30	6	22	11	4	25
i	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
P(i)	48	39	52	53	61	44	60	49	33	47	55	58	37	50	63	42
i	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
P(i)	34	40	56	46	64	59	35	41	51	45	62	38	54	43	36	57

Предлагаемые изменения в классической схеме криптопреобразований DES

Увеличение только длины шифруемого блока при стандартной схеме криптопреобразований может повлечь за собой только ухудшение показателей безопасности нового алгоритма, а также создаст возможность применения к нему старых, проверенных на DES-64, атак.

Стараясь радикально не менять стандартную схему шифрования, для защиты от вышеперечисленных криптоаналитических атак мы ввели в каждый цикл алгоритма всего одну дополнительную нелинейную операцию. Эта операция названа нами "параметрический циклический сдвиг".

Порядок осуществления дополнительной операции следующий: в каждом цикле шифрования (дешифрования) после прохождения правого полублока через шифрующую функцию, производится его циклический сдвиг влево на число, определяемое младшими 6-ю битами этого же полублока. В остальном порядок выполнения операций не изменяется.

На рис.1 представлен модифицированный цикл DEA-128.

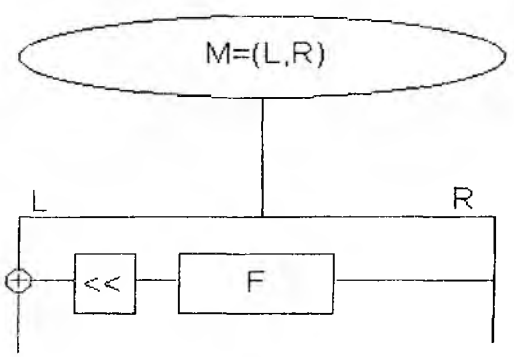


Рис. 1

Как показывает анализ и расчеты, параметрический сдвиг существенно уменьшает вероятности построения дифференциальных и линейных характеристик, что делает атаки дифференциального и линейного криптоанализов неэффективными.

Проверка статистической безопасности алгоритма шифрования данных DEA-128

Проверка статистической безопасности предлагаемой версии алгоритма шифрования выполнялась по методике, изложенной в работе [10].

Использовались четыре основных показателя статистической безопасности, традиционные для симметричных блочных шифров:

1. Число циклов алгоритма, начиная с которого криптограммы, полученные шифрованием двух, отличающихся на один бит блоков данных (открытых текстов), становятся устойчиво независимыми (в том смысле, что при дальнейшем увеличении числа циклов они остаются независимыми). Другими словами, необходимо было определить число циклов в алгоритме шифрования, начиная с которого изменение одного бита открытого текста приводит к изменению криптограммы приблизительно в половине битов. Этот показатель отражает качество, так называемого, лавинного эффекта.

2. Число циклов шифрования, при котором один и тот же открытый текст, зашифрованный на ключах, отличающихся одним битом, порождает устойчиво независимые (некоррелированные) криптограммы.

3. Коэффициент сжатия шифрованного текста при использовании процедуры архивирования Лемпела-Зива.

4. Корреляция входа-выхода, т.е. степень статистической связи открытого текста с соответствующим ему закрытым текстом.

Результаты, полученные в ходе статистических экспериментов, иллюстрируют табл. 5, 6. В табл. 5 приведены типичные значения математического ожидания числа единичных (ненулевых) бит в побитной сумме по модулю 2 пар, полученных на k-том цикле криптограмм, открытые тексты или ключи которых отличаются одним битом.

Таблица 5

№ цикла	Изменение 1 бита текста	Изменение 1 бита ключа
0	4,563	2,155
1	17,108	17,220
2	37,207	43,393
3	54,330	59,917
4	62,138	63,521
5	**63,911**	63,962
6	64,065	64,075
7	64,135	63,829
8	64,283	63,758
9	64,084	63,898
10	63,843	63,843
11	63,912	63,826
12	63,943	63,765
13	64,041	64,081
14	63,951	64,227
15	63,870	63,914
16	63,870	63,914

При исследовании показателей сжатия текстов с помощью процедуры архивирования Лемпела-Зива исследовались тексты трех категорий: обычный EXE-файл, текст редактора Microsoft Word и обычный текстовый файл. Результаты этих исследований, иллюстрирует табл. 6.

Таблица 6

Тип файла	Сжатие, %
Exe-файл	1,3-12,9
Документ Microsoft Word	43
Обычный текст	1,60

Статистическая независимость выходных битов от входных битов, в соответствии с полученными результатами, наблюдается при использовании 6 и более циклов шифрования, как и в случае стандартного шифра DES.

Во всех случаях обеспечивается сжатие шифрованного текста по Лемпелу-Зиву, аналогичное сжатию, достигаемому при использовании стандартной процедуры шифрования DES.

Коэффициент корреляции открытого и соответствующего ему шифрованного текста не превышает значения 0,001.

Рассматриваемая версия DES подобной процедуры шифрования обеспечивает удельную скорость криптопреобразований, превышающую более чем в два раза тройной DES и 128-битный блочный шифр DEAL–A, строящийся на основе DES, предлагаемый кандидатом на NIST AES стандарт. Этот момент может оказаться принципиальным для многих приложений.

Остается отметить, что использование процедуры параметрического сдвига в принципе позволяет по-новому подойти к обеспечению характеристик стойкости алгоритмов шифрования к атакам дифференциального и линейного криптоанализов. В частности, для рассмотренного варианта построения шифра отпадает необходимость выполнения одного из наиболее жестких требований к S-блокам стандарта DES – запрета однобитных переходов [12]. Вместе с тем, в рассмотренном варианте применения процедуры параметрического сдвига сохраняется необходимость запрета при отборе S-блоков переходов обнуляющего типа [11]. Более детальный анализ защищенности рассмотренной модифицированной процедуры шифрования от атак дифференциального и линейного криптоанализа выходит за рамки настоящей работы. Этому вопросу мы посвятим отдельное исследование.

Мы продолжаем изучать устойчивость нашей версии построения шифра и от других возможных атак. Однако основная идея – введение в процедуру криптографических преобразований DES подобных шифров дополнительной операции параметрического циклического сдвига – открывает возможность использования в качестве S-блоков в этих шифрах таблиц подстановок случайного типа (прошедших тесты на случайность) и вселяет в нас уверенность, что развиваемый подход может оказаться

достаточно эффективным и в других случаях. В частности, одним из перспективных решений в свете развиваемого подхода может стать использование таблиц *S*-блоков как еще одного секретного параметра шифра (подобно алгоритму ГОСТ 28147-89), что позволит добиться дальнейшего наращивания показателей безопасности.

Список литературы: 1. *Wiener M.J.* Efficient DES key search. Technical Report TR-244. School of Computer Science. Carleton University. Ottawa. Canada. May 1994. Presented at the Rump Session of CRUPTO'93. 2. *Wiener M.J.* Efficient DES key search - an update. CryptoBytes, 3(2): 6-6, 1998. 3. *Diffie W., Hellman M.* Exhaustive cryptanalysis of the NBS data encryption standard. Computer. P. 74-84. 1977. 4. *Вербицький О.В.* Вступ до криптології. Видавництво наук.-техн. літератури. Львів. 1998. 5. *Knudsen L.R.* DEAL – A 128-bit Block Cipher. 1998. P. 1-9. 6. *ANSI X9.F.1.* TDEA modes of operation. Draft 5.5. X9.52. March 29, 1996. 7. *NIST. AES announcement.* Draft. June 15, 1997. 8. *Барсуков В.С., Дворянкин С.В., Шеремет И.А.* Безопасность связи в каналах телекоммуникаций. М. Россия. 1993. Т.20. 123 с. 9. *Горбенко И.Д., Лисицкая И.В.* Критерии отбора случайных таблиц подстановок для алгоритма шифрования по ГОСТ 2847-89 // Радиотехника. 1997. Вып 103. С. 121–130. 10. *Горбенко И.Д., Лисицкая И.В., Коряк А.С.* Анализ стойкости алгоритма ГОСТ 28147-89 при использовании подстановок случайного типа. // Радиоэлектроника и информатика. 1998. №1 (02). С. 39–43. 11. *Schneier B.* Applied Cryptography. Second Edition: protocols, algorithms and source code in C. Published by John Wiley & Sons. Inc, New York: Chichester Brisbane Toronto Singapore, 1996 – 158 p. 12. *Лисицкая И.В., Коряк А.С., Олейников Р.В.* К вопросу построения случайных *S*-блоков для алгоритма DES. Критерии отбора *S*-блоков. // Радиоэлектроника и информатика. 1999. №3. С. 94–100.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 6.03.2001

УДК 681.3.06: 519.248.681

М. Ф. БОНДАРЕНКО, д-р техн. наук, А. С. КОРЯК, В. И. РУЖЕНЦЕВ

ПОВЫШЕНИЕ УСТОЙЧИВОСТИ ШИФРА DES К АТАКАМ ДИФФЕРЕНЦИАЛЬНОГО КРИПТОАНАЛИЗА

В работе [1] изучаются и обсуждаются принципы построения таблиц подстановок (*S*-блоков) для шифра DES, использованные разработчиками. Показано, что на момент принятия стандарта предложенные разработчиками шифра ограничения действительно позволяли считать его достаточно надежным для практического применения. Сегодня, однако, найдены атаки на шифр, сложность которых оказалась меньше сложности прямого перебора всех возможных ключей (атаки грубой силой). Сначала Эли Бихамом и Ади Шамиром был разработан дифференциальный криптоанализ [2], а немного позже Мицуру Мацуи была обоснована атака на шифр DES, названная линейным криптоанализом [3]. Известно, что стойкость к названным атакам определяется свойствами таблиц *S*-блоков. Естественным, поэтому, является интерес, проявленный в последнее время к разработке подходов в том числе и методов отбора таблиц подстановок, позволяющих повысить показатели защищенности шифра DES [4-6 и др.].

В этой работе речь будет идти о защите от известной атаки на шифр DES, предложенной Э. Бихамом и А. Шамиром [2]. Она основана на использовании двухцикловых характеристик, допускающих итеративное продолжение. Центральная идея, использованная при обосновании излагаемого подхода, состоит в нахождении таких *S*-блоков, которые исключают всякую возможность (ненулевую вероятность) построения "обнуляющего" разностного преобразования (так в [1] названо выполнение одноциклового преобразования, при котором ненулевая разность на входе цикловой функции F преобразуется в нулевую разность на ее выходе).

Целесообразно будет начать с того, что напомнить методику построения трехблочных обнуляющих характеристик , изложенную в работе [1].

При построении трехблочных характеристик рассматриваются 14-битные входы цикловой функции вида (0, 0, x, y, z, 1, t, p, 1, q, l, m, 0, 0) трех смежных (соседних) *S*-блоков. При этом учет требований, предъявленных к *S*-блокам разработчиками стандарта [7], позволяет выделить два принципиальных момента:

- из всего множества входов (0, 0, x, y, z, 1, t, p, 1, q, l, m, 0, 0) трех (смежных) активных *S*-блоков реально могут быть использованы только те, у которых биты z и q одновременно не равны нулю (т.к. переход в 0 для входной разности вида (0, x, y, z, t, 0) в таблицах, составленных из перестановок, невозможен);

- в атаке, использующей несколько активных *S*-блоков в обнуляющем цикле, в принципе могут участвовать только связанные (смежные) *S*-блоки.

Все допустимые варианты входов *S*-блоков, которые могут участвовать в формировании трехблочных характеристик (характеристик с тремя активными *S*-блоками), представлены в табл. 1.

Напомним, что связь обозначений входов в *S*-блоки, представленных в табл.1, с таблицами стандарта определяется правилами, оговоренными в [8]: вход по строкам таблицы ab_x соответствует 6-битному вектору входной разности $\Delta=(\Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5, \Delta_6)$, где a представляет собой шестнадцатеричную запись двоичного числа $\Delta_1\Delta_6$, а b – числа $\Delta_2\Delta_3\Delta_4\Delta_5$.

Анализ представленных в табл. 1 композиций входов в *S*-блоки позволяет заключить, что всего возможно 64×3=192 варианта атак для каждой тройки таблиц, а для шифра в целом 192×8=1536 вариантов.

Для перекрытия этих характеристик, как показано в [1], разработчики шифра DES пошли двумя путями. Основную массу трехблочных обнуляющих характеристик они просто запретили с помощью требования 6 (входы 28_x, $2A_x$, $2C_x$, $2E_x$ во всех таблицах побитовых разностей выбраны с нулевыми вероятностями переходов в выходную разность ноль) и требования 4 (для входов 10_x и 20_x нет переходов в ноль). В результате из общего числа 192 вариантов остается защититься от 48 оставшихся характеристик для каждой тройки смежных *S*-блоков. Разработчики постарались перекрыть их за счет ограничений на допустимые значения вероятностей одноблочных переходов, участвующих в формировании этих трехблочных характеристик [1]. Однако для защиты от атак, предложенных Э. Бихамом и А. Шамиром, этого оказалось недостаточно. Указанные атаки на 16-цикловый DES построены именно на возможности реализации трехблочных обнуляющих характеристик. Э. Бихам и А. Шамир используют две наиболее вероятные итерационные характеристики, в которых участвуют первые три

S-блока: одна – для входной разности 19 60 00 00 00 00 00 00_x (ей соответствует трехблочная входная разность 00001100101100 или входы таблиц разностей 11_x, 29_x, 26_x), вторая – для входной разности $1B$ 60 00 00 00 00 00 00_x (трехблочная входная разность 00001101101100 или, соответственно, входы таблицы разностей 11_x, $2B_x$, 26_x).

Таблица 1

Участие S-блоков в формировании входной разности (0, 0, x, y, z, 1, t, p, 1, q, l, m, 0, 0) при трехблочной характеристике					
z	Входные разности первого S-блока (0, 0, x, y, z, 1)	z q	Входные разности второго S-блока (z, 1, t, p, 1, q)	q	Входные разности третьего S-блока (1, q, l, m, 0, 0)
0	(0, 0, 0, 0, 0, 1) → 10_x		(0, 1, 0, 0, 1, 1) → 19_x		(1, 0, 0, 0, 0, 0) → 20_x
	(0, 0, 0, 1, 0, 1) → 12_x	0	(0, 1, 0, 1, 1, 1) → $1B_x$		(1, 0, 0, 1, 0, 0) → 22_x
	(0, 0, 1, 0, 0, 1) → 14_x	1	(0, 1, 1, 0, 1, 1) → $1D_x$	0	(1, 0, 1, 0, 0, 0) → 24_x
	(0, 0, 1, 1, 0, 1) → 16_x		(0, 1, 1, 1, 1, 1) → $1F_x$		(1, 0, 1, 1, 0, 0) → 26_x
1	(0, 0, 0, 0, 1, 1) → 11_x		(1, 1, 0, 0, 1, 0) → 29_x		(1, 1, 0, 0, 0, 0) → 28_x
	(0, 0, 0, 1, 1, 1) → 13_x	1	(1, 1, 0, 1, 1, 0) → $2B_x$		(1, 1, 0, 1, 0, 0) → $2A_x$
	(0, 0, 1, 0, 1, 1) → 15_x	0	(1, 1, 1, 0, 1, 0) → $2D_x$	1	(1, 1, 1, 0, 0, 0) → $2C_x$
	(0, 0, 1, 1, 1, 1) → 17_x		(1, 1, 1, 1, 1, 0) → $2F_x$		(1, 1, 1, 1, 0, 0) → $2E_x$
			(1, 1, 0, 0, 1, 1) → 39_x		
		1	(1, 1, 1, 0, 1, 1) → $3D_x$		
		1	(1, 1, 0, 1, 1, 1) → $3B_x$		
			(1, 1, 1, 1, 1, 1) → $3F_x$		

В то же время, анализ приведенных выше вариантов возможных входных разностей для трехблочной атаки и их распределения по S-блокам позволяет сделать вывод, что разработчики стандарта при формировании требований к отбору S-блоков, по-видимому, не ставили перед собой задачу добиться максимально возможной защищенности шифра от атак дифференциального криптоанализа, а стремились просто гарантировать некоторый уровень защиты, который, по их мнению, обеспечит достаточную его надежность, так как существуют возможности улучшения показателей надежности шифра. Легко убедиться, что на таблицы дифференциальных разностей, а соответственно, на таблицы S-блоков можно наложить дополнительные ограничения, которые делают шифр DES неуязвимым по крайней мере к дифференциальным атакам, использующим трехблочные обнуляющие характеристики. Для этого достаточно потребовать, например, чтобы в таблицах дифференциальных разностей дополнительно были запрещенными переходы в нуль еще для четырех входов, а именно, для входов 29_x, $2B_x$, $2D_x$, $2F_x$. Действительно, как уже отмечалось ранее, в стандарте DES S-блоки построены так, что для входных разностей 28_x, $2A_x$, $2C_x$, $2E_x$ не существует переходов в нуль. Это означает, что в соответствии с табл.1 в трехблочной характеристике для последнего (третьего в связке) S-блока "работают" только четыре входа, для которых $q = 0$. Этому же значению бита во входной разности соответствуют четыре значения входов для второго S-блока связки, а именно, входы 29_x, $2B_x$, $2D_x$, $2F_x$ (все другие входы срабатывают только при $q = 1$). Но тогда трехблочную обнуляющую характеристику можно сделать нереализуемой, если, как уже отмечалось выше, сделать запрещенными переходы в ноль для входов 29_x, $2B_x$, $2D_x$, $2F_x$. В результате, для повышения устойчивости к атакам дифференциального криптоанализа необходимо S-блоки строить так, чтобы в таблицах дифференциальных разностей отсутствовали нулевые выходные разности одновременно для восьми входов в эти таблицы: 28_x, 29_x, $2A_x$, $2B_x$, $2C_x$, $2D_x$, $2E_x$, $2F_x$.

Отметим здесь, что к этому же результату, как оказалось, раньше нас пришли и исследователи группы Кваджио Ким [5]. В их интерпретации изложенные выше дополнительные требования к отбору S-блоков сформулированы более компактно: необходимо чтобы для любого S-блока $S(\mathbf{x}) \neq S(\mathbf{x} \oplus 11ef10)$.

Здесь мы хотим привлечь внимание к тому, что представленный результат не является единственно возможным решением задачи перекрытия трехблочных обнуляющих характеристик. Можно предложить еще несколько вариантов правил отбора S-блоков, позволяющих защитить шифр DES от известных атак дифференциального криптоанализа (перекрыть наиболее вероятные атаки), чему и посвящаются дальнейшие результаты

Сразу можно отметить, что для наших целей подходят только таблицы S-блоков, для которых не "срабатывают" и двухблочные обнуляющие характеристики, а для этого, как показано в [1], должны быть равными нулю вероятности переходов в ноль (должны иметь нулевые значения ячейки таблиц дифференциальных разностей, соответствующие нулевым выходам) для входов S-блоков хотя бы одного из столбцов табл. 2. (В обозначениях корейских исследователей – это либо требование $S(\mathbf{x}) \neq S(\mathbf{x} \oplus 00ef11)$, либо требование $S(\mathbf{x}) \neq S(\mathbf{x} \oplus 11ef00)$).

Таблица 2

Участие S-блоков в формировании входной разности $(0, 0, x, y, 1, 1, t, p, 0, 0)$ при двухблочной характеристике	
Входные разности первого S-блока $(0, 0, x, y, 1, 1)$	Входные разности второго S-блока $(1, 1, t, p, 0, 0)$
$(0, 0, 0, 0, 1, 1) \rightarrow 11_x$	$(1, 1, 0, 0, 0, 0) \rightarrow 28_x$
$(0, 0, 0, 1, 1, 1) \rightarrow 13_x$	$(1, 1, 0, 1, 0, 0) \rightarrow 2A_x$
$(0, 0, 1, 0, 1, 1) \rightarrow 15_x$	$(1, 1, 1, 0, 0, 0) \rightarrow 2C_x$
$(0, 0, 1, 1, 1, 1) \rightarrow 17_x$	$(1, 1, 1, 1, 0, 0) \rightarrow 2E_x$

Кроме того, если идти дальше, введенные ограничения должны обеспечивать и невозможность построения "обнуляющего" разностного преобразования для большего количества активных S-блоков (многоблочных обнуляющих характеристик), так как для характеристики обнуляющего типа с 4-мя активными S-блоками при одноцикловом преобразовании с вероятностью $\frac{16}{64} = \frac{1}{4}$ на один S-блок (граничное значение, определяемое элементами таблиц дифференциальных разностей S-блоков, установленное требованием 7 разработчиков стандарта [1]), для вероятности результирующей 13-ти цикловой характеристики получаем оценку

$$\left(\frac{1}{4}\right)^{4\cdot 6} = \left(\frac{1}{4}\right)^{24} = 2^{-48},$$

что меньше, чем вероятность прямого перебора всех ключей (2^{-55}).

Участие S-блоков в формировании входной разности при построении четырехблочных характеристик иллюстрирует табл. 3.

Анализ представленных результатов позволяет заключить, что для всего набора возможных переходов, представленных в табл.1 и табл. 2, приведенным ограничениям удовлетворяют, включая отмеченный выше, пять вариантов задания нулевых вероятностей переходов в нулевую выходную разность (нас будут интересовать в первую очередь ограничения минимального типа, под которыми будем понимать минимальное число ячеек (выходов) таблиц дифференциальных разностей, нулевые значения которых обеспечивают нереализуемость многоцикловых обнуляющих характеристик). Они показаны входами в таблицы дифференциальных разностей, представленными в табл. 4.

Как показывает более тщательный анализ, из этих характеристик необходимо исключить последнюю, так как в отличие от остальных она не запрещает характеристики обнуляющего типа с числом активных S-блоков в каждом цикле, большем, чем три (см. табл. 3). Нулевые значения переходов в ноль для первых четырех сочетаний входов обеспечивают перекрытие не только всех двухблочных, трехблочных, но и всех других обнуляющих характеристик с числом активных S-блоков до 7 включительно в каждом цикле.

Заметим также, что во втором и четвертом случаях мы отходим от требования 6 стандарта, вводя вместо него другое (вместо $S(\mathbf{x}) \neq S(\mathbf{x} \oplus 11ef00)$ вводится требование $S(\mathbf{x}) \neq S(\mathbf{x} \oplus 00ef11)$ с дополнительным ограничением на еще одну четверку входов).

Таблица 3

Участие S-блоков в формировании входной разности (0, 0, x, y, z, 1, p, q, r, s, t, l, 1, m, n, k, 0, 0) при четырехблочной характеристике							
z	Входные разности первого S-блока (0, 0, x, y, z, 1)	z r s	Входные разности второго S-блока (z, 1, p, q, r, s),	r s m	Входные разноси третьего S-блока (r, s, t, l, 1, m)	m	Входные разности четвертого S-блока (1, m, n, k, 0, 0)
0	(0, 0, 0, 0, 0, 1) → 10_x		(1, 1, 0, 0, 0, 0) → 28_x		(0, 1, 0, 0, 1, 1) → 19_x	0	(1, 0, 0, 0, 0, 0) → 20_x
	(0, 0, 0, 1, 0, 1) → 12_x	1	(1, 1, 0, 1, 0, 0) → $2A_x$	0	(0, 1, 0, 1, 1, 1) → $1B_x$		(1, 0, 0, 1, 0, 0) → 22_x
	(0, 0, 1, 0, 0, 1) → 14_x	0	(1, 1, 1, 0, 0, 0) → $2C_x$	1	(0, 1, 1, 0, 1, 1) → $1D_x$		(1, 0, 1, 0, 0, 0) → 24_x
	(0, 0, 1, 1, 0, 1) → 16_x	0	(1, 1, 1, 1, 0, 0) → $2E_x$	1	(0, 1, 1, 1, 1, 1) → $1F_x$		(1, 0, 1, 1, 0, 0) → 26_x
1	(0, 0, 0, 0, 1, 1) → 11_x		(0, 1, 0, 0, 1, 1) → 19_x		(1, 1, 0, 0, 1, 0) → 29_x	1	(1, 1, 0, 0, 0, 0) → 28_x
	(0, 0, 0, 1, 1, 1) → 13_x	0	(0, 1, 0, 1, 1, 1) → $1B_x$	1	(1, 1, 0, 1, 1, 0) → $2B_x$		(1, 1, 0, 1, 0, 0) → $2A_x$
	(0, 0, 1, 0, 1, 1) → 15_x	1	(0, 1, 1, 0, 1, 1) → $1D_x$	1	(1, 1, 1, 0, 1, 0) → $2D_x$		(1, 1, 1, 0, 0, 0) → $2C_x$
	(0, 0, 1, 1, 1, 1) → 17_x	1	(0, 1, 1, 1, 1, 1) → $1F_x$	0	(1, 1, 1, 1, 1, 0) → $2F_x$		(1, 1, 1, 1, 0, 0) → $2E_x$
			(1, 1, 0, 0, 1, 1) → 39_x		(1, 1, 0, 0, 1, 1) → 39_x		
		1	(1, 1, 1, 0, 1, 1) → $3D_x$	1	(1, 1, 1, 0, 1, 1) → $3D_x$		
		1	(1, 1, 0, 1, 1, 1) → $3B_x$	1	(1, 1, 0, 1, 1, 1) → $3B_x$		
		1	(1, 1, 1, 1, 1, 1) → $3F_x$	1	(1, 1, 1, 1, 1, 1) → $3F_x$		
			(1, 1, 0, 0, 1, 0) → 29_x		(1, 0, 0, 0, 1, 0) → 21_x		
		1	(1, 1, 0, 1, 1, 0) → $2B_x$	1	(1, 0, 0, 1, 1, 0) → 23_x		
		1	(1, 1, 1, 0, 1, 0) → $2D_x$	0	(1, 0, 1, 0, 1, 0) → 25_x		
		0	(1, 1, 1, 1, 1, 0) → $2F_x$	0	(1, 0, 1, 1, 1, 0) → 27_x		
			(1, 1, 0, 0, 0, 1) → 38_x		(1, 0, 0, 0, 1, 1) → 31_x		
		1	(1, 1, 0, 1, 0, 1) → $3A_x$	1	(1, 0, 0, 1, 1, 1) → 33_x		
		0	(1, 1, 1, 0, 0, 1) → $3C_x$	0	(1, 0, 1, 0, 1, 1) → 35_x		
		1	(1, 1, 1, 1, 0, 1) → $3E_x$	1	(1, 0, 1, 1, 1, 1) → 37_x		
			(0, 1, 0, 0, 0, 1) → 18_x		(0, 0, 0, 0, 1, 1) → 11_x		
		0	(0, 1, 0, 1, 0, 1) → $1A_x$	0	(0, 0, 0, 1, 1, 1) → 13_x		
		0	(0, 1, 1, 0, 0, 1) → $1C_x$	0	(0, 0, 1, 0, 1, 1) → 15_x		
		1	(0, 1, 1, 1, 0, 1) → $1E_x$	1	(0, 0, 1, 1, 1, 1) → 17_x		

Таблица 4

1) Входы в таблицы дифференциальных разностей S-блоков	2) Входы в таблицы дифференциальных разностей S-блоков	3) Входы в таблицы дифференциальных разностей S-блоков	4) Входы в таблицы дифференциальных разностей S-блоков	5) Входы в таблицы дифференциальных разностей S-блоков
(1, 1, 0, 0, 0, 0)→28_x	(0, 0, 0, 0, 1, 1)→11_x	(1, 0, 0, 0, 0, 0)→20_x	(0, 0, 0, 0, 0, 1)→10_x	(0, 0, 0, 0, 1, 1)→11_x
(1, 1, 0, 1, 0, 0)→$2A_x$	(0, 0, 0, 1, 1, 1)→13_x	(1, 0, 0, 1, 0, 0)→22_x	(0, 0, 0, 1, 0, 1)→12_x	(0, 0, 0, 1, 1, 1)→13_x
(1, 1, 1, 0, 0, 0)→$2C_x$	(0, 0, 1, 0, 1, 1)→15_x	(1, 0, 1, 0, 0, 0)→24_x	(0, 0, 1, 0, 0, 1)→14_x	(0, 0, 1, 0, 1, 1)→15_x
(1, 1, 1, 1, 0, 0)→$2E_x$	(0, 0, 1, 1, 1, 1)→17_x	(1, 0, 1, 1, 0, 0)→26_x	(0, 0, 1, 1, 0, 1)→16_x	(0, 0, 1, 1, 1, 1)→17_x
(1, 1, 0, 0, 1, 0)→29_x	(0, 1, 0, 0, 1, 1)→19_x	(1, 1, 0, 0, 0, 0)→28_x	(0, 0, 0, 0, 1, 1)→11_x	(1, 1, 0, 0, 0, 0)→28_x
(1, 1, 0, 1, 1, 0)→$2B_x$	(0, 1, 0, 1, 1, 1)→$1B_x$	(1, 1, 0, 1, 0, 0)→$2A_x$	(0, 0, 0, 1, 1, 1)→13_x	(1, 1, 0, 1, 0, 0)→$2A_x$
(1, 1, 1, 0, 1, 0)→$2D_x$	(0, 1, 1, 0, 1, 1)→$1D_x$	(1, 1, 1, 0, 0, 0)→$2C_x$	(0, 0, 1, 0, 1, 1)→15_x	(1, 1, 1, 0, 0, 0)→$2C_x$
(1, 1, 1, 1, 1, 0)→$2F_x$	(0, 1, 1, 1, 1, 1)→$1F_x$	(1, 1, 1, 1, 0, 0)→$2E_x$	(0, 0, 1, 1, 1, 1)→17_x	(1, 1, 1, 1, 0, 0)→$2E_x$

Четыре отмеченные варианта ограничений были проверены моделированием и подтвердили свою эффективность. Как показывают результаты эксперимента, вычислительные затраты на построение таблиц подстановок для DES с дополнительными ограничениями 1 или 2 получаются на порядок меньшими, чем соответствующие затраты при использовании дополнительных ограничений 3 или 4 (в табл. 4 варианты 1-4 расположены в порядке увеличения вычислительной сложности процедуры построения таблиц подстановок). Поэтому ограничение в виде $S(\mathbf{x}) \neq S(\mathbf{x} \oplus 11ef10)$, на ко-

тором остановились и корейские ученые, можно действительно рассматривать в качестве одного из предпочтительных.

Отметим также, что дополнительное ограничение $S(\mathbf{x}) \neq S(\mathbf{x} \oplus 11ef10)$ рассматривается корейскими учеными как достаточное для полной защиты шифра DES от атак дифференциального криптоанализа. Вычислительные эксперименты, проведенные нами, подтверждают этот результат.

Список литературы: 1. *Долгов В.И., Лисицкая И.В., Олейников Р.В.* Принципы защиты алгоритма DES от атак дифференциального криптоанализа. // Радиотехника. 2000. Вып 113. С. 145-157. 2. *Biham E., Shamir A.* Differential Cryptanalysis of the DES- like Cryptosystems, Journal of Cryptology. Vol. 4. P. 3-72. 1991. 3. *Matsui M.* Linear Cryptanalysis Metod for DES Cipher // Pros. Eurocrypt'93. P. 386-397. Norway. 1993. 4. *Kim K.* Construction of DES-like S-boxes Based on Boolean Function Satisfying the SAK. Pros. Of Asiacrypt'91. P. 59-72. Fujiyoshida. Japan. 1991. 5. *Kim K., Park S., Lee S.* Reconstruction of s^2DES *S*-boxes and their Immunity to Differential Cryptanalysis. Pros. of 1993 Korea-Japan Joint Workshop on Information Security fnd Cryptology (JW-ISC'93). Oct. 24-36. Seoul. 1993. 6.. *Knudsen L* Iterative Characteristics of DES and s^2DES. Proc. of Crypto'92. UCSB. 1992. 7. *B. Schneier.* Applied Cryptography. *S*econd Edition: protocols, algorithms, and Source code in C. Published by John Wiley & *S*on*S*. Inc. New York: Chiche*S*ter Bri*S*bane Toronto *S*ingapore. 1996. 758 p. 8. *Барсуков В.С., Дворянкин С.В., Шеремет И.А.* Безопасность связи в каналах телекоммуникаций. М.: Россия. 1993. Т.20. 123 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 6.03.2001

УДК 681.3.06: 519.248.681

И. В. ЛИСИЦКАЯ, канд. техн. наук, А. С. БОНДАРЕНКО, А. И. КОЛЫБЕЛЬНИКОВ

ОБЕСПЕЧЕНИЕ СТОЙКОСТИ ШИФРА DES К АТАКАМ ЛИНЕЙНОГО КРИПТОАНАЛИЗА. ТРЕБОВАНИЯ К ОТБОРУ S-БЛОКОВ, ЗАЩИЩЕННЫХ ОТ АТАК НА ХАРАКТЕРИСТИКИ ОБНУЛЯЮЩЕГО ТИПА, ЧЕТЫРЕХЦИКЛОВЫЕ И ШЕСТИЦИКЛОВЫЕ ИТЕРАТИВНЫЕ АППРОКСИМАЦИИ

Идея построения S-блоков для шифра DES с улучшенными показателями безопасности по сравнению с таблицами, предложенными разработчиками стандарта, и сегодня многими специалистами в Украине воспринимается с известной долей скептицизма. Вместе с тем, исследования, проведенные в научном коллективе АО Институт Информационных Технологий [1-4], показали, что разработчики стандарта реализовали в своих таблицах далеко не потенциальные характеристики защищенности как в отношении дифференциального (ДК), так и линейного криптоанализа (ЛК). Однако наши проработки по этим вопросам, как оказалось, не являются полностью оригинальными. Практически с момента опубликования идей ДК и ЛК подобные исследования начала вести группа корейских ученых [5-10]. Анализ результатов этих публикаций и сопоставление их с результатами наших исследований показывает, что они во многом пересекаются, хотя, как оказалось, при определении критериев отбора таблиц S-блоков, защищенных от атак ЛК, мы не смогли избежать погрешностей. В этой работе мы хотим возвратиться к анализу уже известных результатов в рассматриваемом направлении (по линейному криптоанализу) и привести наши аргументы по их обоснованию. В дальнейшем полагается, что читатель знаком с работами [11-13], в которых излагаются основополагающие идеи ДК и ЛК, а также с отмеченными выше работами корейских ученых.

Подчеркнем здесь, что общей методологической основой формирования корейскими учеными условий повышения сопротивляемости шифра DES к атакам ЛК, стало выполнение требований к S-блокам, сформулированных разработчиками стандарта [14,5], и решалась задача определения дополнительных ограничений к отбору таблиц подстановок для шифра DES. И мы придерживаемся этой же точки зрения.

В качестве основы дальнейших рассуждений приведем формулировки дополнительных требований к отбору S-блоков в редакции, предложенной корейскими учеными. В работах [7-9] предлагается пять дополнительных критериев (условий) обеспечения невосприимчивости шифра DES к атакам ЛК.

Условие *L*-1 (условие единообразия таблиц распределения линейных аппроксимаций): Значения выходов таблиц распределения линейных аппроксимаций (ТРЛА) для всех S-блоков должны быть меньше, чем 16;

Условие *L*-2 (условие перекрытия 3R итеративных аппроксимаций): $S(x) \neq S(x \oplus 11efg0)$ для любых *efg*.

Здесь $S(x)$ –значение 6-битного входа x произвольного S-блока;

Условие *L*-3 (условие перекрытия 4R итеративных аппроксимаций): Для ТРЛА S-блоков необходимо выполнить следующие (общее число 18 случаев) условия:

- *S*1-блок : $NS_1(4, 4) = NS_1(2, 2) = 0$,
- *S*2-блок : $NS_2(4, 4) = NS_2(2, 1) = 0$,
- *S*3-блок : $NS_3(8, 4) = NS_3(4, 8) = 0$,
- *S*4-блок : $NS_4(8, 4) = NS_4(2, 2) = 0$,
- *S*5-блок : $NS_5(16, 1) = NS_5(8, 8) = NS_5(2, 4) = 0$,
- *S*6-блок : $NS_6(16, 4) = NS_6(4, 8) = NS_6(2, 2) = 0$,
- *S*7-блок : $NS_7(4, 8) = NS_7(2, 1) = 0$,
- *S*8-блок : $NS_8(16, 1) = NS_8(2, 4) = 0$.

Здесь $NS(\alpha, \beta)$ – элементы ТРЛА, которые определяются следующим образом:

$$NS(\alpha,\beta) = \{x \in GF(2)^6 \mid x \bullet \alpha = S(x) \bullet \beta\} - 32,$$

где $\alpha \in GF(2)^6$; $NS(\alpha,\beta)$ – маски входов (по строкам и столбцам) в ТРЛА *S*-блока; # обозначает число раз выполнения соотношения в фигурных скобках. Для элементов $x, y \in GF(2)^n$, $x \bullet y$, обозначает скалярное произведение x и y.

Из остальных вариантов ограничений, рассмотренных корейскими учеными, мы представим их окончательные предложения.

Условие *L*-4 (условие не обнаружения 4R итеративных аппроксимаций): Для $W(\alpha)$, $W(\beta) \le 2$, необходимо, чтобы $|NS(\alpha,\beta)| \le 10$, где, как и ранее, $W(\alpha)$, и $\beta \in GF(2)^4$, $W(\alpha)$,-вес по Хэммингу α.

Условие *L*-5 (условие перекрытия 5R итеративных аппроксимаций): Если $\alpha \neq 10_x$, то для $W(\alpha) = 1$, и $W(\beta_1 \oplus \beta_2) = 1$:

$$|NS_k(\alpha,\beta_1) \cdot NS_k(\alpha,\beta_2)| \le 48$$

Если же $\alpha = 10_x$, то

$$|NS_k(\alpha,\beta_1) \cdot NS_k(\alpha,\beta_2)| \le 48 \text{ для } \beta_1 \oplus \beta_2 = 1,$$

$$|NS_k(\alpha,\beta_1) \cdot NS_k(\alpha,\beta_2)| \le 48, \text{ для } \beta_1 \oplus \beta_2 = 4$$

для $k = 5,8$ и $l = 6$.

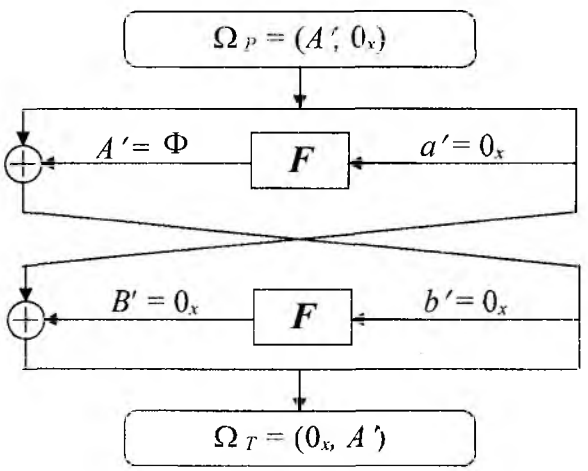
В этой статье мы остановимся на обосновании этих условий.

Прежде всего хотелось бы здесь отметить, что представленные в нашей работе [4] результаты не являются полным решением поставленной в ней задачи. Мы в целом правильно исходили из условия, что в аппроксимационные характеристики не должны входить цикловые преобразования, когда нулевая входная маска 0_x (здесь α_x: шестнадцатеричное значение α) сочетается с выходной маской ненулевого типа (вероятность перехода нулевого входа ТРЛА в ненулевой равна нулю). Но если для одноблочной характеристики это действительно справедливо, то для двух активных S-блоков (S-блоков с ненулевыми входными масками), участвующих в линейном соотношении, может возникнуть ситуация, когда входы соседних S-блоков, за счет расширяющей *E*-подстановки имеют совпадающие биты и тогда для масок, пропускающих эти биты, они в результирующем линейном соотношении будут компенсировать друг друга (входная маска будет эквивалентна нулевой). Именно этот эффект и использован в атаке Девиса, описанной в [12]. Но тогда можно говорить о формировании двухцикловой итеративной характеристики с нулевым входом активного цикла. В обозначениях корейских ученых речь идет об итеративной характеристике типа

$$\Phi \leftarrow 0_x,$$

которая названа ими одноцикловым итеративным линейным выражением. Ее лучше представить в естественном двухцикловом изображении (рис. 1), как это сделано в работе [12].

И здесь мы подходим к обоснованию одного из представленных выше дополнительных ограничений (условий). Нетрудно убедиться, что выполнение условия *L*-2 позволяет рассматривать эти характеристики как не реализуемые (имеющие нулевую вероятность).



$$\Phi \leftarrow 0_x$$
$$0_x \leftarrow 0_x$$

$$17 \oplus 18 \leftarrow 4 \oplus 5$$
$$0_x \leftarrow 0_x$$

Рис. 1

Действительно, в рассматриваемом случае речь идет об использовании трех сочетаний ненулевых входов двух смежных (соседних) S-блоков:

$$S_i(1_x,\beta'), S_{i+1}(10_x,\beta''),$$
$$S_i(2_x,\beta'), S_{i+1}(20_x,\beta''),$$
$$S_i(3_x,\beta'), S_{i+1}(30_x,\beta'').$$

Характеристики, использующие первые два варианта входов, нереализуемы ввиду того, что для ТРЛА всех S-блоков, отобранных по требованиям разработчиков стандарта, выполняется условие:

$$NS_i(1_x,\beta)=NS_i(20_x,\beta)=0 \qquad (1)$$

при любых β. Перекрытие третьей характеристики как раз и обеспечивается выполнением условия L-2, в котором наряду с отмеченными ограничениями разработчиков стандарта ТРЛА S-блоков удовлетворяют дополнительному условию:

$$NS_i(30_x,\beta)=0.$$

При выполнении указанных условий одновременно "перекрываются" и все другие характеристики, использующие нулевой вход активного цикла (для этих характеристик выходы (маски выходов) Г – это выходы S-блоков, формирующие общие входы смежных S-блоков очередного цикла). Четырехцикловая итеративная характеристика такого типа представлена на рис. 2 под номером 0. В результате итеративные характеристики с нулевым входом активного цикла при выполнении требования L-2 запрещены.

Отметим здесь, что рассмотренное ограничение использовалось и в наших исследованиях. Но даже, если исключать из рассмотрения характеристики с нулевым входом активного цикла, то утверждение в работе [4] о том, что в ней сформирована характеристика минимального типа в том смысле, что не существует итеративных характеристик, содержащих большее число циклов тождественного типа. все равно является неверным.

Действительно, кроме рассмотренных выше, существуют итеративные характеристики с числом циклов меньшим, чем 8, в том числе и характеристики, содержащие циклы тождественного типа. Возможные варианты таких характеристик, удовлетворяющие условиям сшивки аппроксимаций соседних циклов, изображенные в манере публикации [5], представлены на рис. 2. И здесь мы подходим к обоснованию одного из представленных выше дополнительных ограничений (условий). Нетрудно убедиться, что выполнение условия L-2 позволяет рассматривать эти характеристики как не реализуемые (имеющие нулевую вероятность).

1	2	3	4
$\Gamma \leftarrow \Phi$ $\Phi \leftarrow 0_x$ $\Gamma \leftarrow \Phi$ $0_x \leftarrow 0_x$	$\Gamma \leftarrow \Phi$ $\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$ $\Psi \leftarrow \Phi$ $\Theta \leftarrow \Gamma \oplus \Psi$	$\Gamma \leftarrow \Phi$ $\Phi \leftarrow \Gamma$ $0_x \leftarrow 0_x$ $\Phi \leftarrow \Gamma$ $\Gamma \leftarrow \Phi$ $0_x \leftarrow 0_x$	$\Gamma \leftarrow \Phi$ $\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$ $\Psi \leftarrow \Theta$ $\Phi \leftarrow \Gamma$ $\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$ $\Theta \leftarrow \Psi$
4-х цикловая итеративная характеристика	*4-х цикловая итеративная характеристика3*	*6-ти цикловая итеративная характеристика*	*6-ти цикловая итеративная характеристика*

Рис. 2

Предложенные в работе [4] ограничения касаются восьмицикловых итеративных характеристик. Они перекликаются с требованиями, предлагаемыми корейскими учеными, но являются, как теперь стало понятно, недостаточными. Необходимо еще защититься и от атак, построенных на использовании итеративных характеристик с меньшим числом циклов, которые не рассмотрены нами. Основной задачей этой работы и является изучение условий "перекрытия" характеристик, представленных на рис. 2.

Прежде всего заметим, что, как следует из рис. 2, в принципе возможны как итеративные характеристики, состоящие только из активных циклов, у которых S-блоки всех циклов участвуют в построении линейной аппроксимации, так и характеристики с переходами $0_x \leftarrow 0_x$, содержащие циклы "тождественного" типа. Другая особенность рассматриваемых характеристик заключается в том, что значения входов в циклы (левых частей характеристик) фиксированы, в то время, как значения

выходов (правых частей характеристик) являются свободными (в пределах используемой композиции выходов, задействованных S-блоков). Общим для всех итеративных характеристик с числом циклов большим 2, так как возможны линейные итеративные аппроксимационные характеристики только с четным числом циклов, является использование при их построении циклических переходов между одноименными S-блоками.

Изучим сначала возможности и требования по перекрытию четырехцикловой итеративной характеристики под номером 1 (рис. 2). Расчеты показывают, что необходимо "перекрыть" характеристики такого типа с числом активных S-блоков (приходящихся на четырехцикловую характеристику) меньшим семи (шесть и меньше):

$$\left[\left(\frac{16}{64}\right)^6 \cdot 2^5\right]^4 \cdot 2^3 = 2^{-25}, \left[\left(\frac{16}{64}\right)^7 \cdot 2^6\right]^4 \cdot 2^3 = 2^{-29}.$$

Как следует из рис. 2, главной особенностью четырехцикловых характеристик рассматриваемого типа является использование двух пар циклов с идентичными входами. Это значит, что нас должны интересовать итеративные характеристики 1 (рис. 1), составленные из двух пар одноблочных циклов или пары одноблочных и пары двухблочных циклов. Еще одной особенностью рассматриваемых характеристик следует считать то, что они допускают с точностью до порядка следования циклов еще два варианта представления, которые вместе с соответствующими им графами переходов приведены на рис. 3.

Рассмотрим сначала условия реализации характеристики 1.1 (рис. 3). Отметим, что для ее осуществления, как следует из рис. 3, одновременно должны выполняться два перехода: $\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$ и $\Theta \leftarrow \Gamma \oplus \Psi$, а, следовательно, должен выполняться один из переходов $\Phi \leftarrow \Psi$ и $\Phi \leftarrow \Gamma$ или оба эти перехода должны выполняться вместе. В результате, с учетом существования для данной характеристики переходов $\Gamma \leftarrow \Phi$ и $\Psi \leftarrow \Phi$, приходим к выводу, что для получения ЛАХ рассматриваемого типа должен выполняться один из циклических переходов $\Phi \leftarrow \Gamma \leftarrow \Phi$, $\Phi \leftarrow \Psi \leftarrow \Phi$ или оба вместе, и при этом должны быть допустимыми переходы $\Theta \leftarrow \Gamma$ и $\Theta \leftarrow \Psi$ (Θ не входит в циклический переход и поэтому может выбираться из выходных битов S-блока, имеющего вход $\Gamma \oplus \Psi$). Именно из этих соображений построен граф переходов для этой характеристики, представленный под соответствующим номером на рис. 3.

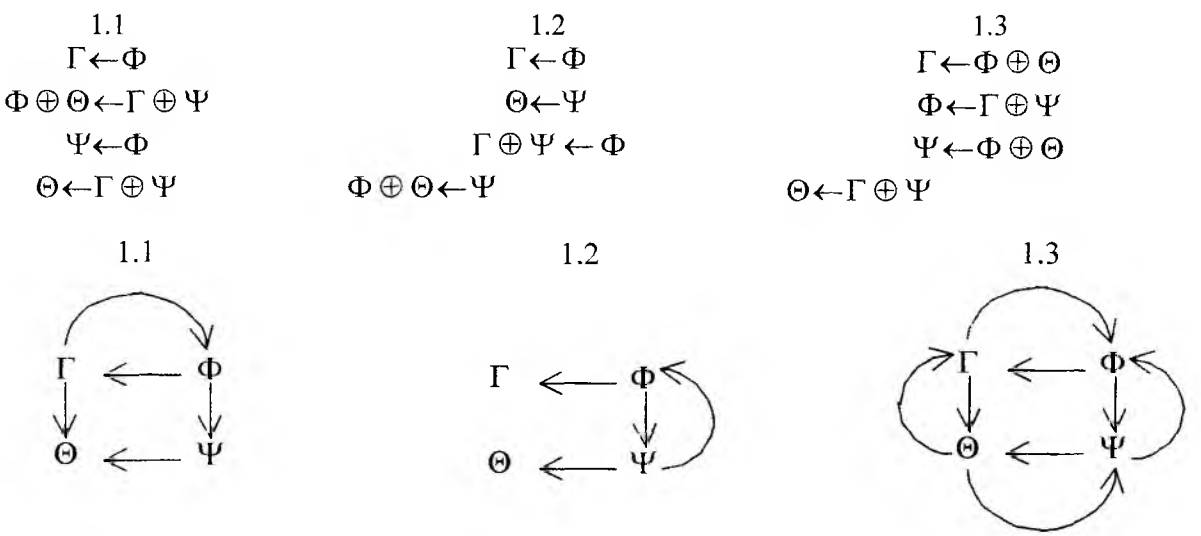


Рис. 3

Будем сначала считать, что каждый из символов в обозначении характеристик представляет собой один бит циклового входа или выхода. Назовем такие характеристики минимальными, только теперь в это понятие будем вкладывать тот смысл, что они формируются с помощью минимального числа ненулевых битов. Очевидно, что для характеристик минимального типа переходы $\Gamma \leftarrow \Phi$ и $\Psi \leftarrow \Phi$ одно-

блочные. Тогда, если Ф - это входы однотипных S-боков, то из-за *P*-перестановки, использованной в шифре DES, выходы Г и Ψ могут стать входами только разных S-блоков, и, следовательно, переходы $\Theta \leftarrow \Gamma \oplus \Psi$ и $\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$ являются двухблочными. Получается, что должен быть справедливым двухблочный переход (цикл) $\Theta \leftarrow \Gamma \oplus \Psi$, для которого два различных (разнесенных) S-блока имеют общий однобитный выход Θ, что для шифра DES может быть только в том случае, если Θ – это выход одного из S-блоков, а выход второго S-блока имеет нулевое значение. Такие переходы для шифра DES имеют вероятность, равную нулю. Но выход Θ в рассматриваемой характеристике имеет свободу выбора, и поэтому он может быть выбран многобитным (для характеристики не минимального типа). Отметим теперь, что многобитными (в пределах выходов задействованных S-блоков) могут быть и выходы (маски выходов) одноблочных циклов. Результирующая характеристика получается шестиблочной и, следовательно, должна быть запрещена. Примеры построения четырехцикловых характеристик типа 1.1 приведены на рис. 4. Справа для каждой из характеристик построчно приведены значения числа активных S-блоков, участвующих в формировании цикла.

1.1.1

(14,25),3 ← 17 1

(2,9,13,18,23,31),17,28,31 ← 3,8 2

(14,25),8 ← 17 1

(2,9,13,18,23,31),17,28,31 ← 3,8 2

1.1.2

(14,25),3 ← 17,18 1

(2,9,13,23,31),17,18,28,31 ← 3,8 2

(14,25),8 ← 17,18 1

(2,9,13,23,31),28,31 ← 3,8 2

1.1.3

28 ← 5 1

(21,27),5,15 ← 28,31 1

31 ← 5 1

(21,27),15 ← 28,31 1

1.1.4

(2,9,13,17,18,23),28 ← 5 2

(21,27),5,15 ← 28,31 1

(2,9,13,17,18,23),31 ← 5 2

(21,27),15 ← 28,31 1

1.2.1

(9,31),23 ← 3 1

(14,25),8 ← 17 1

(9,31),17,23 ← 3 1

(14,25),3,8 ← 17 1

1.2.2

(9,31),23 ← 3 1

(1,10,14,20,25,26),8 ← 17 2

(9,31),17,23 ← 3 1

(1,10,14,20,25,26),3,8 ← 17 2

1.2.3

4 ⊕11 ⊕ 19 ← 21 1

5 ⊕ 15 ⊕ 27 ← 29 1

4 ⊕11 ⊕ 19 ⊕ 29 ← 21 1

5 ⊕ 15 ⊕ 21 ⊕27 ← 29 1

1.2.4

4 ⊕11 ⊕ 19 ← 21 (20.22,23,24,25) 1

5 ⊕ 15 ⊕ 27 ← 29 1

4 ⊕11 ⊕ 19 ⊕ 29 ← 21 1

5 ⊕ 15 ⊕ 21 ⊕27 ← 29 (28,30,31,32,1)1

Рис. 4

Если Ф – это входы разных S-боков, где Ф – это общий бит входа двух смежных S-блоков, то выходы Г и Ψ могут быть только входами в однотипные S-блоки и тогда четырехцикловая характеристика минимального типа будет одноблочной. В этом случае при построении характеристик используются пары циклических (однобитных) переходов, имеющих общий бит (Ф). Пример построения такой характеристики также приведен на рис. 4 (см. пример 1.1.3). При использовании для построения характеристик циклических переходов с большим, чем в случае однобитного циклического перехода числом S-блоков, они могут содержать и больше четырех задействованных S-блоков (см. пример 1.1.4).

Рассмотрим теперь условия реализации характеристики 1.2 (рис. 3). Из графа переходов, соответствующего этой характеристике, также представленного на рис. 3, следует, что эта характеристика строится с использованием циклического перехода $\Phi \leftarrow \Psi \leftarrow \Phi$. Легко убедиться, что для однобитного циклического перехода $\Phi \leftarrow \Psi \leftarrow \Phi$ существует характеристика минимального типа, которая будет одноблочной. Примеры построения подобных характеристик также приведены на рис. 4.

Именно одноблочные характеристики будут представлять наибольшую опасность во всех рассмотренных случаях, и поэтому они должны быть перекрыты в первую очередь. Как показывает анализ, ограничения, предложенные корейскими учеными, не учитывают рассмотренные характеристики. Некоторые из возможных характеристик, правда, попадают под ограничение *L*-5, которое введено для 10-цикловых характеристик.

Для перекрытия указанных выше четырехцикловых итеративных характеристик минимального и не минимального типа можно воспользоваться тем, что все они, как уже было отмечено выше, состоят из двух пар циклов с идентичными входами. Причем, как следует из приведенных рассуждений и примеров (рис.4), входы в пары циклов могут быть как однобитными, так и двухбитными. С учетом принципа формирования линейных аппроксимаций для перекрытия таких характеристик можно воспользоваться ограничением, подобным первой части Условия *L*-5, которое предлагается корейскими учеными для перекрытия десятицикловых характеристик (5R итеративных линейных аппроксимаций). Это условие, однако, мы переформулируем в виде двух новых. В дальнейшем будем пользоваться обозначениями дополнительных условий в виде символов У (Условие) со своими порядковыми номерами в отличии от символов *L*, использованных корейскими учеными. Считая, что первые два условия полностью повторяют предложения корейских ученых, закрепим за ними номера У-1 (*L*-1) и У-2 (*L*-2). Тогда для последующих двух ограничений, относящихся к четырехцикловым итеративным характеристикам, пользуясь оговоренной символикой, можем записать:

Условие У-3′ - условие перекрытия четырехцикловых итеративных характеристик с однобитными входами в однотипные S-блоки. Элементы ТРЛА для пар S-блоков, имеющих входные и выходные маски, удовлетворяющие условию $W(\alpha) = 1$, $W(\beta_1 \oplus \beta_2) = 1$, должны подчиняться ограничению:

$$\left|NS_k(\alpha,\beta_1)\cdot NS_k(\alpha,\beta_2)\right| \le 96.$$

Это условие практически только по форме напоминает условие *L*-5, однако по содержанию является совершенно другим.

Условие У-3″ (условие перекрытия четырехцикловых итеративных характеристик с однобитными входами в различные S-блоки). S-блоки для шифра DES должны выбираться так, чтобы пары элементов ТРЛА, имеющих входные и выходные маски, удовлетворяющие условию $W(\alpha) = 2$, $W(\beta_1 \oplus \beta_2) = 2$, подчинялись ограничению:

$$\left|NS_k(\alpha,\beta_1)\cdot NS_k(\alpha,\beta_2)\right| \le 96.$$

При выполнении указанных условий оценка максимального значения вероятности для 16-ти цикловой характеристики, использующей одноблочные четырехцикловые характеристики рассматриваемого типа, приводит к результату:

$$\left[\left(\frac{96}{64^2}\right)^2\cdot 2^3\right]^4\cdot 2^3 = 2^{-28,3}.$$

Но здесь рассмотрен случай, когда итеративная характеристика повторяется соответствующее число раз без изменений. В то же время в приведенных расчетах не учтена возможность свободного выбора значений правой части ЛАХ для ее начального и заключительного циклов.

Действительно, шестнадцатицикловая характеристика (первая и последняя четырехцикловые характеристики) допускает свободу выбора правой части (входа и выхода), как это проиллюстрировано характеристикой 1.2.4 на рис. 4. Здесь одновременно в одной четырехцикловой характеристике в скобках учтены возможности произвольного задания начального и конечного циклов всей шестнадцатицикловой характеристики и, следовательно, приведенное выше ограничение для первого и последнего циклов не "сработает".

Расчет вероятности такой 16-ти цикловой характеристики (с учетом свободного выбора значений ее входа и выхода) приводит к результату:

$$\left[\left(\frac{96}{64^2}\right)^2\cdot 2^3\right]^3\cdot\left[\left(\frac{96}{64^2}\right)\cdot\left(\frac{16}{64}\right)^2\cdot 2^3\right]\cdot 2^3 = 2^{-26,9},$$

чего уже оказывается недостаточным для защиты от атак на подобные характеристики. Поэтому условия У-3′ и У-3″ необходимо ужесточить. Вместо этих условий предлагается воспользоваться ограничением вида:

Условие У-3 (условие перекрытия четырехцикловых итеративных характеристик). S-блоки для шифра DES должны выбираться таким образом, чтобы для пары элементов ТРЛА, имеющих входные и выходные маски, удовлетворяющие условиям $W(\alpha) = 1$, $W(\beta_1 \oplus \beta_2) = 1$ или $W(\alpha) = 2$, $W(\beta_1 \oplus \beta_2) = 2$, подчинялись ограничению:

$$\left|NS_k(\alpha,\beta_1) \cdot NS_k(\alpha,\beta_2)\right| \le 80.$$

В этом случае, учитывая свободу в выборе значений ЛАХ на ее начальном и заключительном циклах, приходим к оценке вероятности результирующей характеристики:

$$\left[\left(\frac{80}{64^2}\right)^2 \cdot 2^3\right]^3 \cdot \left[\left(\frac{80}{64^2}\right) \cdot \left(\frac{16}{64}\right)^2 \cdot 2^3\right] \cdot 2^3 = 2^{-28,7}.$$

Этот результат уже является вполне приемлемым.

Что касается четырехцикловой характеристики под номером 3 на рис. 3, то здесь не возникает проблем. Из графа ее переходов, приведенного на этом же рисунке, видно, что она для шифра DES не реализуема (для нее нужно, чтобы одновременно выполнялись четыре однобитных циклических перехода с общими битами).

Перейдем к изучению условий перекрытия шестицикловых характеристик, представленных на рис. 2 под номерами 2 и 3.

Легко убедиться, что в этом случае представляют интерес характеристики с числом активных S-блоков, приходящихся на трехцикловую характеристику (половину шестицикловой итеративной характеристики), не превышающем шести (пять и меньше).

И здесь мы приходим к одному из вариантов обоснования условия L-1.

Действительно, трехблочные переходы в каждом из активных циклов не опасны, так как даже при максимально возможном значении вероятности перехода (значения элемента ТРЛА) для каждого из задействованных S-блоков, равном $\frac{16}{64}$, для вероятности всей 15-цикловой характеристики (пятикратном повторе трехцикловой характеристики, а еще более точно – двухкратном повторе шестицикловой итеративной характеристики, продолженной еще на три цикла) имеем:

$$2^4 \cdot \left[2^5 \cdot \left(\frac{16}{64^2}\right)^6\right]^5 = 2^{-7 \cdot 5 + 4} = 2^{-31}$$

в то время как для пятиблочной трехцикловой характеристики рассматриваемого типа соответственно получим:

$$2^4 \cdot \left[2^4 \cdot \left(\frac{16}{64^2}\right)^5\right]^5 = 2^{-6 \cdot 5 + 4} = 2^{-26},$$

что уже является недопустимым значением. Как известно [13], граничным значением вероятности результирующей характеристики, обеспечивающим более высокую эффективность линейного криптоанализа по сравнению с переборной атакой, является $\sqrt{2^{-56}} = 2^{-28}$.

Заметим далее, что при значении вероятности перехода (значения элемента ТРЛА) для каждого из задействованных S-блоков, равном $\frac{18}{64}$, для 16-цикловой характеристики (пятикратном повторении трехцикловой характеристики, а более точно – двукратном повторении шестицикловой итеративной характеристики, продолженной еще на три цикла с дополнительным циклом тождественного типа) с тремя S-блоками в каждом из активных циклов, имеем:

$$2^4 \cdot \left[2^5 \cdot \left(\frac{18}{64}\right)^6\right]^5 = 2^{-25,9}.$$

Это значение уже следует рассматривать как превышающее допустимое (т.е. значение вероятности $\frac{16}{64} = \frac{1}{4}$ для элементов ТРЛА в этом случае действительно является граничным).

Отметим, наконец, что при использовании одноблочной 16-цикловой характеристики рассматриваемого вида для максимально допустимых значений соответствующих элементов ТРЛА S-блоков можно получить оценку:

$$2^4\left[2(p)^2\right]^5 \le 2^{-28} \rightarrow p \le \sqrt[10]{2^{-37}} = 2^{-3,7}. \qquad (2)$$

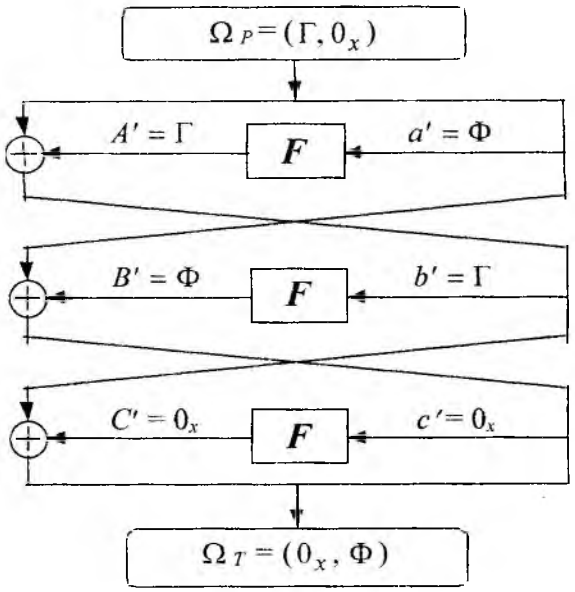
Это значит, что вероятность одноцикловой одноблочной характеристики в рассматриваемом случае ограничена значением 4.

Изучение возможностей атак, строящихся с использованием шестицикловых итеративных характеристик, начнем с рассмотрения трехцикловой характеристики под номером 2 на рис. 2 (половину шестицикловой характеристики). Графическое представление этой характеристики и варианты ее компактного изображения приведены на рис. 5. Сначала рассмотрим характеристики минимального типа, под которыми будем, как уже говорилось ранее, понимать характеристики, использующие для своего построения минимальное число битов (каждый символ – один бит входа или выхода). При такой договоренности характеристика под номером 2.1, рис. 5 будет состоять из одноблочных ненулевых циклов (1+1+0), характеристика 2 имеет вид 1+2+0, т.е. один из активных циклов является одноблочным, а второй – двухблочным, а характеристика 3 может включать 3 или четыре S-блока (имеет вид 1+2+0, 2+1+0 или 2+2+0). Естественно, существуют трехцикловые (шестицикловые) характеристики и с большим числом S-блоков, которые мы рассмотрим позднее.

Прежде всего отметим, что характеристики, интересующего нас типа, могут быть построены на основе использования однобитных переходов S-блоков (переходов одного входного бита маски в один выходной). Напомним, что из-за *P*-перестановки, используемой в шифре DES каждый S-блок может сформировать (инициировать) только однобитные входы S-блоков очередного цикла. Эти S-блоки, в свою очередь, могут только своими однобитными выходами сформировать двухбитный вход исходного (одного) S-блока (сформировать переход $\Gamma \oplus \Psi \leftarrow \Phi$ или переход $\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$).

На рис. 6 приведено распределение битов 32-битного полублока в каждом из циклов преобразования шифра DES, а возможные варианты однобитных переходов, построенных на основе изучения распределения битов рис. 6, перечислены в табл. 1.Пользуясь данными табл. 1 можно рассмотреть для каждого S-блока все переходы, которые могут быть задействованы при построении трехцикловых (шестицикловых) итеративных характеристик. Заметим далее, что переходы с входными масками 1_x и 20_x могут быть сразу исключены из рассмотрения, так как для S-блоков, отобранных по требованиям разработчиков, выполняется условие (1). Тогда для первого S- блока возможны три однобитных перехода $S_1 \Leftrightarrow S_1$:

$$3 \rightarrow 17 \rightarrow 3 : S_1(4_x, 4_x) \Leftrightarrow S_5(10_x, 1_x),$$
$$4 \rightarrow 23 \rightarrow 4 : S_1(2_x, 2_x) \Leftrightarrow S_6(4_x, 8_x),$$
$$5 \rightarrow 31 \rightarrow 5 : S_1(1_x, 1_x) \Leftrightarrow S_8(4_x, 8_x).$$



2.1

$\Gamma \leftarrow \Phi$

$\Phi \leftarrow \Gamma$

$0_x \leftarrow 0_x$

2.2

$\Gamma \oplus \Psi \leftarrow \Phi$

$\Phi \leftarrow \Gamma \oplus \Psi$

$0_x \leftarrow 0_x$

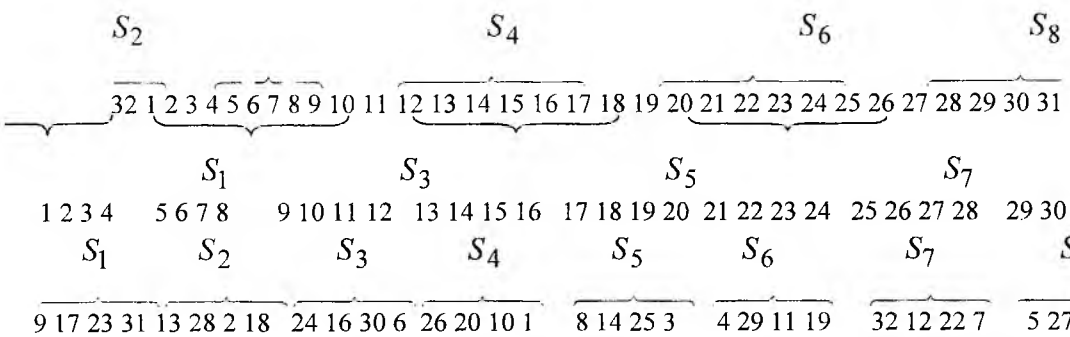
2.3

$\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$

$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$

$0_x \leftarrow 0_x$

Рис. 5



Биты на входах S-блоков и после расширения E

Биты на выходах S-блоков

P- перестановка

Рис. 6

Таблица 1

S-блок S_1	Задействованные переходы	S-блок S_2	Задействованные переходы	S-блок S_3	Задействованные переходы	S-блок S_4	Задействованные переходы
17→1	$S_4(1_x,1_x)$	13→6	$S_3(1_x,1_x)$	16→8	$S_5(20_x,8_x)$	26→12	$S_7(8_x,4_x)$
9→2	$S_2(1_x,2_x)$	28→7	$S_7(2_x,1_x)$	16→10	$S_4(2_x,2_x)$	20→14	$S_5(2_x,4_x)$
17→3	$S_5(10_x,1_x)$	28→5	$S_8(20_x,8_x)$	24→11	$S_6(2_x,2_x)$	1→15	$S_8(1_x,2_x)$
23→4	$S_6(4_x,8_x)$	18→8	$S_5(8_x,8_x)$	24→12	$S_7(20_x,4_x)$	10→16	$S_3(8_x,4_x)$
31→5	$S_8(4_x,8_x)$	2→9	$S_1(8_x,8_x)$	6→13	$S_2(8_x,8_x)$	1→17	$S_1(10_x,4_x)$
S-блок S_5	Задействованные переходы	S-блок S_6	Задействованные переходы	S-блок S_7	Задействованные переходы	S-блок S_8	Задействованные переходы
8→16	$S_3(20_x,4_x)$	29→21	$S_8(10_x,1_x)$	12→24	$S_3(2_x,8_x)$	5→28	$S_2(10_x,4_x)$
3→17	$S_1(4_x,4_x)$	29→22	$S_7(1_x,2_x)$	12→26	$S_4(20_x,8_x)$	21→29	$S_6(10_x,4_x)$
8→18	$S_2(2_x,1_x)$	4→23	$S_1(2_x,2_x)$	32→27	$S_8(2_x,4_x)$	5→31	$S_1(1_x,1_x)$
25→19	$S_6(1_x,1_x)$	11→24	$S_3(4_x,8_x)$	7→28	$S_2(4_x,4_x)$	27→32	$S_7(4_x,8_x)$
14→20	$S_4(8_x,4_x)$	19→25	$S_5(4_x,2_x)$	22→29	$S_6(8_x,4_x)$	15→1	$S_4(4_x,1_x)$

Для того, чтобы сделать одноблочные и двухблочные характеристики, построенные с помощью этих переходов не опасными, достаточно наложить ограничения на два перехода: $S_5(10_x,1_x)$, $S_6(4_x,8_x)$. Остающиеся однобитные переходы включают в себя вход в S-блок с маской 1_x и, следовательно, как уже отмечалось выше, вероятность этого перехода равна нулю.

Для второго S-блока остаются три однобитных перехода $S_2 \Leftrightarrow S_2$:

$$7 \to 28 \to 7: S_2(4_x,4_x) \Leftrightarrow S_7(2_x,1_x),$$
$$8 \to 18 \to 8: S_2(2_x,1_x) \Leftrightarrow S_5(8_x,8_x),$$
$$9 \to 2 \to 9: S_2(1_x,2_x) \Leftrightarrow S_1(8_x,8_x),$$

и здесь по аналогии с предыдущим случаем достаточно наложить ограничения (запретить) на два перехода $S_7(2_x,1_x)$, $S_5(8_x,8_x)$.

Аналогично для переходов $S_3 \Leftrightarrow S_3$ имеем:

$$10 \to 16 \to 10: S_3(8_x,4_x) \Leftrightarrow S_4(2_x,2_x),$$
$$11 \to 24 \to 11: S_3(4_x,8_x) \Leftrightarrow S_6(2_x,2_x),$$
$$13 \to 6 \to 13: S_3(1_x,1_x) \Leftrightarrow S_2(8_x,8_x),$$

и здесь достаточно наложить ограничения на переходы $S_4(2_x,2_x)$, $S_6(2_x,2_x)$.

В случае $S_4 \Leftrightarrow S_4$ получаем четыре варианта переходов:

$$12 \to 26 \to 12: S_4(20_x,8_x) \Leftrightarrow S_7(8_x,4_x),$$
$$14 \to 20 \to 14: S_4(8_x,4_x) \Leftrightarrow S_5(2_x,4_x),$$
$$16 \to 10 \to 16: S_4(2_x,2_x) \Leftrightarrow S_3(8_x,4_x),$$
$$17 \to 1 \to 17: S_4(1_x,1_x) \Leftrightarrow S_1(10_x,4_x).$$

Здесь опять достаточно наложить ограничения на два перехода: $S_5(2_x,4_x)$ и $S_3(8_x,4_x)$.

Действительно, оставшиеся два перехода имеют нулевые вероятности (содержат переходы с входными масками 1_x и 20_x), если их рассматривать как однобитные. Если эти однобитные переходы рассматривать совместно как двухбитный вход $S_4(21_x,8_x)$, то легко убедиться, что для таблиц стандарта выполняется условие: $NS_4(21_x,\beta_x) = 0$, т.е. эта характеристика для S-блоков, отобранных по требованиям разработчиков стандарта, также не реализуема.

Для пятого S-блока $S_5 \Leftrightarrow S_5$ имеем:

$$17 \to 3 \to 17: S_5(10_x,1_x) \Leftrightarrow S_1(4_x,4_x),$$
$$18 \to 8 \to 18: S_5(8_x,8_x) \Leftrightarrow S_2(2_x,1_x),$$
$$20 \to 14 \to 20: S_5(2_x,4_x) \Leftrightarrow S_4(8_x,4_x),$$

Возможные варианты однобитных переходов S-блоков и здесь должны быть перекрыты все три перехода $S_1(4_x,4_x)$, $S_2(2_x,1_x)$ и $S_4(8_x,4_x)$.

Для шестого S-блока получаем однобитные переходы $S_6 \Leftrightarrow S_6$:

$$19 \to 25 \to 19: S_6(1_x,1_x) \Leftrightarrow S_5(4_x,2_x),$$
$$21 \to 29 \to 21: S_6(10_x,4_x) \Leftrightarrow S_8(10_x,1_x),$$
$$23 \to 4 \to 23: S_6(4_x,8_x) \Leftrightarrow S_1(2_x,2_x),$$
$$24 \to 11 \to 24: S_6(2_x,2_x) \Leftrightarrow S_3(4_x,8_x).$$

Достаточно наложить ограничения на три перехода $S_3(4_x,8_x)$, $S_1(2_x,2_x)$ и $S_8(10_x,1_x)$ (оставшийся переход не реализуем).

Для седьмого S-блока однобитные переходы $S_7 \Leftrightarrow S_7$:

$$24 \to 12 \to 24: S_7(20_x,4_x) \Leftrightarrow S_3(2_x,8_x),$$
$$27 \to 32 \to 32: S_7(4_x,8_x) \Leftrightarrow S_8(2_x,4_x),$$
$$28 \to 7 \to 28: S_7(2_x,1_x) \Leftrightarrow S_2(4_x,4_x),$$
$$29 \to 22 \to 29: S_7(1_x,2_x) \Leftrightarrow S_6(8_x,4_x).$$

Достаточно наложить ограничения на два перехода $S_8(2_x,4_x)$ и $S_2(4_x,4_x)$ (по аналогии с четвертым S-блоком).

Совершенно аналогичная ситуация возникает для восьмого S-блока с однобитными переходами $S_8 \Leftrightarrow S_8$:

$$28 \to 5 \to 28: S_8(20_x,8_x) \Leftrightarrow S_2(10_x,4_x),$$
$$29 \to 21 \to 29: S_8(10_x,1_x) \Leftrightarrow S_6(10_x,4_x),$$
$$32 \to 27 \to 32: S_8(2_x,4_x) \Leftrightarrow S_7(4_x,8_x),$$
$$1 \to 15 \to 1: S_8(1_x,2_x) \Leftrightarrow S_4(4_x,1_x).$$

Здесь достаточно наложить ограничения на два перехода $S_6(10_x,4_x)$ и $S_7(4_x,8_x)$ (правда, эти переходы уже защищены введенным ранее ограничением на переходы $S_8(10_x,1_x)$ и $S_8(2_x,4_x)$).

Упорядочивая эти переходы по S-блокам, получим список из 18 значений переходов, подлежащих ограничению:

$$S_1(4_x,4_x),\ S_1(2_x,2_x);\quad S_2(4_x,4_x),\ S_2(2_x,1_x);$$
$$S_3(8_x,4_x),\ S_3(4_x,8_x);\quad S_4(8_x,4_x),\ S_4(2_x,2_x);$$
$$S_5(10_x,1_x),\ S_5(8_x,8_x),\ S_5(2_x,4_x);$$
$$S_6(2_x,2_x),\ S_6(4_x,8_x),\ S_6(10_x,4_x);$$
$$S_7(2_x,1_x),\ S_7(4_x,8_x);\quad S_8(2_x,4_x),\ S_8(10_x,1_x).$$

Нетрудно убедиться, что он полностью совпадает со списком элементов ТРЛА, представленным в Условии *L*-3. Приведенный список, однако, характеризует условия перекрытия характеристик, содержащих и одноблочные и двухблочные циклы. Так, если интересоваться характеристиками, содержащими только один активный S блок в каждом цикле, то, как легко видеть, все используемые в этом случае варианты однобитных переходов содержатся в приведенном списке. Если далее считать, что вероятности этих переходов ограничены значением 4, то для результирующей вероятности 16-цикловой характеристики, составленной из таких однобитных переходов, приходим к оценке:

$$2^4 \cdot \left[\left(\frac{4}{64}\right)^2 \cdot 2\right]^5 \cdot \left(\frac{4}{64}\right) \cdot 2 = 2^{-33}, \qquad 2^3 \cdot \left[\left(\frac{4}{64}\right)^2 \cdot 2\right]^4 \cdot \left(\frac{4}{64}\right) \cdot \left(\frac{16}{64}\right)^2 \cdot 2^3 = 2^{-30}$$

и, следовательно, эти характеристики действительно становятся не опасными (правое выражение учитывает произвол в выборе значений входа и выхода характеристики).

Корейские ученые, задавая нулевые значения вероятностей однобитных переходов (Условием *L*-3), полностью запретили все характеристики рассмотренного типа. Однако реальные ограничения для этих характеристик, как следует из приведенных выше расчетов (2), можно сделать менее жесткими (вместо нуля достаточно ограничиться значением 4).

Если теперь рассматривать характеристики с двухблочными циклами (1+2+0), реализующими переходы $\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$ (переходы $\Phi \leftarrow \Gamma \oplus \Psi \leftarrow \Phi$ для шифра DES не реализуемы), то при вероятности ЛАХ одноблочного цикла, равной максимально возможному значению $\left(\frac{16}{64}\right)$, для 16-цикловой характеристики, построенной из шестицикловых характеристик рассматриваемого типа, получим результирующее значение вероятности равное:

$$2^4 \cdot \left[2^2 \cdot \left(\frac{16}{64}\right) \cdot \left(\frac{4}{64}\right)^2\right]^5 = 2^{-8 \cdot 5+4} = 2^{-36},$$

что является уже достаточным (если также учитывать особенности в начале и конце характеристики).

Полученное нами ограничение с измененным по отношению к условию L-3 граничным значением мы определим как условие У-4.

Условие У-4 (условие перекрытия шестицикловых итеративных аппроксимаций с однобитными переходами): Для ТРЛА S-блоков необходимо выполнить следующие (общее число 18 случаев) условия:

- $S1$-блок: $|NS_1(4_x, 4_x)| \le 4$, $|NS_1(2_x, 2_x)| \le 4$,
- $S2$-блок: $|NS_2(4_x, 4_x)| \le 4$, $|NS_2(2_x, 1_x)| \le 4$,
- $S3$-блок: $|NS_3(8_x, 4_x)| \le 4$, $|NS_3(4_x, 8_x)| \le 4$,
- $S4$-блок: $|NS_4(8_x, 4_x)| \le 4$, $|NS_4(2_x, 2_x)| \le 4$,
- $S5$-блок: $|NS_5(16_x, 1_x)| \le 4$, $|NS_5(8_x, 8_x)| \le 4$, $|NS_5(2, 4)| \le 4$,
- $S6$-блок: $|NS_6(16_x, 4_x)| \le 4$, $|NS_6(4_x, 8_x)| \le 4$, $|NS_6(2_x, 2_x)| \le 4$,
- $S7$-блок: $|NS_7(4_x, 8_x)| \le 4$, $|NS_7(2_x, 1_x)| \le 4$,
- $S8$-блок: $|NS_8(16_x, 1_x)| \le 4$, $|NS_8(2_x, 4_x)| \le 4$.

Трехцикловые характеристики с двумя двухблочными циклами получаются при использовании для их построения двух циклических однобитных переходов, связывающих различные S-блоки, и поэтому сразу становится очевидным, что здесь также "сработает" введенное выше ограничение на однобитные переходы. Они будут включать в себя однобитные переходы из списка ограничений Условия L-3 или нереализуемые переходы.

Особого внимания заслуживают трехцикловые (шестицикловые) характеристики, которые могут быть построены без использования S-блоков с однобитными переходами. Анализ показывает, что существует только два варианта характеристик с пятью S-блоками типа 3+2+0, которые строятся без однобитных переходов: это циклический переход 3,4,8,11 ← 18,23,(24) ← 3,4,8,11 и циклический переход 7,8,12,14 ← 18,(20),26 ← 7,8,12,14. Характеристики с большим числом S-блоков, как указано выше, для атак ЛК уже не опасны.

Для перекрытия этих характеристик можно воспользоваться условием L-4 корейских ученых, которое они предложили для защиты от атак на восьмицикловых характеристик. В новых обозначениях это будет условие У-5.

Условие У-5 (условие защищенности от атак ЛК на шестицикловые итеративные аппроксимации без однобитных переходов): Для $W(\alpha)$, $W(\beta) \le 2$ необходимо, чтобы $|NS(\alpha,\beta)| \le 10$, где, как и ранее, $\alpha \in GF(2)^6$ и $\beta \in GF(2)^4$, $W(\alpha)$ – вес битового входа, а $W(\beta)$ – вес битового выхода S-блока.

Это ограничение здесь даже является черезчур уж жестким, так как для вероятности пятнадцатицикловой характеристики, составленной из трехцикловых аппроксимаций рассматриваемого вида, получаем оценку

$$2^4 \cdot \left\lceil 2^4 \cdot \left(\frac{10}{64}\right)^5 \right\rceil^5 = 2^{-42,9}.$$

Осталось рассмотреть последнюю характеристику под номером 3 (рис. 2). Легко убедиться, что эта характеристика допускает многовариантное представление. Некоторые из возможных вариантов шестицикловых характеристик, содержащих все циклы активного типа, вместе с исходной характеристикой (рис. 2), приведены на рис. 7. Несмотря на их значительное многообразие, они с точностью до обозначений входов и выходов описываются тремя различными графами переходов, которые под номерами, соответствующими характеристикам рис. 7, приведены на рис. 8.

Характеристика с графом переходов 3.1 использует два независимых циклических перехода $\Phi \leftarrow \Gamma \leftarrow \Phi$, $\Psi \leftarrow \Theta \leftarrow \Psi$. Если символы Φ, Θ, Γ и Ψ в обозначениях характеристики – это однобитные входы и выходы различных S-блоков, то получается шестицикловая характеристика (по числу активных S-блоков каждого цикла) типа 1+2+1+1+2+1. Пары однобитных входов Φ, Θ и Γ, Ψ – входы в различные S-блоки и, следовательно, выходы соответствующих пар S-блоков не могут принимать свободные значения. В итоге, рассматриваемые шестицикловые характеристик содержат по четыре однобитных перехода, которые либо содержатся в списке ограничений У-4, либо хотя бы один из переходов является для шифра DES нереализуемым (использует входы в S-блоки с нулевой вероятностью). Для шестицикловых итеративных характеристик рассматриваемого типа достаточно защититься от атак на характеристики с числом активных S-блоков, приходящихся на шестицикловую характеристику, не превышающим десяти:

$$\left[\left(\frac{16}{64}\right)^4 \cdot 2^2\right]^5 \cdot 2^4 \cdot \left(\frac{16}{64}\right) \cdot 2 = 2^{-27}$$

3,1

$\Gamma \leftarrow \Phi$

$\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$

$\Psi \leftarrow \Theta$

$\Phi \leftarrow \Gamma$

$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$

$\Theta \leftarrow \Psi$

3.2

$\Gamma \leftarrow \Phi$

$\Theta \leftarrow \Psi$

$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$

$\Theta \leftarrow \Gamma$

$\Psi \leftarrow \Phi$

$\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$

3.4

$\Gamma \leftarrow \Phi$

$\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$

$\Psi \leftarrow \Phi$

$\Theta \leftarrow \Gamma$

$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$

$\Theta \leftarrow \Psi$

3.5

$\Gamma \leftarrow \Phi$

$\Theta \leftarrow \Psi$

$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$

$\Phi \leftarrow \Gamma$

$\Psi \leftarrow \Theta$

$\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$

Рис. 7

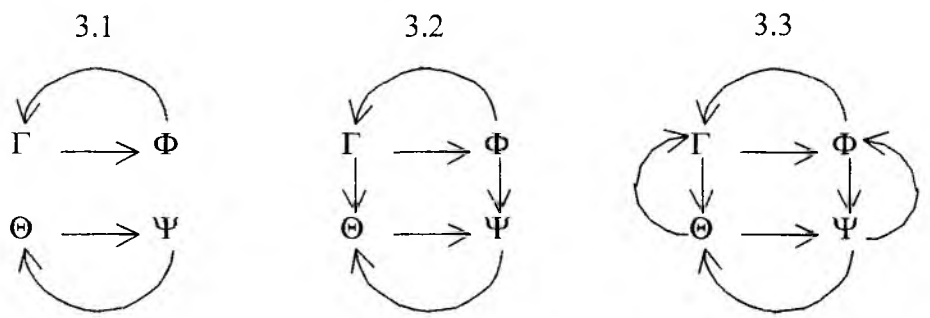


Рис. 8

Если символы Φ, Θ и Γ, Ψ в обозначениях характеристик – это однобитные входы и выходы однотипных S-блоков, то получается шестицикловая характеристика типа 1+1+1+1+2+1 или 1+2+1+1+2+1. В этом случае имеются пары однобитных входов Φ и Θ, Γ и Ψ, которые являются входами в различные S-блоки и, следовательно, только два или три выхода разнесенных циклов может принимать свободные значения. Но все равно и в этих характеристиках сохраняются минимум три однобитных перехода, которые делают ее с при выполнении ограничения У-4 неуязвимой для атак линейного криптоанализа. Примеры построения рассмотренных выше шестицикловых итеративных характеристик иллюстрирует рис. 9.

Если при построении характеристики используется хотя бы один циклический двухбитный переход (см. пример в нижней части рис.9), то и в этом случае оказывается, что в ее формировании принимает участие минимум два однобитных (одноблочных) перехода (выходы соответствующих S-блоков не являются свободными). В результате во всех рассмотренных случаях для вероятности результирующей пятнадцатицикловой характеристики приходим к оценке

$$\left[\left(\frac{4}{64}\right) \cdot \left(\frac{16}{64}\right)^3 \cdot 2^2\right]^5 \cdot 2^4 = 2^{-36},$$

т.е. эти характеристики не подвержены атакам ЛК. Заметим, что при получении последнего результата полагалось, что на каждые три цикла шестицикловой характеристики приходится не менее одного S-блока, удовлетворяющего условию У-4, и трех S-блоков, не попадающих под какие-либо ограничения, кроме максимально допустимого значения элементов ТРЛА. Естественно, что при использова-

нии ограничения L-3, предложенного корейскими учеными, все рассмотренные выше шестицикловые характеристики становятся просто не реализуемыми.

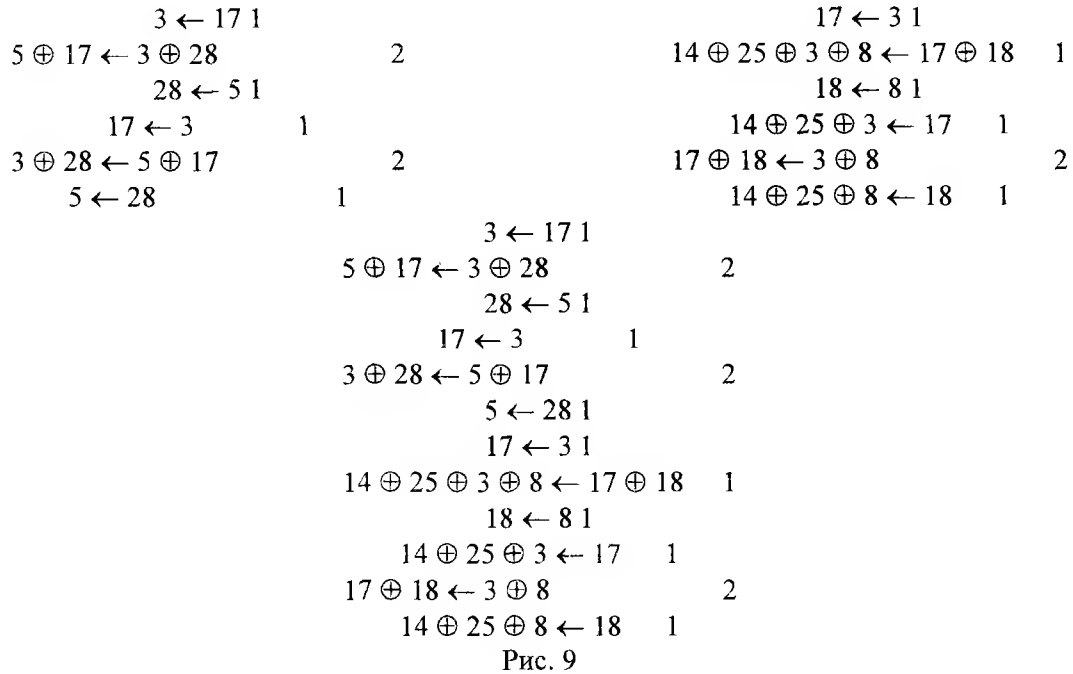


Рис. 9

Что касается характеристик с графами переходов 3.4 и 3.5, рис. 8, то все они, как показывает анализ, для шифра DES не осуществимы.

Таким образом, нам удалось обосновать три первых (L-1÷L-3) условия отбора S-блоков для шифра DES, которые корейскими учеными рассматривались как необходимые и достаточные для защиты четырехцикловых и шестицикловых ЛАХ от атак линейного криптоанализа. Вместе с тем показано, что рассмотренных трех ограничений явно недостаточно для решения этой задачи. Имеются четырехцикловые характеристики, не рассмотренные корейскими учеными, которые требуют использования дополнительного ограничения, сформулированного в виде Условия У-4. Кроме того, для защиты шестицикловых характеристик обосновано Условие У-5. В очередной работе мы рассмотрим условия обеспечения защищенности от атак линейного криптоанализа для итеративных характеристик с большим числом циклов.

Список литературы: 1. Построение таблиц подстановок для стандарта шифрования данных / *И.В. Лисицкая, С.А. Головашич, Р.В. Олейников и др.* // Проблемы бионики. 1999. Вып 50. С. 185–194. 2. Анализ стойкости DES подобных алгоритмов шифрования при использовании таблиц подстановок случайного типа / *И.В. Лисицкая, А.С. Коряк, Р.В. Олейников и др.* // Радиоэлектроника и информатика. 1999. № 1. С. 111-115. 3. The selection criteria of random substitution tables for symmetric enciphering algorithms / *I.V. Lysytska, A.S. Koriak, S.A. Golovashich, O.I. Oleshko, R.V. Oleinik* // Abstracts of XXVIth General Assembly. Toronto, Ontario Canada. 1999. P. 204. 4. Обеспечение стойкости DES-подобных алгоритмов шифрования к атакам линейного криптоанализа при использовании подстановок случайного типа / *В.И. Долгов, И.В. Лисицкая, С.А. Головашич и др.* // Радиотехника. 2000. Вып 114. С. 39-46. 5. *Kim K., Park S., Lee S.* Reconstruction of s^{2}DES S-boxes and their Immunity to Differential Cryptanalysis, Pros. of 1993 Korea-Japan Joint Workshop on Information Security fnd Cryptology (JW-ISC'93), Oct. 24-36, Seoul, 1993. 6. *Lars Ramkilde Knudsen.* Iterative Characteristics of DES and s^{2}DES, Proc. of Crypto'92. UCSB. 1992. 7. *Kim K., Lee S., Park S.* Necessary Conditions to Strengthen DES S-boxes against Linear Cryptanalysis. Pros. of SCIS'94, Biwako, Japan, Pp.15D. 1-11. 1994. 8. *Kim K.* Construction of DES-like S-boxes Based on Boolean Function Satisfying the SAK, Pros. Of Asiacrypt'91, Pp. 59-72, Fujiyoshida, Japan, 1991. 9. *Kim K., Lee S., Park S., Lee D.* DES can be Immune to Linear Cryptanalysis, Workshop Record of SAC '94 (Selected Areas in Cryptography) May 5-6. Queen's Univ. Canada. 1994. 10. *Kim K., Lee S., Park S., Lee D.* How to Strengthen DES against Two Robust Attacks, Joint Workshop on Information Security and Cryptology Inuyata. Japan. January 24-25, 1995. 11. *Biham E., Shamir A.* . Differential Cryptanalysis of the Data Encryption Standard. Springler Verlag. Berlin. 1993. 12. *Biham E., Shamir A.* Differential Cryptoanalysis of the full 16-round DES. Technical Report - Computer Science Department. Technion. Israiel. 1993. 13. *Mitsuru Matsui* Linear Cryptanalysis Method for DES Cipher. Proc. of Eurocrypt'93, Norway, 1993. 14. *Schneier B.* Applied Cryptography. Second Edition: protocols, algorithms, and Source code in C. Published by John Wiley & SonS. Inc, New York: ChicheSter BriSbane Toronto Singapore. 1996 758 p.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 6.03.2001

УДК 681.3.06:519.248.681

А. А. ЗАМУЛА, канд. техн. наук, Ю. С. ПАВЛЕНКО

ЗАЩИТА ИНФОРМАЦИИ В IP-ТЕЛЕФОНИИ

Введение

В настоящее время особое развитие получила цифровая телефония. При этом цифровая телефония способствовала улучшению качества телефонных сетей и дала возможность более эффективного их объединения с сетями передачи данных. Это позволило снизить стоимость услуг, а так же удовлетворить возросшие запросы на пропускную способность. К качеству цифровой телефонии предъявляется ряд противоречивых требований – минимизация скорости передачи данных, сохранение приемлемой узнаваемости, сложности реализации, задержек, а в ряде случаев и требования конфиденциальности передачи информации. Особо проблематичными являются задачи аутентификации, так как современные технологии синтеза речи позволяют подделывать ее с высоким качеством.

В связи с вышеуказанным весьма актуальными являются следующие задачи:

1) проведение сравнительного анализа и выбор методов высококачественного сжатия речи, позволяющего с одной стороны минимизировать необходимую пропускную способность, а с другой – обеспечить требуемое качество узнаваемости;

2) поиск и применение методов аутентификации (своего рода узнаваемости) за счет применения криптографических методов аутентификации;

3) выбор и реализация методов криптографической защиты с целью обеспечения конфиденциальности и аутентичности;

4) разработка принципов организации связи для сети общего пользования, обоснование и разработка состоятельных протоколов работы с ключами и режимов работы системы в целом;

1.Анализ методов сжатия речи

Частотный диапазон человеческого голоса равен примерно 20-20000 Гц и может значительно отличаться у отдельных людей. Исследования показали, что без существенных потерь диапазон может быть уменьшен до 100-6000 Гц. Но в самом начале строительства телекоммуникаций предоставить такую полосу частот каждому абоненту было невозможно, и пришлось ради удешевления услуг связи пойти на большее сокращение частотного диапазона до полосы 200-3400 Гц. Это привело к ухудшению разборчивости, но она осталась все же на достаточном уровне. В этой полосе частота выборок согласно теореме Котельникова должна составить не менее 8 кГц, и если при этом использовать 8-разрядный код, то необходимая пропускная способность должна составить не менее 64 кбит/с.

Для уменьшения этой величины применяются различные методы сжатия речи. Одним из первых стандартов, получивших широкое распространение, был G.726 ADPCM (адаптивная дифференциальная импульсно-кодовая модуляция)[3]. При его использовании оцифровывался не сам сигнал, а только его отклонения от предсказанного значения. На каждый отсчет затрачивалось по 4 бита, что позволяло снизить скорость до 32 кбит/с. Позже появилась его разновидность, позволившая еще снизить скорость до 24 кбит/с. По-настоящему снизить скорость удалось только с распространением кодеков на базе линейного предсказания, которые хоть и появились в конце 60-х, но до сих пор остаются основным способом сжатия речи.

В табл. 1 приведены характеристики кодеков и их применение.

Таблица 1

Скорость передачи, кбит/с	Субъективное качество	Название стандарта	Год выпуска	Алгоритм	Область применения
64	4,1	ITU-T G.711	1960	PCM	Телефонные сети
32	3,8	ITU-T G.726	1984	ADPCM	Телефонные сети
6,4	3,1	INMAR-SAT-M	1990	IMBE	Спутниковая телефония
13	3,3	ETSI GSM	1992	RPE-LTP	Сотовая телефония (Европа)
16	3,6	ITU-T G.728	1992	LD-CELP	Телефонные сети

Продолжение таблицы 1

Скорость передачи, кбит/с	Субъективное качество	Название стандарта	Год выпуска	Алгоритм	Область применения
4,8	3,4	ETSI TETRA	1996	ACELP	Сотовая телефония (Европа)
6,3	3,9	ITU-T G.723.1	1996	MP-MLQ	Телефонные сети
5,3	3,7	ITU-T G.723.1	1996	ACELP	Телефонные сети
8	3,9	ITU-T G.729	1997	CS-ACELP	Телефонные сети
2,4	3,5	США (проект)	1998	MELP	Мин. Обороны США

Из относительно новых алгоритмов стоит назвать совместную разработку фирм Audiocodec (Израиль) и DSP Group (США) - кодек с линейным предсказанием и скоростью выходного потока информации 6,3 кбит/с[4]. Несмотря на значительно большую, чем у АДИКМ, среднеквадратичную погрешность синтеза, получена лучшая (3,9 балла MOS) оценка качества, чем у АДИКМ. Это достигнуто благодаря двум усовершенствованиям алгоритма линейного предсказания алгебраического CELP. Первое: длина сглаживающего окна трехкратно увеличена относительно длины анализируемого сегмента речи, что ослабило искажения, вносимые асинхронностью анализируемых сегментов речи и интервалов основного тона. Второе: метод более точного формирования сигнала возбуждения синтезирующего фильтра, названный авторами MP-MLQ (Multipulse Maximum Likelihood Quantisation). Еще одним важным преимуществом этого алгоритма является более надежное, по сравнению с предшественниками, вычисление параметров основного тона голоса. Алгоритм MP-MLQ относится к семейству analysis-by-synthesis алгоритмов. В кодере построенном на MP-MLQ используется линейное прогнозирование 10-го порядка и может работать на скоростях 4.8, 6.4, 7.2 и 8.0 кбит/с. Гибкость алгоритма дает возможность выполнять перепрограммирование в ходе разговора при ухудшении связи, а также снизить скорость до 4.0 кбит/с. Все эти особенности, а также заявленное качество речи, склонили нас к выбору именно этого кодека.

Для существующих и распространенных реализаций кодеков авторами было проведено исследование по качеству сжатия и размерам полученного сжатого файла. Для этого была написана программа, позволяющая преобразовать WAV файл в формате PCM в WAV файл, где речь записана в другом формате[1]. В качестве исходного файла был взят файл длиной 4,25 секунды и размером 76612 байт. Результаты приведены в табл. 2.

Таблица 2

Алгоритм	Размер, байтов	Время, сек	Примечание
PCM	76612	4,25	1
CCITT A-Law	38342	4,50	1
CCITT u-Law	38342	4,50	1
DSP Group TrueSpeech™	5178	4,74	1
GSM 6.10	7860	4,46	1
IMA ADPCM	19260	4,98	1
Lernout & Hauspie CELP 4.8kbit/s	2926	4,66	3
Lernout & Hauspie SBC 16kbit/s	11358	4,48	3
Lernout & Hauspie SBC 12kbit/s	6470	2,98 ?	4
Lernout & Hauspie SBC 8kbit/s	3676	1,54 ?	4
Microsoft ADPCM	19694	4,97	1
MPEG Layer-3	9356	4,00	2

В графе примечания указаны особенности преобразования речи:

1. Отличная разборчивость, помех практически нет. Фраза звучит полностью.
2. Отличная разборчивость, но фраза звучит не полностью (несколько последних букв отсутствуют).
3. Разборчивость хорошая, но есть заметные (не мешающие восприятию) искажения. Фраза звучит полностью.
4. Достаточная разборчивость, но искажения почти на всем протяжении фразы. Фраза звучит не полностью (несколько последних букв отсутствуют). Непонятны значения длительности записи, так как фраза реально звучала в обоих случаях более 4 секунд.

2. Протоколы обеспечения конфиденциальности и установления подлинности абонентов защищенного канала

Обеспечение высокого уровня аутентичности корреспондентов только за счет узнаваемости на слух нельзя считать эффективным методом решения этой задачи. Она должна решаться комплексно – с учетом принципов узнаваемости, а также с применением несимметричной криптографии на этапах вхождения в связь и ведения связи. Принципы несимметричной криптографии в данном случае предполагает применение личных ключей каждого из абонентов и состоятельных протоколов на комбинации долговременных и сеансовых ключей.

Конфиденциальность цифровой телефонии предпочтительно обеспечивать за счет применения симметричного шифрования. Для этих целей можно использовать такие стандарты, как ГОСТ 28147, IDEA, Rijndael. Кроме непосредственно шифрования речевой информации с помощью вышеуказанных алгоритмов требуется создание защищенной сети со станцией генерирования и распределения ключей. Сгенерированные долговременные ключи распределяются корреспондентам с использованием состоятельных протоколов. Доступ к системе пользователь получает после ввода личного ключа, носителем которого является смарт-карта. После этого производится синхронизация и аутентификация обеих сторон связи и вырабатываются сеансовые ключи.

Передача речевой информации происходит в обоих направлениях и логичной была бы организация полнодуплексного канала связи. Но так как разговор идет обычно по очереди (абоненты слушают друг друга) и вычислительное устройство, производящее кодирование и декодирование, работает в последовательном режиме обработки команд, то здесь нужно использовать многозадачность с разделением по времени, т.е. кодированию и декодированию поочередно выделяют небольшой промежуток времени. Связь при этом получается полудуплексной. Рассмотрим протоколы используемые для шифрования информации во время разговора и установления связи. В начале установления сеанса связи каждый из абонентов вырабатывает личный сеансовый ключ и абоненты производят обмен данными ключами по схеме Диффи-Хелмана.

После этого вырабатывается общий секрет K, который и служит в дальнейшем для образования сеансового ключа симметричной системы шифрования r.

На этом ключе формируется сеансовый ключ для шифра Rijndael функционирующего в потоковом режиме, а также шифруют пакеты с ключевой информацией.

Схема выработки сеансового ключа K для схемы Диффи-Хелмана указана в таблице 3.

Таблица 3

A		B
$Y_A = \Theta^{X_A} \bmod N$	⟷	$Y_B = \Theta^{X_B} \bmod N$
S_A		S_B
$D_A = \Theta^{S_A} \bmod N$	⟷	$D_B = \Theta^{S_B} \bmod N$
$K_{AB} = D_B^{S_A} \cdot Y_B^{X_A} \bmod N = \Theta^{S_B \cdot S_A} \cdot \Theta^{X_B \cdot X_A} \bmod N$		$K_{BA} = D_A^{S_B} \cdot Y_A^{X_B} \bmod N = \Theta^{S_A \cdot S_B} \cdot \Theta^{X_A \cdot X_B} \bmod N$

После этого производят выработку сеансового ключа для алгоритма шифрования Rijndael по правилу:

$$r = Rijndael(H(K), R_{дол г}),$$

т.е. путем шифрованием хеш-функции сеансового ключа K на долговременном ключе для алгоритма Rijndael $R_{долг}$.

Полученный сеансовый ключ r объединяется с возможной другой служебной информацией, шифруется на сеансовом ключе K и отправляется другому абоненту.

В результате получается трехуровневый протокол управления ключами:

1) долговременные ключи X,R;
2) сеансовый ключ для схемы Диффи-Хелмана K;
3) сеансовый ключ для потокового режима передачи речевой информации r;

При этом долговременные ключи используются для формирования сеансовых ключей и никогда – для передачи речевой информации, что существенно снижает возможность перехвата и накопления статистического материала.

Заключение

Существующие разработки в области защиты речевой информации основываются на различных методах скремблирования речи, однако применение таких методов не позволяет обеспечить аутентификацию участников обмена информацией. Предложенная выше схема совместного использования преобразования речи в цифровой вид и цифровых методов шифрования и аутентификации позволяет обеспечить как требуемый уровень надежности аутентификации и сокрытия смысла информации, так и снизить требования к пропускной способности, а значит увеличить количество одновременных разговоров по стандартной линии передачи информации.

Список литературы: **1.** *Секунов Н.Ю.* Обработка звука на PC. С-Пб.: БХВ-Петербург, 2001. 1248 с. **2.** *Назаров М.В., Прохоров Ю.Н.* Методы цифровой обработки и передачи речевых сигналов. М: Радио и связь, 1985. 176 с. **3.** *American National Standards Institute*, Inc. 1987. American National Standard for Telecommunications: Digital Processing of Voice-Band Signals-Algorithm and Line Format for 32kbit/s. Adaptive Differential Pulse-Code Modulation (ADPCM). New York: ANSI, Inc. **4.** *ITU-T Recommendation G.723.1* Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 27.03.2001

УДК 519.7:007.52

Н. С. ЛЕСНАЯ, канд. техн. наук, Т. Б. ШАТОВСКАЯ, канд. техн. наук, В. Б. РЕПКА

МЕТОД ВЫБОРА ЭФФЕКТИВНЫХ ПРОЦЕДУР ОЦЕНИВАНИЯ ПАРАМЕТРОВ МОДЕЛЕЙ КВАЗИСТАЦИОНАРНЫХ ПРОЦЕССОВ В НЕЙРОСЕТЕВОЙ ЭКСПЕРТНОЙ СИСТЕМЕ

В последнее время стремительно возрастает значение информационного обеспечения различных производственных технологий, оно становится критическим фактором развития практически во всех областях знания. В связи с этим разработка и внедрение новых информационных технологий является на сегодняшний день одной из самых актуальных задач.

Представление знаний и механизма рассуждений в информационной системе возможно благодаря современному направлению развития науки – искусственному интеллекту (ИИ), продуктами которого являются интеллектуальные системы (ИС) [1]. Несмотря на значительные успехи в области ИИ, существует еще определенный разрыв между имеющимися аппаратными и программными средствами ИИ и возможностями их практического применения на производстве. В этой связи актуальной является попытка активного внедрения систем ИИ на всех уровнях управления производством [1].

В связи с высокими требованиями к точности математических моделей процессов управления в различных областях техники проблема идентификации становится исключительно важной. Невозможно обеспечить качественное управление системой, если ее математическая модель не известна с достаточной точностью. Для построения математической модели могут быть использованы как теоретические, так и экспериментальные методы. Опыт, накопленный при проектировании систем управления, убедительно свидетельствует о том, что нельзя построить математическую модель, адекватную реальной системе, только на основе теоретических исследований физических процессов в системе. Сформированная таким образом математическая модель, как правило, значительно отличается от реальной системы, что приводит соответственно к снижению качества управления. Поэтому в процессе проектирования интеллектуальных систем управления одновременно с теоретическими исследованиями проводятся многочисленные опыты по определению и уточнению математической модели системы.

Следует отметить, что точность синтезируемой математической модели зависит не только от метода ее построения, но и от подхода к выбору метода оценивания коэффициентов модели [2].

Определяющим фактором в выборе метода идентификации является характер изменения во времени входных воздействий исследуемого объекта. В случае нестационарности входов часто используют подход выделения периодов их квазистационарности. Получаемые при этом выборки, как правило, имеют недостаточный объем для получения надежных оценок. Кроме того, как показывает опыт, на малых участках квазистационарности сильнее проявляется эффект мультиколлинеарности [2].

В случае нестационарности входов объекты описываются системой моделей, построенных на различных подвыборках, соответствующих квазистационарным периодам функционирования объекта. Это обстоятельство существенно усложняет выбор базы методов идентификации с одной стороны, а с другой – ставит вопрос о выявлении идентичности регрессионных моделей, получаемых на различных подвыборках. Таким образом, задача сводится к идентификации стационарных объектов, осложненной высокой коррелированностью входов, наличием "загрязненности", выбросов в исходных данных и малыми объемами подвыборок.

Как известно [2, 4], при мультиколлинеарности наилучшее решение обеспечивается одним из методов смещенного оценивания, а при загрязненности и выбросах в исходных данных – определенным методом робастного оценивания. Существующая на сегодняшний день достаточно обширная алгоритмическая база методов смещенного и робастного оценивания предлагает широкое множество алгоритмов $a_i \in A$, $i=1,\ldots, N$ (более 60 алгоритмов в классе смещенного оценивания и более 20 – в классе робастного оценивания), которые могут иметь в конкретных условиях идентификации различную эффективность $\varphi(a_i,S_j)$. В связи с этим актуальной является задача выбора наиболее эффективного метода идентификации для каждого исследуемого участка, решаемая посредством автоматической классификации методов оценивания, что позволит для каждой ситуации S_j отобрать некоторые подмножества наиболее эффективных в S_j алгоритмов a_i. Следовательно, основной идеей создания высококачествен-

ных моделей идентификации является осуществление выбора наиболее эффективного алгоритма из подмножества соответствующей классификационной ячейки в зависимости от доступных контролю характеристик объекта и возмущений, т.е. выбор алгоритма, обеспечивающего max $\varphi(a_i, S_i)$.

Отметим, что самым распространенным показателем эффективности алгоритмов идентификации является, конечно, среднеквадратичная ошибка предсказания как в режиме прогноза, так и в режиме управления [5]. В силу указанной специфики исходной информации и используемых в данном случае методов построения статистических моделей оценку их эффективности предлагается осуществлять на основе некоторого подмножества специфических критериев и их комбинаций, что позволит не только выявить области доминирования каждого метода, но и выбрать наиболее эффективный из них в зависимости от характеристик исследуемого объекта.

Задачу классификации исследуемых методов по набору характеристик, учитывающих специфику объекта в режиме самообучения, предлагается осуществлять на основе аппарата нейронных сетей (НС). При практической работе с НС, как правило, приходится экспериментировать с большим числом различных типов сетей, порой обучая каждую из них по несколько раз и сравнивая полученные результаты. Главным показателем качества результата является контрольная ошибка [3, 5]. При этом в соответствии с общенаучным принципом, согласно которому при прочих равных условиях следует предпочесть более простую модель, из двух НС с приблизительно равными контрольными ошибками имеет смысл выбрать сеть меньшего объема.

Таким образом, предлагается принципиально новый подход к выбору эффективного метода идентификации, состоящий в выявлении соответствия между исходными данными объекта, методами идентификации в классах смещенного и робастного оценивания и типами используемых нейронных сетей.

Рассмотрим совокупность статистических критериев, которые будем использовать для оценки эффективности методов идентификации квазистационарных процессов. Точность методов смещенного оценивания будем характеризовать совокупностью следующих критериев:

– критерий относительной среднеквадратичной погрешности, характеризующей эффективность исследуемого метода относительно МНК:

$$E(L_1^2) = \frac{1}{p}\sum_{i=1}^{p}(\hat{\beta}_j - \beta_j)^2 / \beta_j^2, \qquad (1)$$

где $\hat{\beta}$ – оценка параметров методами смещенного оценивания; β – истинное значение параметров; $E(\cdot)$ – символ математического ожидания:

– критерий, характеризующий дисперсию оценок коэффициентов модели:

$$E(L_2^2) = tr\ E[(\hat{\beta} - E\hat{\beta})(\hat{\beta} - E\hat{\beta})']; \qquad (2)$$

– критерий, характеризующий смещение оценок модели относительно моделируемых истинных значений оценок:

$$E(L_3^2) = (E(\hat{\beta}) - \beta)'(E(\hat{\beta}) - \beta); \qquad (3)$$

– стандартный критерий среднеквадратичной ошибки модели:

$$E(L_4^2) = E(\hat{\beta} - \beta)'(\hat{\beta} - \beta); \qquad (4)$$

– критерий максимальной абсолютной координатной ошибки метода:

$$E(L_5^2) = \max_j \left|\hat{\beta}_j - \beta_j\right|^2, \quad j = \overline{1, p}; \qquad (5)$$

– критерий максимальной координатной относительной ошибки метода:

$$E(L_6^2) = \max_j \left|\frac{\hat{\beta}_j - \beta_j}{\beta_j}\right|^2, j = \overline{1, p}. \qquad (6)$$

В качестве меры оценки эффективности робастных методов оценивания используются следующие критерии:

– критерий относительной среднеквадратической ошибки модели:

$$E(L^2 1) = \sum_{j=1}^{p} (\hat{\beta}_j - \beta_j)^2 \Big/ \sum_{j=1}^{p} (\hat{\beta}^*_j - \beta_j)^2, \qquad j = \overline{1,p}, \tag{7}$$

где $\hat{\beta}^*$ – оценки параметров модели регрессии, полученные МНК; $\hat{\beta}$ – оценки параметров модели регрессии, полученные с помощью робастного метода; β – истинные значения параметров регрессионной модели.

– критерий относительной медианы абсолютных отклонений:

$$E(L^2 2) = \sum_{j=1}^{p} med(|\hat{\beta}_j - \beta_j|) \Big/ \sum_{j=1}^{p} med(|\hat{\beta}^*_j - \beta_j|), \qquad j = \overline{1,p}. \tag{8}$$

Необходимо отметить, что критерии оценки точности методов смещенного оценивания инвариантны относительно уровня мультиколлинеарности, в то время как при оценке точности методов робастного оценивания использование критерия относительной медианы абсолютных отклонений имеет более высокий приоритет в случае "сильной загрязненности" исходных данных, чем критерий относительной среднеквадратической ошибки модели.

В качестве обобщенной меры оценки точности методов смещенного и робастного оценивания предлагается использование критериев степени рассеяния и индекса ошибки рассмотренных ранее критериев, что позволит оценить эффективность методов идентификации в среднем по всей совокупности критериев. Критерий индекса ошибки представим в виде:

$$S_i(A,H) = \frac{\left|L_i - L_i^*\right|}{\frac{1}{l}\sum_{j=1}^{l}\left|L_{ij} - L_{ij}^*\right|}, \quad \bar{S}(A,H) = \frac{1}{k}\sum_{i=1}^{k} S_i(A,H), \tag{9}$$

где A – метод оценивания; H – выборка исходных данных; L_i^* – наилучшее значение L_i.

Определим для индекса ошибки меру рассеивания в виде:

$$q(S) = \sqrt{\frac{1}{k-1}\sum_{i=1}^{k}[S_i(A,H) - \bar{S}(A,H)]^2}, \tag{10}$$

которая определяет вариации характеристики $S(\cdot)$ от набора к набору исходных данных и неустойчивость метода в области, определяемой H. Небольшие значения $q(S)$ характеризуют метод оценивания как устойчивый.

Также необходимо рассматривать ранговые оценки эффективности метода в области наборов исходных данных H, учитывающие лучшие и худшие результаты при испытаниях. Для получения ранга оценки располагаются в порядке возрастания $S(\cdot)$ для каждого набора. Если m альтернатив получили равные $S(\cdot)$, то каждому из m методов присваивается среднее из m рангов. Аналогично определяется индекс

$$\bar{p}(A,H) = \frac{1}{k}\sum_{i=1}^{k} p_i(A,H) \tag{11}$$

и рассеяние

$$q(p) = \sqrt{\frac{1}{k-1}\sum_{i=1}^{k}[p_i(A,H) - \bar{p}(A,H)]^2}. \tag{12}$$

При исследовании связи между типом нейросети и методами оценивания была промоделирована работа следующих типов НС: многослойный перцептрон, сеть встречного распространения, вероятностная нейронная сеть (ВНС), сеть радиальной базисной функции, самоорганизующиеся карты признаков Кохонена [2,3]. Для каждого типа исследуемых сетей предварительно был задан уровень на-

дежности, а также рассчитан коэффициент уверенности вида:

$$КУ = (Max1\text{-}Max2)/R * 100\ \%, \quad (13)$$

где Max1 – ответ выходного нейрона, отвечающего за класс "победитель"; Max2 – ответ выходного нейрона, выдавшего следующий по максимальной величине сигнал; R – уровень надежности.

При проведении экспериментов наилучший результат по показателю контрольной ошибки с поправкой на размер сети был получен вероятностной НС.

Рассмотрим решение поставленной ранее задачи на примере классификации смещенных и робастных методов оценивания с использованием ВНС. Для этого типа сети применяется режим обучения "с учителем". Для ВНС необходима обучающая выборка, которая включает наборы характеристик объекта исследования и их принадлежность к одному из классов методов оценивания. Входными характеристиками для ВНС являются: объем выборки, количество независимых переменных, показатель мультиколлинеарности, коэффициент корреляции между независимыми переменными, степень загрязнения независимых переменных, степень "засоренности" "загрязненного" закона распределения ошибок модели регрессии, качественный параметр, определяющий форму выбросов в независимых переменных, длина хвоста "загрязненного" распределения независимых переменных или величина выбросов в случае остаточных выбросов.

Один из показателей качества обучения, прогностическая способность нейросети, состоит в подсчете процента правильно распознанных примеров. При сравнении качества обучения двух нейросетей в случае, когда обе сети дают одинаковую прогностическую способность, можно подсчитывать средний процент уверенности при тестировании выборки. Он рассчитывается как среднее арифметическое процентных величин уверенности, полученных при тестировании каждого примера.

Зашумленность данных, представленных в обучающей выборке, неизбежно приводит к отдельным ошибкам классификации. Целесообразно считать, что некоторые виды ошибок обходятся "дороже" других. В такой ситуации относительная цена ошибки классификации определяется как вероятность принадлежности к определенному классу, умноженная на коэффициент потерь. При проведении исследования в вероятностную нейронную сеть был добавлен дополнительный слой, содержащий матрицу потерь. Таким образом, матрица умножается на вектор оценок вероятностей, полученный в выходном слое, после чего в качестве ответа выбирается класс, имеющий наименьшую оценку потерь. Исследования показали, что если уровни мультиколлинеарности и засоренности исходных данных высокие, то классам назначаются одинаковые приоритеты. В случае, если отношение уровня мультиколлинеарности к уровню засоренности исходных данных больше единицы, классу методов смещенного оценивания присваивается более высокий приоритет.

Вероятностная нейронная сеть имеет единственный управляющий параметр обучения, значение которого должно выбираться пользователем, степень сглаживания. Требуемое значение было найдено опытным путем и подобрано таким образом, чтобы контрольная ошибка ВНС была как можно меньше. В качестве параметров нейронов радиального слоя выбрана радиальная функция активации, а нейронов выходного слоя – функция активации SOFTMAX.

Таким образом, на основании выше изложенного выбора типа нейронной сети, настройки ее параметров, режима обучения была установлена связь между входными характеристиками исследуемых квазистационарных объектов и классами статистических методов оценивания параметров моделей. Экспериментальным образом был определен наилучший тип нейросети для классификации методов оценивания параметров модели в зависимости от характеристик исследуемого объекта. Предложенный подход позволяет не только повысить точность синтезируемых математических моделей исследуемых процессов при высокой коррелируемости и зашумленности исходной информации за счет выбора наиболее эффективного метода смещенного и робастного оценивания, а также упростить и ускорить сам процесс выбора метода идентификации.

Список литературы: 1. Производственные системы с искусственным интеллектом / *Р.А. Алиев и др.* М: Радио и связь, 1990. 264 с. 2. *Лесная Н.С., Шамша Т.Б., Репка В.Б.* Об одном подходе к оценке качества исходной информации при обработке данных // Проблемы бионики. 1999. № 50. С. 71–74. 3. *Горбань А.Н.* Обучение нейронной сети. М.: СССР – США СП "Paragraph", 1990. 160 с. 4. *Timothy Masters.* "Advanced Algorithms for Neural Networks". Wiley, New York. 1995. Chapter 6. 5. *Шамша Б.В., Антонов В.А.* Об одном подходе к разработке модифицированных алгоритмов робастного оценивания. // Вестник ХГПУ. 2000. №80. С.24 – 26. 6. Нейронные сети. STATISTICA Neural Networks: Пер. с англ. М. Горячая линия – Телеком. 2000. 182 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 24.04.2001

УДК 621.396.67

А. И. ЛУЧАНИНОВ, д-р. физ.-мат. наук, А. А. КОНОВАЛЬЦЕВ, канд. техн. наук,
Ю. А. ЛУЧАНИНОВ, М. А. ОМАРОВ, канд. техн. наук, В. М. ШОКАЛО, д-р. техн. наук

АЛГОРИТМ АНАЛИЗА ЭКВИДИСТАНТНОЙ РЕШЕТКИ ЛЕНТОЧНЫХ МИКРОПОЛОСКОВЫХ ИЗЛУЧАТЕЛЕЙ ПРОИЗВОЛЬНОЙ ГЕОМЕТРИИ, АДАПТИРОВАННЫЙ К РАСЧЕТУ КРУПНОАПЕРТУРНЫХ АНТЕНН С НЕЛИНЕЙНЫМИ ЭЛЕМЕНТАМИ

3. ОСОБЕННОСТИ ЧИСЛЕННОЙ РЕАЛИЗАЦИИ АЛГОРИТМА

6. Выбор системы базисных функций и формирование матрицы импедансов

Одним из исходных факторов, определяющих эффективность решения системы интегральных уравнений (ИУ) типа (17, [1]) методом Галеркина, является выбор типа ортогонального базиса, по которому выполняется разложение искомой функции тока (1, [2]). При построении универсальных алгоритмов, позволяющих рассчитывать характеристики излучателей произвольной конфигурации, наиболее целесообразно использовать в качестве базисных так называемые функции подобластей [3], определенные на всей длине излучателя, но отличные от нуля только на отдельных его участках. Применение функций такого типа, сохраняющих непрерывность в точках соединения двух и более проводников, автоматически обеспечивает выполнение закона Кирхгофа в узлах. В случае разложения функции, описывающей распределение тока, по функциям полной области в систему уравнений, полученную в соответствии с изложенной в [1] процедурой, необходимо дополнительно включать уравнения, которые устанавливают соотношения между токами входящих в каждый узел излучателя ветвей. Очевидно, что при этом для каждой новой конфигурации излучателя система уравнений будет принимать новый вид, а следовательно, не будет соблюден принцип универсальности разрабатываемого алгоритма.

Достоинством кусочного базиса является также возможность уже на этапе описания модели анализируемого излучателя выделить в нем подпоследовательности базисных функций, характеризующих токи разомкнутых ветвей, а также внутренних и внешних узлов, что, как будет показано ниже, позволит более оптимально организовать процедуру решения исходной системы ИУ и расчета внешних параметров антенны.

Из функций указанного типа наиболее предпочтительными оказываются кусочно-синусоидальные. Использование их позволяет ускорить процесс решения ИУ по сравнению со случаем кусочно-треугольного и кусочно-постоянного базисов соответственно в 3...5 и 6...8 раз. Кроме того, если в качестве базиса выбраны кусочно-синусоидальные функции, на свободных концах проводников излучателя автоматически обеспечивается выполнение условия обращения в нуль протекающих по ним токов. Заметим также, что использование кусочно-синусоидальных базисных функций (БФ) позволяет получить достаточно простые аналитические выражения для спектральных плотностей собственных и взаимных сопротивлений и сократить объем и время вычислений.

Определим n-ую БФ кусочно-синусоидального базиса следующим образом:

$$\vec{\varphi}_n(\xi) = \frac{\sin(\xi_n - l_n)}{\sin l_n}\vec{\xi}_0^{(n)} + \frac{\sin(l_{n+1} - \xi_{n+1})}{\sin l_{n+1}}\vec{\xi}_0^{(n+1)}, \qquad (1)$$

где $\vec{\xi}_0^{(n)}$, $\vec{\xi}_0^{(n+1)}$ - ориентированные вдоль оси излучателя единичные векторы, общим началом которых является центр БФ (общая точка соседних сегментов, на которых определена функция $\vec{\varphi}_n$); ξ_n, ξ_{n+1} - координаты, отсчитываемые в местной системе координат, связанной с n-ой БФ, в направлении векторов $\vec{\xi}_0^{(n)}$ и $\vec{\xi}_0^{(n+1)}$ соответственно; l_n, l_{n+1} - длины сегментов, на которых определена n-ая базисная функция тока. При этом очевидно, что в наиболее общем случае для полного формирования матрицы импедансов $[Z]$ уравнения (17, [2]) достаточно определить взаимное сопротивление между двумя элементарными токами, которые характеризуются функциями в виде четверти периода синусоиды (одно из слагаемых в (1)). Для этого подставим выражения для $\vec{\varphi}_k(\xi)$ и $\vec{\varphi}_l(\xi')$ в (3, [2]) и (9, [2]) и выполним соответствующие операции интегрирования и суммирования, в результате чего

получим следующее выражение для спектральной плотности взаимного сопротивления указанных элементов тока:

$$\tilde{\mathbf{Z}}_{kl}(\alpha,\beta)=\frac{jZ_c^i}{d_p d_q \sin\alpha_r \sin l_k \sin l_l}\times$$
$$\times\sum_{\nu=-\infty}^{\infty}\sum_{\mu=-\infty}^{\infty}\left\{\left[\cos(\varphi_k-\varphi_l)-\delta_k^{\nu\mu}\delta_l^{\nu\mu}\right]\cdot Q_{\nu\mu}-\gamma_i^{\nu\mu}\delta_k^{\nu\mu}\delta_l^{\nu\mu}F_{\nu\mu}\right\}\times$$
$$\times\frac{\exp\left(-j\delta_k^{\nu\mu}l_k\right)-\cos(l_k)+i\delta_k^{\nu\mu}\sin l_k}{1-\left(\delta_k^{\nu\mu}\right)^2}\cdot\frac{\exp\left(i\delta_l^{\nu\mu}l_l\right)-\cos(l_l)-i\delta_l^{\nu\mu}\sin l_l}{1-\left(\delta_l^{\nu\mu}\right)^2}\times \tag{2}$$
$$\times\frac{\sin\left(\Delta_k^{\nu\mu}b_s\right)}{\Delta_k^{\nu\mu}b_s}\cdot\frac{\sin\left(\Delta_l^{\nu\mu}b_s\right)}{\Delta_l^{\nu\mu}b_s}\cdot\exp\left\{j\left[\chi_1^{\nu}(x_k-x_l)+\chi_2^{\nu\mu}(y_k-y_l)\right]\right\},$$

где Z_c^i - характеристическое сопротивление материала i-го слоя подложки, в котором находится точка наблюдения; l_k, l_l - длины сегментов, на которых определены функции $\vec{\varphi}_k(\xi)$ и $\vec{\varphi}_l(\xi)$; φ_k, φ_l - углы между осью x, введенной в разделе 1 прямоугольной системы координат, и векторами $\vec{\xi}_0^k$ и $\vec{\xi}_0^l$ соответственно; $Q_{\nu\mu}$, $F_{\nu\mu}$ - коэффициенты, значения которых определяются геометрией и материальными параметрами подложки АР; x_k, y_k, x_l, y_l - координаты точек, принятых за начало k-ой и l-ой токовых функций;

$$\chi_1^{\nu}=S_x+\frac{2\pi\nu}{d_p};\ \chi_2^{\nu\mu}=S_y-\frac{2\pi\nu}{d_p\operatorname{tg}\alpha_r}+\frac{2\pi\mu}{d_p\sin\alpha_r}- \tag{3}$$

коэффициенты распространения в плоскости решетки;

$$\gamma_i^{\nu\mu}=\sqrt{\left(\chi_1^{\nu}\right)^2+\left(\chi_2^{\nu\mu}\right)-1}\ - \tag{4}$$

коэффициент распространения в направлении нормали к плоскости решетки в i-ом слое подложки;

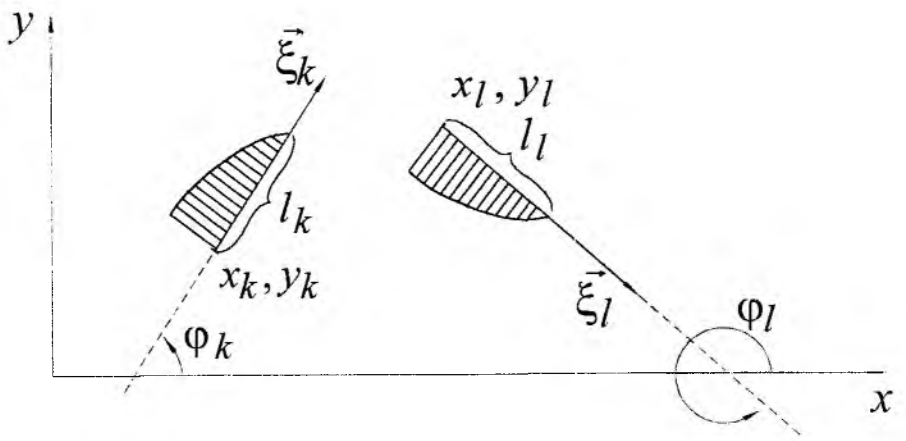
$$\delta_m^{\nu\mu}=\chi_1^{\nu}\cos\varphi_m+\chi_2^{\nu\mu}\sin\varphi_m,\quad m=k,l, \tag{5}$$

$$\Delta_m^{\nu\mu}=\chi_1^{\nu}\sin\varphi_m-\chi_2^{\nu\mu}\cos\varphi_m,\quad m=k,l\,. \tag{6}$$

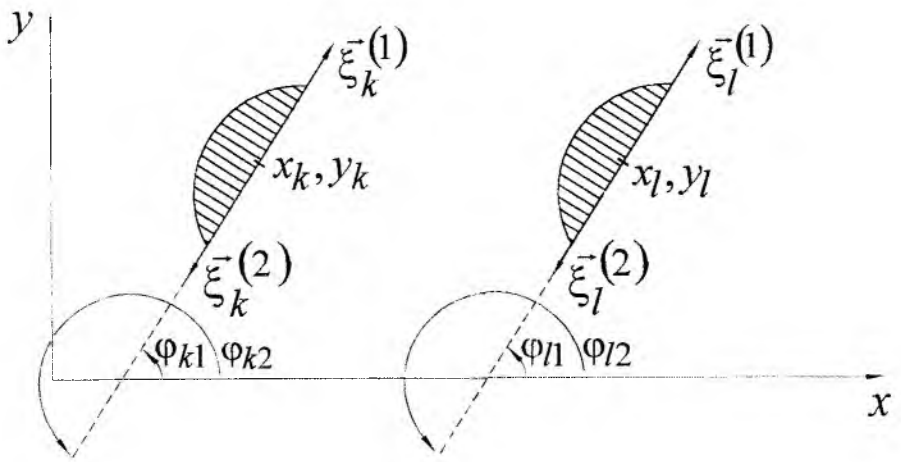
Однако непосредственно по (2) целесообразно рассчитывать только взаимные сопротивления токов в местах соединений проводников излучателя. Для вычисления взаимного сопротивления между элементарными токами, определенными на прямолинейных участках ветвей, целесообразно рассмотреть взаимное сопротивление токов, показанных на рис. 1 б,в. Очевидно, что спектральная плотность их взаимного сопротивления может быть определена как

$$\tilde{Z}_{kl}(\alpha,\beta)=-\tilde{Z}_{kl}^{11}(\alpha,\beta)-\tilde{Z}_{kl}^{22}(\alpha,\beta)+\tilde{Z}_{kl}^{12}(\alpha,\beta)+\tilde{Z}_{kl}^{21}(\alpha,\beta). \tag{7}$$

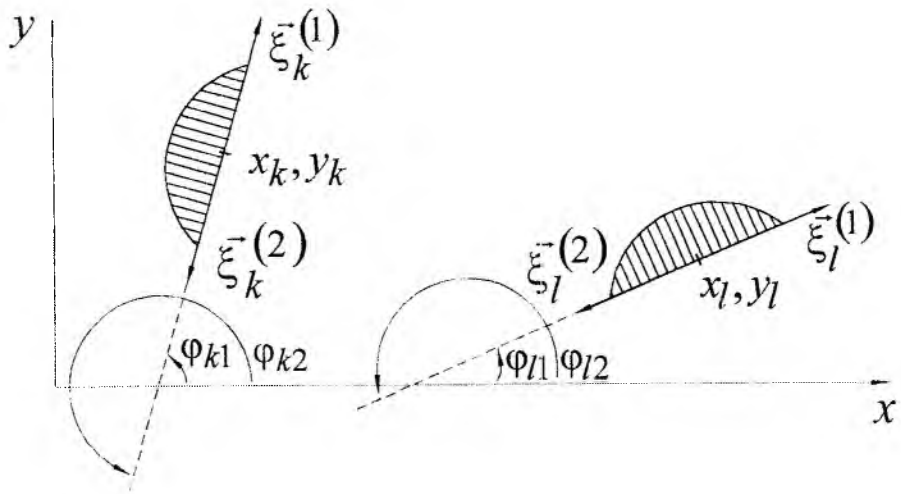
Слагаемые $\tilde{Z}_{kl}^{mn}(\alpha,\beta)$, $m=1,2$; $n=1,2$ в (7) представляют собой спектральные плотности собственных (при $m=n$) и взаимных (при $m\neq n$) сопротивлений между соответствующими элементами токов $\vec{\varphi}_k(\xi)$ и $\vec{\varphi}_l(\xi)$ и рассчитываются по (2). Выполнив суммирование, получим:


y
$\vec{\xi}_k$
x_l, y_l
l_l
l_k
x_k, y_k
$\vec{\xi}_l$
φ_l
φ_k
x

а)


y
$\vec{\xi}_k^{(1)}$
$\vec{\xi}_l^{(1)}$
x_k, y_k
x_l, y_l
$\vec{\xi}_k^{(2)}$
$\vec{\xi}_l^{(2)}$
φ_{k1} φ_{k2}
φ_{l1} φ_{l2}
x

б)


y
$\vec{\xi}_k^{(1)}$
x_k, y_k
$\vec{\xi}_k^{(2)}$
$\vec{\xi}_l^{(2)}$
$\vec{\xi}_l^{(1)}$
x_l, y_l
φ_{k1} φ_{k2}
φ_{l1} φ_{l2}
x

в)

Рис. 1

$$\tilde{\mathbf{Z}}_{kl}(\alpha,\beta)=\frac{-iZ_c^i}{d_p d_q \sin\alpha_r \sin l_k \sin l_l}\times$$
$$\times\sum_{\nu=-\infty}^{\infty}\sum_{\mu=-\infty}^{\infty}\left\{\left[\cos(\varphi_k-\varphi_l)-\delta_k^{\nu\mu}\delta_l^{\nu\mu}\right]\cdot Q_{\nu\mu}-\gamma_i^{\nu\mu}\delta_k^{\nu\mu}\delta_l^{\nu\mu}F_{\nu\mu}\right\}\times$$
$$\times\frac{\cos\left(\delta_k^{\nu\mu}l_k\right)-\cos(l_k)}{1-\left(\delta_k^{\nu\mu}\right)^2}\cdot\frac{\cos\left(\delta_l^{\nu\mu}l_l\right)-\cos(l_l)}{1-\left(\delta_l^{\nu\mu}\right)^2}\cdot\frac{\sin\left(\Delta_k^{\nu\mu}b_s\right)}{\Delta_k^{\nu\mu}b_s}\cdot\frac{\sin\left(\Delta_l^{\nu\mu}b_s\right)}{\Delta_l^{\nu\mu}b_s}\times \tag{8}$$
$$\times\exp\left\{i\left[\chi_1^{\nu}(x_k-x_l)+\chi_2^{\nu\mu}(y_k-y_l)\right]\right\}.$$

В частном случае, когда токи $\bar{\varphi}_k(\xi)$ и $\bar{\varphi}_l(\xi)$ имеют одинаковую амплитуду и определены на идентичных параллельных сегментах (рис. 1, б), выражение (8) принимает вид:

$$\tilde{\mathbf{Z}}_{kl}(\alpha,\beta)=\frac{-i4Z_c^i}{d_p d_q \sin\alpha_r \sin^2 l}\sum_{\nu=-\infty}^{\infty}\sum_{\mu=-\infty}^{\infty}\left\{\left[1-\delta_{\nu\mu}^2\right]\cdot Q_{\nu\mu}-\gamma_i^{\nu\mu}\delta_{\nu\mu}^2 F_{\nu\mu}\right\}\times$$
$$\times\left[\frac{\cos\left(\delta_{\nu\mu}l\right)-\cos l}{1-\delta_{\nu\mu}^2}\right]^2\left[\frac{\sin\left(\Delta_{\nu\mu}b_s\right)}{\Delta_{\nu\mu}b_s}\right]^2\exp\left\{i\left[\chi_1^{\nu}(x_k-x_l)+\chi_2^{\nu\mu}(y_k-y_l)\right]\right\}. \tag{9}$$

При необходимости определения характеристик излучателя на сетке значений α и β (то есть для различных направлений прихода возбуждающей плоской волны) объем, а следовательно и время вычислений, может быть сокращено за счет частичного вычисления сумм в выражениях (2), (8) и (9). В работе [4] приведены данные по расчетам такого типа для решетки излучателей с прямоугольной сеткой. В случае решетки с произвольной ячейкой алгоритм вычисления элементов матрицы обобщенных импедансов (МОИ) на сетке значений α и β имеет особенности, суть которых заключается в необходимости рассчитывать значение границы области суммирования по переменной μ для каждого определенного значения переменной ν.

Суммы в (2), (8) и (9) можно представить в виде нескольких частичных сумм, которые, в свою очередь, объединяются в две подгруппы:

$$S_i(\theta_0,\varphi_0)=\sum_{\nu=-\infty}^{\infty}\sum_{\mu=-M_{min}}^{M_{min}}+\sum_{\nu=-N_{min}}^{N_{min}}\sum_{\mu=-\infty}^{-M_{min}-1}+\sum_{\nu=-N_{min}}^{N_{min}}\sum_{\mu=-M_{min}-1}^{\infty} \tag{10}$$

и

$$S_2(\theta_0,\varphi_0)=\sum_{\nu=-\infty}^{-N_{min}-1}\sum_{\mu=M_{min}+1}^{\infty}+\sum_{\nu=N_{min}+1}^{\infty}\sum_{\mu=M_{min}+1}^{\infty}+$$
$$+\sum_{\nu=-\infty}^{-N_{min}-1}\sum_{\mu=-\infty}^{-M_{min}-1}+\sum_{\nu=N_{min}+1}^{\infty}\sum_{\mu=-\infty}^{-M_{min}-1}. \tag{11}$$

Выражения под знаками сумм опущены для краткости записей.

Таким образом, в $S_2(\theta_0,\varphi_0)$ группируются слагаемые с индексами $|\nu|>N_{\min}$ и $|\mu|>M_{\min}$, а в $S_1(\theta_0,\varphi_0)$ - все остальные. Поскольку зависимость элементов МОИ $\tilde{z}_{kl}(\alpha,\beta)$ от θ_0 и φ_0 определяется через параметры χ_1^{ν} и $\chi_2^{\nu\mu}$ (см. выражения (3)), значения $N_{\min}$ и $M_{\min}$ можно выбрать такими, что будут справедливы соотношения

$$1 << \frac{2\pi\nu}{d_p}, \tag{12}$$

$$1 << \left| \frac{2\pi\mu}{d_q tg\alpha_r} - \frac{2\pi\nu}{d_p tg\alpha_r} \right|. \tag{13}$$

При записи (12) и (13) учтено, что максимальное значение параметров S_x и S_y не может быть больше единицы. Кроме того, условие (13) может выполняться при любых μ и ν, если $d_q \sin\alpha_r \neq d_x \operatorname{tg}\alpha_r$. В противном случае может иметь место ситуация, когда правая часть этого неравенства обращается в нуль. Следовательно, элементы с индексами $\nu = \left(d_p / d_q \cos\alpha_r\right)$ также должны быть включены в сумму $S_1(\theta_0, \varphi_0)$.

Если выполняются условия (12) и (13), изменение значений S_x и S_y не оказывает существенного влияния на величину параметров χ_1^{ν} и $\chi_2^{\nu\mu}$. Очевидно, что при этом сумму $S_2(\theta_0, \varphi_0)$ достаточно вычислить только один раз при $\theta_0 = \varphi_0 = 0$, а приближенные значения элементов матрицы импедансов определять по формуле

$$\tilde{z}_{kl}(\alpha, \beta) = S_1(\theta_0, \varphi_0) + S_2(0,0).$$

Численные эксперименты по оценке точности такого алгоритма вычисления элементов матрицы импедансов показали, что погрешность расчета не превышает 1%, если значения $N_{\min}$ и $M_{\min}$ составляют 20% от числа слагаемых по μ и ν, учитываемых при точном определении элементов $\tilde{z}_{kl}(\alpha, \beta)$. При этом время вычислений может быть уменьшено до 2,5 раз.

7. Определение тензорной функции Грина плоскослоистой диэлектрической среды

Одним из существенных моментов при расчете микрополосковых антенн является определение тензорной функции Грина (ТФГ) многослойной плоскопараллельной диэлектрической среды, в виде которой моделируется их подложка. Скалярные элементы ТФГ могут быть представлены различными способами [5, 6, 7 и др.], в частности – как соответствующие компоненты электрического и/или магнитного векторного потенциалов некоторого локального распределения тока в одном из слоев диэлектрика.

Рассмотрим задачу возбуждения электромагнитного поля в среде, геометрия которой показана на рис. 2 из [1]. Пусть в i–ом слое в точке $M(x', y', z')$ расположен элементарный электрический диполь длиною dl, характеризуемый моментом $\vec{p} = j \cdot dl \cdot \vec{p}_0$. где j - плотность тока диполя, $\vec{p}_0$ - единичный вектор, характеризующий направление оси диполя. На бесконечности выполняются нулевые условия излучения, а на границах слоев – условия непрерывности векторов электромагнитного поля:

$$B_{\perp}^1 = 0; \; H_{\parallel}^1 = 0; \; E_{\parallel}^1 = 0 \quad \text{при} \quad z = 0 \tag{14}$$

и

$$B_{\perp}^{n-1} = B_{\perp}^n; \; H_{\parallel}^{n-1} = H_{\parallel}^n; \; E_{\parallel}^{n-1} = E_{\parallel}^n \quad \text{при} \quad z = z_n, \; 2 \le n \le N, \tag{15}$$

где индексы $\perp$ и $\parallel$ соответственно обозначают нормальные и тангенциальные составляющие векторов.

Будем характеризовать поле в рассматриваемой среде электрическим векторным потенциалом $\vec{A}$, который удовлетворяет системе волновых уравнений

$$\Delta\vec{A}^n + k_n^2 \vec{A}^n = \begin{cases} 0 & , \text{при } n \neq i; \\ -j \cdot \delta(x - x') \cdot \delta(y - y') \cdot \delta(z - z'), & \text{при } n = i, \end{cases} \tag{16}$$

где $k_n = \omega\sqrt{\tilde{\varepsilon}_n \mu_n}$ - волновое число, ω - круговая частота.

Из анализа геометрии задачи очевидно, что компоненты тока j_x и j_y порождают по две компоненты векторного потенциала: A_x, A_z и A_y, A_z соответственно, тогда как компонента тока j_z - только одну компоненту векторного потенциала: A_z. Следовательно, в наиболее общем виде искомую ТФГ можно представить следующей матрицей:

$$\overline{G} = \begin{bmatrix} G_{xx} & 0 & 0 \\ 0 & G_{yy} & 0 \\ G_{xz} & G_{yz} & G_{zz} \end{bmatrix}, \tag{17}$$

элементы G_{kl} которой представляют собой l -тые компоненты векторного потенциала $\vec{A}$, обусловленные k-ыми компонентами тока $\vec{j}$. В приложении к анализу плоской решетки микрополосковых излучателей непосредственный интерес представляют две пары элементов тензора: G_{xx}, G_{xz} и G_{yy}, G_{yz}. Последние имеют идентичный вид, поскольку среда неограниченна в плоскости $X0Y$. Таким образом, для решения задачи достаточно определить компоненты векторного потенциала, обусловленные током j_x.

С учетом изложенного систему волновых уравнений (16) можно записать в виде системы дифференциальных уравнений в частных производных относительно неизвестных компонент векторного потенциала A_x и A_z:

$$\begin{cases} \dfrac{\partial^2 A_x^n}{\partial x^2} + \dfrac{\partial^2 A_x^n}{\partial y^2} + \dfrac{\partial^2 A_x^n}{\partial z^2} + k_n^2 A_x^n = \begin{cases} 0 & \text{, при } n \neq i; \\ -j \cdot \delta(x-x') \cdot \delta(y-y') \cdot \delta(z-z'), & \text{при } n \neq i; \end{cases} \\ \dfrac{\partial^2 A_x^n}{\partial x^2} + \dfrac{\partial^2 A_x^n}{\partial y^2} + \dfrac{\partial^2 A_x^n}{\partial z^2} + k_n^2 A_x^n = 0 \qquad \text{, при } n = \overline{1, N}. \end{cases} \tag{18}$$

Для представления граничных условий (14), (15) в терминах векторного потенциала воспользуемся известными из электродинамики соотношениями [8]

$$\begin{cases} \vec{H}^n = rot\vec{A}^n; \\ \vec{E} = \dfrac{1}{i\omega\tilde{\varepsilon}_n}\left(graddiv + k_n^2\right)\vec{A}^n; \quad n = \overline{1, N}, \end{cases} \tag{19}$$

где $\vec{H}^n$ и $\vec{E}^n$ - векторы напряженности магнитного и электрического поля в n -ом слое диэлектрической структуры. В рассматриваемом случае $\vec{A}^n = A_x^n \vec{x}_0 + A_z^n \vec{z}_0$ и, следовательно, (19) принимает вид:

$$\begin{cases} \vec{H}^n = \dfrac{\partial A_z^n}{\partial y} \cdot \vec{x}_0 + \left(\dfrac{\partial A_x^n}{\partial z} - \dfrac{\partial A_z^n}{\partial x} \right) \cdot \vec{y}_0 - \dfrac{\partial A_x^n}{\partial y} \cdot \vec{z}_0, \quad n = \overline{1, N} \\ \vec{E} = \dfrac{1}{i\omega\tilde{\varepsilon}_n} \left[\left(\dfrac{\partial}{\partial x} \cdot \vec{x}_0 + \dfrac{\partial}{\partial y} \cdot \vec{y}_0 + \dfrac{\partial}{\partial z} \cdot \vec{z}_0 \right) \left(\dfrac{\partial A_x^n}{\partial x} + \dfrac{\partial A_z^n}{\partial z} \right) + k_n^2 \left(A_x^n \cdot \vec{x}_0 + A_z^n \cdot \vec{z}_0 \right) \right], \end{cases} \tag{20}$$

где $\vec{x}_0, \vec{y}_0, \vec{z}_0$, -орты осей x, y, z введенной прямоугольной системы координат.

На основании (20) получим следующие выражения для условий непрерывности векторов поля на границах слоев структуры:

$$\tilde{\mu}_1 \frac{\partial A_x^1}{\partial y} = 0;\ \frac{\partial A_x^1}{\partial z} = \frac{\partial A_z^1}{\partial x};\ \frac{\partial A_z^1}{\partial y} = 0;\ \frac{\partial}{\partial x}\left(\frac{\partial A_x^1}{\partial x} + \frac{\partial A_z^1}{\partial z}\right) + k_1^2 A_x^1 = 0;\ \frac{\partial}{\partial y}\left(\frac{\partial A_x^1}{\partial x} + \frac{\partial A_z^1}{\partial z}\right) = 0 \text{ при } z = 0 \quad (21)$$

и

$$\tilde{\mu}_{n-1} \frac{\partial A_x^{n-1}}{\partial y} = \tilde{\mu}_n \frac{\partial A_x^n}{\partial y};\ \frac{\partial A_x^{n-1}}{\partial z} - \frac{\partial A_z^{n-1}}{\partial x} = \frac{\partial A_x^n}{\partial z} - \frac{\partial A_z^n}{\partial x};\ \frac{\partial A_z^{n-1}}{\partial y} = \frac{\partial A_z^n}{\partial y},$$

$$\frac{1}{i\omega\tilde{\varepsilon}_{n-1}}\left[\frac{\partial}{\partial x}\left(\frac{\partial A_x^{n-1}}{\partial x} + \frac{\partial A_z^{n-1}}{\partial z}\right) + k_{n-1}^2 A_x^{n-1}\right] = \frac{1}{i\omega\tilde{\varepsilon}_n}\left[\frac{\partial}{\partial x}\left(\frac{\partial A_x^n}{\partial x} + \frac{\partial A_z^n}{\partial z}\right) + k_n^2 A_x^n\right],$$

$$\frac{1}{i\omega\tilde{\varepsilon}_{n-1}}\left[\frac{\partial}{\partial y}\left(\frac{\partial A_x^{n-1}}{\partial x} + \frac{\partial A_z^{n-1}}{\partial z}\right)\right] = \frac{1}{i\omega\tilde{\varepsilon}_n}\left[\frac{\partial}{\partial y}\left(\frac{\partial A_x^n}{\partial x} + \frac{\partial A_z^n}{\partial z}\right)\right] \text{ при } z = z_n,\ n = \overline{2, N}. \quad (22)$$

Таким образом, поле, создаваемое элементарным током j_x в слоистом диэлектрике, может быть определено в результате решения задачи Коши для системы дифференциальных уравнений (18) при выполнении условий (21) и (22). Указанное решение будем искать в спектральной области, для чего представим компоненты векторного потенциала в виде интегрального разложения Фурье в однородном поперечном сечении структуры (плоскости $X0Y$):

$$\overset{=n}{A}_\alpha = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int A_\alpha^n(x, y, z)\cdot\exp(-j\chi_1 x - j\chi_2 y)\cdot dxdy, \quad (23)$$

где индекс α может принимать значение x или y, а параметры χ_1 и χ_2 имеют смысл постоянных распространения в поперечном сечении. Неоднородность структуры в данном случае будет учитываться в характеристической части решения, связанной с координатой z.

В результате применения двукратного преобразования Фурье краевые задачи для уравнений в частных производных сводятся к более простым краевым задачам для обыкновенных дифференциальных уравнений следующего типа:

$$\begin{cases} \dfrac{d^2\overset{=n}{A}_x}{dz^2} - \gamma_n^2\overset{=n}{A}_x = \begin{cases} 0 & \text{, при } n \neq i; \\ P_x\cdot\delta(z - z') & \text{, при } n = i; \end{cases} \\ \dfrac{d^2\overset{=n}{A}_z}{dz^2} - \gamma_n^2\overset{=n}{A}_z = 0 \quad \text{, при } n = \overline{1, N} \end{cases} \quad (24)$$

с соответствующими граничными условиями

$$\overset{=1}{A}_x = 0,\ \frac{\overset{=1}{A}_x}{dz} = 0,\ \overset{=1}{A}_z = 0,\ \frac{\overset{=1}{A}_z}{dz} = 0, \text{ при } z = 0 \quad (25)$$

и

$$\tilde{\mu}_{n-1}\overset{=n-1}{A}_x = \tilde{\mu}_n\overset{=n}{A}_x,\ \frac{\overset{=n-1}{A}_x}{dz} = \frac{\overset{=n}{A}_x}{dz},\ \overset{=n-1}{A}_z = \overset{=n}{A}_z;$$

$$\frac{1}{\tilde{\varepsilon}_{n-1}}\left(\frac{d\overset{=n-1}{A}_z}{dz} + i\chi_1\overset{=n-1}{A}_x\right) = \frac{1}{\tilde{\varepsilon}_n}\left(\frac{d\overset{=n}{A}_z}{dz} + i\chi_1\overset{=n}{A}_x\right) \text{ при } z = x_n,\ n = \overline{2, N}. \quad (26)$$

В уравнениях (24) $\gamma_n^2 = \chi_1^2 + \chi_2^2 - k_n^2$ имеет смысл постоянной распространения вдоль оси z, а $P_x = -\frac{j_x}{2\pi}\exp(-i\chi_1 x' - i\chi_2 y')$.

Решения для уравнений системы (24) хорошо известны и имеют следующий вид:

$$\begin{cases} \overset{=n}{A}_x = B_1^n \exp(\gamma_n z) + B_2^n \exp(-\gamma_n z), \quad n \neq i; \\ \overset{=i}{A}_x = B_1^i \exp(\gamma_i z) + B_2^i \exp(-\gamma_i z) - \dfrac{P_x}{2\gamma_i}\exp\left[-\gamma_i (z - z')m\right], \quad n = i; \\ \overset{=n}{A}_z = C_1^n \exp(\gamma_n z) + C_2^n \exp(-\gamma_n z), \quad 1 \le n \le N, \end{cases} \tag{27}$$

где $m = \dfrac{(z - z')}{|z - z'|}$. С учетом нулевых условий излучения на бесконечности решения для слоя $n = N$ содержат только одно слагаемое:

$$\begin{cases} \overset{=N}{A}_z = B_2^N \exp(-\gamma_n z), \\ \overset{=N}{A}_z = C_2^N \exp(-\gamma_n z). \end{cases} \tag{28}$$

Неизвестные коэффициенты B_j^n и C_j^n $(1 \le n \le N,\ j = 1,2)$ определяются на основе заданных граничных условий путем подстановки выражений (27), (28) в уравнения (25), (26). В результате получаем следующую систему уравнений:

$$B_1^1 + B_2^1 = 0; \quad \gamma_1 B_1^1 + \gamma_1 B_2^1 = j\left(1 + \frac{\chi_1}{\chi_2}\right),$$

$$\mu_1\left(B_1^1 \exp(\gamma_1 z_2) + B_2^1 \exp(-\gamma_1 z_2)\right) = \mu_2\left(B_1^2 \exp(\gamma_2 z_2) + B_2^2 \exp(-\gamma_2 z_2)\right),$$

$$\gamma_1\left(B_1^1 \exp(\gamma_1 z_2) + B_2^1 \exp(-\gamma_1 z_2)\right) = \gamma_2\left(B_1^2 \exp(\gamma_2 z_2) + B_2^2 \exp(-\gamma_2 z_2)\right),$$

$$\bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet$$

$$\bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet$$

$$\mu_{i-1}\left(B_1^{i-1} \exp(\gamma_{i-1} z_i) + B_2^{i-1} \exp(-\gamma_{i-1} z_i)\right) =$$

$$= \mu_i\left(B_1^i \exp(\gamma_i z_i) + B_2^i \exp(-\gamma_i z_i) - \frac{P_x}{2\gamma_i}\exp(-\gamma_i z')\exp(\gamma_i z_i)\right),$$

$$\gamma_{i-1}\left(B_1^{i-1} \exp(\gamma_{i-1} z_i) + B_2^{i-1} \exp(-\gamma_{i-1} z_i)\right) =$$

$$= \gamma_i\left(B_1^i \exp(\gamma_i z_i) + B_2^i \exp(-\gamma_i z_i) - \frac{P_x}{2\gamma_i}\exp(-\gamma_i z')\exp(\gamma_i z_i)\right),$$

$$\mu_i\left(B_1^i \exp(\gamma_i z_{i+1}) + B_2^i \exp(-\gamma_i z_{i+1}) - \frac{P_x}{2\gamma_i}\exp(-\gamma_i z')\exp(\gamma_i z_{i+1})\right) =$$

$$= \mu_{i+1}\left(B_1^{i+1} \exp(\gamma_{i+1} z_{i+1}) + B_2^{i+1} \exp(-\gamma_{i+1} z_{i+1})\right),$$

$$\gamma_i\left(B_1^i \exp(\gamma_i z_{i+1}) + B_2^i \exp(-\gamma_i z_{i+1}) - \frac{P_x}{2\gamma_i}\exp(-\gamma_i z')\exp(\gamma_i z_{i+1})\right) =$$

$$= \gamma_{i+1}\left(B_1^{i+1} \exp(\gamma_{i+1} z_{i+1}) + B_2^{i+1} \exp(-\gamma_{i+1} z_{i+1})\right),$$

$$\bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet$$

$$\bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet$$

$$\mu_{N-2}\left(B_1^{N-2}\exp(\gamma_{N-2}z_{N-1})+B_2^{N-2}\exp(-\gamma_{N-2}z_{N-1})\right)=$$
$$=\mu_{N-1}\left(B_1^{N-1}\exp(\gamma_{N-1}z_{N-1})+B_2^{N-1}\exp(-\gamma_{N-1}z_{N-1})\right),$$
$$\gamma_{N-2}\left(B_1^{N-2}\exp(\gamma_{N-2}z_{N-1})+B_2^{N-2}\exp(-\gamma_{N-2}z_{N-1})\right)=$$
$$=\gamma_{N-1}\left(B_1^{N-1}\exp(\gamma_{N-1}z_{N-1})+B_2^{N-1}\exp(-\gamma_{N-1}z_{N-1})\right),$$
$$\mu_{N-1}\left(B_1^{N-1}\exp(\gamma_{N-1}z_N)+B_2^{N-1}\exp(-\gamma_{N-1}z_N)\right)=\mu_N B_1^N\exp(-\gamma_N z_N),$$
$$\gamma_{N-1}\left(B_1^{N-1}\exp(\gamma_{N-1}z_N)-B_2^{N-1}\exp(-\gamma_{N-1}z_N)\right)=-\gamma_N B_2^N\exp(-\gamma_N z_N),$$

$$C_1^1+C_2^1=-i\frac{j}{\chi_2};\qquad \gamma_1C_1^1+\gamma_1C_2^1=0;$$

$$C_1^1\exp(\gamma_1z_2)+C_2^1\exp(-\gamma_1z_2)=C_1^2\exp(\gamma_2z_2)+C_2^2\exp(-\gamma_2z_2),$$

$$\frac{1}{\tilde{\varepsilon}_1}\left(\gamma_1C_1^1\exp(\gamma_1z_2)-\gamma_1C_2^1\exp(-\gamma_1z_2)+i\chi_1\overset{=1}{A}_x\Big|_{z=z_2}\right)=$$
$$=\frac{1}{\tilde{\varepsilon}_2}\left(\gamma_2C_1^2\exp(\gamma_2z_2)-\gamma_2C_2^2\exp(-\gamma_2z_2)+i\chi_1\overset{=2}{A}_x\Big|_{z=z_2}\right),$$

$$\bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet$$
$$\bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet$$

$$C_1^{N-1}\exp(\gamma_{N-1}z_N)+C_2^{N-1}\exp(-\gamma_{N-1}z_N)=C_2^N\exp(-\gamma_Nz_N),$$

$$\frac{1}{\tilde{\varepsilon}_{N-1}}\left(\gamma_{N-1}C_1^{N-1}\exp(\gamma_{N-1}z_N)-\gamma_{N-1}C_2^{N-1}\exp(-\gamma_{N-1}z_N)+i\chi_1\overset{=N-1}{A}_x\Big|_{z=z_N}\right)=$$
$$=\frac{1}{\tilde{\varepsilon}_N}\left(-\gamma_NC_2^N\exp(-\gamma_Nz_N)+i\chi_1\overset{=N}{A}_x\Big|_{z=z_N}\right). \qquad (29)$$

Из рассмотрения структуры записанной СЛАУ очевидно, что наиболее рационально определить сперва значения коэффициентов $B_j^n\left(j=1,2\right)$, решив первые $2\left(N-1\right)$ уравнений, а затем, подставив полученные при этом значения в оставшиеся уравнения системы, найти значения коэффициентов $C_j^n\left(j=1,2\right)$. Кроме того, диагональный характер матрицы записанной СЛАУ позволяет использовать при определении вектора неизвестных коэффициентов последней рекуррентный алгоритм, при помощи которого исходная задача сводится к решению последовательности систем двух уравнений с двумя неизвестными. Физическая интерпретация такого алгоритма заключается в пересчете параметров всех выше- и нижележащих слоев подложки к слою, в котором расположен рассматриваемый точечный источник тока.

Полученные в результате решения системы уравнений (29) значения коэффициентов B_j^n и $C_j^n\left(j=1,2\right)$ подставляются в соответствующие выражения для Фурье-образов компонент векторного потенциала (27) и (28). После применения к последним обратного преобразования Фурье имеем:

$$A_{x,y}^{n}=\frac{j_{x,y}}{4\pi^2}\int_{-\infty}^{+\infty}\int VQ_n(\chi_1,\chi_2,z)\cdot\exp\left(i\chi_1(x-x')+i\chi_2(y-y')\right)\cdot d\chi_1 d\chi_2\,, \tag{30}$$

где

$$VQ_n(\chi_1,\chi_2,z)=\begin{cases}-\dfrac{\mu_i}{\mu_n}Q_1\dfrac{1-\phi_n\exp\left[2\gamma_n(z_{n+1}-z)\right]}{1-\phi_n}\exp\left[-\gamma_n(z_{n+1}-z)\right]\prod\limits_{k=n+1}^{i-1}T_k, & 1\le n\le i-1;\\ \dfrac{\left(1-\nu_i^x\right)\left[\phi_i^e\exp\left[2\gamma_i(z-z_i)+1\right]\right]}{1+\nu_i^x\phi_i^e}\dfrac{\exp\left[-\gamma_i(z-z_i)\right]}{2\gamma_i}, & n=i;\\ -\dfrac{\mu_i}{\mu_n}Q_2\dfrac{1+\phi_n\exp\left[2\gamma_n(z_{n+1}-z)\right]}{1+\phi_n}\exp\left[-\gamma_n(z-z_n)\right]\prod\limits_{k=i+1}^{n-1}T_k, & i+1\le n\le N,\end{cases} \tag{31}$$

и

$$A_z^n(x,y,z)=\frac{j_{x,y}}{4\pi^2}\int_{-\infty}^{+\infty}\int i\chi_{1,2}VF_n(\chi_1,\chi_2,z)\cdot\exp\left(i\chi_1(x-x')+i\chi_2(y-y')\right)\cdot d\chi_1 d\chi_2\,, \tag{32}$$

где

$$\begin{aligned}VF_n(\chi_1,\chi_2,z)=&-\frac{\exp\left[-\gamma_n(z-z_n)\right]}{1+\xi_n}\left\{R_n d_{n+1}\left[\xi_n+\exp\left[-2\gamma_n(z_{n+1}-z)\right]\right]+\right.\\ &+\left.\left[1-\exp\left[-2\gamma_n(z_{n+1}-z)\right]\right]\sum_{k=1}^{n-1}(1+\xi_k)S_k\prod_{l=k+1}^{n-1}R_l\exp\left(-\gamma_l h_l\right)-\exp\left[-\gamma_n(z_{n+1}-z)\right]\right\}\cdot\\ &\cdot\left[1+\xi_n\exp\left[2\gamma_n(z_{n+1}-z)\right]\right]\sum_{k=n}^{N-1}S_k d_{k+1}\prod_{l+n+1}^{k}t_l,\quad 1\le n\le N.\end{aligned} \tag{33}$$

Переменные в (31) и (33) определяются следующим образом:

$$Q_1=\frac{1}{2\gamma_i}\frac{\left(\nu_i^x-1\right)\left(\phi_i^e+1\right)}{1+\nu_i^x\phi_i^e}, \tag{34}$$

$$Q_2=\frac{\exp\left(-\gamma_i h_i\right)}{2\gamma_i}\frac{\left(\nu_i^x-1\right)\left(\phi_i^e+1\right)}{1+\nu_i^x\phi_i^e}, \tag{35}$$

$$T_k=\begin{cases}\dfrac{1-\nu_k}{1-\phi_k}\exp\left(-\gamma_k h_k\right), & 1\le k\le i-1;\\ \dfrac{1+\nu_k}{1+\phi_k}\exp\left(-\gamma_k h_k\right), & i+1\le k\le N-1,\end{cases} \tag{36}$$

$$\left.\begin{aligned}&\nu_1=1\\ &\phi=\nu\exp\left(-2\gamma h\right)\\ &\nu_{n+1}=\frac{\mu_{n+1}\gamma_n\left(1+\phi_n\right)-\mu_{n+1}\gamma_n\left(1-\phi_n\right)}{\mu_{n+1}\gamma_n\left(1+\phi_n\right)+\mu_{n+1}\gamma_n\left(1-\phi_n\right)}\\ &\nu_i=\nu_i^e\end{aligned}\right\}\ 1\le n\le i-1, \tag{37}$$

$$\left.\begin{aligned}
&\phi_N = 0 \\
&\phi_n = \nu_n \exp(-2\gamma_n h_n) \\
&\nu_n = \frac{\mu_{n+1}\gamma_n(1+\phi_{n+1}) - \mu_n\gamma_{n+1}(1-\phi_{n+1})}{\mu_{n+1}\gamma_n(1+\phi_{n+1}) + \mu_n\gamma_{n+1}(1-\phi_{n+1})} \\
&\nu_i = \nu_i^{\sigma}, \quad \phi_i = \phi_i^{\sigma}
\end{aligned}\right\} \quad i \le n \le N-1, \tag{38}$$

$$\left.\begin{aligned}
&\varphi_1 = 1 \\
&\xi_n = \varphi_n \exp(-2\gamma_n h_n) \\
&\varphi_{n+1} = \frac{\widetilde{\varepsilon}_n\gamma_{n+1}(1+\xi_n) - \widetilde{\varepsilon}_{n+1}\gamma_n(1-\xi_n)}{\widetilde{\varepsilon}_n\gamma_{n+1}(1+\xi_n) + \widetilde{\varepsilon}_{n+1}\gamma_n(1-\xi_n)} \\
&\xi_N = 0
\end{aligned}\right\} \quad i \le n \le N-1, \tag{39}$$

$$t_n = \frac{1+\varphi_n}{1+\xi_n}\exp(-\gamma_n h_n), \quad 2 \le n \le N-1, \tag{40}$$

$$S_n = \begin{cases}
\dfrac{\mu_i\left(\dfrac{\widetilde{\varepsilon}_n}{\mu_{n+1}} - \dfrac{\widetilde{\varepsilon}_{n+1}}{\mu_n}\right)Q_1 \prod\limits_{m=n+1}^{i-1} T_m}{\widetilde{\varepsilon}_n\gamma_{n+1}(1+\xi_n) + \widetilde{\varepsilon}_{n+1}\gamma_n(1-\xi_n)}, & 1 \le n \le i-1; \\
\dfrac{\mu_i\left(\dfrac{\widetilde{\varepsilon}_n}{\mu_{n+1}} - \dfrac{\widetilde{\varepsilon}_{n+1}}{\mu_n}\right)Q_1 \prod\limits_{m=i+1}^{n} T_m}{\widetilde{\varepsilon}_n\gamma_{n+1}(1+\xi_n) + \widetilde{\varepsilon}_{n+1}\gamma_n(1-\xi_n)}, & i \le n \le N-1,
\end{cases} \tag{41}$$

$$R_n = \frac{2\widetilde{\varepsilon}_{n+1}\gamma_n}{\widetilde{\varepsilon}_n\gamma_{n+1}(1+\xi_n) + \widetilde{\varepsilon}_{n+1}\gamma_n(1-\xi_n)}, \quad 1 \le n \le N-1, \tag{42}$$

$$\left.\begin{aligned}
&d_N = 1 \\
&d_n = R_n d_{n+1}\frac{1+\varphi_n}{1+\xi_n}\exp(-2\gamma_n h_n) + \frac{1-\exp(-2\gamma_n h_n)}{1+\xi_n}
\end{aligned}\right\} \quad 1 \le n \le N-1, \tag{43}$$

В частном случае, когда точка наблюдения находится на границе раздела n-го и $(n-1)$-го слоев, выражения для коэффициентов $VQ_{\nu\mu}$ и $VF_{\nu\mu}$ значительно упрощаются:

$$VQ_{\nu\mu}(\chi_1,\chi_2,z_n)=\begin{cases}-Q_1\dfrac{\mu_i}{\mu_n}\prod\limits_{k=n}^{i-1}T_k, & 1\le n\le i-1;\\ -Q_1, & n=i;\\ -Q_1\dfrac{\mu_i}{\mu_n}\prod\limits_{k=n}^{i-1}T_k, & i+1\le n\le N,\end{cases} \tag{44}$$

$$VF_{\nu\mu}(\chi_1,\chi_2,z_n)=-d_n\sum_{k=n}^{n-1}(1+\xi_n)S_k\prod_{l=k+1}^{n-1}R_l\exp(-\gamma_l h_l)-$$

$$-\exp(-\gamma_n h_n)(1+\varphi_n)\sum_{k=n}^{N-1}S_k d_{k+1}\prod_{l=n+1}^{k}t_l. \tag{45}$$

Таким образом, при определении элементов матрицы импедансов $[Z]$ в (2), (8), (9) используются выражения(44) и (45) для коэффициентов $VQ_{\nu\mu}$ и $VF_{\nu\mu}$ соответственно. Более общие соотношения (30) и (32) могут быть использованы при необходимости определять поле в произвольной точке пространства, моделируемого в виде слоисто-однородной диэлектрической среды с потерями.

Список литературы: **1** *Шокало В.М., Лучанинов А.И., Коновальцев А.А., Лучанинов Ю.А., Омаров М.А.* Алгоритм анализа эквидистантной решетки ленточных микрополосковых излучателей произвольной геометрии, адаптированный к расчету крупноапертурных антенн с нелинейными элементами 1. Модель, описание геометрии и система интегральных уравнений для токов ленточных микрополосковых излучателей сложной геометрии в составе бесконечной решетки // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2001. Вып. 117. **2** *Шокало В.М., Лучанинов А.И., Коновальцев А.А., Лучанинов Ю.А., Омаров М.А.* Алгоритм анализа эквидистантной решетки ленточных микрополосковых излучателей произвольной геометрии, адаптированный к расчету крупноапертурных антенн с нелинейными элементами 2. Решение системы интегральных уравнений // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2001. Вып. 118. **3** Вычислительные методы в электродинамике/ Под ред. Р.Митры. М.: Мир, 1977. 486 с. **4** *Татарников Д.В., Павлов С.А.* Повышение эффективности численного алгоритма задач анализа периодических антенных решеток/ Изв. высш. учеб. завед. МВ и ССО СССР. Радиоэлектроника. Киев, 1988. 8 с. (Рус.) Библиогр.: 4 назв. Деп. в ВИНИТИ, 20.07.88, № 6395-В88. **5** *Панченко Б.А., Нефедов Е.И.* Микрополосковые антенны. М.: Радио и связь, 1986. 144 с. **6** *Тозони О.В.* Метод вторичных источников в электротехнике. М.: Энергия, 1975. 296 с. **7** *Das N.K., Pozar D.M.* A generalized spectral-domain Green's function for multilayer dielectric substrates with application to multilayer transmission lines// IEEE Trans. Microwave Theory and Tech. 1987. Vol. 35. PP. 326 – 335. **8** *Марков Г.Т., Чаплин А.Ф.* Возбуждение электромагнитных волн. М.: Радио и связь, 1983. 295 с.

Харьковский государственный технический университет радиоэлектроники

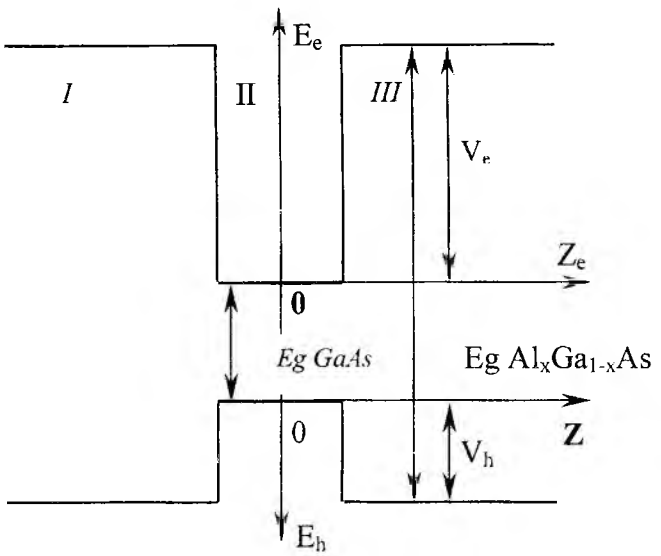
Поступила в редколлегию 11.12.2000

УДК 621.373.826

А. Г. ПАЩЕНКО, канд. физ.-мат. наук.

ВЛИЯНИЕ ВНЕШНЕГО ПЕРЕМЕННОГО ЭЛЕКТРИЧЕСКОГО ПОЛЯ НА ЭНЕРГЕТИЧЕСКИЕ СОСТОЯНИЯ ЧАСТИЦ И КВАЗИЧАСТИЦ В КВАНТОВОРАЗМЕРНОЙ СТРУКТУРЕ

Для решения ряда практических задач по использованию полупроводниковых светодиодов и лазеров на основе квантоворазмерных структур возникает необходимость исследования влияния внешних электрических или магнитных полей, изменяющихся во времени, на энергетические состояния частиц и квазичастиц (электронов, лёгких и тяжёлых дырок), которые находятся в активной области прибора и участвуют в процессах излучательной рекомбинации. Для решения этой проблемы в данной работе используется теория возмущений, развитая в фундаментальных работах по квантовой механике [1 – 4].



Рассмотрим квантоворазмерную структуру, энергетический профиль которой приведен на рисунке. Если на эту систему не действуют возмущения то частицы и квазичастицы (электроны и дырки) в этой квантоворазмерной структуре находятся в стационарном состоянии, описываемом стационарным уравнением Шредингера:

$$\hat{H}_0\Psi^0 = \mathrm{E}_0\Psi^0. \tag{1}$$

В этом уравнении $\hat{H}_0$ и Ψ^0 являются функцией только координаты, т.е. $\hat{H}_0 = \hat{H}_0(z)$ и $\Psi^0 = \Psi^0(z)$. Собственные значения энергии частиц и квазичастиц E_0 и их волновые функции $\Psi'^0(z)$ определяются при решении уравнения (1). В том случае, если энергетический профиль квантоворазмерной структуры представляет собой прямоугольную квантовую яму для электронов и дырок, то решение уравнения (1) является каноническим и приведено в работах [5 – 7].

Если на рассматриваемую квантоворазмерную структуру действует возмущение, зависящее от времени, оператор которого – $\hat{V}(z,t)$, то уравнение Шредингера в этом случае может быть записано так:

$$i\hbar\frac{\partial\Psi(z,t)}{\partial t} = \left[\hat{H}_0 + \hat{V}(z,t)\right]\Psi(z,t). \tag{2}$$

Пусть в некоторый момент времени $t = t_0$ в системе частиц и квазичастиц в рассматриваемой квантоворазмерной структуре возмущение отсутствует, т.е. $\hat{V}(z,t) = 0$. Тогда уравнение (2) можно переписать в виде:

$$i\hbar\frac{\partial\Psi^0(z,t)}{\partial t} = \hat{H}_0(z)\Psi^0(z,t). \tag{3}$$

При этом:
$$\Psi(z,t) = \Psi^0(z,t)\Big|_{t=t_0}. \tag{4}$$

Решением этого уравнения является функция:

$$\Psi_n^0(z,t) = \Psi_n^0(z)\cdot e^{-\frac{i}{\hbar}E_n^0 t}. \tag{5}$$

Функции $\Psi_n^0(z,t)$ образуют полную замкнутую систему собственных функций оператора $\hat{H}_0(z)$ и, следовательно, любую функцию $\Psi(z,t)$ можно разложить в ряд Фурье по функциям $\Psi_n^0(z)$. Тогда общим решением уравнения Шредингера (2) при $\hat{V}(z,t)=0$ будет:

$$\Psi(z,t)=\sum_{n=1}^{\infty}C_n\Psi_n^0(z)\cdot e^{-\frac{i}{\hbar}E_n^0 t}, \qquad (6)$$

где C_n – произвольные постоянные, удовлетворяющие условию нормировки.

Практически это соответствует случаю, когда система частиц и квазичастиц (электронов и дырок) находится в каком то одном определённом состоянии n так что волновая функция этого состояния определяется выражением (6). При этом необходимо предположить, что коэффициенты разложения – C_n в (6) не являются функциями времени, т.е. при $t=t_0$ все C_n равны нулю за исключением одного, например, C_m, который, согласно условию ортонормировки, можно принять равным единице.

Для моментов времени больших $t=t_0$ оператор возмущения $\hat{V}(z,t)\neq 0$. В этом случае волновая функция, удовлетворяющая уравнению (2), может быть записана так:

$$\Psi(z,t)=\sum_{n=1}^{\infty}C_n(t)\Psi_n^0(z)\cdot e^{-\frac{i}{\hbar}E_n^0 t}. \qquad (7)$$

Подстановка решения (7) в уравнение (2) преобразует его к виду:

$$i\hbar\sum_{n=1}^{\infty}\frac{dC_n(t)}{dt}\Psi_n^0(z)\cdot e^{-\frac{i}{\hbar}E_n^0 t}=\sum_{n=1}^{\infty}\hat{V}(z,t)C_n(t)\Psi_n^0(z)\cdot e^{-\frac{i}{\hbar}E_n^0 t}. \qquad (8)$$

После умножения уравнения (8) скалярно на функцию $\Psi_m^{0*}(z)\cdot e^{\frac{i}{\hbar}E_m^0 t}$ и интегрирования по z, оно преобразуется в систему дифференциальных уравнений относительно коэффициентов разложения $C_n(t)$:

$$i\hbar\sum_{n=1}^{\infty}\frac{dC_n(t)}{dt}\cdot\delta_{mn}\cdot e^{i\omega_{mn}t}=\sum_{n=1}^{\infty}C_n(t)e^{i\omega_{mn}t}\hat{V}_{mn}(t), \qquad (9)$$

где $\delta_{mn}=\int_{-\infty}^{+\infty}\Psi_m^{0*}(z)\Psi_n^0(z)dz$.

Так как $\delta_{mn}=1$ при $m=n$ и $\delta_{mn}=0$ при $m\neq n$, то уравнение (9) принимает вид:

$$i\hbar\frac{dC_m(t)}{dt}=\sum_{n=1}^{\infty}C_n(t)\hat{V}_{mn}(t)e^{i\omega_{mn}t}, \qquad (10)$$

где $\hat{V}_{mn}(t)=\int_{-\infty}^{+\infty}\Psi_m^{0*}(z)\hat{V}(z,t)\Psi_n^0(z)dz$; $\omega_{mn}=\frac{E_m-E_n}{\hbar}$.

Выражение (10) представляет собой бесконечную систему линейных, однородных дифференциальных уравнений первого порядка, в которых неизвестными функциями будут $C_n(t)$. Уравнение (10) отражает тот факт, что переход системы в состояние m зависит от всех состояний системы, которые при действии данного возмущения комбинируют с состоянием m. Следовательно, если один из коэффициентов, например $C_m(t)$, изменился, то должны измениться и другие коэффициенты, но так, чтобы сумма:

$$\sum_{n=1}^{\infty}|C_n(t)|^2 = 1.$$

Решение уравнения (10) может быть представлено в виде ряда:

$$C_m(t) = \sum_{k=0}^{\infty}\lambda^k C_m^{(k)}(t). \tag{11}$$

Этот ряд сходится, при $0 < \lambda \le 1$.

Подстановка (11) в (10) даёт:

$$i\hbar \sum_{k=0}^{\infty}\lambda^k \frac{dC_m^{(k)}(t)}{dt} = \sum_{k=0}^{\infty}\lambda^{k+1}\sum_{n=1}^{\infty}C_n^{(k)}(t)\hat{V}_{mn}(t)e^{i\omega_{mn}t}. \tag{12}$$

Приравнивая коэффициенты при равных степенях λ, находим:

$$i\hbar\frac{dC_m^{(0)}(t)}{dt} = 0, \tag{13}$$

$$i\hbar\frac{dC_m^{(1)}(t)}{dt} = C_n^{(0)}(t)\hat{V}_{mn}(t)e^{i\omega_{mn}t}. \tag{14}$$

В общем виде, уравнение для нахождения коэффициентов $C_m^{(k)}(t)$, записывается так:

$$i\hbar\frac{dC_m^{(k)}(t)}{dt} = \sum_{n=1}^{\infty}C_n^{(k-1)}(t)\hat{V}_{mn}(t)e^{i\omega_{mn}t}. \tag{15}$$

Коэффициенты $C_m(t)$ могут быть определены с точностью до любого порядка с помощью метода последовательных приближений.

Итак, пусть рассматриваемая система находится в одном из собственных стационарных состояний $\Psi_n^0(z)\cdot e^{-\frac{i}{\hbar}E_n^0 t}$ при $t = t_0$. Тогда $C_m(t_0) = C_m^0(t) = \delta_{mn}$ даёт решение задачи в нулевом приближении:

$$i\hbar\frac{dC_m^{(0)}(t)}{dt} = 0. \tag{16}$$

Очевидно, что решение уравнения (16) будет:

$$C_m^{(0)}(t) = const. \tag{17}$$

При $m = n$ $C_m^{(0)}(t) = 1$, при $m \ne n$ $C_m^{(0)}(t) = 0$.

Поправка первого порядка получается из уравнения (14):

$$i\hbar\frac{dC_m^{(1)}(t)}{dt} = e^{i\omega_{mn}t}\hat{V}_{mn}(t). \tag{18}$$

Пусть возмущение действует на рассматриваемую систему частиц и квазичастиц в течение определённого времени, начиная от t_0 до некоторого τ. Положив $t_0 = 0$, проинтегрируем (19) по t в заданном интервале и получим:

$$C_m^{(1)}(t) = -\frac{i}{\hbar}\int_0^{\tau}e^{i\omega_{mn}t}\hat{V}_{mn}(t)dt. \tag{19}$$

Это решение системы (14) в первом приближении для моментов времени t, лежащих в пределах $0 \le t \le \tau$. Подставляя первое приближение для $C_m^{(1)}(t)$ в правую часть (15), положив $k = 2$, найдём

уравнение для второго приближения:

$$i\hbar\frac{dC_m^{(2)}(t)}{dt}=\sum_{n=1}^{\infty}C_n^{(1)}(t)\hat{V}_{mn}(t)e^{i\omega_{mn}t}. \qquad (20)$$

Так как $C_n^{(1)}(t)$ известные функции времени (19), то, интегрируя (20) по времени, можно найти $C_m^{(2)}(t)$, т.е. второе приближение. Эту процедуру можно продолжить и дальше, и она ведёт к точному решению для $C_m(t)$. В большинстве случаев достаточно ограничиться первым или вторым приближением.

Оператор возмущения в уравнении (2) для большинства практически важных случаев изменяется по гармоническому закону со временем.

В случае, если квантоворазмерная структура, показанная на рисунке, помещена во внешнее переменное магнитное поле $\vec{H}\cdot\cos(\omega t)$, направленное вдоль оси z, оператор возмущения $\hat{V}(z,t)$ равен:

$$\hat{V}(z,t)=-\mu\cdot\vec{H}\cdot z\cdot\cos(\omega t), \qquad (21)$$

где $\mu=-\frac{e\cdot\hbar}{2\cdot m^{*}\cdot c}$ (m^{*} – приведенная масса частицы или квазичастицы (электрона или дырки); c – скорость света).

Если рассматриваемая квантоворазмерная структура помещена во внешнее переменное электрическое поле, которое направлено вдоль оси z, то оператор возмущения $\hat{V}(z,t)$ записывается так:

$$\hat{V}(z,t)=e\cdot U\cdot z\cdot\cos(\omega t), \qquad (22)$$

где e – заряд электрона; U – напряженность приложенного электрического поля.

В дальнейшем будет рассматриваться задача о влиянии переменного электрического поля на энергетические состояния частиц и квазичастиц в одномерной прямоугольной квантоворазмерной структуре, показанной на рисунке. В этом случае матричные элементы оператора возмущения $V_{mn}(t)$ равны:

$$V_{mn}(t)=e\cdot U\cdot\cos(\omega t)\cdot\int_{-\infty}^{+\infty}\Psi_m^{0*}(z)\cdot z\cdot\Psi_n^{0}(z)dz. \qquad (23)$$

Поскольку для рассматриваемого случая возмущение изменяется во времени по гармоническому закону, то можно утверждать, что по окончании его действия и прошествии времени τ система снова возвращается в стационарное состояние, являющееся суперпозицией стационарных состояний невозмущённой системы:

$$\Psi(z,t)=\sum_{n=1}^{\infty}C_n(t)\Psi_n^{0}(z)\cdot e^{-\frac{i}{\hbar}E_n^{0}t}. \qquad (24)$$

Для рассматриваемого в работе случая собственные значения гамильтониана в уравнении (2) определяются аналогично тому, как это было сделано в работах [6,7], с учётом зависимости матричных элементов оператора возмущения от времени.

Список литературы: 1. *Ландау Л.Д., Лифшиц Е.М.* Квантовая механика (нерелятивистская теория) М.: Физматгиз. 1963. 704 с. 2. *Бом Д.* Квантовая теория. М.: Физматгиз, 1965. 727 с. 3. *Киттель Ч.* Квантовая теория твёрдых тел. М.: Наука, 1967. 491 с. 4. *Борисоглебский Л. А.* Квантовая механика: Учеб. пособие для физ. спец. вузов Минск: Университетское, 1988. 623 с. 5. *Пащенко А.Г., Ванцан В.М.* Исследование стационарных энергетических состояний экситонов Ванье-Мотта в полупроводниковых инжекционных лазерах на основе квантоворазмерных структур // Радиотехника. 1997. Вып. 102. С. 85-92. 6 *Пащенко А.Г.* Влияние внешнего стационарного электрического поля на энергетические состояния частиц и квазичастиц в квантоворазмерной структуре. Часть 1. Постановка задачи // Радиотехника. 2001. Вып. 117. С. 117 – 120. 7. *Пащенко А.Г.* Влияние внешнего стационарного электрического поля на энергетические состояния частиц и квазичастиц в квантоворазмерной структуре. Часть 2. Обсуждение результатов // Радиотехника. 2001. Вып. 118. С. 55 -60.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 20.04.2001

УДК 621.371

С. И. БАБКИН, канд.техн. наук, Г. В. ГРУША, канд. физ.-мат. наук,
Е. Г. ПРОШКИН, д-р техн.наук, Н. И. СЛИПЧЕНКО, канд.техн.наук

НЕКОТОРЫЕ ИССЛЕДОВАНИЯ ФАЗОВОЙ ПОВЕРХНОСТИ АКУСТИЧЕСКИХ ВОЛН ДЛЯ ЗАДАЧ РАДИОМЕТЕОРОЛОГИИ

Введение

Влажность атмосферного воздуха наряду с такими метеорологическими параметрами, как температура и давление воздуха, является основной характеристикой при исследовании хода показателя преломления радиоволн в тропосфере. В задачах радиометеорологии для УКВ-диапазона используется формула Дебая, хорошо совпадающая с эмпирической зависимостью. В то же время модель экспоненциально убывающей высотной зависимости стандартной радиоатмосферы характеризует только изменения показателя преломления в целом. Тонкая структура показателя преломления радиоволн в тропосфере создается ее неоднородностью: инверсионными слоями, аэрозолями, термиками, турбулентными движениями атмосферы, а также слоями повышенной влажности воздуха.

Опыт разработки и создания аппаратурных комплексов дистанционного измерения влажности воздуха показывает, что при этом требуется решить проблему искажающего влияния атмосферной турбулентности [1,2]. В то время как амплитудные методы измерения многочастотными системами акустического (АЗ) [3], и радиоакустического (РАЗ) [2] зондирования требуют привлечения информации о структурных характеристиках температуры воздуха и скорости ветра, фазовые методы [1,4] согласно теоретическим расчетам позволяют избавиться от влияния турбулентности (в пределах аппаратурной погрешности). Однако, фазовому методу свойственны трудности экспериментальной методики (поскольку должны измеряться очень малые разности фаз, создаваемые дисперсией скорости звука). Поэтому весьма важным является учет влияния вертикальной изменчивости параметров атмосферы.

1. Предварительный анализ задачи

Некоторые результаты расчета влияния градиентов основных метеорологических величин для измерений влажности воздуха двухчастотной системой АЗ были приведены в [4]. Важность задачи подтверждалась эмпирическими и расчетными оценками: при сухоадиабатическом градиенте температуры воздуха вариация скорости звука $\delta_c \approx 0{,}6\,\text{мс}^{-1}$, уже на высоте 100 м превышая дисперсионную разность скоростей звука $\Delta c \approx 0{,}12\,\text{мс}^{-1}$ (при температуре воздуха +20 ^{0} С). Еще больше превышение может наблюдаться в неустойчиво стратифицированной атмосфере или при сильных температурных инверсиях. В данной статье дан последовательный вывод соотношений для фазовой поверхности акустических волн с учетом молекулярной релаксациии и вертикальной изменчивости атмосферного пограничного слоя, пригодных для измерения влажности воздуха двухчастотными системами акустического (АЗ) и радиоакустического (РАЗ) зондирования с применением целочисленного преобразователя частоты [1,4].

2. Обоснование применимости геомегрико-акустических фазовых поверхностей в условиях молекулярной релаксации

При выводе дисперсионной формулы для разности акустических фаз сигналов, принимаемых системами АЗ и связанных с ними разностей фаз между доплеровскими сдвигами принимаемых радиочастот системы РАЗ, ранее [1,4] использовалось простейшее выражение для фазы плоской волны в изотермической атмосфере:

$$\varphi = 2\pi f \frac{z}{C}, \qquad (1)$$

где f – акустическая частота; z – высота зондирования; C – фазовая скорость звука.

В условиях вертикальной изменчивости атмосферных параметров для строгого определения фазы волны требуется решение уравнений акустики движущейся неоднородной среды. Предложенный в классических работах подход [6] основан на предположении “беззвучности” воздушной среды (отсутствия нелинейного возбуждения пульсаций, звучания нестационарного потока), в результате чего действие среды сводится к сносу акустической волны ветром со средней скоростью потока и взаимо-

действию с турбулентными неоднородностями. При стандартных условиях применимости высокочастотного разложения по степеням $1/k = c_n/2\pi f$, c_n – нормальная скорость звука,

системы уравнений акустики в первом приближении по $1/k$, названном в [6] геометрической акустикой, рабочее уравнение для расчетов имеет вид уравнения эйконала Θ:

$$\left|\vec{\nabla}\Theta\right|^2 = \frac{q^2}{c^2}, \tag{2}$$

где $q = c_n - \left(\vec{v}, \vec{\nabla}\Theta\right)$; $\vec{v}$ – скорость ветра; c – скорость звука в неподвижной среде. Значение эйконала Θ связано с фазой Φ колебаний быстро осциллирующих атмосферных параметров при медленной изменчивости их амплитуд на протяжении длины волны:

$$\Phi = \omega t - k\Theta, \quad \varphi = k\Theta, \tag{3}$$

где $\omega = 2\pi f$; $k = \omega / c_n$.

Уравнение эйконала в акустике (2), полученное для адиабатических процессов [6], дает возможность оценить степень температурной или ветровой рефракции [3,6], а также изучить структуру фазовой поверхности в зависимости от частоты.

Двухчастотные фазовые методы измерения атмосферной влажности системами АЗ и РАЗ, разработанные в Лаборатории зондирования атмосферы ХТУРЭ, опираются на явление дисперсии скорости звука во влажном воздухе, возникающее в результате молекулярной релаксации при распространении акустической волны [1,4]. В обзоре [7] доказано, что релаксационный процесс в акустической волне является адиабатическим. Поэтому для изучения направляющих косинусов нормали к фазе волны, характеризующих рефракцию, а также для изучения двухчастотных разностей акустических фаз, требующихся данным методом, применимо эйкональное уравнение (2). Как известно, решение нелинейного уравнения (2) относительно $\vec{\nabla}\Theta$ для случая, когда одна из осей направлена вдоль нормали $\vec{n}$ к поверхности постоянной фазы, превращает задачу в линейную [6]:

$$\frac{\partial\Theta}{\partial n} = \frac{c_n}{c + \vec{v}\vec{n}}, \quad \vec{C} = \left(c + \vec{v}\vec{n}\right)\vec{n}. \tag{4}$$

При вертикальном излучении акустических волн в условиях вертикальной изменчивости метеопараметров при горизонтальной однородности атмосферы можно исходить из постоянства градиентов вдоль осей OX и OY ($\frac{\partial\Theta}{\partial x} = \cos\varphi_0, \frac{\partial\Theta}{\partial y} = \cos\psi_0$). В результате при $\varphi_0 = \psi_0 = \pi/2$ для направляющих косинусов

$$\alpha = \frac{\partial\Theta}{\partial x} / \left|\vec{\nabla}\Theta\right|, \quad \beta = \frac{\partial\Theta}{\partial y} / \left|\vec{\nabla}\Theta\right|, \quad \gamma = \frac{\partial\Theta}{\partial z} / \left|\vec{\nabla}\Theta\right| \tag{5}$$

получим $\alpha = \beta = 0$, $\gamma = 1$, $\left|\vec{\nabla}\Theta\right| = \frac{\partial\Theta}{\partial z}$. Таким образом, в этом случае не происходит движения нормали в плоскости ветра и вертикали, а также не влияет поворот ветра, причем этот вывод справедлив для обеих частот.

3. Двухчастотные разности фаз с поправками на вертикальную изменчивость атмосферы

Переходя от эйконала к фазам (3), составим двухчастотную разность акустических фаз на высоте z на основании решения уравнения (4) в виде:

$$\Delta\varphi = 2\pi\left[f_1\int_0^z \frac{dz'}{c_1 + w_1} - f_2\int_0^z \frac{dz'}{c_2 + w_2}\right], \tag{6}$$

где $w = \vec{v}\vec{n}$ – вертикальная проекция скорости ветра. Соотношение (6) может считаться базовым для получения регистрируемых разностей фаз при акустическом зондировании, когда отражателем явля-

ется турбулентность атмосферы, и, вообще говоря, при радиоакустическом зондировании, когда отражателем радиоволн является акустическая волна. Приведем последовательный вывод рабочих формул для аппаратуры двухчастотного акустического зондирования , комплексированной с одночастотной системой РАЗ для уточнения высотного хода температурных профилей [8]. При одновременном излучении акустических волн двух частот можно считать $w_1 = w_2 = w$. Учитывая применение целочисленного преобразователя частоты на выходе акустического приемника и возможность появления сравнимых по порядку градиентов скоростей ветра и звука при малости реально наблюдающихся в АПС скоростей вертикального ветра ($w / c_i << 1$), приведем соотношение для разности акустических фаз к виду:

$$\Delta\varphi = 4\pi f \int_0^z \frac{\Delta C}{C_1 C_2} dz', \tag{7}$$

где $\Delta C = C_2 - C_1$ – разность фазовых скоростей звука в движущемся воздухе; $f = f_2 = mf_1$; m – кратность частот.

Заменим все интегралы вида $J = \int_0^z C^{-1}(z')dz'$ суммой интегралов:

$$J = \sum_{i=1}^{N} \int_{z_i}^{z_{i+1}} C^{-1}(z')dz'. \tag{8}$$

Применим линейное разложение фазовой скорости внутри слоя при малых значениях толщин слоев Δz_i:

$$C(z') \approx C(z_i) + \left(\frac{\partial C}{\partial z}\right)_i (z' - z_i). \tag{9}$$

Подставив (8) и (9) в (7) и учитывая, что в этом случае интегралы (8) аппроксимируются выражениями:

$$J^{(1,2)} = \sum_{i=1}^{N} \frac{\Delta z_i}{C_i}\left(1 - \alpha_i^{(1,2)}\right),$$

где $\alpha_i^{(1,2)} = \frac{1}{C_i}\left(\frac{\partial C}{\partial z}\right)_i \frac{\Delta z_i}{2}$, можно привести разность фаз (7) к виду:

$$\Delta\varphi = 4\pi f \sum_{i=1}^{N} \frac{\Delta z_i \Delta C_i}{C_i^2}(1 - \beta_i), \tag{10}$$

где C_i – среднее геометрическое скоростей звука двух частот в слое.

Чтобы выразить коэффициенты β_i через $\alpha_i^{(1,2)}$, воспользуемся в (7) заменой:

$$\int_0^z \frac{\Delta C}{C_1 C_2} dz' = J^{(1)} - J^{(2)},$$

а в коэффициенты $\alpha_i^{(1,2)}$ подставим $\frac{\partial C}{\partial z} = \gamma_c + \gamma_w$. В результате вычислений получим:

$$\beta_i = \frac{\Delta z_i}{2}\left[\frac{2\gamma_c}{c} + \frac{2\gamma_w}{c} - \frac{\gamma_{\Delta c}}{\Delta c}\right]_i. \tag{11}$$

Данная формула соответствует аппроксимации интегралов в (8) методом трапеций. При обработке результатов измерений удобно перейти от формулы (10) к рекуррентным соотношениям:

$$\Delta\varphi_{i+1} = \Delta\varphi_i + 4\pi f \frac{\Delta z_i}{C_i^2}(1-\beta_i)\Delta C_i . \qquad (12)$$

Тогда разность набегов акустических фаз между слоями равна:

$$\Delta\varphi_{i+1} - \Delta\varphi_i = 4\pi f \frac{\Delta z_i}{C_i^2}(1-\beta_i)\Delta C_i . \qquad (13)$$

4. О возможности уточнения результатов восстановления влажности воздуха градиентными измерениями метеопараметров

Исходя из полученных рабочих формул (12) и (13) нетрудно доказать, что для учета вертикальной изменчивости метеопараметров полученные ранее на основе упрощенной фазы акустической волны (1) формулы восстановления относительной влажности воздуха по данным зондирования следует видоизменить , произведя замену:

$$H = \frac{p}{e_0} F\left(\frac{\Delta\varphi}{z}, c\right) \to H_i = \frac{p}{e_0} F\left[\frac{\Delta\varphi_{i+1} - \Delta\varphi_i}{(1-\beta_i)\Delta z_i}, c_i\right], \qquad (14)$$

где p – атмосферное давление; e_0 – упругость насыщенного водяного пара; F – функциональная зависимость от измеряемых величин. Поскольку точность измерения скорости звука современной радиоакустической аппаратурой $\sigma_c \approx (0{,}06\text{-}0{,}17)$ мс$^{-1}$,что соответствует точности измерения температуры $\sigma_T \approx (0{,}1-0{,}3)^0 K$, не позволяет оценить дисперсионную разность скоростей звука $\Delta c \approx 0{,}12$ мс$^{-1}$ в прямых измерениях, для численной оценки коэффициентов β_i необходимо в соотношении (11) перейти к градиентам метеорологических величин.

Исходя из явного выражения для Δc [5,7]

$$\Delta c = \varepsilon \frac{c_0^2}{2c}\left(\frac{f_2^2}{f_r^2 + f_2^2} - \frac{f_1^2}{f_r^2 + f_1^2}\right),$$

где f_r – релаксационная частота; $\varepsilon = \left(c_\infty^2 - c_0^2\right)/c_0^2$ – зависящая от температуры релаксационная сила; c_∞ и c_0 – скорости звука на частотах $f >> f_r$ и $f << f_r$, приведем градиент дисперсионной разности скоростей звука к виду:

$$\frac{\gamma_{\Delta c}}{\Delta c} \approx \frac{\gamma_\varepsilon}{\varepsilon} - \frac{2\gamma_{f_r}}{f_r} B(f_r) + \frac{\gamma_c}{c}, \qquad (15)$$

где коэффициент $B(f_r) \to 1$ при $f_r >> f_1$, $f_r >> f_2$ и произведена замена $\frac{2\gamma_{c_0}}{c_0} - \frac{\gamma_c}{c} \approx \frac{\gamma_c}{c}$. На основе результатов теории молекулярной релаксации во влажном воздухе [7] и формулы Лапласа для скорости звука вычислим:

$$\frac{\gamma_{f_r}}{f_r} \approx 1{,}3\frac{\gamma_h}{h}, \quad \frac{\gamma_\varepsilon}{\varepsilon} \approx \frac{\gamma_T}{T}\left(\frac{\theta}{T} - 2\right), \quad \frac{\gamma_c}{c} \approx \frac{\gamma_T}{2T}, \qquad (16)$$

где $\theta \approx 2230^0 K$ – характеристическая температура кислорода; T – абсолютная температура воздуха, и выразим коэффициенты β_i через градиенты метеорологических величин T, w, h :

$$\beta_i = \frac{\Delta z_i}{2}\left[\frac{\gamma_T}{2T} + \frac{2\gamma_w}{c} + 2{,}6B(f_r)\frac{\gamma_h}{h} - \frac{\gamma_T}{T}\left(\frac{\theta}{T} - 2\right)\right]_i, \qquad (17)$$

где $h = (e/p)*100\%$ – молярная концентрация водяного пара; e – его парциальное давление. Данная рабочая формула получена без каких-либо ограничений температурного и влажностного диапазона. В приземном слое удобной метеорологической характеристикой является относительная влажность $H = (p/e_0)h$. Поскольку упругость насыщенного водяного пара e_0 зависит от температуры воздуха [9], градиент γ_h, вносящий основной вклад в коэффициенты β_i на основе численных оценок, приведенных в п.6, можно выразить через градиенты γ_T и γ_H. При практическом восстановлении влажности воздуха на основе измерений скорости звука на двух соседних уровнях высоты (доплеровской системой РАЗ) может вычисляться γ_T , а по доплеровским измерениям АЗ – γ_w, в то время как градиент относительной влажности может вводиться в коэффициент β_i формулы (14) только как поправка, исходя из вычисленного начального значения влажности на верхнем уровне высоты.

5. Численные оценки в различных метеоусловиях

Для численной оценки влияния вертикальной изменчивости атмосферы на основе рабочей формулы (17) необходимы эмпирические данные о градиентах метеовеличин. По результатам многих сезонных наблюдений усредненная относительная влажность H часто медленно убывает с высотой в слое от 100 м до 1000 м [9]. В подобных ситуациях градиент молярной концентрации водяного пара возникает в результате вертикальной изменчивости температуры воздуха, поскольку $h = H(e_0/p)$, где e_0 – упругость насыщенного пара , зависящая от температуры воздуха. На основе эмпирических формул из [9]:

$$e_0 = 6{,}11*10^A, \quad A = \frac{7{,}63t}{241{,}9 + t}, \quad t = T - 273{,}15,$$

при температуре воздуха $t = +20^0 C$, например,

$$\frac{\gamma_h}{h} \approx 18{,}2\frac{\gamma_T}{T}; \quad \frac{\gamma_{f_r}}{f_r} \approx 23{,}7\frac{\gamma_T}{T}; \quad \frac{\gamma_\varepsilon}{\varepsilon} \approx 5{,}6\frac{\gamma_T}{T}.$$

Сравнивая градиенты физических величин в (11) и учитывая, что градиенты усредненных скоростей вертикального ветра γ_w и звука γ_c сравнимы по порядку, можно убедиться, что основное влияние на коэффициенты β_i должна оказывать сильная зависимость релаксационной частоты от температуры воздуха, указанная еще в первоначальных исследованиях по молекулярной релаксации [7]. В результате, при температурах воздуха от $t = +20^0 C$ до $t = 0^0 C$ на основе (11) и (15) (либо (17)) получены численные оценки (при $\gamma_w \Rightarrow 0$)

$$\frac{\gamma_{\Delta c}}{\Delta c} \approx -(82{,}4\text{K}\,90{,}6)\frac{\gamma_c}{c}, \quad \beta_i \approx (21\ldots23)\frac{\gamma_T}{T}\Delta z_i, \qquad (18)$$

где толщина зондируемого слоя $\Delta z_i \approx 17$ м [8]. В реальной атмосфере, например, при наблюдавшихся отклонениях от сухоадиабатических условий в инверсионных слоях и при содарных наблюдениях туманов можно установить пределы $\gamma_T \le 3^0/100$ м,[9] в которых $|\beta_i| \le 0{,}04$ в слое $\Delta z = 17$ м. Другие метеорологические условия нередко наблюдаются на низких высотах (до 50 м) при сильном испарении с поверхности земли, создающем сильное влияние градиента влажности. Например, на основании наблюдений суточного хода парциального давления водяного пара e в [9] приведены значения $\gamma_e \le (0{,}065\text{K}\,0{,}1)$ мб м$^{-1}$ на высотах до 20 м. Для этого случая на основе рабочей формулы (17) и $\gamma_h / h = \gamma_e / e$ получены численные оценки поправочных коэффициентов $|\beta_i| \approx$ 0,06-0,2 в слое.

Численные оценки поправочных коэффициентов β_i можно применить для характеристики методической погрешности, возникающей в том случае, если они не учитываются. Так, на основе рекуррентной формулы (13) с ростом высоты зондирования должно происходить накопление методической линейной погрешности $\sigma_{\Delta\varphi} = \Delta\varphi - (\Delta\varphi)_{is}$, где $(\Delta\varphi)_{is}$ – разность фаз в изотермической атмосфере. При рассмотренных условиях $\sigma_{\Delta\varphi} / \Delta\varphi$ может меняться от 5% до (24...120)% на границе приземного слоя $z = 100$ м, достигая в предельных случаях (120...600)% в середине атмосферного пограничного слоя $z = 500$ м, что привело бы к значительной погрешности восстановления относительной влажности [4]. Для практических измерений метеорологических параметров (температура и влажность воздуха, скорость вертикального ветра) и их градиентов достаточно иметь аппаратуру двухчастотного АЗ, комплексированную с одночастотной системой РАЗ с точностью измерения $\sigma_T \approx 0{,}1^0$ К; $\sigma_W \approx 0{,}1$ мс$^{-1}$ и $\sigma_H / H \approx 3\%$, достижимой на основе расчета методической и аппаратурной погрешности.

О возможности появления градиентов температуры и влажности воздуха, значительно превышающих пороговый уровень их учета, создаваемый суммарной погрешностью измерения, может свидетельствовать ряд наблюдений изменчивости метеопараметров, не относящейся к воздействию турбулентных движений атмосферы. Так, в практических измерениях, проводившихся в ПНИЛ ЗА, отмечались колебания в температурных профилях до $1{,}5^0$/17 м [8], связывавшиеся со структурой приморского тумана. Можно предположить, что превышение в 4,5 раза фазовых сдвигов, измеренных двухчастотной системой АЗ на высотах до 50 м, над рассчитанными в изотермической атмосфере [4] соответствует температурному градиенту в указанных выше пределах и градиенту влажности, в несколько раз превышающему значение $\gamma_e \approx$ 1,7 мб в слое 17 м.

Заключение

В данной статье на основе традиционных соотношений акустики движущейся вертикально-изменчивой атмосферы и теории молекулярной релаксации во влажном воздухе получены практически применимые соотношения, позволяющие дополнить дистанционные измерения высотного хода влажности фазовым методом [4] градиентными измерениями метеопараметров. Градиентные измерения, представляющие собой совокупность измерений метеовеличины на нескольких уровнях высоты, целесообразно проводить при заметных отклонениях, создающих достаточно сильную рефракцию радиоволны. Модернизация фазовых методов измерения влажности воздуха позволяет повысить точность измерений, что представляется важным при радиометеорологических исследованиях в движущейся неоднородной атмосфере.

Список литературы: 1. *Babkin S.I., Delov I.A., Grusha G.V., Proshkin E.G., Slipchenko N.I.The possibility of radioacoustic sounding application for the remote measuring of the air humidity//* Proc. 10th Intern.Sympos. on Acoustic Remote Sensing and Associated Techniques of the Atmosphere and Oceans.Auckland,2000.New Zealand. P.296-298. 2. *Babkin S.I., Grusha G.V., Proshkin E.G.* The measurement of the air humidity by two-frequency radioacoustic sounding// 4th Intern.Sympos. on Tropospheric Profiling. Extended Abstracts. V.1.1998.Snowmass. Colorado. USA. P.25-27. 3. *Красненко Н.П.* Акустическое зондирование атмосферы. Новосибирск: Наука, Сибирское отделение,1986. 166 с. 4.*Babkin S.I., Delov I.A., Grusha G.V., Proshkin E.G.* The influence of atmosferic parameters variability on the accuracy of air humidity measurement by the phase method// Proc. 10th Intern.Sympos. on Acoustic Remote Sensing and Associated Techniques of the Atmosphere and Oceans. Auckland,2000. New Zealand. P.119-122. 5. *Бабкин С.И., Груша Г.В.* Оценка погрешности определения влажности в турбулентной атмосфере по разности фаз при радиоакустическом зондировании// Оптика атмосферы и океана. 1995. №4. С.60-66. 6. *Блохинцев Д.И.* Акустика неоднородной движущейся среды. 2 изд. М.:Наука, 1981. 206 с. 7.*Кнезер Г.* Релаксационные процессы в газах. В кн. Физическая акустика/ Под ред. У.Мэзона. М.: Мир,1968. Т.2, ч.А. С.155-221. 8. *Бабкин С.И., Делов И.А., Прошкин Е.Г.* Комплекс аппаратуры для совмещенного зондирования пограничного слоя атмосферы электромагнитными и акустическими волнами// Радиотехника. 1998. Вып. 106-107. С. 23-28. 9. *Хргиан А.Х.* Физика атмосферы. Л.: Гидрометеоиздат, 1969. 642 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 26.03.2001

УДК 621.396.96'06

В. М. КАРТАШОВ, канд. техн. наук

ВЛИЯНИЕ ВЗАИМНОГО ЭНЕРГЕТИЧЕСКОГО СПЕКТРА ЗОНДИРУЮЩИХ СИГНАЛОВ НА ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ СИСТЕМ ЗОНДИРОВАНИЯ АТМОСФЕРЫ

Задача выбора видов и параметров зондирующих сигналов является одной из наиболее важных при проектировании локационных систем, так как ее результат предопределяет, по существу, значения основных информационных характеристик, а также структуру проектируемой системы. Для радиоакустических и акустических систем зондирования атмосферы в настоящее время не существует приемлемой процедуры, позволяющей осуществлять анализ характеристик сигналов, а следовательно, производить аргументированный и акцентированный их выбор. В данной статье рассматривается подход, позволяющий анализировать свойства и особенности различных видов зондирующих сигналов, в том числе особенности различных комбинаций из акустической и электромагнитной волн для радиоакустических систем. Предлагаемая методика позволяет производить также синтез структур сигналов.

Как показано в [1], при рассеянии излученных волн на объектах, используемых в акустических и радиоакустических системах зондирования атмосферы, формируется взаимокорреляционная функция излученного сигнала и рассеивающей неоднородности:

$$y(r_1) = \int_{-\infty}^{\infty} e(2r)s(r - r_1)dr, \tag{1}$$

где $y(r_1)$ – рассеянный сигнал; e – зондирующий сигнал; s – неоднородность атмосферы (для радиоакустических систем – это акустическая волна), далее для удобства s будем называть также сигналом; r – пространственная координата в направлении зондирования.

Правая часть равенства (1) может быть переписана в пространстве волновых чисел k в виде:

$$y(r_1) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_e\left(\frac{k}{2}\right)S_{sr}^*(k)dk, \tag{2}$$

где $S_e(k/2)$, $S_{sr}(k)$ – пространственные спектры соответствующих процессов, индекс r в последнем члене означает, что Фурье-отображение соответствует сдвинутой по дальности на расстояние r_1 неоднородности.

Спектр смещенного в пространстве сигнала, как известно, определяется выражением $S_{sr}(k) = S_s(k)e^{-jkr_1}$. Подставив последнее соотношение в (2), имеем:

$$y(r_1) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_e\left(\frac{k}{2}\right)S_s^*(k)e^{jkr_1}dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_y\left(\frac{k}{2}\right)e^{jkr_1}dk. \tag{3}$$

Величина $S_y(k/2)$ представляет собой взаимный энергетический пространственный спектр процессов e и s:

$$S_y(k/2) = S_e(k/2)S_s^*(k). \tag{4}$$

Выражение (3), определяющее рассеянный сигнал, представляет собой преобразование Фурье взаимного энергетического пространственного спектра зондирующих колебаний. Здесь и далее в отношении e и s будем использовать для удобства терминологию радиоакустических систем. При $r_1 = 0$ в правой части (3) имеем корреляционный интеграл по пространственной частоте спектральных плотностей сигналов, совмещенных в пространстве. При любом другом фиксированном значении r_1 правая часть (3) также представляет собой корреляционный интеграл по частоте k, но уже для сигналов, смещенных на расстояние r_1.

Понятия взаимного энергетического пространственного спектра и корреляционной функции сигналов по пространственной частоте играют ключевую роль в понимании особенностей процессов рассеяния и формирования отраженной волны в системах зондирования атмосферы. По существу можно сказать, что взаимный энергетический пространственный спектр (в форме спектра временных частот) воспроизводится на выходе спектроанализатора радиоакустической системы, выполняющего анализ рассеянной волны. Действительно, сформированный в результате рассмотренных преобразований пространственно-временной сигнал достигает приемной антенны и преобразуется во временное колебание. После соответствующих преобразований и фильтрации в приемнике системы информационный сигнал, как правило, подвергается преобразованию Фурье, т. е., по существу, воспроизводится взаимный энергетический пространственный спектр взаимодействующих сигналов.

Естественно, что форма воспроизведенного спектра, заложенного в сигнал при рассеянии, а также его параметры, характеризующие, например, несимметричность или «скошенность» спектра, кардинальным образом влияют на значения основных информационных характеристик систем зондирования атмосферы.

Рассмотрим рассеяние волн в радиоакустических системах и покажем, что вид, параметры и принципиальная возможность существования рассеянного радиосигнала, полученного от звуковой посылки, полностью определяются особенностями взаимного энергетического пространственного спектра зондирующих акустического и электромагнитного колебаний. В радиоакустических системах $s(t,r)$ – излучаемый акустический сигнал. Достаточно распространено использование в качестве $s(t,r)$ импульсных акустических колебаний с гармоническим заполнением. В этом случае спектр $S_s(k)$ является узкополосным. Соотношение (4) отображает основную особенность рассеяния на такой неоднородности, т. е. его существенную частотную зависимость, которая проявляется как с энергетической, так и с информационной стороны.

Изменение метеопараметров по трассе зондирования приводит к деформации (растяжению или сжатию) звуковой волны вдоль координаты r, а следовательно, к перемещению $S_s(k)$ по оси частот k, вследствие чего максимумы пространственных спектров взаимодействующих сигналов не совпадают (см. рис.), а диапазон перекрытия спектров сужается. Амплитуда результирующего рассеянного сигнала при этом уменьшается. Представленные на рисунке спектры процессов $e(r)$ и $s(r)$ изображены один над другим, причем, ось пространственных звуковых частот для наглядности сжата в два раза в соответствии с условием Брэгга $k'_e = k'_s / 2$.

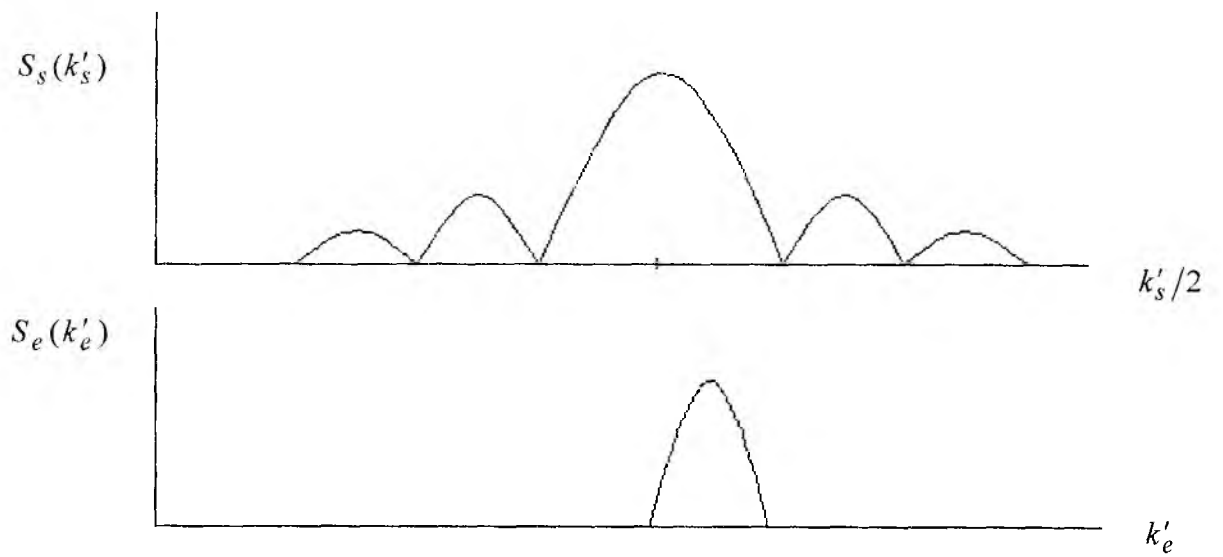


Рис.

Если максимумы спектров $S_e(k/2)$ и $S_s(k)$ не совпадают, т. е., как говорят, наблюдается расстройка по Брэггу, то максимум пространственного спектра $S_y(k/2)$ смещается дополнительно на

величину Δk вследствие неравномерности и несимметричности функции $S_s(k)$ в диапазоне перекрытия спектров. Это приводит к смещению максимума временного спектра рассеянного сигнала на величину $\Delta\omega = c\Delta k$ (где c – скорость распространения электромагнитных волн) и отличию частотного сдвига от чисто доплеровского. Можно рассматривать также смещение центра тяжести $S_y(k/2)$. Следует отметить, что при переходе от пространственных к временным частотам даже очень малое смещение Δk ввиду большого значения c трансформируется в ощутимую разность $\Delta\omega$. $\Delta\omega$ представляет собой систематическую ошибку при определении скорости звука по доплеровскому сдвигу частоты, наличие которой приводит к существенным погрешностям в определении температуры. При определенных условиях пространственный, а следовательно, и временной спектры рассеянного сигнала будут двумодовыми, что ранее отмечалось в работах по зондированию атмосферы.

Проиллюстрируем возможности физической интерпретации известных в литературе результатов анализа погрешностей оценки температуры системами радиоакустического зондирования (РАЗ) с использованием предложенной модели рассеяния. Запишем формулу для частоты рассеянного на звуке радиосигнала, полученную в [2,3] в результате строгого решения волновой задачи при гауссовых формах огибающих звукового импульса и пространственного окна:

$$\omega_p = \omega - 2k_e c_s + q c_s l_s / \left(l_p^2 + l_s^2\right), \tag{5}$$

где ω и $k_e = 2\pi/\lambda_s$ – частота и волновое число излучаемых электромагнитных колебаний; $l_s = c_s \tau_s$ – пространственная протяженность акустического волнового пакета (АВП); $q = 2k_e - k_s$ – параметр расстройки условия Брэгга; $k_s = 2\pi/\lambda_s$ – волновое число для звука; l_p – продольный размер области (пространственного окна), из которой возможен прием рассеянного излучения или, другими словами, размер области пересечения диаграмм направленности передатчика и приемника. Поскольку формула (5) получена для бистатической схемы зондирования и непрерывного радиосигнала, характеризующегося большим значением пространственной протяженности l_e, то роль l_e здесь выполняет параметр l_p.

При $q = 0$ частота рассеянного сигнала совпадает с доплеровской частотой ω_g, а доплеровский сдвиг $\Omega_g = 2k_e c_s$ равен частоте звука Ω. При наличии расстройки (когда $q \neq 0$) третий член в (5) описывает ошибку частотного сдвига $\Delta\omega$, зависящую от параметров l_p и l_s. С увеличением l_p и уменьшением l_s значение $\Delta\omega$ уменьшается. При $l_p \to \infty$ ошибка $\Delta\omega$ стремится к нулю.

Данные результаты получают четкую физическую интерпретацию в пространстве волновых частот, где уменьшение l_s приводит к расширению пространственного спектра АВП и уменьшению "скошенности" спектра в области перекрытия спектров взаимодействующих сигналов. Увеличение l_p (l_e) сопровождается сужением спектра радиосигнала и, соответственно, уменьшением влияния "скошенности" спектра АВП. При l_p (l_e) $\to \infty$ пространственный спектр радиосигнала стремится к δ-функции и ошибка $\Delta\omega \to 0$. Ширина спектральной линии рассеянного сигнала, как следует из предложенного рассмотрения рассеяния, также будет определяться значениями пространственных параметров l_p, l_e, l_s, что подтверждается результатами строгого решения соответствующих волновых уравнений [2,3].

При существовании в атмосфере высотного градиента температуры (если предположить, что градиент является постоянным) акустическая решетка превращается в линейно-частотно модулированный (ЛЧМ) сигнал [4,5], имеющий, как известно, прямоугольный амплитудный и квадратичный фазовый спектры. Влияние последнего приводит к дополнительному изменению формы рассеянного сигнала. Заметим, что в [3], вследствие неправильного физического толкования параметра l_p, получены завышенные значения погрешностей оценки температуры, вызванных влиянием градиента.

Если под воздействием метеопараметров спектр $S_s(k)$ смещается настолько, что спектры $S_e(k/2)$ и $S_s(k)$ в пространстве волновых чисел не перекрываются вовсе, то такие сигналы в соответствии с (3) становятся ортогональными, т.е. акустический сигнал в этом случае полностью прозрачен для радиосигнала. Совпадения максимумов спектров $S_e(k/2)$ и $S_s(k)$ на практике достигают адаптивным изменением частот зондирующих акустического и (или) электромагнитного сигналов.

Соотношения (1) и (3) отображают также изменения формы зондирующего радиосигнала при рассеянии. Увеличению длительности рассеянного радиосигнала вследствие конечной протяженности АВП, определяемому (1), соответствует сужение его пространственного спектра, как следует из (3).

Представление процесса рассеяния в области пространственных спектров наглядно отображает также особенности взаимодействия различных видов радио и акустических зондирующих сигналов. Для иллюстрации рассуждений будем использовать рисунок.

Если используются импульсный акустический и непрерывный электромагнитный сигналы, то спектр $S_e(k/2)$ стягивается в δ – функцию. Условие пространственного резонанса «выбирает» из совокупности бесконечных бегущих плоских волн, суперпозиция которых составляет акустический пакет, единственную спектральную компоненту, отвечающую условию Брэгга.

Если радиосигнал будет иметь конечную длительность (пространственную протяженность), а акустическое излучение – непрерывное, то на рисунке спектр $S_s(k)$ будет представлен δ – функцией. В этом случае звуковая волна с частотой k_s самостоятельно выделяет рассеиваемую пространственную гармонику, соответствующую условию брэгговского резонанса (если, конечно, таковая окажется, т.е. если спектры перекрываются).

Показателен случай, когда короткий радиоимпульс (например, наносекундной длительности) рассеивается на продолжительной акустической посылке. При этом $S_s(k)$ на рис. будет выглядеть в виде δ – функции, а $S_e(k/2)$ значительно расширится. Амплитуда рассеянного радиосигнала слабо зависит от изменения метеопараметров, вызывающих перемещение $S_s(k)$ по оси частот, а вот смещение максимума частотного спектра рассеянного сигнала по абсолютной величине может быть значительно большим, чем привычное значение доплеровского сдвига. Величина сдвига может достигать значений порядка ширины спектра зондирующего импульса $\Delta k \sim 2\pi / l_e$, $\Delta\omega \sim 2\pi c / l_e$, т.е. единиц, десятков, сотен мегагерц в зависимости от длительности радиосигнала.

Аналогично представляется рассеяние на звуке при использовании зондирующих электромагнитного и акустического сигналов в виде периодических последовательностей импульсов, имеющих дискретный спектр, в том числе квазинепрерывного зондирующего радиосигнала.

Изменения при рассеянии формы и параметров зондирующих сигналов акустических и радиоакустических систем по оси запаздывания и по оси частот определяются характеристиками и протяженностью области рассеяния в направлении зондирования, как следует из выражений (1) и (3).

В соответствии с этим становится возможной целенаправленная деформация рассеянного сигнала с целью достижения его требуемой формы путем соответствующего подбора (оптимизации) форм и параметров зондирующего сигнала и области рассеяния. В радиоакустических системах область рассеяния, как правило, задается акустическим сигналом, поэтому здесь подбираются характеристики зондирующих электромагнитного и акустического сигналов. Такие исследования представляют определенный теоретический интерес и в ряде случаев могут обеспечить необходимый практический результат.

Проиллюстрируем возможности синтеза необходимого вида сигнала, приходящего на вход приемника в акустических системах зондирования атмосферы. Рассмотрим бистатическую схему акустического зондирования [6]. В таких системах, как правило, используется непрерывный зондирующий сигнал, позволяющий повысить энергетический потенциал системы, но он создает помехи на входах приемников при приеме рассеянных сигналов в виде прямого зондирующего сигнала передатчика.

Влияние паразитного сигнала передатчика на работу приемных устройств можно устранить, если излучать импульсный акустический сигнал, имеющий малую скважность, с длительностью τ и периодом T, а принимать рассеянное излучение импульсами с длительностью $T - \tau$ и периодом T. Причем, прием в каждой точке пространства следует осуществлять тогда, когда в этой точке отсутствует прямой сигнал передатчика.

Такой режим работы становится возможным вследствие того, что рассеяние импульсных сигналов на целях, протяженных в направлении вектора рассеяния, сопровождается увеличением длительности рассеянных сигналов по сравнению с зондирующими импульсами, а при определенных условиях импульсный зондирующий сигнал при рассеянии превращается в непрерывное колебание. Рассмотрим более подробно эти условия, используя пространственные и временные характеристики сигналов. Изменения формы сигнала при рассеянии могут быть отображены также в области волновых чисел с использованием пространственных спектров сигналов и области рассеяния.

Если протяженность по вертикали l_v области рассеяния $l_v > \frac{c_s \tau}{2} \sin\frac{\theta}{2}$, где θ – угол рассеяния, то при использовании прямоугольного зондирующего импульса рассеянный сигнал будет иметь трапециевидную форму. Длительность плоской части сигнала

$$\tau_p = \frac{2l_v}{c_s \sin(\theta/2)} - \tau,$$

длительности фронтов – τ, а протяженность сигнала на уровне 0,5 составит

$$\tau_{0,5} = \frac{2l_v}{c_s \sin(\theta/2)}.$$

Если выбрать период T следования зондирующих импульсов $T \le \tau_{0,5}$, то рассеянные сигналы перекрываются во времени и образуют непрерывное колебание. Так, при c_s =340 м/с, l_v =15 м, θ =150^0, τ = 10 мс, имеем τ_p =81 мс, а $\tau_{0,5}$ =91мс. Следовательно, при $T \le 91$ мс рассеянный сигнал будет непрерывным.

Осуществляя прием сигнала в промежутках времени, когда зондирующий сигнал отсутствует в точке приема, полностью устраняем влияние последнего на процесс обработки рассеянного сигнала и измерения его информативных параметров. Способы обработки и измерения параметров импульсных сигналов, имеющих малую скважность, известны [7].

Анализ энергетических соотношений показывает, что при использовании данного способа зондирования (по сравнению с использованием непрерывного зондирующего сигнала) следует ожидать увеличения отношения сигнал-шум на входе приемников на несколько порядков, что существенно повышает качественные показатели системы.

Таким образом, представление процесса рассеяния волн на неоднородностях атмосферы в пространстве волновых чисел с использованием понятия взаимного энергетического пространственного спектра значительно облегчает анализ характеристик зондирующих и рассеянных сигналов систем акустического и радиоакустического зондирования атмосферы

Полученные результаты позволяют объяснить особенности взаимодействия различных видов электромагнитных и акустических сигналов, а также специфические погрешности оценивания температуры, свойственные методу РАЗ. Предложенная модель достаточно проста, конструктивна, адекватно описывает многие особенности рассеяния волн на решетках и может использоваться при решении задач, имеющих место при синтезе и анализе систем акустического и радиоакустического зондирования атмосферы.

Список литературы: 1. *Карташов В.М., Сакало С.Н.* Модель рассеяния волн на неоднородностях атмосферы // Радиоэлектроника и информатика. 2001. №2 (в печати). 2. *Гурвич А.С., Кон А.И., Татарский В.И.* Рассеяние электромагнитных волн на звуке в связи с задачами зондироваия атмосферы // Изв. вузов. Радиофизика. 1987. Т. 30, №4. С. 451-472. 3. *Каллистратова М.А., Кон А.И.* Радиоакустическое зондирование атмосферы. М.: Наука, 1985. 200с. 4. *Steinhagen H., Neisser J.* Improvement of the Altitude coverage of Temperature Measurements using RASS // Proc. 9 Int. Symposium on Acoustic Remote Sensing of the Atmosphere. 1996. P. 329-334. 5. *Goupil P., Klaus V., Cherel G., Durbe R.* On the Use of the Wavelet-Packet Transform to Improve the Measurement of the RASS Temperature Profiles // Proc. 9 Int. Symposium on Acoustic Remote Sensing of the Atmosphere. Vienna, 1998. P. 76-79. 6. *Принципы построения автоматизированных систем метеорологического обеспечения авиации*/ Под ред. Г.Г. Щукина. Л.: Гидрометеоиздат, 1991. 321 с. 7. *Справочник по радиолокации* в 4-х т. / Под ред. М. Сколника. М.: Сов. радио, 1979. Т. 3. 528 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 4.04.2001

УДК 532. 59

А. Ю. ПАНЧЕНКО, канд.физ.-мат.наук

К ОЦЕНКЕ ИНТЕНСИВНОСТИ ВТОРИЧНЫХ ИСТОЧНИКОВ ПОЛЯ ПРИ АКУСТИЧЕСКОМ ЗОНДИРОВАНИИ ТУРБУЛЕНТНЫХ ДВИЖУЩИХСЯ СРЕД

Практическая потребность дистанционного зондирования движущихся потоков заставляет скрупулезнее подойти к анализу известных положений. Рассеяние акустических волн турбулентной движущейся средой до настоящего времени имеет ряд невыясненных нюансов, что допускает разночтения и неадекватную интерпретацию экспериментальных результатов. Единственной причиной этому можно назвать лишь неоптимальное составление моделей реальных процессов, нерациональную или недостаточно удобную запись известных выражений. Это приводит к последовательности незначительных отклонений, которые на каждом отдельном этапе кажутся несущественными и легко исправимыми, но в итоге они могут дать столь существенное различие между теорией и экспериментом, что в некоторых случаях приводится постулировать дополнительные предположения и на их основе строить дополнительную теорию, с единственной целью в основных чертах восстановить соответствие. Однако, такой путь часто приводит в тупик. Поэтому при расширении сферы применения известных законов необходимо вновь возвращаться на начальные позиции, расширяя смысл исходных положений, и скрупулезнее проводить дальнейшие преобразования. В случае зондирования движущихся потоков это требует более детального анализа процессов, основные модели которых и математическое описание были предложены задолго до появления возможностей технической реализации дистанционного неконтактного зондирования сплошных сред.

При расчете параметров потоков реальных жидкостей и газов, как правило, требуется учитывать вязкость, но в случае анализа распространения акустических волн с высокой степенью точности можно ограничится приближением невязкой, нетеплопроводной среды. Поэтому для линеаризации системы уравнений гидродинамики и последующего получения волнового уравнения можно заранее исключить из рассмотрения вязкость, и вместо весьма сложного уравнения Навье-Стокса использовать более простое уравнение Эйлера, описывающее движение идеального газа.

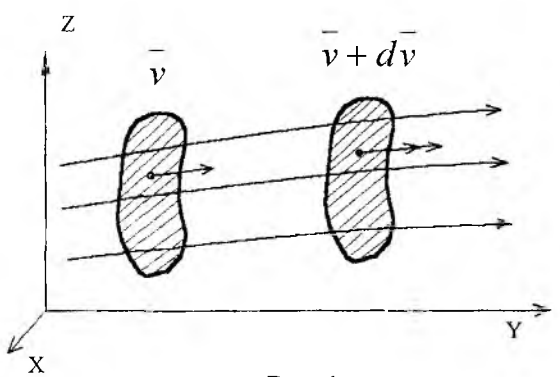


Рис. 1

Традиционный вывод уравнения Эйлера основан на втором законе Ньютона. Иногда в качестве исходной посылки берут закон сохранения импульса, который является его прямым следствием. Для вывода уравнения второй закон Ньютона записывают для малой массы газа, движущейся вместе с потоком (рис.1). В процессе вывода совершают предельный переход при стремлении массы к нулю, что позволяет представить конечное соотношение через удельные параметры среды и точечные параметры потока. В результате получается уравнение вида

$$\rho \frac{d\vec{v}}{dt} = -\operatorname{grad} p + \vec{f}\,. \qquad (1)$$

Производную скорости в левой части называют полной производной скорости, отнесенной к некоторому малому объему среды, движущемуся вместе с потоком. В правой части стоит сумма удельных значений сил, действующих на эту точку. Она состоит из градиента давления в самом потоке и внешней силы, если ее источники существуют.

Первое уравнение систем уравнений гидродинамики и акустики – уравнение непрерывности – записывается также для малого объема. Но этот объем неподвижен относительно выбранной системы координат (рис.2). Первое уравнение имеет вид:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{v}) = \omega \quad , \qquad (2)$$

где ω – удельная мощность внешних источников вещества.

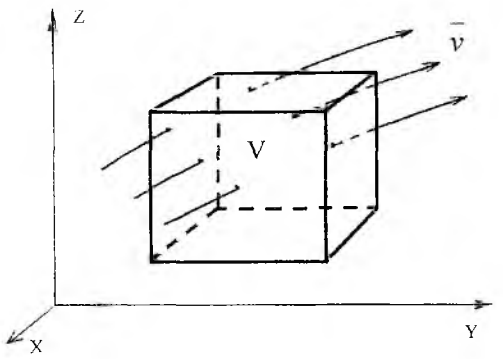


Рис. 2

Составлено оно для элементарного объема пространства (рис.2), в котором изменение плотности определяется двумя причинами: во-первых, перенос плотности потоком из соседних точек пространства; во вторых, действием внешних источников ω, если такие имеются.

Таким образом, уравнение (1) составлено для производной скорости, отнесенной к малому объему, движущемуся вместе с потоком, а уравнение (2) – для объема, неподвижного относительно выбранной системы координат. Совместно использовать эти уравнения для описания одного процесса нельзя, так как они относятся к разным объектам. Это противоречие обусловлено тем, что исходные физические законы существенно отличаются и, соответственно, их математическая интерпретация, также существенно разная. Иногда, если это допускают условия конкретной задачи, противоречие разрешается выбором системы координат, движущейся вместе с потоком. Но для большинства задач такое представление неудобно.

В известных в настоящее время источниках указывается, что Эйлер привел эти уравнения к системе с помощью следующего приема. Для совмещения объектов, для которых записаны уравнения, необходимо в уравнении (1) выделить производную скорости в локальной точке пространства. Для этого полная производная при перемещении массы от одной точки пространства к другой представлена в виде суммы производной, действующей в локальной точке, и существующего в потоке изменения скоростей по пространству – конвективной производной. Конвективная производная по отношению к движущейся массе по сути определяется переносом пространственного изменения скорости, точно так же как в уравнении непрерывности перенос градиента плотности. разницей скоростей между двумя близкорасположенными точками пространства:

$$\frac{d\vec{v}}{dt} = \frac{\partial \vec{v}}{\partial t} + v_x \frac{\partial \vec{v}}{\partial x} + v_y \frac{\partial \vec{v}}{\partial y} + v_z \frac{\partial \vec{v}}{\partial z} \quad . \tag{3}$$

Естественно, что при этом иной смысл должен быть вложен и в слагаемые правой части (1). В этом случае $\operatorname{grad} p$ и $\vec{f}$ должны быть отнесены также к неподвижной точке пространства, а не к движущейся вместе с потоком. Но учитывая, что предельный переход приводит к бесконечно малым расстояниям, эти величины также будут отличаться на бесконечно малую величину.

В уравнении Эйлера отсутствуют иные источники, кроме источников внешней силы. Рассмотрим другой подход для вывода уравнения Эйлера, который можно провести непосредственно для точки, фиксированной относительно выбранной системы координат.

В качестве исходной переменной воспользуемся кинетической энергией, которая так же как и масса обладает свойством неуничтожимости. Ранее [1,2] использовалась полная энергия, что создавало значительные трудности при получении конечных выражений. Поэтому подобный подход для анализа процессов рассеяния звука широкого применения не нашел, по сути его анализ и использование до получения конечных выражений ограничились лишь [1]. Более того, эти трудности привели к тому, что впоследствии этот анализ был признан как грубый.

Вероятно, раздельное использование кинетической и потенциальной энергии, для которой аналогичный подход показан в [3], не делалось ранее из-за неуверенности в возможности получения векторного уравнения для скорости на основании исходного скалярного уравнения.

Для удельного значения кинетической энергии можно записать

$$\frac{\partial e_c}{\partial t} + \operatorname{div}\left(e_c \vec{v}\right) = \omega_c \quad . \tag{4}$$

Удельное значение кинетической энергии можно выразить как:

$$e_c = \rho \frac{v^2}{2} = \rho \frac{\left(\vec{v} \cdot \vec{v}\right)}{2} \quad . \tag{5}$$

Подставляя (5) в (4), имеем:

$$\frac{v^2}{2}\frac{\partial \rho}{\partial t}+\rho\left(\vec{v}\cdot\frac{\partial \vec{v}}{\partial t}\right)+\frac{v^2}{2}\operatorname{div}(\rho\vec{v})+\left(\rho\vec{v}\cdot\operatorname{grad}\frac{v^2}{2}\right)=\omega_c \quad . \tag{6}$$

Второе и третье слагаемое в левой части представляют собой левую часть уравнения (2), умноженную на половину квадрата скорости, их можно заменить на $\omega v^2/2$. Градиент в четвертом слагаемом можно представить на основании известной формулы векторного анализа

$$\operatorname{grad}\frac{v^2}{2}=(\vec{v}\cdot\nabla)\vec{v}+[\vec{v}\times\operatorname{rot}\vec{v}] \quad . \tag{7}$$

Тогда

$$\frac{v^2}{2}\omega+\rho\left(\vec{v}\cdot\frac{\partial \vec{v}}{\partial t}\right)+\rho(\vec{v}\cdot(\vec{v}\cdot\nabla)\vec{v})+\rho(\vec{v}\cdot[\vec{v}\times\operatorname{rot}\vec{v}])=\omega_c \quad . \tag{8}$$

Четвертое слагаемое тождественно равно нулю. Объединяя под знаком скалярного произведения второе и третье слагаемые, получим:

$$\frac{v^2}{2}\omega+\rho\left(\vec{v}\cdot\left(\frac{\partial \vec{v}}{\partial t}+(\vec{v}\cdot\nabla)\vec{v}\right)\right)=\omega_c \quad . \tag{9}$$

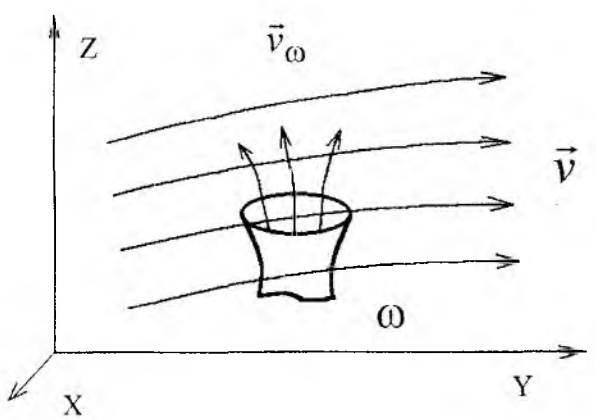


Рис. 3

Второй сомножитель в скалярном произведении представляет сумму первых двух слагаемых уравнения Эйлера. Но в уравнении Эйлера не присутствует источник массы ω, появившийся в левой части. Тем не менее его появление имеет ясный физический смысл. Задавая источник вещества в уравнении непрерывности, ничего не было сказано относительно скорости, которую сообщает веществу источник в момент ее поступления в поток. Умолчание подразумевает нулевую скорость, значит, для того, чтобы вновь поступившее вещество стало двигаться совместно с остальным веществом потока, ему должна перейти часть кинетической энергии потока. Первое слагаемое в данном уравнении представляет этот механизм (рис.3). Однако, сделанное в начале вывода приближение относительно среды, как об идеальном газе, накладывает более строгие ограничения на действие внешних источников массы. Приближение идеального газа частично предполагает отсутствие взаимодействие между его молекулами. А именно, такого взаимодействия, которое обеспечивает передачу тепла, и такого, которое обеспечивает передачу момента количества движения через вязкость. Оставлена только та часть, которая обуславливает давление. Поэтому, строго говоря, поток не может передать поступившему веществу свою скорость. Для этого газ должен обладать вязкостью. В идеальном газе в этом случае будут существовать два потока молекул, не изменяющие своего направления. Таковы издержки сделанного предположения. Поэтому, говоря о действии источника вещества в потоке идеального газа, необходимо сразу определять его так, чтобы скорость вносимого вещества была равна скорости потока, или решать локальную задачу передачи моментов количества движения с учетом вязкости.

Далее рассмотрим источники кинетической энергии. Здесь нужно рассмотреть два механизма: первый – действие силы на движущуюся массу, находящуюся в заданной точке пространства. Этот механизм может содержать две составляющих – обусловленную градиентом давления в самом потоке и обусловленную некоторой внешней силой. Второй механизм возникнет в том случае, если действующий в потоке источник вещества будет сообщать ему некоторую начальную скорость. Для пол-

ного исключения этого механизма из рассмотрения нужны основания, а при анализе его действия необходимо учитывать издержки приближения идеального газа, которые отмечены выше. Тогда

$$\omega_c = (\operatorname{grad} p \cdot \vec{v}) + (\vec{f} \cdot \vec{v}) + \omega \frac{(\vec{v}_\omega \cdot \vec{v}_\omega)}{2} \quad . \tag{10}$$

Подставляя (10) в (9), получим:

$$\left(\vec{v} \cdot \left(\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} - \frac{1}{\rho} \operatorname{grad} p - \frac{1}{\rho} \vec{f} \right) \right) + \frac{\omega}{\rho} \frac{(\vec{v} \cdot \vec{v})}{2} = \frac{\omega}{\rho} \frac{(\vec{v}_\omega \cdot \vec{v}_\omega)}{2} \quad . \tag{11}$$

При выполнении для источника массы вышеперечисленных условий слагаемые, которые его содержат, скомпенсируют друг друга. Тогда скалярное произведения в правой части должно быть равно нулю, что при ненулевой скорости потока потребует равенства нулю второго сомножителя, представляющего собой уравнение Эйлера в каноническом виде.

Представленный подход инвариантен тому, с помощью которого получено уравнение непрерывности. Кроме того. инвариантность подходов состоит и в том, что в обоих случаях исходными переменными являются скалярные величины, для которых справедлив закон сохранения, и в обоих случаях совершался переход к удельным значениям при стремлении к нулю объема, фиксированного относительно выбранной системы координат. Последующий переход от уравнения сохранения кинетической энергии к уравнению Эйлера является лишь формальными преобразованиями.

Можно отметить, что представленный подход является более строгим, поскольку заставил обратить внимание на действие источников вещества. Кроме того, прослеживая путь возникновения каждого из слагаемых, можно заметить, что конвективная производная $(\vec{v} \cdot \nabla)\vec{v}$ возникла из второго слагаемого в левой части уравнения сохранения – $\operatorname{div}(e_c \vec{v})$, таким образом, по аналогии с уравнением непрерывности, можно сказать, что конвективное слагаемое само представляет механизм изменения скорости в выбранной точке пространства за счет переноса этого свойства из соседней точки. Это положение полностью соответствует строгим математическим представлениям. Конвективная производная является произведением вектора скорости $\vec{v}$ на дифференциальную диаду $\nabla\vec{v}$ – тензор второго ранга, которую иногда называют градиентом вектора $\operatorname{Grad}\vec{v}$ [2].

Действующие силы $\operatorname{grad} p$ и $\vec{f}$ в этом случае также оказываются отнесенными к выбранной точке пространства.

Сделанные выше выводы относительно источников вещества необходимы для корректного представления источников вторичного поля акустических волн, образовавшегося на неоднородностях среды для решения задач в приближении однократного рассеяния. Эта соответствует случаю, когда источником вторичного поля является сама турбулезированная среда, облучаемая звуковыми волнами. При этом не имеет значения, введены эти источники в исходной системе акустических уравнений, или представлены в совокупности с другими в конечном волновом уравнении. Конечно же, при представлении вида источников, как в случае однократного рассеяния, так и, тем более, в случае генерации звука турбулентным потоком, необходимо учесть возможные последствия приближения невязкого газа.

Проведенный анализ позволяет сделать вывод о высокой эффективности инвариантного подхода и математического аппарата, которые дают возможность строго, без дополнительных предположений получить уравнения для вектора, используя исходные положения для скалярной величины.

Список литературы: 1. *Блохинцев Д.И.* Акустика неоднородной движущейся среды. М.: Наука, 1981. 208 с. 2. *Лойцянский Л.Г.* Механика жидкости и газа. М.: Наука, 1978. 736 с. 3. *Панченко А.Ю.* Уравнение состояния в системе уравнений акустики для неоднородной движущейся среды // Радиотехника. 1997. Вып. 103. С. 169–174.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 4.04.2001

ИНСТИТУТ
ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ

Украина, 61166, г. Харьков, ул. Бакулина 12
Тел./Факс. (0572) 14-22-04(05), 30-24-52

АО «Институт информационных технологий» (АО «ИИТ») является одним из ведущих предприятий в Украине в области создания высокоэффективных систем и средств защиты информации в автоматизированных информационных и телекоммуникационных системах (сетях) различного класса и назначения (ИТСС). АО «ИИТ» получил право на проведение широкого спектра видов работ, а также предоставления услуг в области криптографической защиты информации (лицензии ДСТСЗИ СБУ № КВ1191200507 от 08.07.1999г и № КВ1421234567 от 23.06.2000г.). В рамках АО «ИИТ» создана и получила аккредитацию лаборатория по сертификации средств криптографической защиты информации (аттестат аккредитации Госстандарта № UA 6.001.Т 685 от 22.09.2000г.).

Коллективом сотрудников АО «ИИТ», а это – 5 докторов, 8 кандидатов технических наук, высококвалифицированные специалисты в области безопасности информационных технологий, накоплен большой научный и практический опыт создания комплексных систем защиты информации (КСЗИ) ИТСС.

Создаваемые коллективом АО «ИИТ» многоуровневые системы защиты информации реализуются как программно, так и аппаратно-программно с использованием криптомодулей семейства «Гряда», выполняющих функции криптографических преобразований информации, генерации и хранения ключей в защищенном виде. Применяемые аппаратные, аппаратно-программные и программные средства криптографической защиты предоставляют услуги по обеспечению конфиденциальности, целостности, доступности и наблюдаемости данных с использованием криптоалгоритмов ГОСТ 28147-89, ГОСТ 34.310-95, ГОСТ 34.311-95, DES, TDES, IDES, RC6, RIJNDAEL, MARS, DSS(DSA), ECDSA, ISO-11166 и др.

Программные средства криптографической защиты обеспечивают защиту на прикладном, транспортном и сетевом уровнях и могут интегрироваться в ИТСС, работающие под различными операционными системами – Windows-95, Windows-98, Windows-NT, Windows-2000, Unix, Linux, OS/2 и реализующие различные службы и протоколы (X-400, TCP/IP, SSL/TSL, IPSec и др.). Криптомодули семейства «Гряда», обеспечивают комплексную защиту на прикладном, сеансовом, сетевом или транспортном уровне и построены на основе семейства сигнальных процессоров ADSP-21061, 21062, 21065 и др. На все криптомодули семейства «Гряда» получены патенты Украины. Разработаны и практически апробированы аппаратно-программные средства защиты радиолиний «земля-борт КА» и «борт КА-земля».

В настоящее время разработаны и находятся на различных этапах внедрения 7 систем защиты банковской и коммерческой информации. Применяемые системы и средства криптографической защиты сертифицированы установленным порядком на соответствие их реализации ГОСТ 28147-89, ГОСТ 34.310-95, ГОСТ 34.311-95 (сертификат соответствия № UA 1.112.24334-00 от 15.11.2000г). Разработанные коллективом сотрудников АО «ИИТ» проекты ведомственных документов «Политика безопасности информации в АС» и «Концепция безопасности информации в АС» получили положительный отзыв Заказчиков и специалистов.

Накопленный многолетний опыт, высокий уровень квалификации, системный подход и добросовестное отношение к работам, современная научно-техническая база и владение современными информационными технологиями, служат лучшими гарантиями создания АО «ИИТ» высокоэффективных КСЗИ для автоматизированных информационных и телекоммуникационных систем (сетей).

СОДЕРЖАНИЕ

Блочные симметричные криптоалгоритмы: особенности построения и криптоанализ

Обработка сигналов, теория антенн, полупроводниковые КРС, акустические волны

CONTENTS

УДК 681.3.06:519.248.681

Методологические основы концепции и политики безопасности информационных технологий / М.Ф. Бондаренко, С.П. Черных, И.Д. Горбенко, А.А. Замула, А.А. Ткач // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 5-16.

Рассматривается методология подготовки и разработки основополагающих документов, используемых при создании комплексной системы защиты информации в автоматизированных системах и сетях различного класса и назначения. Излагаются принципы построения документов, их структура и требования к содержанию разделов.

Библиогр.: 8 назв.

УДК 681.3.06:519.248.681

Методологічні основи концепції і політики безпеки інформаційних технологій / М.Ф. Бондаренко, С.П. Черних, І.Д. Горбенко, О.А. Замула, О.О. Ткач // Радіотехніка Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 5-16.

Розглядається методологія підготовки і розробки основних документів, які використовуються при створенні комплексної системи захисту інформації в автоматизованих системах і мережах різного класу і призначення. Викладаються принципи побудови документів, їхня структура та вимоги до змісту розділів.

Библиогр.: 8 назв.

UDC 681.3.06:519.248.681

Methodological bases of the concept and policy of information security technologies/ M.F.Bondarenko, S.P. Chernyh, I.D.Gorbenko, A.A.Zamula, A.A.Tkach // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 5-16.

The methodology of preparation and development of the basic documents used in creation of the complex information security system in automated systems and networks for a various class and purpose. Principles of documents construction, their structure and requirements to the contents of sections are stated.

Ref.: 8 items.

УДК 681.3.06:519.248.681

Гибкость и специализация профиля защиты автоматизированной системы / А.В. Потий, Л.В. Скрыпник // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 17-21.

В существующих нормативных документах к профилю защиты автоматизированной системы предъявляются такие требования, как функциональная полнота, непротиворечивость (согласованность), реализуемость, адекватность, эффективность и корректность. В работе обосновывается необходимость предъявления к профилю защиты такого требования, как гибкость. В статье гибкость рассматривается как экономическая категория, вводится показатель гибкости профиля защиты и системы защиты информации.

Библиогр.: 7 назв.

УДК 681.3.06:519.248.681

Гнучкість та спеціалізація профілю захисту автоматизованої системи / О.В. Потій, Л.В. Скрипник // Радіотехніка: Всеукр. межвід. наук.-техн. зб. 2001. Вип. 119. С. 17-21.

У діючих нормативних документах до профілю захисту автоматизованої системи пред'являються такі вимоги, як функціональна повнота, несуперечність, (погодженість), реалізованість, адекватність, ефективність та коректність. У роботі обґрунтовується необхідність пред'явлення до профілю захисту такої вимоги, як гнучкість. У статті гнучкість розглядається як економічна категорія, пропонується показник гнучкості профілю захисту та системи захисту інформації.

Бібліогр.: 7 назв.

UDC 681.3.06:519.248.681

Methodolosical bases of the concept and policy of information security technologies / M.F. Bondarenko. S.P. Chernyh. I.D. Gorbenko, A.A. Zamula. A.A. Tkach // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. 119. P. 17-21.

The methodology of preparation and development of the basic documents used in creation of the complex information security system in automated systems and networks for a various class and purpose. Principles of documents construction, their structure and requirements to the contents of sections are staled.

Ref.: S items.

УДК 681.3.06:519.248.681

Стандарт симметричного шифрования 21 века: свойства, режимы работы, реализация / И.Д. Горбенко, Л.В. Скрыпник, С.А. Головашич, Т.А. Гриненко // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 22-35.

Дается описание режимов применения, порядка организации и выполнения прямых и обратных криптографических преобразований, а также формирования цикловых ключей в стандарте 21 века – криптоалгоритме Rijndael. Приводятся результаты анализа основных характеристик и свойств криптоалгоритма Rijndael, режимов работы и применения. Указывается на ряд особенностей криптоалгоритма и необходимость совершенствования алгоритма разворачивания цикловых ключей, используемого в Rijndael.

Табл. 4. Ил. 9. Библиогр.: 6 назв.

УДК 681.3.06:519.248.681

Стандарт симетричного шифрування 21 століття: властивості, режими роботи, реалізація / І.Д. Горбенко, Л.В. Скрипник, С.О. Головашич, Т.О. Гріненко // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вид. 119. С. 22-35.

Дається опис режимів застосування, порядку організації та виконання прямих і зворотних криптографічних перетворень, а також формування циклових ключів у стандарті 21 століття – криптоалгоритмі Rijndael. Приводяться результати аналізу основних характеристик і властивостей криптоалгоритму Rijndael, режимів роботи і застосування. Указується на ряд особливостей криптоалгоритму і необхідність удосконалювання алгоритму розгортання циклових ключів, який використовується в Rijndael.

Табл. 4. Іл. 9. Бібліогр.: 6 назви.

UDC 681.3.06:519.248.681

21 century symmetric encryption standard: properties, modes of operation, implementation / I.D. Gorbenko, L.V. Skripnik, S.A. Golovashich, T.A. Grinenko // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 22-35.

The modes of operation, organizational procedures, encryption and decryption transformations as well as the key schedule in the XXI century standard – the cryptographic algorithm Rijndael are given. The main characteristics and properties analysis results, modes of operation and application are given. Several features of the cryptographic algorithm and necessity of Rijndael's key schedule improving are pointed out.

4 tab. 1 fig. Ref.: 6 items.

УДК 681.3.06:519.248.681

Свойства и возможности оптимизации криптографических преобразований в AES – RIJNDAEL / И.Д. Горбенко, Д.А. Чекалин // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 36-42.

Рассматриваются возможности повышения быстродействия программной реализации алгоритма RIJNDAEL.

Табл. 4. Библиогр.: 1 назв.

УДК 681.3.06: 519.248.681

Властивості та можливості оптимізації криптографічних перетворень в AES – RIJNADAEL / І.Д. Горбенко, Д.О. Чекалін // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 36-42

Розглядаються можливості підвищення швидкодії програмної реалізації алгоритма RIJNDAEL.

Табл. 4. Бібліогр.: 1 назв.

UDC 681.3.06: 519.248.681

Optimization characteristics and possibilities in AES – RIJNDAEL cryptographic transformation. / I.D. Gorbenko, D.A. Chekalin // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 36-42.

The possibilities to increase the speed of program implementation of Rijndael algoritm are regarded.

4 tab. Ref.: 1 items.

УДК 681.3.06:519.248.681

Криптоанализ криптографических преобразований в группах точек эллиптических кривых методом Полларда / И.Д. Горбенко, С.И. Збитнев, А.А. Поляков // Радиотехника: Всеукр. межвед. научю-тех. сб 2001. Вып. 119. С. 43-50.

Рассматриваются наиболее эффективные алгоритмы криптоанализа ρ и λ - методы Полларда. Приводится оценка стойкости криптосистем, основанных на эллиптических кривых, к криптоанализу с использованием ρ и λ - методов. Приводятся подробные примеры использования ρ - метода. Оценивается сложность криптоанализа практически применяемых ЭК над $GF(2^m)$.

Табл. 4. Библиогр.: 4 назв.

УДК 681.3.06: 519.248.681

Криптоаналіз криптографічних перетворень у групах крапок еліптичних кривих засобом Полларда. / І.Д. Горбенко, С.І. Збітнєв, А.А. Поляков // Радіотехніка: Всеукр. міжвід. наук.-тех. зб.2001. Вип. 119. С. 43-50.

Розглядаються найбільш ефективні алгоритми криптоаналізу ρ і λ – методи Полларда. Наводиться оцінка стійкості криптосистем, базованих на еліптичних кривих, до криптоаналізу з використанням ρ и λ - методів. Наводяться докладні приклади використання ρ-методів. Оцінюється складність криптоаналізу ЕК над $GF(2^m)$, які практично використовуються.

Табл. 4. Бібліогр.: 4 назв.

UDC 681.3.06: 519.248.681

Pollard method cryptanalysis of operations over group of elliptic curve points /I.D. Gorbenko, S.I. Zbitnev, A.A. Polyakov // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 43-50.

The most efficient algorithms of cryptanalysis using ρ and λ – methods of Pollard are considered. Cryptosystem stability evaluation, based on elliptic curve, to cryptanalysis using ρ and λ methods is shown. Detailed examples of ρ method use are given. Cryptanalysis difficulty practically used is estimated.

4 tabs. Ref.: 4 items.

УДК 681.3.06:519.248.681

Сложность арифметических операций в группах точек эллиптических кривых для криптографических операций / И.Д. Горбенко, С.И. Збитнев, А.А. Поляков// Радиотехника: 2001. Вып. 119. С. 51-55.

Рассматриваются сложность арифметических операций над элементами в поле $GF(2^m)$, сложность преобразований над эллиптической кривой с использованием аффинных и проективных координат. Проводиться подробный сравнительный анализ производительности аффинных и проективных преобразований. Оценивается общая производительность аффинных и проективных координат.

Табл. 1. Ил. 10. Библиогр.: 4 назв.

УДК 681.3.06: 519.248.681

Складність арифметичних операцій у групах крапок еліптичних кривих для криптографічних операцій / І.Д. Горбенко, С.І. Збітнєв, А.А. Поляков // Радіотехніка: 2001. Вип. 119. С. 51-55.

Розглядаються складність арифметичних операцій над елементами в полі $GF(2^m)$, складність перетворень над еліптичній кривій з використанням афінних та проективних координат. Проводиться докладний порівняльний аналіз потужності афінних та проективних координат. Оцінюється загальна потужність афінних та проективних координат.

Табл. 1. Іл. 10. Бібліогр.: 4 назв.

UDC 681.3.06: 519.248.681

Complexity of arithmetic operations in points group of elliptic curve for crypto operations /I.D. Gorbenko, S.I. Zbitnev, A.A. Polyakov // Radiotehnika: All-Ukr. Sci. Interdep. Mag. 2001. 119. P. 51-55.

Complexity of arithmetic operations over elements in field $GF(2^m)$, transformation complexity over elliptic curve using affine and projective coordinates are considered. Detailed comparative analysis of productivity of affine and projective coordinates are carried out. Total productivity affine and projective coordinates are estimated.

1 tabs. 10 figs. Ref.: 4 items.

УДК 681.3.06

Оптимально расширенные поля в алгоритмах для эллиптических кривых / Д.И. Лавриненко // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 56-61.

Рассматривается применение оптимально расширенных полей Галуа для уменьшения вычислительной сложности криптографических преобразований на основе эллиптических кривых. Показаны результаты теоретических расчётов вычислительной сложности алгоритма умножения в ОРП и результаты вычислительного эксперимента.

Табл. 1. Ил. 1. Библиогр.: 11 назв.

УДК 681.3.06

Оптимально розширені поля у алгоритмах для еліптичних кривих / Д.І. Лавриненко // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 56-61.

Розглядається застосування оптимально розширених полів Галуа для зменшення обчислювальної складності криптографічних перетворень на основі еліптичних кривих. Показано результати теоретичних вимірювань обчислю-

вальної складності алгоритму множення з ОПП і результати обчислювального експерименту.

Табл. 1. Іл. 1. Бібліогр.: 11 назв.

UDC 681.3.06

Optimal extended fields in algorithms for elliptic curves / D.I. Lavrinenko // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 55-60.

The application of optimal extended Galois fields for computing complexity reduction of cryptographic transformations based on elliptic curves is considered. The results of theoretical estimates of computing complexity of multiplication algorithm with OEF and results of computing experiment are shown.

1 tab. 1 fig. Ref.: 11 items.

УДК 681.3.06: 519.248.681

Оценка стойкости RSA-систем, в которых открытые ключи или параметры являются личными / И.Д. Горбенко, П.В. Колесников // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 62-68.

В статье рассматриваются криптографические алгоритмы на базе RSA. Предложены новые схемы работы и хранения открытых параметров системы. Проанализирована возможность хранения открытых параметров наравне с личными ключами. Рассчитана криптостойкость RSA-системы для различных вариантов хранения открытых параметров. Проведен сравнительный анализ предложенной схемы хранения параметров и существующей на сегодняшний момент.

Табл. 6. Ил. 2. Библиогр.: 5 назв.

УДК 681.3.06: 519.248.681

Оцінка стійкості RSA систем, в яких відкриті ключі та параметри є особистими / І.Д. Горбенко, П.В. Колесніков // Радіотехніка: Всеукр. між. від. наук.-техн. зб. 2001. Вип. 119. С. 62-68.

У статті розглядаються криптографічні алгоритми на базі RSA-систем. Запропоновані нові схеми роботи та зберігання відкритих параметрів системи. Проаналізовано можливість зберігання відкритих параметрів на рівні з особистими ключами. Розрахована криптостійкість RSA-системи для різних варіантів зберігання відкритих параметрів. Проведено порівняльний аналіз запропонованої системи зберігання параметрів та системи, існуючої на сьогодні .

Табл. 6. Іл. 2. Бібліогр.: 5 назв.

UDC 681.3.06: 519.248.681

Reliability estimation of RSA systems where the open keys or parameters are in the private form. / I.D. Gorbenko, P.V. Kolesnikov // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 62-68.

The cryptographic algorithms based on RSA system are described. New schemes of public parameters operation and storage are proposed in the article. The possibility to store public parameters equally to private keys is analyzed. Estimation of cryptographic reliability was performed for various cases of the open keys storing. Corporative analysis of the offered scheme of the parameters security and the one available at present is given.

6 tab. 2 fig. Ref.: 5 items.

УДК 681.3.06:519.248.681

Методи автентифікації в безумовно стійких криптосистемах / О.А. Замула, Г.М. Гулак, Є.В. Попович // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 69-76.

Розглядаються методи та умови забезпечення автентичності в безумовно стійких криптосистемах. Формулюються та доводяться ствердження, які дозволяють отримати необхідні та достатні умови забезпечення автентичності в системі Вернама.

Іл. 4. Бібліогр.: 3 назв.

УДК 681.3.06:519.248.681

Методы аутентификации в безусловно стойких криптосистемах / А.А. Замула, Г.Н. Гулак, Е.В. Попович // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119.С. 69-76.

Рассматриваются методы и условия обеспечения аутентичности в безусловно стойких криптосистемах. Формируются и доказываются утверждения, которые позволяют получить необходимые и достаточные условия обеспечения аутентичности в системе Вернама.

Ил. 4. Библиогр.: 3 назв.

UDC 681.3.06:519.248.681

Authentication methods in certainly proof cryptosystems / A.A. Zamula, G. M. Gulak, E.V. Popovich // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 69-76.

Methods and conditions of authentication maintenance in certainly proof cryptosystems are considered. Statements that allow deriving necessary and sufficient conditions of authentication maintenance in Vernam's system are formed and proved.

4 fig. Ref.: 3 items

УДК 681.3.06:519.248.681

Методы обеспечения аутентификации с введением избыточности/ И.Д. Горбенко, А.А. Замула, Е.В. Попович, Т.А. Гриненко // Радиотехника: Всеукр. межвед. научн.-техн. сб. 2001. Вып. 119. С. 77-80.

Проводится сравнительный анализи методов обеспечния аутентичности, а именно: с использованием однонаправленных хеш-функций, с использованием ключевых хеш-функций, цифровой подписи, контрольных сумм. Сравнение осуществляется с точки зрения вероятности обмана (показателя аутентичности), а также вычислительной сложности.

Табл. 2. Библиогр.: 3 назв.

УДК 681.3.06:519.248.681

Методи забезпечення автентичності з введенням надлишковості/ І.Д. Горбенко, О.А. Замула, Є.В. Попович, Т.О. Гріненко // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 77-80.

Проводиться порівняльний аналіз методів забезпечення автентичності, а саме: з використанням однонаправлених хеш-функцій, з використанням ключових хеш-функцій, цифрового підпису, контрольних сум. Порівняння здійснюється з точки зору ймовірності обману (показника автентичності), а аткож очислювальної складності.

Табл. 2. Бібліогр.: 3 назв.

UDC 681.3.06:519.248.681

Methods of authentication maintenance with redundancy inclusion/ I.D. Gordenko, A.A. Zamula, E.V. Popovich, T.A. Grinenko // Radiotekhnika: All-Ukr. Sci. Inderdep. Mag. 2001. № 119. P. 77-80

The comparative analysis of authentication maintenance methods is carried out, namely: with use of the unidirectional hash-functions, with use of keying hash-functions, the digital signature, the control sums. Comparison is carried out from the point of view of probability of a deceit (an authenticity parameter), and also computing complexity.

2 tab. Ref.: 3 items.

УДК 681.3.06

Аутентификация с применением алгеброгеометрических кодов / Г.З. Халимов, А.А. Кузнецов // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 81-87.

Изложены общетеоретические вопросы построения аутентификации с применением кодовых конструкций. Предложены практические схемы с оценкой их параметров. Рассматривается теория применения алгеброгеометрических кодов для целей универсального хеширования.

Ил. 3. Библиогр.: 15 назв.

УДК 681.3.06

Автентифікація із застосуванням алгеброгеометричних кодів / Г.З. Халімов, О.О. Кузнецов // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 81-87.

Викладені загальнотеоретичні питання побудови автентифікації із застосуванням кодових конструкцій. Запропоновані практичні схеми з оцінкою їх параметрів. Розглядається теорія застосування алгеброгеометричних кодів для цілей універсального хешування.

Іл. 3. Бібліогр.: 15 назв.

UDC 681.3.06

Authentication with algebraic geometric codes application / G.Z. Khalimov, A.A. Kuznetsov // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 81-87.

General theoretical authentication construction problems with application of codes are stated. Practical schemes with their parameters estimation are offered. Theory of algebraic geometric codes application for universal hashing is considered.

3 fig. Ref.: 15 items.

УДК 681.3.06

Аутентификация и универсальное хеширование / Г.З. Халимов, А.А. Кузнецов // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 88-94.

Изложены общетеоретические вопросы построения аутентификации с применением строго универсального хеширования на основе ортогональных массивов. Приведены определения универсальных и строго универсальных классов хеш-функций для построения кодов аутентификации. Рассмотрены практические схемы аутентификации с оценкой их параметров.

Табл. 4. Библиогр.: 11 назв.

УДК 681.3.06

Автентифікація та універсальне хешування / Г.З. Халімов, О.О. Кузнецов // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 88-94

Викладені загальнотеоретичні питання побудови автентифікації із застосуванням суворо універсального

хешування на основі ортогональних масивів. Приведені визначення універсальних і суворо універсальних класів хеш-функцій для побудови кодів автентификації. Розглянуті практичні схеми автентификації з оцінкою їх параметрів.

Табл. 4. Бібліогр.: 11назв.

UDC 681.3.06

Authentication and universal hashing / G.Z. Khalimov, A.A. Kuznetsov // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 88-94.

General theoretical problems of authentication construction questions with a strongly universal hashing based on orthogonal arrays are outlined. Determinations of universal and strongly universal classes of hash-functions for authentication codes construction are brought. Practical authentication schemes with their parameters estimation are considered.

4 tab. Ref.: 11 items.

УДК 681.3.06:519.248.681

Условия и возможности создания безусловно стойких криптосистем / А.А. Замула, Е.В. Попович, Ю.И. Горбенко // Радиотехника: Всеукр. межвед. научн.-техн. сб. 2001. Вып. 119. С. 95-100.

Рассматривается математическая модель взаимодействия пользователей в информационно телекоммуникационной системе. Формулируются и доказываются теоремы, которые позволяют получить необходимые и достаточные условия создания безусловно стойких криптосистем. Приводятся оценки стойкости безусловно стойких криптографических систем для различных значений ключей шифрования.

Ил. 1. Табл. 1. Библиогр.: 3 назв

УДК 681.3.06:519.248.681

Умови та можливості створення безумовно стійких криптосистем/ О.А. Замула, Є.В. Попович, Ю.І. Горбенко // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 95-100.

Розглядається математична модель взаємодії користувачів в інформаційно телекомунікаційній системі. Формулюються та доводяться теореми, які дозволяють отримати необхідні та достатні умови створення безумовно стійких криптосистем. Наводяться оцінки стійкості безумовно стійких криптографічних систем для різних значень ключів щифрування.

Іл. 1. Табл. 1. Бібліогр.: 3 назв.

UDC 681.3.06:519.248.681

Conditions and opportunities of certainly proof cryptosystems creation / A.A. Zamula, E.V. Popovich, U.I. Gorbenko // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 95-100.

The mathematical model of users interaction in the information telecommunication system is considered. Theorems that allow to derive necessary and sufficient conditing of creation certainly proof cryptosystems are formulated and proved. Resistance estimations of certainly proof cryptographic systems for various encryption key values are presented.

1 fig. 1 tab. Ref.: 3 items

УДК 681.3.06:519.248.681

Принципы функционирования протокола IPSec / А.В. Свинарев, А.Н. Лепеха, Р.В. Олейников, А.И. Шумов, А.А. Гайович, А.А. Казьмин // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 101-107.

В статье описывается возможность создания корпоративных сетей на основании защищенного сетевого протокола IPSec, работающего на сетевом уровне модели взаимодействия открытых систем. Описываются его возможности, варианты реализации и их надежность.

Ил. 9. Библиогр.: 5 назв.

УДК 681.3.06:519.248.681

Принципи функціонування протоколу IPSec / А.В. Свинарьов, О.М. Лепеха, Р.В. Олійников, О.І. Шумов, О.А. Гайович, О.О. Казьмін // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 101-107.

У статті описується можливість створення корпоративних мереж на підставі захищеного мережного протоколу IPSec, що працює на мережному рівні моделі взаємодії відкритих систем. Описуються його можливості, варіанти реалізації та їх надійність.

Іл. 9. Бібліогр.: 5 назв.

UDC 681.3.06:519.248.681

The functioning principles of the IPSec protocol/ A.V.Svinarev, A.N.Lepeha, R.V.Oleynikov, A.I.Shumov, A.A.Gayovich, A.A.Kazmin // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 101-107.

The possibility of the corporate networks creation based on the protected network protocol IPSec corresponding to

the opened system interaction model is described. Its abilities, the implementation versions and reliability are described.

9 fig. Ref.: 5 items.

УДК 681.324.067

Генерация равновероятных случайных последовательностей на основе физических датчиков / А.А. Торба, С.Г. Елаков, А.З. Степченко // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 108-113.

Предложены схемы генераторов случайных последовательностей на основе физических датчиков шума. Проведен анализ причин нарушения равновероятного закона распределения генерируемых случайных чисел. Описаны алгоритмы выравнивания вероятностей случайных чисел и повышения надежности работы генераторов. Предложен способ повышения быстродействия генераторов равновероятных случайных последовательностей.

Табл.1. Ил.7. Библиогр.: 2 назв.

УДК 681.324.067

Генерація рівноімовірних випадкових послідовностей на основі фізичних давачів / О.О. Торба, С.Г. Єлаков, О.З. Степченко // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 108-113.

Запропоновано схеми генераторів випадкових послідовностей на основі фізичних давачів шуму. Проведено аналіз причин порушення рівноімовірного закону розподілу випадкових чисел, що генеруються. Описано алгоритми вирівнювання імовірностей випадкових чисел і підвищення надійності роботи генераторів. Запропоновано спосіб підвищення швидкодії генераторів рівноімовірних випадкових послідовностей.

Табл.1. Іл.7. Бібліогр.: 2 назв.

UDK 681.324.067

Generation of equiprobable random sequences on the basis of physical sensors / A.A. Torba, S.G. Yelakov, A.Z. Stepchenko // Radiotekhnika: All-Ukr. Interdep. Mag.2001. № 119. P. 108-113.

Circuits of random sequences generators based on the noise physical sensors are offered. Analysis of equiprobable distribution law violation causes is carried out for the generated random numbers. Algorithms of random numbers probabilities alignment and the generator operation reliability are described. The method of equiprobable random sequences generator speed increase is offered.

1 tab. 7 fig. Ref.: 2 items.

УДК 681.3.06:519.248.681

Алгоритмы и средства тестирования случайных и псевдослучайных последовательностей / М.А. Кривошлык, Ю.И. Горбенко, В.Н. Вервейко // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 114-119.

Рассматриваются методики тестирования случайных и псевдослучайных последовательностей.

Табл. 1. Библиогр.: 3 назв.

УДК 681.3.06:519.248.681

Алгоритми та засоби тестування випадкових та псевдовипадкових послідовностей / М.А. Кривошлик, Ю.І. Горбенко, В.М. Вервейко // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 114-119.

Розглядаються методики тестування випадкових та псевдовипадкових послідовностей.

Табл. 1. Бібліогр.: 3 назви.

UDC 681.3.06: 519.248.681

Algorithms and methods for testing random and pseudo-random sequences / M.A. Krivoshlyk, Y.I. Gorbenko, V.N. Verveiko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 114-119.

The analysis of methods for testing random and pseudo-random sequences is provided in this paper.

1 tab. Ref.: 3 items.

УДК 681.3.06

- **Метод формирования и свойства псевдослучайных последовательностей на эллиптических кривых** / Т.А. Гриненко, Ю.И. Горбенко, С.Ю. Орлова // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 119-123.

Рассматривается метод формирования псевдослучайных последовательностей на эллиптических кривых. Описан алгоритм построения генератора для формирования таких последовательностей. Приведены результаты тестирования и дана оценка выходной последовательности генератора.

Табл. 4. Библиогр.: 9 назв.

УДК 681.3.06

Метод формування та властивості псевдовипадкових послідовностей на еліптичних кривих / Т.О. Гріненко, Ю.І. Горбенко, С.Ю. Орлова // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 119-123.

Розглядається метод формування псевдовипадкових послідовностей на еліптичних кривих. Описано алгоритм побудови генератора для формування таких послідовностей. Наведені результати тестування та надана оцінка вихідної послідовності генератора.

Табл. 4. Бібліогр.: 9 назв.

UDC 681.3.06

Method of formation and properties of pseudo-random sequences on the elliptic curves / T.A. Grinenko, Y.I. Gorbenko, S.Y. Orlova // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 119-123.

The method of pseudo-random sequences formation on elliptical curves is considered. The algorithm of the generator construction for such sequences formation is described. The outcomes of testing are adduced and the estimation of the generator output sequence is given.

4 tab. Ref.: 9 items.

УДК 681.32

Исследование методов генерации многосвязных марковских цепей в задаче сертификации микропроцессорных компонентов / В.А. Горбачев, М.А. Волк, С.Н. Саранча, В.В. Степаненко // Радиотехника: Всеукр. межвед научно-техн. сб. 2001. Вып. 119. С. 124-129.

Данная статья посвящена проблеме сертификации микропроцессорных устройств методом псевдослучайного функционального тестирования. Рассмотрены методы оптимизации длины тестирующей последовательности путем применения генератора многосвязных марковских последовательностей. Предложен метод программной и аппаратной генерации многосвязной марковской последовательности.

Ил. 6. Библиогр.: 5 назв.

УДК 681.32

Дослідження методів генерації багатозв'язкових марківських ланцюгів у задачі сертифікації мікропроцесорних компонентів / В.О. Горбачов, М.О. Волк, С.М. Саранча, В.В. Степаненко // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 124-129.

Стаття присвячена проблемі сертифікації мікропроцесорних пристроїв методом псевдовипадкового функціонального тестування. Розглянуто методи оптимізації довжини послідовності, що тестує, шляхом застосування генератора багатозв'язкових марківських послідовностей. Запропоновано метод програмної й апаратної генерації багатозв'язкових марківських послідовностей

Іл. 6. Бібліогр.: 5 назв.

UDC 681.32

Research of multiply connected Markov chains generation methods in the microprocessor components certification task / V.A. Gorbachev, M.A. Volk, S.N. Sarancha, V.V. Stepanenko // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 124-129.

The given paper is devoted to a problem of certification of microprocessor devices with a method of pseudorandom functional testing. The methods of the testing sequence length optimization are considered using of the generator of multicoupling Markov sequences. The method of program and hardware generation of the multiply connected Markov sequence is offered.

6 fig. Ref.: 5 items.

УДК 681.04

Методы реализации модульных операций в системах цифровой обработки информации / В.А.Краснобаев // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 130-134.

Рассматриваются методы реализации арифметических операций в непозиционной системе счисления в остаточных классах

Табл. 5. Библиогр.: 9 назв.

УДК 681.04

Методи реалізації модульних операцій у системах цифрової обробки інформації / В.А.Краснобаєв // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 130-134.

Розглядаються методи реалізації арифметичних операцій у непозиційній системі числення у залишкових класах.

Табл. 5. Бібліогр.: 9 назв.

UDC 681.04

Methods of the modulus operation realization in the digital information processing systems / V.A.Krasnobaev // Radiotekhnika: All – Ukr. Sci Interdep. Mag. 2001. № 119. P. 130-134.

Methods of arithmetic operations realization in the system of residual classes are considered.

Tabl. 5. Ref.: 9 items.

УДК 681.3.06

Безопасность режимов блочного шифрования / С.А. Головашич // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 135-145.

В статье дан анализ основных режимов применения блочных симметричных шифров. Проанализированы достоинства и недостатки каждого из режимов, предложены способы устранения обнаруженных недостатков. Приведены две схемы режимов поточного шифрования, удовлетворяющие предложенным требованиям.

Ил. 2. Библиогр. 8 назв.

УДК 681.3.06

Безпека режимів блокового шифрування / С.О. Головашич // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 135-145.

У статті дано аналіз основних режимів застосування блокових симетричних шифрів. Проаналізовано переваги та недоліки кожного з режимів, запропоновані засоби усунення виявлених недоліків. Наведено дві схеми режимів потокового шифрування, які задовольняють запропонованим вимогам.

Іл. 2. Бібліогр. 8 назв.

UDC 681.3.06

Block ciphers modes of operations security / S.A. Golovashich // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 135-145.

The symmetric block ciphers standard modes of operations analysis is carried out. The advantages and disadvantages of each mode is analyzed, the improvement methods are suggested. The schemes of two new modes for stream encryption are proposed. These schemes completely satisfy the suggested requirements.

2 fig. Ref.: 8 items.

УДК 681.3.06: 519.248.681

Дифференциальный криптоанализ алгоритма шифрования ГОСТ 28147-89 / Р.В. Олейников // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 146-152.

Рассматриваются условия, при которых возможна эффективная дифференциальная атака на алгоритм шифрования ГОСТ 28147-89. Приводится пример слабой подстановки и анализируются её свойства с точки зрения дифференциального криптоанализа. Подробно описывается методика учёта влияния ключевого сумматора при построении дифференциальных характеристик. Приводятся методы анализа вероятностей характеристик и вычисления значений ключа при выполнении криптоанализа. Оценивается сложность атаки на ГОСТ 28147-89.

Табл. 5. Ил. 1. Библиогр.: 3 назв.

УДК 681.3.06: 519.248.681

Диференційний криптоаналіз алгоритму шифрування ГОСТ 28147-89 / Р.В. Олійников // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 146-152.

Розглядаються умови, при яких можлива ефективна диференційна атака на алгоритм шифрування ГОСТ 28147-89. Наводиться приклад слабкої підстановки та аналізуються її властивості з точки зору диференційного криптоаналізу. Докладно описується методика підрахунку впливу ключового суматора при побудові диференційних характеристик. Наводяться методи аналізу ймовірностей характеристик та обчислення значень ключа при здійсненні криптоаналізу. Оцінюється складність атаки на ГОСТ 28147-89.

Табл. 5. Іл. 1. Бібліогр.: 3 назв.

UDC 681.3.06: 519.248.681

Differential cryptanalysis of GOST 28147-89 encryption algorithm / R.V. Oleynikov // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. 119. P. 146-152.

The conditions of the effective differential attack on the GOST 28147-89 encryption algorithm are considered. The example of a weak substitution is shown and its differential properties are analyzed. The methods for the key adder influence calculation when building differential characteristics are described in detail. The characteristics probability analysis and key extraction methods are given. The complexity of the attack on the GOST 28147-89 is estimated.

5 tabs. 1 figs. Ref.: 3 items.

УДК 681.3.06: 519.248.681

Дополнительные требования к отбору таблиц подстановок для ГОСТ 28147-89 / В.И. Долгов, И.В. Лисицкая, Р.В. Олейников, С.А. Головашич, А.С. Коряк // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 153-159.

Производится исследование стойкости алгоритма шифрования ГОСТ 28147-89 к атакам дифференциального и линейного криптоанализа. Рассматриваются примеры дифференциальных характеристик и обсуждаются условия их выполнения. Решается проблема прохождения разностей через ключевой сумматор и выполняется оценка вероятности нужного преобразования. Рассматривается влияние битов сеансового ключа на успешное прохождение дифференциальной характеристики. Приводится оценка количества «слабых» подстановок для ГОСТа. Формулируются требования к долговременным ключам, позволяющим защитить шифр от атак дифференциального и линейного криптоанализа.

Табл. 2. Ил. 2. Библиогр.: 13 назв.

УДК 681.3.06: 519.248.681

Додаткові вимоги до вибору таблиць підстановок для ГОСТ 28147-89 / В.І. Долгов, І.В. Лисицька, Р.В. Олійников, С.О. Головашич, О.С. Коряк // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 153-159.

Проводиться дослідження стійкості алгоритму шифрування ГОСТ 28147-89 до атак диференційного та лінійного криптоаналізу. Розглядаються приклади диференційних характеристик та обговорюються вимоги їх виконання. Вирішується проблема проходження різниць через ключовий суматор та виконується оцінка ймовірності потрібного перетворення. Розглядається вплив бітів сеансового ключа на вдале проходження диференційної характеристики. Наводиться оцінка кількості «слабких» підстановок для ГОСТу. Формулюються вимоги до довгострокових ключів, які дозволяють захистити шифр від атак диференційного та лінійного криптоаналізу.

Табл. 2. Іл. 2. Бібліогр.: 13 назв.

UDC 681.3.06: 519.248.681

Additional requirements to the substitution table selection for GOST 28147-89 / V.I. Dolgov, I.V. Lisitskaya, R.V. Oleynikov, S.A. Golovashich, A.S. Koryak // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. 119. P. 153-159.

The analysis of the GOST 28147-89 resistance to the differential and linear cryptanalysis is carried out. The examples of differential characteristics are considered and the conditions of their implementation are discussed. The problem of difference transformations through the key adder is solved and the required transformation probability is estimated. The session key bits influence on a successful differential transformation is considered. The number of weak substitutions for the GOST is estimated. The requirements to S boxes for the protection from the differential and linear cryptanalysis are formulated.

2 tabs. 2 figs. Ref.: 13 items.

УДК 681.3.06: 519.248.681

Исследование возможностей модернизации шифра ГОСТ 28147-89 с целью дальнейшего повышения его безопасности / И.В. Лисицкая, Т.В. Цепурит, В.В. Лесняк, М.В. Пинчук, А.П. Мелецкий // Радиотехника: Всеукр. межвед. наук.-техн. сб. 2001. Вып. 119. С. 160-165.

Рассматриваются возможности модернизации шифра ГОСТ 28147-89 с повышения его стойкости к атакам дифференциального криптоанализа. Предлагаются и обосновываются применения в алгоритме дополнительного нелинейного преобразования – параметрического циклического сдвига и возможность увеличения длины шифруемого блока в два раза.

Табл. 2. Ил. 7. Библиогр.: 4 назв.

УДК 681.3.06: 519.248.681

Дослідження можливостей модернізації шифра ГОСТ 28147-89 з метою подальшого підвищення його безпеки / І.В. Лисицька, Т.В. Цепуріт, В.В. Лєсняк, М.В. Пінчук, О.П. Мелецький // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 160-165.

Розглядаються можливості модернізації шифру ГОСТ 28147-89 с метою підвищення його стійкості до атак диференційного криптоаналізу. Пропонується і обгунтовується застосування у алгоритмі додаткового нелінійного перетворення – параметричного циклічного зсуву та можливість збільшення в два рази довжини блока, який шифрується.

Табл. 2. Іл. 7. Бібліогр.: 4 назв.

UDC 681.3.06: 519.248.681

Research of possibilities of modernisation of the cipher GOST 28147-89 on purpose to further increase his security / I.V. Lisitskaya, T.V. Tsepurit, V.V. Lesnyak, M.V. Pinchuk, A.P. Meletsky // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 160-165.

The possibilities of the GOST 28147-89 cipher modernization to increase its resistance to attacks of the differential cryptoanalysis are considered. The additional non-linear transformation - parametrical cyclical shift and possibility

of increase length of ciphered block are proposed and justified.

2 tab. 7 fig. Ref.: 4 items.

УДК 681.3.06: 519.248.681

MDES-128 с таблицами подстановок случайного типа / В.И. Долгов, В.И. Руженцев, М.А. Федотов, А.П. Мелецкий, М.В. Пинчук // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 166-171.

Внимание сосредотачивается на устранении таких слабых сторон алгоритма DES как недостаточная длина ключа и блока данных, а также подверженность атакам дифференциального и линейного криптоанализов. В предлагаемом шифре размер секретного ключа составляет 128 битов, размер информационного блока – также 128 бит. В целях защиты от атак дифференциального и линейного криптоанализов в цикловую шифрующую функцию введена дополнительная нелинейная операция параметрического циклического сдвига. Приводится теоретическое и экспериментальное обоснование этого нововведения.

Табл. 6. Ил. 1. Библиогр.: 12 назв.

УДК 681.3.06: 519.248.681

MDES-128 з таблицями підстановок випадкового типу / В.І. Долгов, В.І. Руженцев, М.А. Федотов, О.П. Мелецький, М.В. Пинчук // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 166-171.

Увага приділяється винищенню таких слабких сторін алгоритму DES як недостатня довжина ключа та блока даних, а також незахищенність від атак диференційного та лінійного криптоанализів. В запропонованому шифрі розмір секретного ключа складає 128 бітів, розмір информаційного блока – також 128 бітів. З метою захисту від атак диференційного та лінійного криптоанализів в циклову шифруючу функцію введена додаткова нелінійна операція параметричного цикличного зсуву. Наводиться теоретичне і експериментальне обгрунтування цього нововведення.

Табл. 6. Іл. 1. Бібліогр.: 12 назв.

UDC 681.3.06: 519.248.681

MDES-128 with the substitution tables of a random type / V.I. Dolgov, V.I. Ruzhentsev, M.A. Fedotov, A.P. Meletskiy, M.V. Pintchuk // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 166-171.

The attention is focused to elimination of such weak parties of algorithm DES as unsufficient length of a key and block of the data, and also susceptibility to attacks differential and linear cryptanalisys. In the offered cipher the size of a confidential key is 128 bits and the size of the information block is also 128 bits. The additional nonlinear operation of parametrical cyclic shift was added into the round ciphering function. The theoretical and experimental substantiation of this innovation is resulted.

6 tab. 1 fig. Ref.: 12 items.

УДК 681.3.06: 519.248.681

Повышение устойчивости шифра DES к атакам дифференциального криптоанализа / М.Ф. Бондаренко, А.С. Коряк, В.И. Руженцев // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 172-176.

Рассматриваются возможности повышения стойкости алгоритма DES к дифференциальному криптоанализу. Обосновываются дополнительные критерии отбора таблиц *S* блоков, для которых может быть доказана устойчивость шифра DES против известных атак дифференциального криптоанализа.

Табл. 4. Библиогр.: 8 назв.

УДК 681.3.06: 519.248.681

Підвищення стійкості шифра DES до атак диференційного криптоаналізу / М.Ф. Бондаренко, А.С. Коряк, В.І. Руженцев // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 172-176.

Розглядаються можливості підвищення стійкості алгоритма DES до диференційного криптоаналізу. Обгрунтовуються додаткові критерії відбору таблиць *S*-блоків, до яких може бути доведена стійкість шифра DES проти відомих атак диференційного криптоаналізу.

Табл. 4. Бібліогр.: 8 назв.

UDC 681.3.06: 519.248.681

Rise of resistance of the cipher DES to attacks of differential cryptoanalysis / M.F. Bondarenko, A.C. Koryak, V.I. Ruzhentsev // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 172-176.

The possibilities to rise the resistance of algorithm DES to differential cryptoanalysis are considered. The additional criteria of selection of the tables S of blocks are justified. For such tables the stability of the cipher DES against the known attacks of differential cryptoanalysis can be proved.

4 tab. Ref.: 8 items.

УДК 681.3.06: 519.248.681

Обеспечение стойкости шифра DES к атакам линейного криптоанализа.Требования к отбору S-блоков, защищенных от атак на характеристики обнуляющего типа, четырехцикловые и шестицик-

ловые итеративные аппроксимации / И.В. Лисицкая, А.С.Бондаренко, А.И.Колыбельников // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 177-190.

Рассматриваются возможности повышения стойкости шифра DES к атакам линейного криптоанализа. Предлагаются и обосновываются новые требования к отбору S-блоков, защищенных от атак на характеристики обнуляющего типа, четырехцикловые и шестицикловые итеративные характеристики.

Табл. 1. Ил. 9. Библиогр.: 14 назв.

УДК 681.3.06: 519.248.681

Забезпечення стійкості шифру DES до атак лінійного криптоаналізу. Вимоги до відбору S-блоків, захищених від атак на характеристики обнуляючого типу, чотирициклові та шестициклові ітеративні апроксимації / І.В. Лисицька, А.С.Бондаренко, О.І.Колибельніков // Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 177-190.

Розглядаються можливості підвищення стійкості шифру DES до атак лінійного криптоаналізу. Пропонуються та обгрунтовуються нові вимоги до відбору S-блоків, захищених від атак на характеристики обнуляючого типу, чотирициклові та шестициклові ітеративні характеристики.

Табл. 1. Іл. 9. Бібліогр.: 14 назв.

UDC 681.3.06: 519.248.681

Provision of the cipher DES resistance to the linear cryptoanalysis attacks. Requirements to selection of S-bloks protected against the attacks on zeroing type characteristics, 4-cyclic and 6-cyclic iterate approximations / I.V. Lisitskay, A.S. Bondarenko, O. I. Kolybelnikov // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 177-190.

The possibilities of the cipher DES resistance to the linear cryptoanalysis attacks are considered. New requirements to the selection of S-bloks protected against the attack on zeroing type characteristics, 4-cyclic and 6-cyclic iterate approximations are proposed and justified.

1 tab. 9 fig. Ref.: 14 items.

УДК 681.3.06: 519.248.681

Защита информации в IP-телефонии / А.А. Замула, Ю.С. Павленко // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 191-194.

В статье рассматриваются вопросы, связанные с IP-телефонией. Проанализированы существующие стандарты сжатия речевой информации. Проведен анализ доступных реализаций некоторых алгоритмов сжатия. Предложен возможный вариант построения системы защиты речевой информации. Рассмотрены необходимые для этого алгоритмы и протоколы.

Табл. 3. Библиогр.: 3 назв.

УДК 681.3.06: 519.248.681

Захист інформації в IP-телефонії / А.А. Замула, Ю.С. Павлєнко // Радітехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 191-194.

У статті розглядаються питання, зв'язані з IP-телефонією. Проаналізовані існуючі стандарти стиснення мовної інформації. Проведено аналіз доступних реалізацій деяких алгоритмів стиснення. Запропоновано можливий варіант побудування системи захисту мовної інформації. Розглянуто необхідні для цього алгоритми та протоколи.

Табл. 3. Бібліогр.: 3 назв.

UDC 681.3.06: 519.248.681

Information security in IP-telephony / A.A. Zamula, Y.S. Pavlenko // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 199. P. 191-194.

The problems associated with IP-telephony are considered in the article. Existing standards of speech information compression are analyzed. Some compression algorithms available for realization are analyzed. Possible version of the speech information security system building is offered. The required algorithms and protocols are considered.

3 tab. Ref.: 3 items.

УДК 519.7:007.52

Метод выбора эффективных процедур оценивания параметров моделей квазистационарных процессов в нейросетевой экспертной системе / Н.С. Лесная, Т.Б. Шатовская, В.Б. Репка // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 195-198.

Рассмотрен принцип выбора эффективного метода смещенного и робастного оценивания параметров моделей квазистационарных процессов в условиях мультиколлинеарности и зашумленности исходных данных. Выбор методов оценивания базируется на совокупности критериев оценки их точности и применении нейросетевого подхода. Исследована проблема чувствительности типов нейросетей, настройки их параметров, режимов обучения к входным характеристикам исследуемых квазистационарных объектов и классам статистических ме-

тодов оценивания параметров моделей.

Библиогр.: 6 назв.

УДК 519.7:007.52

Метод вибору ефективних процедур оцінювання параметрів моделей квазістаціонарних процесів у нейромережевій експертній системі / Н.С.Лєсна, Т. Б. Шатовська, В.Б.Рєпка // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 195-198.

Розглянуто принцип вибору ефективного методу зміщеного та робасного оцінювання параметрів моделей квазістаціонарних процесів в умовах мультиколінеарності та зашумленості вихідних даних. Вибір методів оцінювання базується на сукупності критеріїв оцінки їх точності та застосуванні нейромережевого підходу. Досліджено проблему чутливості типів нейромереж, настроювання їх параметрів, режимів навчання до вхідних характеристик квазістаціонарних процесів, що досліджуються, та класів статистичних методів оцінювання параметрів моделей.

Бібліогр.: 6 назв.

UDC 519.7:007.52

The method of a choice of effective procedures models parameters estimation for kvazistationary processes in neuralexpertsystem / N.S. Lesnaya, T.B. Shatovskaya, V.B. Repka // Radiotekhnika. All-Ukraine Sci. Interdep. Mag. 2001. № 119. P. 195-198.

The method a choice's of an effective method's of biased and robust parameters models estimation for kvazistationary processes in conditions of multicollinearity and noise in the initial data is considered. The choice of estimation methods is based on set of accuracy estimation criteria and application of neural networks approach. The problem of sensitivity of neural networks types, adjustment of their parameters, modes of training to the entrance characteristics researched kvazistationary processes and classes of statistical methods of models parameters estimation is investigated.

Ref.: 6 items.

УДК 621.396.67

Алгоритм анализа эквидистантной решетки ленточных микрополосковых излучателей произвольной геометрии, адаптированный к расчету крупноапертурных антенн с нелинейными элементами 3. Особенности численной реализации алгоритма / В.М. Шокало, А.И. Лучанинов, А.А. Коновальцев, Ю.А. Лучанинов, М.А. Омаров // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 119. С. 199-210.

Описывается быстродействующий алгоритм расчета электродинамических характеристик бесконечных периодических решеток линейных излучателей произвольной конфигурации. Получены аналитические выражения для определения элементов матрицы обобщенных импедансов в методе моментов. Предложен способ приближенного вычисления последних для случая, когда характеристики антенн определяются в зависимости от направления прихода возбуждающей плоской волны.

Ил. 1. Библиогр.: 8 назв.

УДК 621.396.67

Алгоритм аналізу еквідистантної решітки стрічкових мікросмужкових випромінювачів довільної геометрії, адаптований до розрахунку великоапертурних антен з нелінійними елементами 3. Особливості чисельної реалізації алгоритму / В.М. Шокало, А.І. Лучанінов, А.О. Коновальцев, Ю.А. Лучанінов, М.А. Омаров // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 199-210.

Описано швидкодіючий алгоритм розрахунку електродинамічних характеристик нескінчених періодичних решіток лінійних випромінювачів довільної конфігурації. Отримано аналітичні вирази для визначення елементів матриці узагальнених імпедансів у методі моментів. Запропоновано спосіб наближеного розрахунку останніх для випадку, коли характеристики антени визначаються в залежності від напрямку приходу збуджуючій плоскої хвилі.

Іл. 1. Бібліогр.: 8 назв.

UDC 621.396.67

The analysis algorithm of equidistant arrays of tapered microstrip radiators of an arbitrary geometry adapted to designing large-aperture antennas with non-linear elements 3. Features of the algorithm numerical realization / V.M. Shokalo, A.I. Luchaninov, A.A. Konovaltsev, Yu.A. Luchaninov, M.A.Omarov // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 199-210.

The high-speed algorithm of electromagnetic characteristics calculation of infinite periodic arrays of linear radiators of an arbitrary configuration is described. The analytical expressions for determining the elements of generalized impedance matrix in the method of moments are obtained. The way of approximated calculation of the elements when the antenna characteristics are determined depending on the exciting plane wave arrival direction is proposed.

1 fig. Ref.: 8 items.

УДК 621.373.826

Влияние внешнего переменного электрического поля на энергетические состояния частиц и квазичастиц в квантоворазмерной структуре / А.Г. Пащенко // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 211-214.

Рассматривается влияние внешнего переменного электрического поля на энергетические состояния частиц и квазичастиц в квантоворазмерной структуре, созданной на основе $GaAs/Al_xGa_{1-x}As$, с использованием теории возмущений.

Ил. 1. Библиогр.: 7 назв.

УДК 621.373.826

Вплив зовнішнього змінного електричного поля на енергетичні стани частинок і квазічастинок в квантоворозмірній структурі / О.Г. Пащенко // Радіотехника: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 211-214.

Розглянуто вплив зовнішнього змінного електричного поля на енергетичні стани частинок і квазічастинок в квантоворозмірній структурі, створеній на основі $GaAs/Al_xGa_{1-x}As$, з використанням теорії збурень.

Іл. 1. Бібліогр.: 7 назв.

UDC 621.373.826

The influence of the external time-varying electric field on the particles and quasi-particles power state in the quantum-well structure / A.G. Pashchenko // Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 211-214.

The influence of the external time-varying electric field on the particles and quasi-particles power state in the quantum-well structure manufactured on the basis $GaAs/Al_xGa_{1-x}As$ using excitation theory is considered in this paper.

Fig. 1. Ref.: 7 items

УДК 621.371

Некоторые исследования фазовой поверхности акустических воли для задач радиометеорологии / С.И. Бабкин, Г.В. Груша, Е.Г. Прошкин, Н.И. Слипченко // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 215-220.

На основе применения геометрико-акустических фазовых поверхностей в движущейся неоднородной атмосфере изучается частотная зависимость накопления фазы акустической волны с учетом дисперсии скорости звука. Получены рекуррентные формулы для послойного определения разностей фаз на двух кратных частотах, позволяющие учесть вертикальную изменчивость метеопараметров при двухчастотном акустическом или радиоакустическом зондировании. Для практических применений при радиометеорологических измерениях коэффициента преломления радиоволн указаны метод повышения точности фазового восстановления влажности воздуха градиентными измерениями и требования к точности аппаратуры.

Библиограф.: 9 назв.

УДК 621.371

Деякі дослідження фазової поверхні акустичних хвиль для задач радіометеорології / С.І. Бабкін, Г.В. Груша, Є.Г. Прошкін, М.І. Сліпченко // Радіотехніка:Всеукр. міжвід. наук.-техн. зб. 2001.Вип. 119. С. 215-220.

На основі застосування геометрико-акустичних фазових поверхень у рухомій неоднорідній атмосфері вивчається частотна залежність накопичення фази акустичної хвилі з урахуванням дисперсії швидкості звуку. Одержані рекуррентні формули для пошарового визначення різниць фаз на двох кратних частотах, які дозволяють врахувати вертикальну мінливість метеопараметрів при двохчастотному акустичному або радіоакустичному зондуванні. Для практичних застосувань при радіометеорологічних вимірюваннях коефіцієнта заломлювання радіохвиль вказані метод підвищення точності фазового відтворювання вологості повітря градієнтними вимірюваннями і вимоги до точності апаратури.

Бібліог.: 9 назв.

UDC 621.371

Some investigations of the acoustic waves phase surface for radiometeorology problems / S.I.Babkin, G.V.Grusha, E.G.Proshkin, N.I.Slipchenko// Radiotekhnika: All.-Ukr.Sci.Interdep.Mag.2001. № 119. P. 215-220.

On the base of geometrical acoustic approximation for the phase surfaces in moving unhomogeneous atmosphere the frequency dependence is investigated for the acoustic phase accumulation with account for the sound velocity dispertion. The recurrence formulae for layer by layer definition of the phase differences in two multiple frequencies are obtained, this allows to take into account the vertical variability of meteorological parameters in two-frequency acoustic or radioacoustic sounding. For practical applications in radiometeorological measurements of the radio wave refraction index the accuracy rise method of the air humidity phase restoration by gradient measurements and the requirements to the equipment accuracy are indicatid.

Ref.: 9 items.

УДК 621.396.96'06

Влияние взаимного энергетического спектра зондирующих сигналов на информационные характеристики систем зондирования атмосферы / В. М. Карташов // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2001. Вып. 119. С. 221-225.

Показано, что вид, параметры и принципиальная возможность существования радиосигнала, рассеянного звуковой посылкой, определяются особенностями взаимного энергетического спектра зондирующих акустического и электромагнитного колебаний.

Ил. 1. Библиогр.: 7 назв.

УДК 621.396.96'06

Вплив взаємного енергетичного спектра зондуючих сигналів на інформаційні характеристики систем зондування атмосфери / В.М. Карташов / Радіотехніка: Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 221-225.

Показано, що вид, параметри і принципова можливість існування радіосигналу, розсіяного звуковою посилкою, визначаються особливостями взаємного енергетичного спектра зондуючих акустичного і електромагнітного коливань.

Ил. 1. Бібліогр.: 7 назв.

UDC 621.396.96'06

Action of the sounding signals mutual power spectrum on the information characteristics of the atmosphere sounding systems / V.M. Kartashov //Radiotekhnika: All-Ukr. Sci. Interdep. Mag. 2001. № 119. P. 221-225.

It is shown that the form, parameters and conceptual probability of existence of the scattered radio signal, received from the sound transmission, are defined by the singularities of the mutual power spectrum of sounding acoustic and electromagnetic oscillations.

1 fig. Ref.: 7 items

УДК 532. 59

К оценке интенсивности вторичных источников поля при акустическом зондировании турбулентных движущихся сред / А.Ю. Панченко // Радиотехника: Всеукр. межвед. науч. техн. сб. 2001. Вып. 119. С. 226-229.

Рассматривается инвариантный подход для получения уравнения движения идеального газа. Вывод основан на законе сохранения кинетической энергии. Определяется влияние источников отраженного поля представленных в уравнении непрерывности.

Ил.3.Библиогр.: 3 назв.

УДК 532. 59

До оцінки потужности вторинних джерел поля при акустичному зондуванні турбулентних середовищ, які рухаються / О.Ю.Панченко// Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2001. Вип. 119. С. 226-229.

Розглядається інваріантний підхід для отримання рівняння руху ідеального газу. Висновки базуються на законі збереження кінетичної енергії. Визначено вплив джерел відбитого поля, які подані в рівнянні неперервності.

Іл.3.Бібліогр.: 3 назв.

UDC 532.59

On estimation of the field secondary sources intensity with acoustic sounding of the turbulent moving medi / A. Yu. Panchenko // Radiotekhnika: All.Ukr. Sci. Interdep. Mag. 2001. № 119. P. 226-229

Invariant approach to obtaining of the ideal gas movement equation is considered. The derivation is based on th law of conservation of the kinetic energy. The action of the reflected field sources presented in the continuity equatic is defined.

3 fig. Ref.: 3 items.