МИНИСТЕРСТВО ОБРАЗОВАНИЯ УКРАИНЫ
ХАРЬКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ РАДИОЭЛЕКТРОНИКИ

РАДИОТЕХНИКА

Всеукраинский межведомственный научно-технический сборник

Основан в 1965 г.

ВЫПУСК 114

Харків

Харківський державний технічний
університет радіоелектроніки

2000

УДК 621.3

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

Рассматриваются проблемы защиты информации в различных информационных технологиях. Приводятся анализ и оценка свойств проектируемых и черновых версий криптографических алгоритмов XXI века. Излагаются методы и приводятся примеры их применения для криптоанализа симметричных блочных криптоалгоритмов. Анализируется сложность несимметричных криптографических преобразований, указываются пути ее уменьшения. Обсуждаются системные вопросы защиты информации.

Изложены результаты теоретических исследований в области формирования и обработки радиосигналов. Проанализированы возможности беспроводных систем передачи энергии СВЧ-лучом.

Для научных работников, специалистов, преподавателей ВУЗов, аспирантов.

Ответственный за выпуск д-р техн. наук, проф. *И.Д. Горбенко.*

Рекомендовано Ученым советом университета, протокол №21 от 31.03.2000.

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

ЗБІРНИК НАУКОВИХ ПРАЦЬ
РАДІОТЕХНІКА
Випуск 114

СБОРНИК НАУЧНЫХ ТРУДОВ
РАДИОТЕХНИКА
Выпуск 114

Виконавці комп'ютерної верстки *С.Я. Захарченко, М.В. Цмугун*

Підп. до друку з ориг.-макета 10.04.2000. Формат 60×84/8.
Папір офсет. Друк офсет. Ум. друк. арк. 11,8. Обл.-вид. арк. 18,7.
Тираж 300 пр. Зам. №456. Ціна договір.

Харківський державний технічний університет радіоелектроніки (ХТУРЕ).
61166 Харків, просп. Леніна, 14.
Надруковано в ТОВ «Техно-АРТ»
61024, Харків, вул. Артема, 32.

УВАЖАЕМЫЕ ЧИТАТЕЛИ!

Настоящий выпуск сборника «Радиотехника» в основном посвящен проблемным вопросам теории и практики информационной безопасности. На наш взгляд, публикуемые в нем статьи отражают остро стоящие в мире проблемы защиты информации в различных информационных технологиях.

По мнению ведущих специалистов мира процесс информатизации мирового сообщества, прежде всего его государственных и общественных институтов, развивается чрезвычайно стремительно и в ряде случаев недостаточно управляемо. Большинство государств мира с большим опозданием осмысливают и оценивают политические, экономические, военные, социальные и другие последствия информатизации. Безусловно, использование новых информационных технологий в различных отраслях, в производстве, управлении, общественной жизни, образовании и т.п. является добром. Создание же единого мирового информационного пространства, в котором накапливается, распределяется, передается, принимается, преобразуется и уничтожается информация, ускоряет процесс интеграции мирового сообщества, прежде всего за счет оперативного обмена научно-технической, экономической, учебной и другой информацией. Но подобно тому, как достижения ядерной физики породили опасность ядерных войн, катастроф подобных Чернобылю, так и создание мирового информационного пространства породило значительное число неразрешимых противоречий в сложности обеспечения информационной безопасности. По существу, в этом пространстве уже сегодня идут непрерывные информационные войны, которые только на первый взгляд менее ужасны, чем обычные. Поэтому проведение широких исследований и выполнение практических разработок в области информационной безопасности является чрезвычайно актуальным.

На наш взгляд, под информационной безопасностью необходимо понимать защищенность информации и поддерживающей инфраструктуры от преднамеренных и случайных воздействий естественного и природного характера, в результате которых наносятся убытки владельцам или пользователям информации и поддерживающей инфраструктуре. Информационная безопасность обеспечивается за счет реализации комплекса мероприятий, которые составляют суть защиты информации в различных информационных технологиях, функционирование которых базируется на использовании компьютерных систем и сетей, систем телекоммуникаций, радиотехнических систем и т.п. Поэтому неудивительно, что вопросам защиты информации сегодня уделяется особое внимание, на это направляются существенные материальные и финансовые средства.

Разрешение основных противоречий в области информационной безопасности требует решения ряда научных и прикладных проблем прежде всего криптографической и технической защиты информации. Сегодня в мире сделаны значительные усилия в этих направлениях. Прежде всего, исследования и разработки ведутся в направлении создания систем и средств криптографической защиты информации. Очень проблематичными являются задачи создания, анализа и отбора мировых стандартных криптографических алгоритмов 21 века. Сегодня в мире уже ведется 11 проектов по созданию стандартов шифрования, цифровой подписи, хеширования, аутентификации, направленного шифрования, идентификации и генерации псевдослучайных последовательностей. Первый из них – создание открытого стандарта блочного симметричного шифрования подходит к завершению, очевидно 19 - 25 апреля в Нью-Йорке будет выбран лучший криптоалгоритм. В связи с этим в настоящем сборнике в первой статье представлены наши оценки и высказывается предположение о предпочтительности того или иного кандидата.

Все статьи, которые публикуются в сборнике, можно разделить на пять направлений. Первое – это анализ кандидатов в стандарты AES и рабочие версии проектов стандартов X9.62 и X9.63. Относительно них получен ряд оценок, которые позволяют определить или даже сравнить показатели качества, прежде всего по реальной криптостойкости, статистической безопасности, сложности преобразований и др. Ввиду особой важности в сборник включена статья по критериям и методологии оценки безопасности информационных технологий.

Второе направление – это рассмотрение методов криптоанализа симметричных криптоалгоритмов и получение оценок по стойкости применительно к блочным симметричным криптоалгоритмам.

В настоящее время широкое распространение получили несимметричные криптоалгоритмы и криптографические протоколы, на них базирующиеся. Проблемными, а то и дискуссионными, являются вопросы стойкости и сложности преобразований. В связи с этим в сборник включен ряд статей, связанных с оценкой и разработкой методов и алгоритмов уменьшения вычислительной сложности несимметричных криптопреобразований.

Мы также считаем необходимым опубликовать ряд статей, которые посвящены исследованию системных вопросов и анализу криптографических протоколов в системах защиты информации. Эти статьи составляют по тематике четвертое направление.

Редколлегия предлагает Вашему вниманию также ряд статей радиотехнической тематики.

Издание сборника такой направленности позволит довести до специалистов и интересующихся проблемами защиты информации ряд новых сведений и достижений, развернуть дискуссии.

Конечно же, статьи отражают мнение авторов по рассматриваемым задачам. Публикация статей такой направленности позволит в определенной мере выработать взгляд на пути создания и развития методов и средств защиты информации.

Учитывая важность решения проблемы защиты информации в различных информационных технологиях, редколлегия приглашает специалистов к публикации в сборнике «Радиотехника» результатов исследований в этой области.

С уважением и благодарностью к читателям и специалистам

ректор ХТУРЭ, профессор **М.Ф. Бондаренко**

заведующий кафедрой БИТ,
профессор ХТУРЭ **И.Д. Горбенко**

УДК 681.3.06: 519.248.681

М.Ф. БОНДАРЕНКО, д-р техн. наук, И.Д. ГОРБЕНКО, д-р техн. наук, А.В. ПОТИЙ, канд. техн. наук, О.И. ОЛЕШКО, С.А. ГОЛОВАШИЧ, А.С. БОНДАРЕНКО

УЛУЧШЕННЫЙ СТАНДАРТ СИММЕТРИЧНОГО ШИФРОВАНИЯ XXI ВЕКА: КОНЦЕПЦИЯ СОЗДАНИЯ И СВОЙСТВА КАНДИДАТОВ

Введение

Сегодня в области информационной безопасности решается важная проблема создания улучшенного стандарта симметричного шифрования – криптографического алгоритма XXI века. Ее решение инициировал Национальный Институт Стандартов и Технологий (NIST) США [1, 2]. Для этого в 1997 году NIST объявил конкурс алгоритмов, претендующих на то, чтобы стать стандартом, а также сформировал минимальные требования, которые должны быть выполнены при разработке и представлении на конкурс таких алгоритмов. Предполагается, что в следующем столетии стандарт симметричного шифрования AES (Advanced Encryption Standard) будет описывать общедоступный, распространяемый без ограничений по всему миру симметричный криптоалгоритм, который сможет осуществлять надежную защиту правительственной информации.

Было спланировано три этапа рассмотрения представленных алгоритмов. На первой конференции, посвященной стандарту AES, в 1998 г. вынесено решение о принятии к открытому обсуждению 15 пакетов описаний алгоритмов симметричного шифрования. На втором этапе осуществляется публичный анализ соответствия каждого стандарта требованиям и оценка качества криптографических преобразований в целом. В августе 1999 г. проведена вторая конференция, целью которой было обобщение оценок, полученных государственными организациями и в результате открытого обсуждения, уменьшение количества алгоритмов-кандидатов до 5, которые затем уже будут рассматриваться на третьей конференции в апреле 2000 года. Из всего вышеизложенного вытекает, что NIST рассматривает проблему создания перспективного стандарта шифрования как очень важную, сложную и в значительной степени неопределенную.

Целью этого доклада является изложение концепции создания улучшенного стандарта шифрования, анализ хода и противоречий ее реализации, изложение и обсуждение сформированных требований и ограничений, дополнение и уточнение требований к стандарту, а также проведение сравнительного анализа кандидатов в AES и выбор лучших или лучшего. Кроме того, мы хотим доказать, что приобретенный на сегодня опыт указывает на серьезность проблемы и невозможность ее решения без объединения общих усилий, а главное – использование приобретенного опыта и полученных результатов в интересах Украины.

1. Основные требования.
Основные требования к процедуре представления и свойствам кандидатов AES

Первой конференцией были принятые к дальнейшему рассмотрению только те алгоритмы, на которые были представлены «полные» пакеты документации. «Полные» пакеты должны были содержать:

- сопроводительное письмо;
- описание алгоритма и соответствующую документацию;
- магнитный носитель со всеми материалами;
- заявление о принятии к рассмотрению криптографического алгоритма, включая отказ от прав интеллектуальной собственности на алгоритм.

Полное описание должно было включать в себя все необходимые математические уравнения и выражения, таблицы, диаграммы, обоснование выбора и описание параметров, обоснование

количества циклов и т.п. Некоторые сведения об AES приведены в таблице 1. Криптоалгоритмы проранжированы по эффективности авторами.

Разработчик должен представить свои оценки относительно вычислительной сложности программного и аппаратного обеспечения, а также вычислительную эффективность алгоритма, в том числе и для 8-битовых процессоров. Причем для оценки аппаратного обеспечения рекомендуется использовать в качестве показателя число логических элементов, для программного обеспечения – тип процессора, тактовую частоту его работы, объем памяти, операционную систему, и т.п.

Таблица 1

Наименование алгоритма	Страна	Разработчики
RC6	США	RSA Laboratories (represented by Matthew Robshaw)
MARS	США	IBM (represented by Nevenko Zunic)
Twofish	США	Bruce Schneier, John Kelsey, Doug Whiting, David Wagner, Chris Hall, Niels Ferguson
Rijndael	Бельгия	Joan Daemen and Vincent Rijmen
CRYPTON	Корея	Future Systems, Inc. (represented by Chae Hoon Lim)
CAST-256	Канада	Entrust Technologies, Inc. (represented by Carlisle Adams)
E2	Япония	Nippon Telegraph and Telephone Corporation (NTT) (represented by Masayuki Kanda)
SERPENT	Великобритания, Израиль, Норвегия	Ross Anderson, Eli Biham, Lars Knudsen
HPC	США	Rich Schroeppel
DFC	Франция	Centre National pour la Recherche Scientifique (CNRS) (represented by Serge Vaudenay)
SAFER+	США	Cylink Corporation (represented by Lily Chen)
LOKI97	Австралия	Lawrie Brown, Josef Pieprzyk, Jennifer Seberry
DEAL	США	Richard Outerbridge and Lars Knudsen
FROG	Коста-Рика	TecApro Internacional S.A. (represented by Dianelos Georges Georgoudis)
Magenta	Германия	Deutsche Telekom AG (represented by Klaus Huber)

Оценка скорости работы алгоритма должна быть представлена в виде числа тактов работы, необходимой для:

- шифрования одного блока данных;
- дешифрования одного блока данных;
- разворачивания (настройки) ключа;
- настройки алгоритма или его части (например, формирования таблиц);
- смены ключа для каждой из рабочих длин.

Описание должно содержать все возможные компромиссные варианты, позволяющие выбирать необходимую скорость работы при требуемой устойчивости.

- Для каждого алгоритма должны быть разработаны серии тестов – тесты известных ответов и тесты Монте-Карло.

2. Минимальные требования к кандидатам AES

При создании перспективного стандарта определены и зафиксированы требования к нему. Минимальные требования, которым должны были отвечать на первом этапе алгоритмы-кандидаты AES, таковы:

3.1. Криптоалгоритм должен строиться на использовании криптографии симметричных (секретных) ключей. То есть ключи источника криптограмм и получателя должны или совпадать или рассчитываться один из другого не выше чем с полиномиальной сложностью.

3.2. Криптоалгоритм должен строиться с использованием блочного симметричного шифра с длиной блока $l_б$=128 бит.

3.3. Длины начальных ключей должны быть $l_к$=128, 192 и 256 бит.

3.4. Криптоалгоритм должен работать для различных комбинаций длина сообщения/длина ключа, но обязательно для комбинаций длина блока / длина ключа 128/128, 128/192, 128/256 бит.

3. Критерии и показатели оценки качества

Отбор AES алгоритмов осуществляется с учетом выполнения минимальных требований. Кроме того учитывается:

1. Реальная защищенность от криптоаналитических атак. При этом основными методами криптоанализа являются:
   - дифференциальный криптоанализ;
   - расширения для дифференциального криптоанализа;
   - поиск наилучшей дифференциальной характеристики;
   - линейный криптоанализ;
   - интерполяционное вторжение;
   - вторжение с частичным угадыванием ключа;
   - вторжение с использованием связанного ключа;
   - вторжение на основе обработки сбоев;
   - поиск лазеек.
2. Статистическая безопасность криптографических алгоритмов.
3. Надежность математической базы криптографических алгоритмов.
4. Расчетная сложность криптографических алгоритмов для программной и аппаратной реализаций (может оцениваться скоростью преобразований).
5. Требования к памяти при программной и аппаратной реализациях. При аппаратной реализации оценивается числом логических элементов, при программной – количеством необходимой оперативной и постоянной памяти, в том числе для различных платформ и сред.
6. Гибкость алгоритма, то есть:
   - возможность работы с иными длинами начальных ключей и информационных блоков;
   - безопасность реализации в широком диапазоне различных платформ и приложений, включая 8-битовые процессоры;
   - возможность использования криптографического алгоритма в качестве поточного шифра или генератора псевдослучайных чисел, алгоритма хеширования, для обеспечения подлинности сообщений (выработка кодов аутентификации) и т.п.;
   - одинаковой сложности как аппаратной, так и программной реализации, а также программно-аппаратной реализации.

Анализ основных публикаций относительно исследования свойств алгоритмов-кандидатов AES [3] и полученные авторами результаты позволяют сравнить представленные алгоритмы согласно основным требованиям.

4. Реальная защищенность от криптоаналитических атак

В третьем пункте определены основные методы криптоанализа, которые должны применяться к алгоритмам-кандидатам. Это достаточно сложная проблема. Сегодня необходимо рассматривать и оценивать устойчивость алгоритмов к атакам обычного и расширенного дифференциального криптоанализа, поиска наилучшей дифференциальной характеристики, линейного криптоанализа,

интерполяционного и итерационного вторжения, вторжения с частичным угадыванием ключа и с использованием связанного ключа, вторжения на основе обработки сбоев, поиска лазеек и т.д.

Основные результаты оценки защищенности приведены в таблице 2. В столбце «Блок» указана длина блока, в столбце «Ключ» – длина начального ключа, в столбце «Циклы» – число циклов, реализованных в алгоритме, при которых выполнялся криптоанализ.

В колонке «Тип» указаны типы (условия) атак:

K – наилучшая атака при известном тексте с 2^a парами открытый текст/шифрованный текст, для осуществления которой необходимо 2^b операций шифрования и 2^c слов памяти;

C – наилучшая атака при выборе шифртекста с 2^a парами открытый текст/шифртекст, для осуществления которой необходимо выполнить 2^b операций шифрований и 2^c слов памяти;

? – неизвестна никакая атака, кроме «грубой силы»;

CP – наилучшая атака при выбранном открытом тексте с 2^a парами открытый текст/шифрованный текст, для осуществления которой необходимо выполнить 2^b операций шифрования и 2^c слов памяти;

«–» – означает, что соответствующие значения нам неизвестны.

r – дополнительный параметр

Анализ данных из таблицы 2 показывает, что среди алгоритмов-кандидатов AES (с точки зрения стойкости к атакам криптоанализа) наихудшими являются алгоритмы DEAL, Magenta и Frog. Так, для DEAL при длине блока 128 бит и длине ключа 128 бит существует аналитическая атака при выборе шифртекста, для выполнения которой необходимо 2^{70} пар открытый текст/шифртекст и 2^{72} операций шифрования, что намного меньше, чем при атаке «грубая сила» (2^{128}). Для алгоритма Magenta разработан метод криптоанализа при выборе шифртекста, который требует 2^{64} операций шифрования и 2^{64} пар открытый текст/шифртекст. Для алгоритма Frog разработана атака, которая требует $2^{56,7}$ операций шифрования и 2^{36} пар открытый текст/шифртекст. Указанное исключает алгоритмы DEAL, Magenta и Frog как не обеспечивающие стойкости против аналитических атак, сложность которых намного меньше атаки типа «грубая сила».

Необходимо отметить, что эффективные атаки разработаны также и для существующих и используемых криптоалгоритмов DES, IDEA, Loki(90), Safer(93) и RC5(94). Соответствующие данные приведены в таблице 2.

Таблица 2

Название	Версия	Блок	Ключ	Циклы	Атаки				
					Тип	a	b	c	r
DES	77	64	56	16	K	43	19	13	
	3-DES (77)	64	168	48	K	2	112	56	
	2k3-DES (78)	64	112	4	C	56	56	56	
IDEA	(91)	64	128	8,5	C	56	67	32	3,5
Loki	(90)	64	64	16	C	54	–	–	14
	(91)	64	64	16	C	58	–	–	13
Safer	K(93)	64	64, 128	6, 10	S	45	–	32	5
	SK(95)	64	40, 64, 128	8, 10	?				
Blowfish	(93)	64	32-448	16	?				
RC5	32/12/K (94)	64	8S, S<256	12	C	54	–	–	
	64/16/16(94)	128	8S, S<256	16	C	83	–	–	24
CAST-128	(95)	64	40-128	12, 16	?				
SHARK	(96)	64	128	6	?				
SQUARE	(97)	128	128	8	?				
MISTY	1(97)	64	128	8	?				
	2(97)	64	128	16	?				

Продолжение таблицы 2

ICE	(97)	64	64	16	CP	62	62	30	
Rainbow	(98)	128	128	7	?				
RC6	(98)	128	128, 192, 256						
Mars	(98)	128	128-1248	6					
Twofish	(98)	128	128, 192, 256						
Rinjdael	(98)	128	128, 192, 256						
Crypton	(98)	128	128, 192, 256						
CAST	(98)	128	128, 192, 256						
E2	(98)	128	128, 192, 256						
Serpent	(98)	128	128, 192, 256						
HPC	(98)	128	128, 192, 256						
DFC	(98)	128	128, 192, 256						
Safer	(98)	128	128, 192, 256						
Loki	(98)	128	128, 192, 256						
		128	56	6,8	C	70	72		6
		128	56	6,8	C	32,5	144,5		6
DEAL		192			C	70	136		6
		256			C	32,5	145		6
Magenta		128	128	6	C	64	64		2
		128	128	6	C	33	97		2
		128	192	6	C	128	128		2
		128	192	6	C	33	161		2
		128	256	6	C	128	192		2
		128	256	6	C	33	225		2
FROG		128	128			58			
		128				56			
		128	192			64			
		128	256			36	56,7		

5. Статистическая безопасность криптографических алгоритмов

Основными методами определения статистической безопасности криптографических алгоритмов являются методы, связанные с расчетом связанности (корреляции блоков) криптограмм между собой и со входными блоками открытых текстов, а также определение избыточности криптограмм. Избыточность связана с зависимостью символов сообщений соответствующего алфавита и неодинаковой вероятностью их появления.

Пусть M_1, M_2, ..., M_i, ..., M_n - входные блоки открытого текста, а C_1, C_2, ..., C_i, ..., C_n - блоки криптограмм, полученные при применении ключа K. Символы M_i и C_i являются двоичными, то есть принимают значения (0,1). Связанность Mi и Ci определим функцией:

$$F_1 = f(M_i, C_i) = \sum_{j=1}^{l_d} M_i \oplus C_i , \qquad (1)$$

а C_i и C_k

$$F_2 = f(C_i, C_k) = \sum_{j=1}^{l_d} C_i \oplus C_k . \qquad (2)$$

Будем рассматривать значения F_1 и F_2 как случайные числа. Обрабатывая случайные последовательности F_1 и F_2, например вычислением математического ожидания

$m(F_1)$ и $m(F_2)$

и соответствующих моментов

$$m^s(F_1) \text{ и } m^s(F_2),$$

получим количественные показатели зависимости и связанности M_i и C_i, C_i и C_k.

В дальнейшем будем предполагать M_i и C_i, C_i и C_k независимыми (некоррелированными), если

$$F_1 = {}^{l_á}\!/_2 \text{ и } F_2 = {}^{l_á}\!/_2$$

соответственно.

Оценка статистической безопасности для всех AES кандидатов проведена при следующих начальных данных:

1. случайный ключ, случайные данные, ключ постоянный
2. единичный ключ, случайные данные, ключ постоянный
3. случайный ключ, единичные данные, ключ постоянный
4. случайный ключ, нулевые данные, ключ изменяется
5. случайный ключ, случайные данные, ключ изменяется

В процессе оценки F_2 рассчитывалось на основе C_i и C_k, причем M_i и M_k, для блоков криптограмм, отличаются на один бит.

В таблице 3 приведены значения $m(F_1)$, $m(F_2)$, $m^2(F_i)$, $m^2(F_2)$, а также max F_1, max F_2 и min F_1, min F_2, которые рассчитывались на выборке размера n (n = 5000). В таблице 3 оценкой для математического ожидания служит среднее арифметическое наблюдаемых значений случайных величин F_1 и F_2 в n независимых наблюдениях.

Приведенные в таблице 3 некоторые данные подтвердили вывод о том, что все кандидаты AES обеспечивают статистическую безопасность криптографических преобразований. Вместе с тем, значения $m^2(F_1) = 30.75$, $m^2(F_2) = 30.81$ для Twofish, $m^2(F_1) = 30.67$, $m^2(F_2) = 33.47$ для Rijndael, $m^2(F_1) = 30.86$ для DEAL, $m^2(F_2) = 33.64$ для FROG, $m^2(F_2) = 30.19$ для Magenta позволяют сделать вывод о необходимости проведения дополнительных исследований для названных алгоритмов.

Таблица 3

Название AES	$m(F_1)$	$m^2(F_1)$	min F_1	max F_1	$m(F_2)$	$m^2(F_2)$	min F_2	max F_2	Вариант
RC6	63,87	32,27	45	84	64,20	32,96	41	85	5.1
	64,11	31,97	45	88	64,02	32,09	41	86	5.2
	63,97	32,62	45	88	63,98	31,61	41	86	5.3
	63,94	31,87	43	88	63,95	32,28	41	86	5.4
	64,20	32,21	43	89	64,04	31,41	41	86	5.5
Mars	63,95	31,28	43	84	64,06	31,85	45	87	5.1
	64,08	32,28	43	84	63,99	31,81	41	87	5.2
	63,98	33,27	38	86	64,04	31,15	41	87	5.3
	64,04	31,41	38	86	64,02	32,33	41	87	5.4
	63,94	31,77	38	86	64,10	32,82	41	87	5.5
Twofish	64,05	31,97	42	84	64,01	32,35	46	86	5.1
	64,06	30,75	42	84	64,01	30,81	42	86	5.2
	64,21	32,74	40	85	63,91	32,30	40	86	5.3
	64,01	31,52	40	85	64,18	31,52	40	86	5.4
	63,90	31,89	40	85	63,88	32,06	40	86	5.5
Rijndael	64,16	32,25	44	90	63,95	32,12	44	85	5.1
	63,79	31,86	39	90	64,03	31,07	44	85	5.2
	63,99	32,40	39	90	63,97	32,30	42	86	5.3
	63,98	33,47	39	90	64,05	31,92	42	86	5.4
	63,88	31,70	39	90	64,00	30,67	42	86	5.5

	64,17	32,45	45	85	64,14	32,87	44	84	5.1
	63,94	31,63	40	86	64,05	33,04	41	84	5.2
Crypton	63,96	31,88	40	86	63,91	32,84	41	86	5.3
	63,93	32,35	40	88	63,97	31,42	41	86	5.4
	64,01	32,02	40	88	64,01	32,45	40	86	5.5
	63,99	32,49	46	83	63,98	32,48	44	85	5.1
	64,09	31,29	45	86	64,17	32,96	43	85	5.2
DEAL	63,83	31,73	43	86	63,92	31,80	43	85	5.3
	63,94	30,86	41	86	64,07	32,55	43	85	5.4
	64,00	31,42	41	86	63,99	32,26	43	85	5.5
	64,04	32,42	42	87	64,09	33,64	42	86	5.1
	64,02	31,92	42	87	64,09	32,20	42	86	5.2
Frog	64,01	31,52	42	87	64,19	31,34	42	86	5.3
	64,06	31,72	42	88	64,04	31,08	42	86	5.4
	64,07	31,53	42	88	63,87	31,52	42	86	5.5
	63,92	31,10	42	82	63,93	32,82	45	84	5.1
	64,01	31,60	42	86	64,08	32,16	44	85	5.2
Magenta	64,05	32,61	42	86	64,04	32,20	43	86	5.3
	64,12	30,19	42	86	64,07	31,91	43	86	5.4
	64,01	31,92	42	86	63,94	32,35	43	86	5.5

Существование избыточности в зашифрованных текстах исследовалось с использованием программы -архиватора RAR. В качестве входных текстов использовались:

DLL - исполняемый файл WINDOWS;

DOC - документы Microsoft WORD 6.0;

EXE - исполняемый файл MS-DOS;

PDF - документ переносимого формата;

TXT - обычный текст ASCII;

ZIP - файл - архив.

В качестве показателей сжатия при оценке кандидатов в AES использовались:

$$k_1 = \frac{l'_m}{l_m} \tag{3}$$

и

$$k_2 = \frac{l'_c}{l_m}, \tag{4}$$

где

l'_m - длина сжатого открытого текста;

l_m - длина входного открытого текста;

l'_c - длина сжатой криптограммы.

В таблице 4 в качестве примеров приведены значения k_1 и k_2 для RC6 и Magenta. Длина входного открытого текста $l_m=10^6$ байт. Шифрование осуществлялось в блочном режиме.

Таблица 4

Тип AES	Тип файла	l'_m	l'_c	k_1	k_2
RC6	DLL	463802	959901	0.464	0.960
	DOC	187578	608472	0.188	0.608
	EXE	387178	859101	0.387	0.859
	PDF	891286	961224	0.891	0.961
	TXT	290310	948754	0.290	0.949
	ZIP	1000154	1000060	1.000154	1.000060
Magenta	DLL	463802	916490	0.464	0.916
	DOC	187578	533263	0.188	0.533
	EXE	387178	809533	0.387	0.809
	PDF	891286	959658	0.891	0.960
	TXT	290310	880921	0.290	0.881
	ZIP	1000154	1000060	1.000154	1.000060

Проведенные исследования по определению k_2 для всех AES кандидатов и приведенные в таблице 4 данные для RC6 и Magenta свидетельствуют о том, что алгоритм Magenta хуже чем остальные алгоритмы. Худшим по k_2 является также и алгоритм DEAL..

Все AES кандидаты исследовались на способность сжатия криптограмм, полученных в поточном режиме. Сжатию такие криптограммы не поддаются. Результаты по k_2 близки по значениям к результатам, полученным для ZIP-файлов.

Таким образом, исследование коэффициента сжатия k_2 для AES кандидатов позволяет сделать вывод о том, что алгоритмы Magenta и DEAL относятся к «плохим» и могут быть отклонены.

6. Вычислительная сложность развертывания ключа и криптографических преобразований

В большинстве AES криптоалгоритмов рабочие ключи K_p, то есть ключи, которые используются для прямых и обратных криптографических преобразований, создаются на основе развертывания начальных $K_п$ ключей, причем

$$K_p = F_p\left(K_n, I_{AES}\right),$$

где

F_p - является функцией развертывания ключа;

$K_п$ - значение рабочего ключа;

I_{AES} - вектор - параметр функции развертывания ключа.

Анализ показал, что для большинства AES кандидатов скорость (сложность) развертывания ключа и прямых и обратных преобразований не зависит от длины ключа. Это означает, что время развертывания ключа и шифрования/дешифрования блока данных является одинаковым, независимо от длины ключа – 128, 192 или 256 бит. Девять криптоалгоритмов – RC6, MARS, CRYPTON, CAST, E2, SERPENT, HPC, DFC, FROG – являются совсем не зависящими от длины ключа. Сложность шифрования/дешифрования для алгоритмов LOKI и Twofish не зависит от длины ключа, а сложность развертывания ключа зависит от длины ключа. Причем сложность развертывания ключа для Twofish возрастает при увеличении длины ключа, а для LOKI97 наоборот - уменьшается. В алгоритмах DEAL, Magenta, Rijndael и SAFER+ сложность (скорость) шифрования/дешифрования зависит от длины ключа, это связано с разным числом циклов, выполняемых при криптографических преобразованиях.

Проведенные измерения показали, что алгоритмы DEAL и Magenta при длине ключа 128 и 192 бита выполняют криптографические преобразования на 0.33 (а Rijndael – на 0.4) медленнее, чем при

длине 256 битов. Алгоритм SAFER в два раза медленнее при длине ключа 256 битов и на 0.5 при длине ключа 192 бита.

В таблице 5 приведены результаты оценки сложности разворачивания ключа и шифрования/дешифрования в зависимости от длины ключа. Сложность оценивается количеством тактов работы компьютера.

Таблица 5

Крипто-алго-ритм	Развертывание ключа, тактов			Шифрование, тактов			Дешифрование, тактов			Ср. ско-рость,	Мин. число
	128	192	256	128	192	256	128	192	256	М бит/с.	циклов
RC6	1682	1878	1778	270	269	269	231	230	230	102.2	20
MARS	4311	4375	4344	376	376	376	370	370	370	68.6	10
Twofish	14707	20868	26438	381	381	379	379	379	384	67.4	12
Rijndael	2066	2418	2937	436	510	592	429	498	579	59.2	8
Crypton	1269	1291	1343	476	477	473	476	479	479	53.8	11?
CAST	4344	4330	4313	632	632	632	632	632	632	40.5	10
E2	9508	9531	9596	691	689	702	701	698	701	36.8	10
Serpent	2404	2365	2350	950	950	951	895	895	895	27.8	17
HPC	120631	120695	120658	1418	1418	1418	1607	1607	1591	16.9	8?
DFC	7126	6845	6929	1699	1643	1693	1668	1616	1644	15.2	9
SAFER	4278	7426	11310	1720	2553	3391	1721	2536	3374	14.9	7
LOKI	7407	7288	7147	2141	2133	2133	2187	2177	2177	11.8	>36
DEAL	8632	8647	11724	2362	2354	3093	2371	2370	3110	10.8	6
FROG	13802	14155	14107	2521	2515	2392	2282	2190	2288	10.7	8?
Magenta	31	26	36	6552	6535	8703	6550	6549	8715	3.9	>10

Данные, приведенные в таблице 5, подтверждают вывод, что первые семь криптоалгоритмов являются лучшими кандидатами в AES стандарт по сложности (скорости) прямых и обратных криптографических преобразований. Граничной мы выбрали скорость 30 Мбит/с.

Согласно требованиям к криптоалгоритму, он должен применяться и в качестве функции хеширования. Поэтому для оценки эффективности криптоалгоритма в целом необходимо давать оценки сложности хеширования, при условии, что выработка MAC (КАС) осуществляется с использованием соответствующего алгоритма. В таблице 6 приведены соответствующие данные в виде числа тактов (циклов), которые необходимо выполнить для развертывания ключа и выполнения прямого и обратного преобразований.

Таблица 6

Название криптоалгоритма	RC6	MARS	Twofish	Rijndael	CRYPTON	E2	SERPENT
Pentium	146	260	175	34	49	100	205
Pentium PRO	118	246	132	32	46	100	193

Результаты, приведенные в таблице 6, позволяют сделать вывод, что лучшими являются алгоритмы Rijndael и CRYPTON, удовлетворительными E2, RC6 и Twofish.

По сравнению со стандартными функциями хеширования, например SHA, все AES имеют большую сложность. Например SHA требует всего 13 тактов на байт.

7. Некоторые итоги второй конференции

В августе 1999 года состоялся второй раунд обсуждения кандидатов в стандарт 21 века. По результатам работы конференции принято решение оставить 5 кандидатов в AES из 15. Это криптоалгоритмы RC6, MARS, Twofish, Rijndael и Serpent. Как следует из таблицы 1 первые четыре криптоалгоритма, выделенные нами, совпали с выделенными на конференции. По результатам работы конференции на пятое место вышел алгоритм Serpent. По нашей классификации он находится на 8 месте. Дело в том, что в процессе обсуждения, разработчики изменили параметры алгоритма, уменьшив число циклов в 2 раза, что позволило повысить производительность (скорость) шифрования/дешифрования. Благодаря этому алгоритм переместился на пятое место и вошел в число претендентов в стандарт.

На наш взгляд основная борьба среди кандидатов в AES развернется между RC6, MARS и Twofish. Дело в том, что в качестве дополнительных требований, предъявленных к криптоалгоритмам, должны быть требования случайности, равновероятности и независимости развернутых ключей. Им в большей мере отвечают RC6, MARS и Twofish.

Список литературы: 1. *Announcing* Development of Federal Information Processing Standard for Advanced Encryption Standard. – Federal Register. Vol. 62, № 1, 1997, pp.93-94. **2.** *Announcing* Request for Candidate Algorithm Nomination for Advanced Encryption Standard (AES). – Federal Register. Vol. 62, №177, 1997, pp.48051-48058. 3.http://www.nist.gov/aes

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 21.03.2000

УДК 681.3.06: 519.248.681

М.Ф. БОНДАРЕНКО, д-р. техн. наук, И.Д. ГОРБЕНКО, д-р. техн. наук,
Е.Г. КАЧКО, канд. техн. наук, А.В. СВИНАРЕВ, канд. техн. наук, Т.А. ГРИНЕНКО

СУЩНОСТЬ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ СВОЙСТВ ПЕРСПЕКТИВНЫХ СТАНДАРТОВ ЦИФРОВОЙ ПОДПИСИ X9.62-1998 И РАСПРЕДЕЛЕНИЯ КЛЮЧЕЙ X9.63-199X НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ

Введение

В 80-е годы были разработаны, а в 90-е нашли широкое распространение криптографические системы, построенные на использовании методов несимметричной криптографии. Основными из них являются методы, базирующиеся на использовании RSA, Эль-Гамаля и Диффи-Хеллмана криптоалгоритмов [1-4]. Практическое применение их было закреплено в стандартах, основными из них являются X-509, ISO-11166, X9.30, X9.42, ГОСТ 34.10-94, ГОСТ 34.310-95 и др. Эти системы были отнесены к классу вероятно-стойких или доказуемо-стойких, что объясняется тем, что доказательство их стойкости сводилось к доказательству сложности решения определенных математических задач при соответствующих значениях (размерах) общесистемных параметров. Так доказательство стойкости RSA систем сводилось в основном к доказательству сложности решения задач факторизации модуля преобразования N. Доказательство стойкости алгоритмов Эль-Гамаля и Диффи-Хелмана сводилось к доказательству сложности решения дискретного логарифмического уравнения. При этом по мере расширения применения указанных стандартов активизировались усилия по их взлому. Появились совершенно новые разделы математики, позволяющие существенно уменьшить вычислительную сложность решения указанных задач. Например, создание средств решения таких задач на основе общего решета числового поля в сочетании с применением мощных компьютеров сделало возможным взлом систем с параметрами, используемыми на практике. Иначе средства криптоанализа, в смысле математики и производительности криптоаналитических систем, развивались быстрее, чем изменялись версии средств цифровой подписи, направленного шифрования и распределение ключей.

Основным методом защиты стали изменения параметров, в смысле их увеличения, например, модулей преобразования. Так, в Эль-Гамаля и Диффи-Хеллмана системах длина модуля преобразования составляет порядка 1024 и более битов. Но при этом до такой же длины были увеличены длины ключей, как следствие увеличилась вычислительная сложность криптографических преобразований и уменьшилась скорость. В то же время все преобразования необходимо осуществлять все с возрастающими скоростями, как правило, в реальном масштабе времени. Разрешение указанного противоречия было найдено за счет реализации различных несимметричных преобразований на эллиптичных кривых в полях Галуа [4-38]. По существу в 90-е годы криптографы и криптоаналитики разрабатывали и исследовали стойкость криптоалгоритмов на эллиптичных кривых. К настоящему времени уже разработаны, прошли сертификацию и утверждены ряд стандартов. Прежде всего стандарт цифровой подписи X9.62-1998 [3] и стандарт распределения ключей X9.63-1999[4], а также черновые версии ИИЭР P1363/79 стандартных спецификаций для шифрования с открытыми ключами. Основными преимуществами этих стандартов является возможность уменьшения в 5 и более раз длин ключей и общесистемных параметров, большая степень увеличения сложности криптоанализа с ростом размеров общесистемных параметров, а также уменьшение вычислительной мощности всех преобразований. Все это, на наш взгляд, и предопределило переход, а по существу перевод существующих алгоритмов на вычисления на эллиптических кривых над полями Галуа.

1. Математические основы преобразований на эллиптических кривых в полях Галуа

Наиболее общее определение эллиптической кривой дает уравнение Вейерштрасса. Для конечного поля Галуа GF(q), где q>3 и есть простым числом, уравнение Вейерштрасса имеет вид

$$y^2 = x^3 + ax + b(\mathrm{mod}\, q), \tag{1}$$

где a и b есть целые числа над полем GF(q), но такие, что справедливо выражение

$$4a^3 + 27b^2 \neq 0(\mathrm{mod}\, q). \tag{2}$$

Для расширенного поля GF(2^m), уравнение Вейерштрасса имеет вид

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

где a и b являются элементами поля GF(2^m), т.е. полиномами степени m над полем GF(2). Все вычисления в (3) производятся по двойному модулю (f(x),2), где f(x) примитивный полином степени m. При чем в сравнении (3) $b \neq 0$.

Эллиптичная кривая E над конечным полем GF(q) определяется множеством точек на плоскости P=(x_p, y_p),где x_p и y_p являются элементами поля GF(q). Элементы поля a $\in$GF(q) и b $\in$ GF(q) называются коэффициентами эллиптической кривой E. Составляющие точки P называются x_p - координатой точки P и y - координатой точки P.

Основной характеристикой эллиптической кривой есть ее порядок #E. Под порядком эллиптической кривой понимается число различных точек на E, включая точку O, который обозначается как

$$n = \#\, E(GF(q)). \tag{4}$$

При этом под разными мы понимаем точки, которые отличаются хотя бы одной координатой.

На эллиптической кривой введены операции сложения и скалярного умножения.

Операции сложения обладают следующими свойствами.

1. Сложения с нулем P+0=0+P=P, для всех точек P є E(GF(q)).
2. Для каждой точки P=(x_1, y_1), P $\in$ E(GF(q)) существует точка Q=(x_1,-y_1), Q $\in$ E(GF(q)), такая что

P+Q=0.

Точка Q называется обратным элементом и обозначается как (-P).

3.Если P=(x_1,y_1) $\in$ E(GF(q)), то

(x_1,y_1)+ (x_1,-y_1)=0.

4.Операция сложения двух точек.

Если P=(x_1,y_1) $\in$ E(GF(q)) и Q=(x_2,y_2) $\in$ E(GF(q)) и P$\neq$ 0, то R=P+Q=(x_3,y_3).

При этом, если q есть простое число, то

$$x_3 = (\lambda^2 - x_1 - x_2)(\mathrm{mod}\, q) \tag{5}$$

$$y_3 = (\lambda(x_2 - x_3) - y_1)(\mathrm{mod}\, q), \tag{6}$$

где

$$\lambda = \begin{cases} \dfrac{y_1 - y_2}{x_1 - x_2}(\mathrm{mod}\, q), если P \neq Q (x_1 \neq x_2) \\ \dfrac{3x^2 + a}{2y_2}(\mathrm{mod}\, q), если P = Q (x_1 = x_2) \end{cases} \tag{7}$$

Если q=2^mто вместо 3 имеет место 3^1,а вместо 4 имеет место 4^1.

3^1.(x, y)+(x, x+y)=0 для всех (x, y) $\in$ E(GF(2^m)).

4^1.Если P=(x_1,y_1) $\in$ E(GF(2^m)) и Q=(x_2,y_2) E(GF(2^m)),

то

(x_3,y_3)=(x_1,y_1)+(x_2,y_2),

где

$$x_3 = (\lambda^2 + \lambda + x_1 + x_2 + a)(\mathrm{mod}\, f(x),2) \quad (8)$$

$$y_3 = (\lambda(x_1 + x_3) + x_3 + y_2)(\mathrm{mod}\, f(x),2) \quad (9)$$

$$\lambda = \frac{y_1 + y_2}{x_1 + x_2}(\mathrm{mod}\, f(x),2) \quad (10)$$

5.Скалярное умножение определяется для каждой точки E(GF(q)). Если точка P ∈ E(GF(q)), а ∈ N (а целое положительное), то скалярное умножение

$$a \times P = \underbrace{P + P + P + \ldots + P}_{a\ раз},$$

где операция (+) есть операция сложения на эллиптической кривой, определенная в 4 и 4[1].

2. Основные стандарты, применение, характеристика и возможности

Уже предварительный анализ соотношений (5)-(10) показывает, что выполнение операций сложения и скалярного умножения на эллиптической кривой требует значительных вычислительных ресурсов. При этом, очевидно, наибольшей вычислительной сложности требуют вычисления согласно выражению (7) и (9), решение этих задач ввиду значительной сложности, особой важности и необходимости требует отдельного рассмотрения.

К 1998 году разработаны и уже приняты в качестве рабочих версий следующие криптографические стандарты, которые построены на базе математического аппарата эллиптических кривых над простыми и расширенными полями Галуа (двоичными). Это цифровая подпись на эллиптических кривых X9.62-1998 и схемы управления ключами на эллиптических кривых X9.63-199X . По существу стандарт X9.62-1998 есть усовершенствование уже применяемого стандарта ANSI X9.30 часть 1, т.е. цифровой подписи DSA. Стандарт X9.63-199X есть усовершенствование применяемого стандарта ANSI X9.42-1996 управления ключами по схеме Диффи-Хеллмана.

Алгоритм DSA на эллиптических кривых (ECDSA) является аналогом DSA. В таблицах 1-8 устанавливаются соответствия по параметрам DSA и ECDSA.

В таблице 1 приведена информация о DSA и ECDSA.

Таблица 1

Группа	поле F_p*	$E(F_q)$
Группа элементов	Множество целых {1,2,...,p-1}	Точки (x,y), которые удовлетворяют уравнению ЭК, плюс точка бесконечности 0.
Операция в группе	Умножение по модулю p	Сложение точек на ЭК
Группа	поле F_p*	$E(F_q)$
Обозначения	Элементы: g, h Умножение: $g \times h$ Степень: g^a	Элементы: P, Q Сложение: $P + Q$ Умножение точек (так называемое скалярное умножение): aP
Проблема дискретного логарифма	Дано $g \in F_p^*$ и $h = g^x \bmod p$, Найти целое x.	Дано $P \in E(F_q)$ и $Q = aP$, найти целое a.

В таблице 2 указывается соответствие параметров DSA и ECDSA

Таблица 2

Характеристика	DSA Обозначения	ECDSA Обозначения
Порядок	q	n
Порождающий элемент	g	G
Личный ключ	x	d
Открытый ключ	y	Q

В таблице 3 перечислены параметры DSA и ECDSA

Таблица 3

Параметры	Параметры DSA	Параметры ECDSA
1. Общесистемные параметры	1. p и q - простые, q *делит* p-1.	1. E – ЭК над полем Fq.
2. Порождающий элемент	2. g – элемент порядка q в поле F_p*.	2. G точка порядка n в E(Fq).
3. Используемая группа	3. Используемая группа: $\{g^0, g^1, g^2, ..., g^{q-1}\}$.	3. Используемая группа: $\{0, G, 2G, ..., (n-1)G\}$.

В таблице 4 приведен алгоритм генерации ключей в DSA и ECDSA

Таблица 4

Параметр	Генерация ключей в DSA	Генерация ключей в ECDSA
1.Личный ключ x.	1. Выбрать случайное x в интервале [1, q-1].	1. Выбрать случайное целое d в интервале [1, n-1].
2.Открытый ключ.	2. Вычислить $y = g^x \bmod p$.	2. Вычислить $Q = dG$.
3.Ключи	3. Личный ключ x. Открытый ключ y.	3. Личный ключ d. Открытый ключ Q.

В таблице 5 приведен алгоритм выработки подписи в DSA и ECDSA

Таблица 5

DSA	ECDSA
1. Выбрать случайное целое k в интервале [1, q-1].	1. Выбрать случайное d в интервале [1, n-1].
2. Вычислить $g^k \bmod p$.	2. Вычислить $kG = (x_1, y_1)$.
3. Вычислить $r = (g^k \bmod p) \bmod q$.	3. Вычислить $r = x_1 \bmod n$.
4. Вычислить $e = H(M)$.	4. Вычислить $e = H(M)$.
5. Вычислить $s = k^{-1}(e + xr) \bmod q$.	5. Вычислить $s = k^{-1}(e+dr) \bmod n$.
6. Подпись для M - (r, s).	6. Подпись для M - (r, s).

В таблице 6 приведен алгоритм проверки подписи для DSA и ECDSA

Таблица 6

Проверка подписи для DSA	Проверка подписи для ECDSA
1. Вычислить h = $H(M)$.	1. Вычислить h= $H(M)$.
2. Вычислить $s^{-1} \bmod q$.	2. Вычислить $s^{-1} \bmod n$.
3. Вычислить $u_1 = hs^{-1} \bmod q$.	3. Вычислить $u_1 = hs^{-1} \bmod n$.
4. Вычислить $u_2 = rs^{-1} \bmod q$.	4. Вычислить $u_2 = rs^{-1} \bmod n$.
5. Вычислить $v' = g^{u1} y^{u2} \bmod p$.	5. Вычислить $u_1G + u_2Q = (x_1, y_1)$.

Продолжение таблицы 6

Проверка подписи для DSA	Проверка подписи для ECDSA
6. Вычислить $v = v' \bmod q$.	6. Вычислить $v = x_1 \bmod n$.
7. Принять подпись, если $v = r$.	7. Принять подпись, если $v = r$.

В таблице 7 переведена характеристика параметров X9.63-199X.

Таблица 7

Группа	*поле* F_p	$E(F_q)$
Группа элементов	Множество целых {1,2,...,p-1}	Точки (x,y), которые удовлетворяют уравнению ЭК
Операция в группе	Умножение по модулю p	Сложение точек
Обозначения	Элементы: g_1, g_2 Умножение: $g_1 \times g_2$ Степень: g^k	Элементы: P_1, P_2 Сложение: P_1+P_2 Умножение точек (так называемое скалярное умножение): kP
Проблема дискретного логарифма	Дано $g_1 \in F_p^*$ и $h = g_1^{\,k} \bmod p$, Найти целое k.	Дано $P_1 \in E(F_q)$ и $P_2 = kP_1$, найти целое k.
Диффи-Хеллмана проблема	Известны g^{k1}, $g^{k2} \in F_p^*$. Найти g^{k1k2}	Дано $P_1 \in E(F_q)$ и $P_2 = kP_1$, найти целое k_1k_2P.

В таблице 2 указывается соответствие параметров DSA и ECDSA

Таблица 8

X9.42	q	P	g	x	y	z	t
X9.63	n	#E(F_g)	G	d_s	Q_s	d_e	Q_e

3. Оценка криптостойкости

Пусть E - эллиптическая кривая над конечным полем Галуа F(q). Пусть G $\in$ E(F(q)) будет кратной порядку n, где n - простое число и $n \geq 2^{160}$.

Задачу дискретного логарифма эллиптической кривой (ДЛЭК) сформулируем таким образом: для заданных (известных E, G и открытого ключа Q $\in$ E (F(q)) необходимо найти целое число l, $0 \leq l \leq n-1$ такое, что

$$Q = \Omega = lG$$

при условии существования такого числа l.

Известно несколько методов (алгоритмов) решения дискретного логарифма эллиптической кривой. Лучшими на сегодняшний день алгоритмами являются методы семейства Полларда, в частности ρ-метод и λ-метод Полларда.

Показано, что ρ метод требует выполнения порядка

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

шагов, то есть обладает сложностью такого порядка. Каждый шаг, его сложность, является операцией суммирования на эллиптической кривой. Метод Полларда может быть распараллелен. В этом случае каждый из m-процессоров может выполнять часть I ρ операций (шагов), при m-процессорах.

$$\sqrt{\frac{\pi n}{2}} \Big/ m$$

Показано, что сложность λ-метода Полларда I_λ оценивается соотношением

$$I_\lambda = 2\sqrt{n}.$$

Он также может быть распараллелен. При m параллельных процессорах каждый из них должен выполнить

$$2\sqrt{n}/m$$

операции суммирования точки на эллиптической кривой.

Приведенные соотношения справедливы только для решения задач, исключая суперсингулярные и другие кривые. Запрет суперсингулярных кривых связан с тем, что существует метод эффективного сведения задачи дискретного логарифма на эллиптической кривой к задаче дискретного логарифма в конечном поле.

В 1998 г. было показано, что расчетная сложность лучших методов, например ρ метода может быть уменьшена в $\sqrt{2}$ раз. Для этого улучшенного метода ожидаемая сложность может оцениваться как

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

При распараллеливании имеем сложность на один из m процессоров

$$\sqrt{\pi n/4}/m.$$

Это касается эллиптических кривых над простым полем. Для расширенного поля $GF(2^l)$ на кривой порядка

$$E(F(2^{ld}))$$

может быть ускоренный в $\sqrt{2d}$ раз.

Пример. Пусть двоичная аномальная кривая E имеет вид

$$y^2 + xy = x^3 + x^2 + 1.$$

Для нее порядок кривой

$$\#E\left(F\left(2^{163}\right)\right) = 2n,$$

где n-простое 162 разрядное число.

Задача дискретного логарифма в $E(F(2^{163}))$ может быть решена с 2^{77} операций суммирования на эллиптической кривой. Для случайной кривой рассматриваемого вида оценивается величиной 2^{81} операций.

Для защиты от всех известных на сегодня вторжений необходимо, чтобы:

1. Порядок #E(F(q)) был кратным большому простому числу $n > 2^{160}$;
2. Выполнялось условие MOV.
3. Выполнялось условие аномальности.

MOV условие гарантирует, что эллиптическая кривая не поддается атакам с уменьшенной сложностью. Оно (условие) рассмотрено в [3].

Условие аномальности заключается в том, чтобы $\#E(F(q)) \neq q$, то есть порядок поля не должен совпадать с порядком кривой E.

4. Практические результаты криптоаналитических атак на эллиптические кривые

Очевидно все атаки на эллиптические кривые можно разделить на три вида: программные, аппаратные и программно-аппаратные.

Программные вторжения. Пусть одна MIPS машина выполняет $l = 4 \cdot 10^4$ сложений точек эллиптической кривой в секунду. Это достаточно высокий показатель. Так ASIC схема аппаратной реализации (прикладная специализированная интегральная схема) выполняет $4 \cdot 10^4$ операций на эллиптической кривой в поле $F(2^{155})$. При работе на частоте 40 Мгц в поле $F(2^{155})$ она выполняет 40 000 операций добавления точек на эллиптической кривой.

При таких условиях число операций добавления на эллиптической кривой 1 MIPS машиной определяется как

$$L = l \cdot t_{року} = \left(4 \cdot 10^4\right) \cdot (60 \cdot 60 \cdot 24 \cdot 365) = 1{,}15 \cdot 2^{40}.$$

В таблице 9 приведены необходимые значения мощности, которая необходима для вычисления одного дискретного логарифма для различных значений n. При этом считалось, что мощность криптоаналитической системы составляет $S = 4 \cdot 10^4$ добавлений на эллиптической кривой.

Таблица 9

Размер поля q (в разрядах)	Размер n эллиптической кривой (разрядов)	Значение $\sqrt{\frac{\pi n}{4}} = I_{\rho^0}$	L (MIPS-лет)
131	128	$1{,}64 \cdot 10^{19}$	$1{,}3 \cdot 10^7$
163	160	$1{,}07 \cdot 10^{24}$	$8{,}5 \cdot 10^{11}$
197	192	$7{,}05 \cdot 10^{28}$	$5{,}6 \cdot 10^{16}$
229	224	$4{,}62 \cdot 10^{33}$	$3{,}7 \cdot 10^{21}$
261	256	$3{,}03 \cdot 10^{38}$	$2{,}4 \cdot 10^{26}$
325	320	$1{,}30 \cdot 10^{48}$	$1{,}0 \cdot 10^{36}$
518	512	$1{,}03 \cdot 10^{77}$	$8{,}2 \cdot 10^{64}$
1032	1024	$1{,}19 \cdot 10^{144}$	$9{,}5 \cdot 10^{141}$

Значения L в таблице приведены для случая выполнения $4 \cdot 10^4$ операций сложения на кривой.

В [3] приведены данные, которые при использовании криптоаналитической системы из 10000 компьютеров, причем каждый из них имел мощность 1000 MIPS, то для $n = 2^{160}$ дискретный логарифм может быть вычислен за 85 тысяч лет.

По оценкам Одлизко [3], если для криптоанализа использовать 0,1% мировой мощности компьютеров, то в 2004 году можно выполнить на них 10^8 MIPS, а в 2014 году - $\left(10^{10} - 10^{11}\right)$ MIPS.

Отметим, что приведенные данные справедливы для случая использования методов Полларда. Однако, по-видимому, появятся новые математические методы решения задачи нахождения дискретного логарифма. Так уже было при решении задач факторизации RSA модулей и решения дискретного логарифма над простым полем. Тогда после освоения ρ и $(\rho - 1)$ методов Полларда появились методы кривых Ленстра, решето числового поля и общее решето числового поля, которые стали намного эффективнее по сравнению с методами Полларда.

Ван Ушрот и Вайнер [3] исследовали аппаратные вторжения и возможности построения специализированной криптоаналитической системы для решения дискретного логарифма на эллиптической кривой. Они получили интересный результат. Для $n = 10^{36} = 2^{120}$ система из 325000 компьютеров, цена которой составляет не меньше 10 миллионов долларов, нашла бы дискретный логарифм примерно за 35 дней.

5. Аспекты, связанные с нахождением ключа

В нашей постановке личным ключом является целое число l, $0 \le l \le n-1$ такое, что

$$Q = lG(\mathrm{mod}\, q)$$

Рассмотренные выше вторжения обеспечивают при рассмотренных условиях нахождение личного ключа l. При этом принимается, что системные параметры числовые значения точки G и модуля n известные криптоаналитику.

Мы здесь рассмотрим особенности λ метода Полларда. Пусть известна эллиптическая кривая E и базовая точка G. Пусть также время нахождения l, то есть решением дискретного логарифмического уравнения является t. Можно показать, что в этом случае ожидаемое время решения второго дискретного логарифмического уравнения (при тех же E, n и G) рассчитывается как

$$\left(\sqrt{2}-1\right)t = 0{,}41t.$$

Дальше решение третьего примера требует

$$\left(\sqrt{3}-\sqrt{2}\right)t \approx 0{,}32t.$$

Решение четвертого уравнения требует

$$\left(\sqrt{4}-\sqrt{3}\right)t = 0{,}27t$$

времени. И так далее.

Таким образом, при фиксированных значениях E, n, G решение следующих дискретных логарифмических уравнений все легче и легче.

Считается, что в ближайшее время число n (порядок используемой кривой) должно быть не меньше 150 битового числа для обеспечения краткосрочной защиты и не меньше чем 180 битовое число для среднесрочной защиты.

Необходимо отметить, что при использовании симметричных криптоалгоритмов длина ключа должна быть не меньше 75 бит. Надежная стойкость же может быть обеспечена при длине ключа не меньше 90 бит. Сегодня считается, что в 21 столетии длина симметричного ключа должна быть не меньше 128 бит.

Показано, что полный поиск K-битового ключа для симметричного криптоалгоритма примерно равняется сложности поиска согласно методам Полларда личного ключа на эллиптической кривой, порядок n которой равен 2k. В таблице 10 приведена расчетная сложность криптоанализа методов Полларда.

Таблица 10

n	$\sqrt{n}$	I_ρ (команд)	I_λ (команд)	I_{ρ_0} (команд)	I'_ρ (mipsолет)	I'_λ (mipsолет)	I'_{ρ_0} (mipsолет)
2^{128}	$2{,}42\cdot10^{19}(2^{64})$	$3{,}03\cdot10^{19}$	$4{,}84\cdot10^{19}$	$2{,}15\cdot10^{19}$	$0{,}97\cdot10^{6}$	$1{,}55\cdot10^{6}$	$0{,}69\cdot10^{6}$
2^{160}	$1{,}09\cdot10^{24}(2^{80})$	$1{,}36\cdot10^{24}$	$2{,}18\cdot10^{24}$	$0{,}97\cdot10^{24}$	$0{,}44\cdot10^{11}$	$0{,}69\cdot10^{11}$	$0{,}31\cdot10^{11}$
2^{198}	$7{,}76\cdot10^{28}(2^{96})$	$9{,}70\cdot10^{28}$	$1{,}55\cdot10^{29}$	$6{,}9\cdot10^{28}$	$3{,}10\cdot10^{15}$	$0{,}47\cdot10^{15}$	$2{,}21\cdot10^{15}$
2^{224}	$5{,}13\cdot10^{33}(2^{112})$	$6{,}41\cdot10^{33}$	$1{,}03\cdot10^{34}$	$4{,}56\cdot10^{33}$	$2{,}05\cdot10^{20}$	$0{,}33\cdot10^{21}$	$1{,}46\cdot10^{20}$
2^{256}	$3{,}39\cdot10^{38}(2^{128})$	$4{,}23\cdot10^{38}$	$6{,}78\cdot10^{38}$	$3{,}02\cdot10^{38}$	$1{,}35\cdot10^{25}$	$2{,}17\cdot10^{25}$	$0{,}97\cdot10^{25}$
2^{288}	$2{,}19\cdot10^{43}(2^{144})$	$2{,}74\cdot10^{43}$	$4{,}38\cdot10^{43}$	$1{,}95\cdot10^{43}$	$0{,}88\cdot10^{30}$	$1{,}470\cdot10^{30}$	$0{,}62\cdot10^{30}$
2^{320}	$1{,}44\cdot10^{48}(2^{160})$	$1{,}80\cdot10^{48}$	$2{,}88\cdot10^{48}$	$1{,}28\cdot10^{48}$	$0{,}57\cdot10^{35}$	$0{,}74\cdot10^{35}$	$0{,}41\cdot10^{35}$
2^{352}	$9{,}33\cdot10^{52}(2^{176})$	$1{,}17\cdot10^{53}$	$1{,}87\cdot10^{53}$	$8{,}30\cdot10^{52}$	$0{,}37\cdot10^{40}$	$0{,}60\cdot10^{40}$	$2{,}66\cdot10^{40}$
2^{384}	$6{,}17\cdot10^{57}(2^{192})$	$7{,}71\cdot10^{57}$	$1{,}22\cdot10^{58}$	$5{,}49\cdot10^{57}$	$2{,}46\cdot10^{44}$	$0{,}39\cdot10^{45}$	$1{,}75\cdot10^{44}$
2^{416}	$4{,}07\cdot10^{62}(2^{208})$	$5{,}09\cdot10^{62}$	$8{,}14\cdot10^{62}$	$3{,}62\cdot10^{62}$	$1{,}63\cdot10^{49}$	$2{,}6\cdot10^{49}$	$1{,}16\cdot10^{49}$

Продолжение таблицы 10

n	$\sqrt{n}$	I_ρ (команд)	I_λ (команд)	I_{ρ_0} (команд)	I_ρ' (mipsолет)	I_λ' (mipsолет)	I_{ρ_0}' (mipsолет)
2^{448}	$2{,}63\cdot10^{67}(2^{224})$	$3{,}29\cdot10^{67}$	$5{,}26\cdot10^{67}$	$2{,}34\cdot10^{67}$	$1{,}05\cdot10^{54}$	$1{,}68\cdot10^{54}$	$7{,}49\cdot10^{53}$
2^{512}	$1{,}15\cdot10^{77}(2^{256})$	$1{,}43\cdot10^{77}$	$2{,}30\cdot10^{77}$	$1{,}02\cdot10^{77}$	$0{,}46\cdot10^{64}$	$0{,}74\cdot10^{64}$	$0{,}33\cdot10^{64}$
2^{768}	$3{,}80\cdot10^{115}(2^{384})$	$4{,}75\cdot10^{115}$	$7{,}6\cdot10^{115}$	$3{,}38\cdot10^{115}$	$1{,}21\cdot10^{102}$	$25{,}43\cdot10^{102}$	$1{,}08\cdot10^{102}$
2^{1024}	$1{,}29\cdot10^{154}(2^{512})$	$1{,}61\cdot10^{154}$	$2{,}58\cdot10^{154}$	$1{,}15\cdot10^{154}$	$0{,}52\cdot10^{141}$	$0{,}83\cdot10^{141}$	$0{,}39\cdot10^{141}$

Заключение

Следует предположить, что в ближайшие несколько десятилетий получат распространение криптографические алгоритмы и протоколы, в основу построения которых будет положена математика эллиптических кривых в полях Галуа. Основными направлениями деятельности и исследований в Украине в этом направлении являются освоение и анализ существующих стандартов, реализация их на программной, программно-аппаратной или аппаратной основе, а также разработка или доработка стандартов в интересах Украины. Первоочередной задачей, на наш взгляд, является перевод Гост 34.310-95 на эллиптические кривые. Вместе с тем, следует ожидать, что алгоритмы, базирующиеся на эллиптических кривых, пройдут тот же эволюционный путь развития, применения и разочарований, которые мы видели в отношении RSA, Ель-Гамаля, Диффи-Хеллмана и других криптоалгоритмов.

Список литературы: **1.** *ANSI X9.30-1995*, Part 1: Public key cryptography using irreversible algorithms for the financial services industry: The Digital Signature Algorithm (Revised). **2.** *ANSI X9.30-1993*, Part 2: Public key cryptography using irreversible algorithms for the financial services industry: The Secure Hash Algorithm 1 (SHA-1) (Revised). **3.** *ANSI X9.62-1998:* Certificate Management. **4.** *ANSI X9.63-199x*: Elliptic curve key agreement and transport protocols, draft. **5.** *ANSI NWI:* Prime number generation, draft. **6.** *G. AGNEW, T. BETH, R. MULLIN AND S. VANSTONE,* Arithmetic operations in GF(2^m), Journal of Cryptology, 6 (1993), 3-13. **7.** *G. AGNEW, R. MULLIN AND S. VANSTONE,* An implementation of elliptic curve cryptosystems over F_{2155}, IEEE Journal on Selected Areas in Communications, 11 (1993), 804-813. **8.** *G. AGNEW, R. MULLIN, I. ONYSZCHUK AND S. VANSTONE,* An implementation for a fast public-key cryptosystem, Journal of Cryptology, 3 (1991), 63-79. **9.** *M. BLAZE, W. DIFFIE, R. RIVEST, B. SCHNEIER, T. SHIMOMURA, E. THOMPSON, AND M. WIENER,* Minimal key lengths for symmetric ciphers to provide adequate commercial security, January 1996. **10.** *E. BRICKELL, D. GORDON, K. MCCURLEY AND D. WILSON,* Fast Exponentiation with precomputation, Advances in Cryptology - EUROCRYPT '92 Lecture Notes in Computer Science, 658 (1993), Springer-Verlag, 200-207. **11.** *T. ELGAMAL,* A public key cryptosystem and a signature scheme based on discrete logarithms, IEEE Transactions on Information Theory, 31 (1985), 469-472. **12.** *R. GALLANT, R. LAMBERT, AND S. VANSTONE,* Improving the parallelized Pollard lambda search on binary anomalous curves, to appear in Mathematics of Computation. **13.** *ITU-T* Recommendation X.680, Information Technology - Abstract Syntax Notation One (ASN.1): Specification of Basic Notation (equivalent to ISO/IEC 8824-1). **14.** *ITU-T* Recommendation X.681, Information Technology - Abstract Syntax Notation One (ASN.1): Information Object Specification (equivalent to ISO/IEC 8824-2). **15.** *TU-T* Recommendation X.682, Information Technology - Abstract Syntax Notation One (ASN.1): Constraint Specification (equivalent to ISO/IEC 8824-3). **16.** *ITU-T* Recommendation X.683, Information Technology – Abstract Syntax Notation One (ASN.1): Parametrization of ASN.1 Specifications (equivalent to ISO/IEC 8824-4). **17.** *ITU-T* Recommendation X.690, Information Technology - ASN.1 Encoding Rules. Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER) and Distinguished Encoding Rules (DER) (equivalent to ISO/IEC 8825-1). **18.** *ITU-T* Recommendation X.691, Information Technology - ASN.1 Encoding Rules: Specification of Packed Encoding Rules (PER)

(equivalent to ISO/IEC 8825-2). **19.** *D. JUNGNICKEL,* Finite Fields: Structure and Arithmetics, B.I.-Wissenschaftsverlag, Mannheim, 1993. **20.** *D. KNUTH,* The Art of Computer Programming, volume 2, 2nd edition, 1981. **21.** *N. KOBLITZ,* Elliptic curve cryptosystems, Mathematics of Computation, 48 (1987), 203-209. **22.** *R. LERCIER,* Finding good random elliptic curves for cryptosystems defined over F_{2^n}, Advances in Cryptography - EUROCRYPT '97, Lecture Notes in Computer Science, 1233 (1997), Springer-Verlag, 379-392. **23.** *R. LERCIER AND F. MORAIN,* Counting the number of points on elliptic curves over finite fields: strategies and performances, Advances in Cryptology - EUROCRYPT '95, Lecture Notes in Computer Science, 921 (1995), Springer-Verlag, 79-94. **24.** *R. LIDL AND H. NIEDERREITER,* Finite Fields, Cambridge University Press, 1987. **25.** *R.J. MCELIECE,* Finite Fields for Computer Scientists and Engineers, Kluwer Academic Publishers, 1987. **26.** *A. MENEZES,* Elliptic Curve Public Key Cryptosystems, Kluwer Academic Publishers, 1993. **27.** *A. MENEZES, T. OKAMOTO AND S. VANSTONE,* Reducing elliptic curve logarithms to logarithms in a finite field, IEEE Transactions on Information Theory, 39 (1993), 1639-1646. **28.** *V MILLER,* Uses of elliptic curves in cryptography, Advances in Cryptology - CRYPTO '85, Lecture Notes in Computer Science, 218 (1986), Springer-Verlag, 417-426. **29.** *NATIONAL* INSTITUTE OF STANDARDS AND TECHNOLOGY, Secure Hash Standard (SHS), FIPS Publication 180, May 1993. **30.** *NATIONAL* INSTITUTE OF STANDARDS AND TECHNOLOGY, Digital Signature Standard, FIPS Publication 186, 1993. **31.** *A. ODLYZKO,* The Future of Integer Factorization, Cryptobytes, volume 1, number 2, summer 1995, 5-12. **32.** *P. VAN OORSCHOT AND M. WIENER,* Parallel Collision Search With Application To Hash Functions And Discrete Logarithms, in Proceedings of the 2nd ACM Conference on Computer and Communications Security, Fairfax, Virginia, November 2-4, 1994, 210-218. **33.** *J. POLLARD,* Monte Carlo methods for index computation mod p, Mathematics of Computation, 32 (1978), 918-924. **34.** *R. SCHOOF,* Elliptic curves over finite fields and the computation of square roots mod p, Mathematics of Computation, 44 (1985), 483-494. **35.** *T. SATOH AND K. ARAKI,* Fermat quotients and the polynomial time discrete log algorithm for anomalous elliptic curves, preprint, 1997. **36.** *N. SMART,* The discrete logarithm problem on elliptic curves of trace one, to appear in Journal of Cryptology. **37.** *S. VAUDENAY,* Hidden collisions on DSS, Advances in Cryptology - CRYPTO '96, Lecture Notes in Computer Science, 1109 (1996), Springer-Verlag, 83-88. **38.** *M. WIENER AND R. ZUCCHERATO,* Fast attacks on elliptic curve cryptosystems, to appear in Fifth Annual Workshop on Selected Areas in Cryptography – SAC '98, Lecture Notes in Computer Science, Springer-Verlag.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06: 519.248.681

И.Д. ГОРБЕНКО, д-р. техн. наук, А.В. ПОТИЙ, канд. техн. наук, П.И. Терещенко

КРИТЕРИИ И МЕТОДОЛОГИЯ ОЦЕНКИ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Введение

Обоснованию критериев и созданию методологии оценки информационной безопасности уделено значительное внимание. В настоящее время можно выделить следующие документы, которые внесли серьезный теоретический и практический вклад в решение задач обеспечения информационной безопасности.

1. Критерии оценки защищенности компьютерных систем [1], которые известны как «Оранжевая книга».
2. Европейские критерии оценки безопасности информационных технологий [2]. Данные критерии разработаны с учетом выявленных недостатков и ограничений по применению «Оранжевой книги» и являются гармонизированными по отношению к первым.
3. Канадские критерии оценки безопасности надежных компьютерных систем [3].
4. Федеральные критерии США [4], разработанные по заказу правительства США и направленные на устранение ограничений, неудобств практического применения и недостатков «Оранжевой книги»
5. Международный стандарт ISO/IEC 15408 – «Критерии оценки безопасности информационных технологий» [5-7], или Единые критерии
6. Рабочий проект стандарт CEM-97/017 – «Общая методология оценки безопасности информационных технологий» [8].

Перечисленные нормативные документы, и особенно последние два, вносят существенный вклад в формирование единой международной научно-методологической базы решения проблемы обеспечения информационной безопасности в продуктах и различных информационных технологиях. Анализ этих документов подтверждает тот факт, что для решения задач обеспечения информационной безопасности, наряду с формальными методами моделирования процессов и оценки эффективности функционирования систем необходимо широко использовать методы декомпозиции и структуризации компонентов систем и процессов, неформальные методы оценки эффективности функционирования и принятия решений. Это означает, что аппарат системного анализа необходимо использовать на всех этапах жизненного цикла систем защиты информации.

Целью настоящей статьи есть рассмотрение и обсуждение основных положений Канадских критерии оценки безопасности, Единых критериев и Общей методологии оценки безопасности информационных технологий.

Первая часть статьи посвящена рассмотрению основных положений Канадских критериев, вторая часть – рассмотрению основных положений Федеральных критериев. Третья часть – рассмотрению основных положений стандарта ISO/IEC 15408-1 – Критерии оценки безопасности информационных технологий – Часть 1. Общая модель [5]. В четвертой части обсуждается версия документа CEM–97/017 - 1 – Общая методология оценки безопасности информационных технологий – Часть 1. Введение и общая модель [8].

В пятой части делается попытка оценить эти документы с позиций требований системного подхода к решению проблем и определению дальнейших перспектив развития методов и средств обеспечения информационной безопасности.

1. Канадские критерии оценки безопасности надежных компьютерных систем

В Канадских критериях, как одно из основных, введено такое понятие как гарантия. Она представляет собой степень доверия, с которой в системе реализована политика безопасности. Политика безопасности [3] представляет собой набор правил, регулирующих использование информации, включая ее обработку, хранение, распределение и представление в продукте или системе. Гарантии должны обеспечиваться на всех этапах жизненного цикла информационных продуктов. Каждый оцениваемый продукт должен иметь определенный требуемый уровень гарантий. Уровни гарантий организованы в ие-

рархическую систему и отражают доверие к тому, что политика безопасности продукта или системы реализована корректно.

В канадских критериях требования к гарантиям отделены от требований к функциональности. В них принято жесткое ограничение, что политика безопасности не зависит от функциональных возможностей. Под функциональностью в Канадских критериях понимается группирование услуг безопасности в соответствии с различными задачами безопасности, на решение которых и направлены эти услуги. Для этого используется понятие класса. Внутри каждого класса услуги ранжируются в соответствии с реальной стойкостью, функциональными возможностями и избирательностью действий. При этом реализация каждой из услуг обеспечивает защиту от угроз определенного класса.

При разработке Канадских критериев в основу были положены следующие принципы:

- существенная независимость от политики безопасности;
- обязательное измеримое отличие между уровнями услуг;
- безусловность наличия полезности для заказчика и гибкость документа.

Первый принцип требует, чтобы все аспекты проблем безопасности не были привязаны к какой-либо одной политике безопасности. Второй – возможности измерения разности услуг в части стойкости, функциональных возможностей и избирательности действия. Третий принцип требует, чтобы каждая услуга противостояла конкретным существующим или потенциальным угрозам, которые могут возникнуть при эксплуатации компьютерных систем.

Канадские критерии разрабатывались для технологий, в которых основными являются такие услуги как конфиденциальность, целостность, доступность и наблюдаемость.

Конфиденциальность есть свойство, которое гарантирует, что информация не может быть доступна или раскрыта, для неавторизованных (неуполномоченных на то) лиц, объектов или процессов. По существу, угрозами нарушения конфиденциальности являются такие угрозы, которые могут привести или приводят к несанкционированному ознакомлению с защищаемой информацией.

Целостность представляет свойство, которое обеспечивает условия ведения информационных отношений между субъектами и объектами, при которых информация сохраняется для использования и выполняет основные функции по назначению. Угрозы, относящиеся к несанкционированной модификации информации, являются угрозами нарушения целостности. В результате успешной реализации угрозы нарушения целостности объектам и субъектам наносится или может быть нанесен недопустимый ущерб.

Доступность представляет собой услугу по своевременному и качественному доступу к информации и ресурсам информационных технологий систем санкционированных объектов и субъектов. Как одна из услуг обеспечения безопасности она потенциально подвержена атакам, направленным на то, чтобы сделать ресурсы или информацию, а также услуги информационной системы неудовлетворительными или с пониженным качеством. Такие атаки наносят или могут наносить недопустимый ущерб.

Наблюдаемость (управление доступом) заключается в обеспечении возможности доступа к информации и/или ресурсам (системе) только объектам и субъектам, обладающим соответствующими полномочиями, или отслеживании их действий внутри системы. К угрозам нарушения наблюдаемости относятся угрозы, которые приводят к ухудшению управления и контроля доступом, манипулированию системой, ресурсами или информацией. Для управления доступом используется термин тЭг, который обозначает произвольную информацию, которая используется для управления доступом и связана с пользователями, процессами или объектами. Рассмотрим основные критерии более подробно.

В Канадских критериях каждая из услуг - конфиденциальность, целостность, доступность и наблюдаемость разбивается на уровни. Каждый уровень услуги представляет собой определенный перечень требований к избирательности или качеству защиты от специфического набора угроз. При этом с ростом уровня услуги должна предоставляться более надежная защита от соответствующих угроз. Уровни начинаются с нуля (0) и возрастают до «*n*», причем *n* уникально для каждой услуги.

Канадские критерии позволяют поставщику и заказчику точно определить набор услуг, которые требуются в системе (продукте). Для этого предусмотрена возможность создания функциональных профилей безопасности. Профиль представляет собой объединение (набор) услуг, как правило, совместно с описанием Политики безопасности. Профилю присваивается имя и численный идентификатор. Разработка и использование Канадских критериев было существенным шагом в решении проблем информационной безопасности, однако они имели ряд ограничений, особенно с появлением новых корпоративных и глобальных систем и сетей.

2. Федеральные критерии

Дальнейшее развитие американских, европейских и канадских критериев было заложено в Федеральных критериях [4]. Их особенностью является достаточная общность, совместимость с ранее использованными стандартами, соответствие требованиям. Критерии могут быть использованы для оценки различных информационных технологий (ИТ) - от баз данных до операционных систем. Применение критериев дает конкретные и точные рейтинги. Критерии разрабатывались агентством национальной безопасности США, они могут применяться как к коммерческим так и к военным ИТ. При разработке федеральных критериев за основу были приняты Канадские критерии.

В федеральных критериях, в отличие от канадских критериев, в которых избирательность услуг является «атомарной», компоненты управления доступом представляют собой определенную комбинацию услуг, а именно:

- доверительная конфиденциальность;
- административная конфиденциальность;
- доверительная целостность;
- административная целостность;
- повторное использование объекта.

В эту группу входит также услуга типа «откат», т.е. способность эффективно отменять определенные действия или группу действий.

Основной проблемой, которая возникла при введении Федеральных критериев, является совместимость функциональных услуг. В целом совместимость Канадских и Федеральных критериев составляет порядка 75%. Основным недостатком Федеральных критериев является их объемность. Как подчеркивают пользователи, федеральные критерии оказались очень объемными, а также сложными в применении. Кроме того, Федеральные критерии не предоставляют фиксированного набора уровней гарантий, на соответствие которым могут оцениваться продукты. Вместо набора уровней в Федеральных критериях введены наборы компонентов гарантий оценки и гарантий разработки. Компоненты объединяются и с учетом возможностей анализа взаимных зависимостей, вместе образуют уровни доверия. Более того, этот подход распространяется и на гарантии. Причем гарантии разделены на компоненты гарантий разработки (они касаются исключительно разработчика или поставщика), и компоненты гарантий оценки (например, сертификации). Это достаточно сильное решение и оно взято за основу в Единых критериях.

Для совместимости с Оранжевой книгой в Федеральных критериях введено понятие профиля защиты. Профиль защиты характеризуется тремя наборами компонентов: функциональный, гарантий разработки и гарантий оценки. Профиль защиты только тогда принимается, когда результаты анализа Политики безопасности и профиля защиты непротиворечивы. Освидетельствование состоит из этапов анализа и регистрации. Для создания профиля защиты необходим большой круг квалифицированных специалистов, соответствующие методики и значительные временные затраты на проведение анализа. Решить указанные задачи можно только при фиксации соответствующего набора ограничений, иначе анализ зависимостей может стать совершенно субъективным.

После появления в Федеральных критериях концепции профиля защиты было высказано много сомнений. Высказывались даже мнения, что она будет причиной «поражения» Федеральных критериев. Однако, как мы увидим ниже, эта концепция нашла свое развитие в Единых критериях оценки информационной безопасности.

3. Единые критерии оценки безопасности информационных технологий, ISO/IEC 15408

Стандарт ISO/IEC 15408 прошел достаточно долгий эволюционный путь развития. При его разработке учитывались положения таких документов как «Критерии оценки надежных компьютерных систем» (TCSEC) [1] (США, 1985), «Критерии оценки безопасности информационных технологий» (ITSEC) [2] (Европейская комиссия, 1991), «Канадские критерии оценки безопасности надежных компьютерных продуктов» (CTCPCEC) [3] (Канада, 1993), «Федеральные критерии безопасности информационных технологий» (FC) [4] (США, 1993). Также учитывались положения международных стандартов в области защиты информации, например ISO-7498-2, и ряда других документов [9]. Единые критерии (ЕК) информационной безопасности хорошо согласованы с существующими стандартами,

развивают и совершенствуют их путем внедрения новых компонент обеспечения безопасности для перекрытия большего числа угроз, в том числе в новых информационных технологиях.

В разработке стандарта принимали участие специалисты различных организаций, а именно Communications Security Establishment (Канада), Bundesamt fur Sicherheit in der Informationstechnik (BSI, Германия), German Information Security Agency (GISA), Service Central de la Securite des Systemes d'Information (SCSSI, Франция), Centre de Certification de la Securite des Technologies de l'Information (Франция), Netherlands National Communications Security Agency (Нидерланды), Communicstions-Electronics Security Group (Великобритания), National Security Agency, National Institute of Standards and Technology (США).

По мнению специалистов [10] ISO/IEC 15408 или как исторически сложилось называть этот документ Единые критерии, вобрал в себя все лучшее на сегодняшний день в области решения проблемы защиты информации и по всем показателям (актуальность, гибкость, реализуемость, универсальность, гарантированность) существенно превосходит все выше перечисленные документы. На данный момент этот документ представляет собой великолепный образец применения методов системного подхода к решению проблемы защиты информации и полностью соответствует принципу комплексной стандартизации в области обеспечения безопасности информации. Положительной чертой стандарта является то, что он разработан с учетом и использованием новейших достижений в области безопасности информационных технологий 90-х годов. В нем в полной мере учтены результаты анализа и применения всех существующих стандартов.

Стандарт определяет общие критерии, которые используются в качестве основы для оценки свойств безопасности информационных продуктов и технологий. При этом под продуктами и системами информационных технологий понимаются совокупности аппаратных и/или программных средств, которые представляют собой поставляемое конечному потребителю готовое к использованию средство обработки информации [10].

Единые критерии направлены на обеспечение сравнимости результатов оценок, полученных различными экспертами, путем введения общего множества требований к функциям безопасности продуктов и систем информационных технологий, а также к показателям этих функций. Используя стандарт, можно решить задачу выбора соответствующих требований и показателей безопасности информационных технологий.

Основными потенциальными угрозами безопасности и типовыми задачами защиты от них в Единых критериях приняты:

- защита от угроз целостности (несанкционированного изменения) информации;
- защита от угроз конфиденциальности (несанкционированного получения) информации по всем возможным каналам утечки;
- защита от угроз доступности информации, в смысле несанкционированного или случайного ограничения доступа к ресурсам и информации системы;
- защита от угроз аудиту системы (декларируется 12 потенциальных угроз).

Стандарт также может быть использован для решения других вопросов обеспечения безопасности информации, при этом особое внимание уделяется угрозам информации, порождаемым действиями человека.

Одним из основных понятий Единых критериев есть понятие компонента информационной безопасности. Компонентами Единых критериев являются:

- продукт информационных технологий;
- политика безопасности;
- потенциальные угрозы безопасности;
- типовые задачи защиты;
- профиль защиты;
- проект защиты;
- функциональные требования к средствам защиты;
- требования адекватности средств защиты;
- стандартные уровни адекватности средств защиты.

Политика безопасности определена как совокупность законов, норм и правил, регламентирующих порядок обработки, защиты и распространения информации.

Задача защиты – потребность потребителя продуктов информационных технологий в противостоянии множеству угроз безопасности или в необходимости реализации политики безопасности.

Профиль защиты – совокупность задач защиты, функциональных требований, требований адекватности и их обоснования. Оформляется в виде специального нормативного документа. Профиль защиты служит руководством для разработчика информационной технологии (ИТ- продукта), на основании которого и предложенных в нем технических рекомендаций разрабатывается проект защиты.

Проект защиты – совокупность задач защиты, функциональных требований, требований адекватности, общих спецификаций средств защиты и их обоснования. Проект защиты служит руководством для квалификационного анализа и сертификации ИТ- продукта.

Структура этих документов практически совпадает. Основными разделами профиля и проекта защиты являются:

1. Введение, которое содержит информацию, необходимую для идентификации проекта защиты, определения назначения и обзора его содержания. Во введении содержатся идентификатор проекта (профиля) - уникальное имя проекта защиты, используемое для поиска и идентификации проекта защиты и ИТ-продукта, обзор содержания, т.е. аннотация проекта защиты, на основании которой потребитель может определить пригодность ИТ- продукта для применения в своих целях, и заявка на соответствие требованиям CCITSE, в которой описываются все свойства ИТ- продукта, подлежащие квалификационному анализу по CCITSE.

2. Описание ИТ-продукта, которое содержит его краткую характеристику, назначение, принципы работы, методы исследований и др. Эта информация не подлежит анализу и сертификации, но представляется разработчикам и экспертам для пояснения и обоснования безопасности продукта и определения соответствия продукта задачам, решаемым с его использованием.

3. Среда эксплуатации. В данном разделе описываются характеристики среды эксплуатации ИТ-продукта, включая всевозможные угрозы.

4. Задачи защиты, решение которых позволит реализовать Политику безопасности.

5. Требования безопасности проекта защиты. Этот раздел содержит требования безопасности к ИТ-продукту, которыми руководствовался разработчик ИТ-продукта в ходе его разработки. Это позволяет декларировать разработчику факт успешного решения задач защиты. Раздел содержит типовые требования CCITSE и специфические требования для ИТ-продукта и среды его эксплуатации в формате CCITSE и требования адекватности, которые могут содержать уровни адекватности, а они содержат четкое, непротиворечивое описание уровней адекватности с соответствующей детализацией, в формате CCITSE.

6. Общие спецификации ИТ-продукта отражают вопросы реализации требований безопасности с использованием высокоуровневых спецификаций функций защиты, реализующих функциональные требования и требования адекватности CCITSE. Кроме того, в данном разделе содержатся:

– описание функциональных возможностей средств защиты ИТ-продукта, заявленных его разработчиком посредством декларирования требований безопасности. Спецификации должны позволять установить соответствия между требованиями защиты и функциями защиты;

– спецификация уровня адекватности, содержащая заявленный уровень адекватности защиты ИТ-продукта и его соответствие требованиям адекватности посредством представления параметров технологии проектирования и создания ИТ- продукта. Параметры должны быть представлены в формате, позволяющем определить их соответствие стандартным требованиям адекватности по CCITSE.

7. В проекте защиты содержится заявка на соответствие профилю защиты по одному или нескольким уровням. В данном разделе содержатся:

– ссылки на профиль защиты, на который претендует проект безопасности, а также случаи, в которых уровень защиты превосходит требования профиля, но с корректной реализацией всех его требований;

– результаты определения соответствия возможностей ИТ-продукта профилю защиты;

– возможности усовершенствования профиля защиты, в смысле выхода за рамки задач защиты и требований, установленных в профиле защиты.

8. В обосновании показывается, что проект защиты содержит полное и связное множество требований и что реализующий проект ИТ-продукт будет эффективно противостоять угрозам безопасности среды эксплуатации, а общие спецификации функций защиты соответствуют требованиям безопасности. Обоснование также содержит материалы, подтверждающие соответствие реального профиля заявленному и детализируются следующие вопросы:

– показано, что решение задач защиты, заявленных в проекте защиты, позволит эффективно противодействовать угрозам безопасности и реализовать сформулированную под них Политику безопасности;

– обоснование и разъяснение необходимых и достаточных условий обеспечения безопасности, в том числе что: функциональные требования безопасности соответствуют задачам защиты; требования адекватности соответствуют функциональным требованиям и усиливают их; набор всех стандартных и специфических функциональных требований обеспечивает решение задач защиты; все требования безопасности успешно реализованы; все взаимосвязи между требованиями CCITSE учтены либо путем их указания в самих требованиях, либо путем предъявления соответствующих требований к среде эксплуатации; заявленный уровень адекватности может быть подтвержден;

– доказано соответствие функций защиты функциональным требованиям безопасности и задачам защиты. Для этого должно быть показано, что выбранные функции защиты согласуются с заявленными задачами защиты; совокупность выбранных функций защиты обеспечивает эффективное решение совокупности задач защиты; заявленные возможности функций защиты соответствуют действительности.

– осуществляется обоснование уровня адекватности того, что заявленный уровень безопасности соответствует требованиям адекватности;

– обосновывается то, что требования проекта защиты реализуют все требования профиля защиты. Для этого должно быть показано, что все усовершенствования, реализованные в задачах защиты, по сравнению с профилем защиты, корректны, конкретизируют и развивают их; все усовершенствования требований безопасности по сравнению с профилем защиты реализованы корректно, конкретизируют и развивают исходные; все задачи защиты профиля решены и все требования профиля защиты выполнены; дополнительно введенные в проект защиты специальные задачи защиты и требования безопасности не противоречат профилю защиты.

Функциональные требования в Единых критериях разбиты на 9 классов и 76 разделов. Каждый раздел имеет свое уникальное имя и шестибуквенный идентификатор, состоящий из трехбуквенного обозначения раздела. Ранжирование функциональных требований осуществляется по множеству критериев (более 280). Набор этих критериев представляет собой иерархическую структуру в виде неупорядоченного списка сравнимых и несравнимых требований, в котором усиление требований безопасности происходит монотонно от низших уровней к высшим. Структура имеет вид направленного графа, усиление требований безопасности происходит при движении по его ребрам. Набор же принятых функциональных требований обобщает все существующие ранее стандарты информационной безопасности.

В Единых критериях вводится понятие *адекватность* – показатель реально обеспечиваемого уровня безопасности, отражающий степень эффективности и надежности реализованных средств защиты и их соответствия задачам защиты. Требования адекватности средств защиты в Единых критериях структурированы и детально регламентируют все этапы проектирования, создания и эксплуатации ИТ-продукта. Структура требований адекватности аналогична функциональным требованиям.

Всего определено семь стандартизированных уровней адекватности. Каждый из уровней определяет степень соответствия ИТ-продукта каждому требованию адекватности. По существу, названия уровней отражают возможности средств контроля и верификации, применяющихся в процессе разработки, анализа и совершенствования ИТ-продукта. Требования адекватности средств защиты разбиты на 7 классов и 26 требований. Требования адекватности, в смысле контроля и верификации ИТ- продуктов, разбиты на 7 уровней адекватности.

Наконец *Объект оценки* (ОО) в Единых критериях определен как продукт или система информационных технологий, а также связанные с ними управляющая и пользовательская документация, являющиеся объектом процесса оценки безопасности.

Концепция, представленная в стандарте, направлена на удовлетворение интересов трех основных групп – потребителей ОО, разработчиков ОО, и экспертов по оценке безопасности ОО. Необходимо отметить, что применение ЕК создает условия эффективного взаимодействия всех сторон, принимающих участие в разработке, эксплуатации и оценке систем безопасности, в частности и систем информационных технологий вообще. Применение и реализация положения стандартов позволяет различным категориям специалистов решить следующие задачи.

Потребитель, используя общие критерии, решает следующие задачи:

– выбора и формулировки требований по обеспечению безопасности определенного объекта;

– принятия на основе результатов процесса оценки решения о степени удовлетворения оцениваемого продукта или системы предъявленным им требованиям безопасности;

– сравнения различных продуктов и систем выбора адекватного продукта или системы;

– формулировки особых требований к показателям безопасности ОО на основе профиля защиты.

Разработчик, используя общие критерии, решает задачи:

– подготовки и осуществления процесса оценки разрабатываемых продуктов и систем;

– определения полного и непротиворечивого множества требований безопасности, которым должен удовлетворять разрабатываемый продукт или система;

– обоснования адекватности оцениваемого продукта или системы на основе проекта защиты;

– определения степени ответственности за оценку и доказательство необходимости оценки продукта или системы.

Эксперт по оценке решает задачи:

– выработки и принятия решения о степени соответствия (удовлетворения) объекта оценки требованиям безопасности;

– определения мероприятий и комплекса работ, необходимых для осуществления оценки продуктов или систем.

Единые критерии направлены не только на решение задачи оценки свойств объектов оценки, но и на описание этих свойств. Поэтому с использованием этого документа могут решать свои задачи и другие лица, например офицер безопасности, аудиторы, администраторы оценки, лица ответственные за аккредитацию и сертификацию продуктов и систем, и другие.

Для обеспечения наибольшей степени соответствия между результатами процессов оценивания, осуществляемых различными экспертами, очень важно, чтобы оценка осуществлялась на единой методологической основе с использованием надежных и апробированных схем и методик оценки.

Немаловажное значение имеет организация постоянного управления и контроля за процессом оценки. Именно здесь особенно четко проявляется регулятивная роль нормативных документов, которая направлена на обеспечение однозначности и взаимного соответствия результатов оценки.

На рисунке 1 представлена диаграмма, характеризующая взаимосвязь процесса оценки, критериев и методологии оценки безопасности.

Здесь под ***методологией оценки*** понимают систему принципов, процедур и процессов, применяемых при оценке безопасности информационных технологий.

Под ***схемой оценки*** понимают совокупность нормативных и руководящих документов, обеспечивающих интерпретацию и применение критериев оценки администратором оценки в рамках определенной общности экспертов.

Администратор оценки есть лицо, уполномоченное и ответственное за реализацию общих критериев в рамках отдельной общности экспертов через схему оценки и следовательно через совокупность стандартов и других нормативных документов, а также ответственное за организацию и контроль качества оценки.

Из приведенного следует, что разработка общих вопросов относительно критериев и методологии оценки безопасности информационных технологий является прерогативой международной сообщности, в то время как разработка конкретных схем и методик оценки осуществляется национальными и другими организациями конкретного государства. Эти схемы оценки, очевидно, должны определять взаимосвязанную совокупность методов и методик оценки показателей и свойств продуктов/систем, которые разрабатываются, с одной стороны, на единой методологической основе, что обеспечивает повторяемость и объективность результатов оценки, а с другой, на основе правовых и нормативных документов отдельного государства или организации.

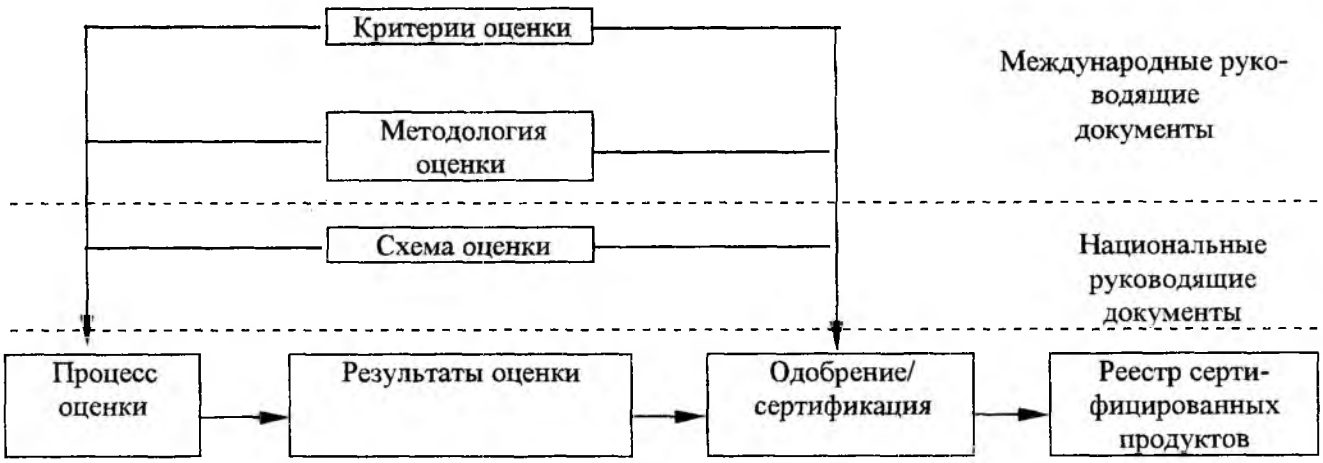


Рис. 1

Необходимо отметить еще одну особенность, а именно наличие такого этапа как одобрение и сертификацию результатов оценки. Дело в том, что большинство критериев оценки требуют привлечения знаний множества экспертов и скорее всего многие показатели могут быть определены только или с применением неформальных методов, в частности методов экспертного опроса. В этом случае неизбежно возникает задача определения степени согласованности мнений экспертов и обеспечения требуемой степени непротиворечивости мнений. Ясно, что не совсем просто обеспечить постоянство и согласованность уровня базовых знаний экспертов. Процесс сертификации в данном случае выступает как средство обеспечения большей степени согласованности мнений и принятых решений экспертов при применении ЕК с последующим оформлением сертификата.

Общий подход критериев безопасности можно охарактеризовать следующим образом.

Уверенность и доказательство безопасности продукта или системы можно получить в процессе разработки, оценки или эксплуатации системы (рис. 2). Разработчики стандарта опираются на общую модель поэтапной разработки системы – от формирования целей функционирования и ограничений к системе до её реального воплощения в "металле". Однако, как показали исследования [11], основной причиной неудач в защите информации является то, что вопросы ЗИ рассматривались без органической связи с проектированием и технологией функционирования систем. Стандарт рекомендует формировать требования по безопасности одновременно и во взаимосвязи с формированием технических, эксплуатационных, экономических и других требований к разрабатываемой системе. На основе сформулированных требований безопасности разрабатываются профиль и проект защиты. В ходе разработки объекта ранее сформулированные требования могут быть уточнены и модифицированы.

Процесс оценки объекта может выполняться либо параллельно с разработкой, либо после неё. Ожидаемыми результатами оценки являются, во-первых, подтверждение того, что объект оценки удовлетворяет требованием безопасности, изложенным в проекте защиты и, во-вторых, обеспечение степени уверенности в полученной оценке, через выполнение требований адекватности и установление уровня адекватности оценки. Полученные результаты оформляются соответствующими документами и могут быть использованы разработчиками и потребителями для решения своих задач.

Необходимо отметить, что процесс оценки оказывает сильное позитивное влияние на формирование требований, процессы разработки и оценки, а также на эксплуатацию продукта. Оценка объекта, прежде всего, предназначена для выявления ошибок и уязвимых мест в системе, которые в дальнейшем будут устранены разработчиком и, тем самым, будет уменьшена вероятность нарушения безопасности в ходе эксплуатации объекта. С другой стороны, разработчик, зная концептуально-методологический подход оценки безопасности, уже на этапе формирования требований и проектирования будет проявлять большое внимание на решение вопросов безопасности.

Этап эксплуатации, с точки зрения обеспечения защиты информации, интересен тем, что здесь могут быть выявлены новые неизвестные ошибки, которые могут появиться при изменении условий эксплуатации. Кроме того, могут появиться и новые угрозы безопасности. Данные ошибки будут учтены разработчиками и экспертами в ходе усовершенствования и модификации объекта.

Стандарт различает три типа оценки: оценку профиля защиты, оценку проекта защиты и оценку объекта защиты.

Целью оценки профиля защиты является подтверждение того, что профиль защиты является полным, согласованным, а также технически применимым и пригодным к использованию в качестве требований для оцениваемого объекта.

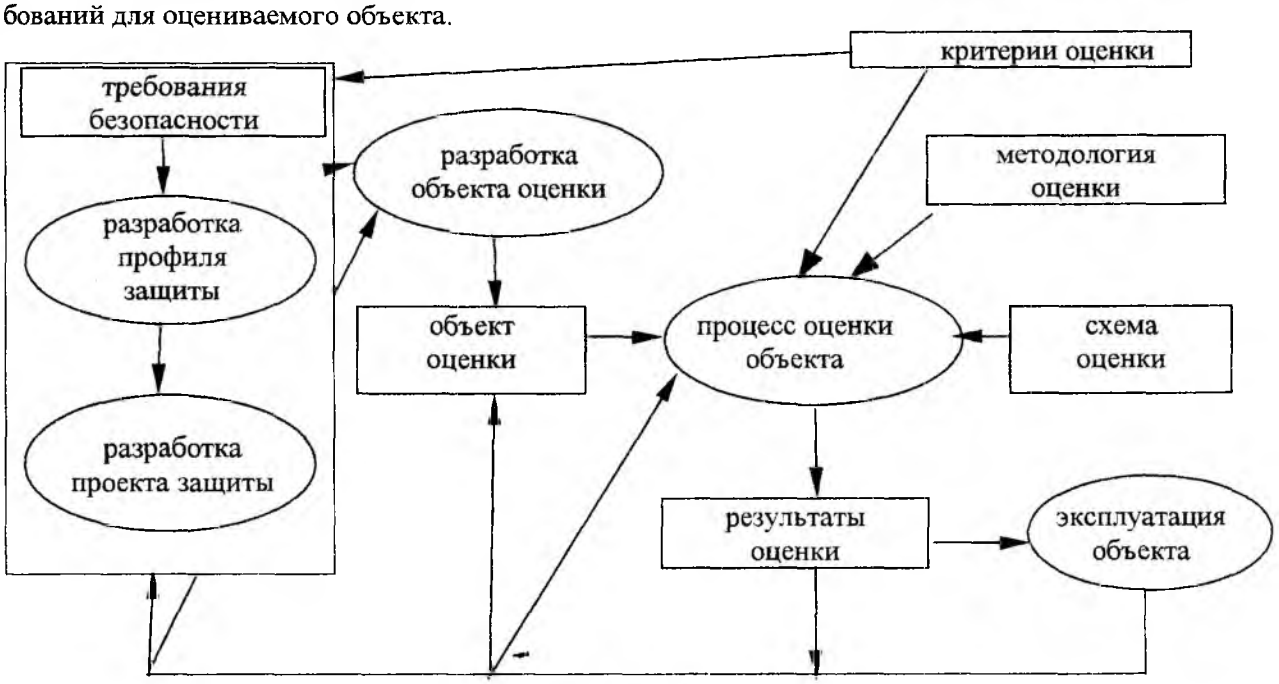


Рис. 2

Целью оценки проекта защиты является: во-первых, подтверждение того, что проект защиты полный, согласованный, а также технически применимый и пригодный для использования его в качестве основы для оценки соответствующего объекта оценки; во-вторых, для подтверждения того, что проект защиты удовлетворяет требованиям профиля защиты (при необходимости).

Целью оценки объекта является подтверждение того, что объект оценки удовлетворяет требованиям безопасности, содержащимся в проекте защиты.

На рисунке 3 представлены возможные варианты использования результатов оценки, которые предлагает стандарт.

Как видно из рисунка, разработка и оценка объекта требует наличия требований безопасности и может опираться на каталоги профилей защиты и продуктов, которые уже были ранее оценены. В зависимости от того, что являлось объектом оценки (продукт или система), результаты оценки используются для формирования каталога продуктов, либо для аккредитации системы. В последнем случае результаты оценки должны быть доступны лицу или организации, ответственным за аккредитацию систем. Важным здесь является то, что предполагается создание международного реестра (каталога) оцененных профилей защиты, проектов защиты, продуктов и сертификатов, которые будут доступны разработчикам и могут быть использованы или при разработке новых, или при усовершенствовании старых систем. Это приведет к значительной экономии материальных, финансовых и людских ресурсов при разработке новых систем, что является одной из основных задач международной стандартизации.

Рис. 3

4. Общая методология оценки безопасности информационных технологий

Общие критерии оценки безопасности могут и должны применяться на единой методологической основе. Поэтому вполне естественно, что сразу же после появления версии ISO/IEC 15408 начались работы по разработке нормативного документа, определяющего общую методологию оценки безопасности информационных технологий. На данный момент таким документом является CEM-97/017 – "Общая методология оценки безопасности информационных технологий" [8].

Данный нормативный документ предназначен в основном для экспертов по оценке безопасности систем, а также необходим разработчикам, заказчикам оценки и контролирующим органам. Именно эти стороны определены в качестве пользователей общей методологии (ОМ).

С точки зрения разработчика профиля защиты применение ОМ позволяет выполнить независимую и последовательную оценку и обоснование профиля защиты.

Для разработчика объекта оценки важно то, что применения ОМ позволит:

– независимо обосновать и проверить задокументированные в профиле и проекте защиты показатели безопасности;

– убедить потребителя в том, что объект оценки обладает заявленными показателями безопасности;

– более эффективно использовать при построении систем безопасности результатов, полученных при оценке других продуктов и систем;

– уменьшить затраты временных и материальных ресурсов на осуществление процесса оценки безопасности.

Заказчик оценки – это организация, которая дает поручение на осуществление оценки безопасности объекта. В роли заказчика могут выступать разработчик, системный интегратор, потребитель (пользователи, аудиторы, системный администратор и т.д.). Здесь применение ОМ позволяет задокументировать, независимо и последовательно обосновать и проверить показатели безопасности и обеспечить возможность сравнения и обоснованного выбора различных объектов оценки.

Для экспертов ОМ выступает как руководство по применения критериев оценки безопасности.

Наконец контролирующий орган, т.е. организация, которая гарантирует, что процесс оценки осуществляется в соответствии с критериями оценки, определяет из CEM-97/017 совокупность документов, их форму и содержание, представляемых экспертом по оценки безопасности продукта или системы.

Таким образом, нормативный документ, определяющий общую методологию оценки безопасности информационных технологий, направлен на обеспечение взаимодействия между различными субъектами, заинтересованными в оценке безопасности объекта, упорядочение процесса оценки безопасности продуктов и систем, всемерное и полное информационное обеспечение заинтересованных сторон о ходе выполнения процесса оценки.

Областью применения положений CEM-97/017 являются принципы, процедуры и процессы оценки безопасности, а также мероприятия и комплекс работ, выполняемые в ходе оценки, разработки и контроля оценки безопасности.

Данный документ определяет следующие общие принципы оценки безопасности.

1. Принцип соответствия прилагаемых усилий и заданного уровня адекватности оценки.

Для обеспечения заданного уровня адекватности оценки все стороны должны выполнять свои задачи с той степенью ответственности и строгости, которая соответствует требованиям уровня адекватности.

2. Принцип беспристрастности оценки.

Любая оценка должна быть получена в условиях, исключающих влияние на нее каких-либо личных предубеждений экспертов.

Ни одна из сторон не должна иметь каких-либо предубеждений к объекту оценки или профилю защиты, которые могут быть основаны на ранее известных результатах оценки других профилей защиты или объектов оценки или на давлении одной стороны на другую. С целью уменьшения взаимных влияний сторон и экспертов друг на друга в процессе оценки продуктов и систем организуется надлежащий организационно-технический надзор и применяются схемы, устраняющие какие-либо конфликты между сторонами и экспертами.

3. Принцип объективности оценки.

Результаты оценки должны быть получены в условиях, обеспечивающих минимальное влияние каких-либо индивидуальных субъективных мнений и решений на общую оценку.

Очевидно, что отдельный эксперт не может быть свободен от влияния каких-либо субъективных факторов при принятии решений. Соответствующий организационно-технический надзор над процессом оценки, основанный на хорошо продуманной методологии, организации процесса оценки и интерпретации результатов оценки, должен обеспечит уменьшение влияния личных взглядов и решений отдельных экспертов на общую оценку до приемлемого уровня.

4. Принцип повторяемости и воспроизводимости.

Повтор процесса оценки одного и того же объекта оценки или профиля защиты с одними и теми же требованиями и при одном и том же информационно-техническом обеспечении должны приводить к одним и тем же результатам.

Любое действие должно приводить к одним и тем же результатом независимо от того, кто выполняет это действие. Воспроизводимость направлена на обеспечение соответствия и согласованности результатов оценки, полученных в различное время (например, на различных этапах жизненного цикла системы безопасности), в то время как повторяемость направлена на обеспечение соответствия и согласованности результатов оценки, полученных различными экспертами и, возможно, при условии использования ими различных схем оценки безопасности.

5. Принцип достоверности.

Результаты оценки должны быть полными и технически корректными.

Результат оценки должен показать высокую степень рассудительности принятого решения и тщательности технической экспертизы объекта оценки и профиля защиты. Процесс оценки и полученные результаты должны быть объектами организационно-технического надзора для того, чтобы гарантировать выполнение требований общих критериев, методологии и схем оценки безопасности.

Реализация вышеперечисленных принципов предполагает выполнение следующих условий.

1. Стоимостная эффективность оценки, заключающаяся в том, что ценность результатов оценки должна компенсировать затраты временных, материальных и других ресурсов на проведение оценки. В процессе оценки баланс между ценностью оценки и затратами на ее проведение должен постоянно отслеживаться. Данное условие порождает ограничения на количество показателей, которые входят в оценку безопасности. То есть в оценку могут входить наиболее весомые показатели. Однако тут же возникает вопрос, каким образом определить степень важности того или иного показателя? Скорее всего это тема отдельного обсуждения.

2. Изменение технических и других условий применения систем безопасности, развития информационных технологий, методов оценки и криптоанализа должны отражаться в методологии оценки. Методология оценки должна иметь возможность адаптации к изменяющимся условиям и быть применимой к развивающимся технологиям в области защиты информации. Это позволит обеспечить требуемый уровень эффективности методов оценки и гарантировать их пригодность к оценке профилей защиты и объектов оценки.

3. Обеспечение возможности эффективного использования известных результатов оценки существующих профилей защиты и систем играет важную роль при выполнении последовательной оценки в одних и тех же условиях. Повторная доступность результатов особенно важна в тех случаях, когда оцениваемые объект оценки или профиль защиты являются интегрированными частями других объектов или профилей защиты.

4. И наконец, важно обеспечить, чтобы все стороны в процессе оценки пользовались единой терминологией. На это и направлена разработка нормативных документов и стандартов.

Каждая из сторон, участвующих в процессе оценки на основе общей методологии, несет определенную ответственность за выполнение определенных задач. СЕМ-97/017 определяет такую ответственность в рамках общих принципов и допускает, что схемы оценки могут вводить дополнительные требования к сторонам, с учетом особенностей национального законодательства и положений руководящих документов. Ответственность распределяется следующим образом.

Заказчик оценки несет ответственность за:

– заключение необходимых соглашений для осуществления оценки;

– обеспечение экспертов необходимыми материально-техническими и информационными ресурсами для осуществления оценки.

Разработчик несет ответственность за:

– поддержку процесса оценки;

– разработку и сохранение необходимых ресурсов для оценки.

Эксперт по оценке несет ответственность за:

– получение необходимых ресурсов для оценки (документация, профиль и проект защиты, копия (образец) объекта оценки);

– выполнение работ по оценке в соответствии с требованиями общих критериев;

– формирование запроса и получение дополнительной помощи или материалов для оценки (обучение у разработчика, интерпретация требований контролирующего органа);

– обеспечение условий для организации надзора за процессом оценки;

– документирование и утверждение промежуточных и окончательного решений;

– создание условий, при которых гарантируется согласованность процесса оценки с общими принципами и требованиями соответствующих схем оценки.

Контролирующий орган несет ответственность за:

– мониторинг процесса оценки;

– получение и рассмотрение материалов контроля;

– создание условий, гарантирующих согласованность процесса оценки с общими принципами положениями СЕМ;

– поддержку процесса оценки через разработку и внедрение схем, методик и правил интерпретации результатов, а так же различного рода руководящих документов;

– одобрение или опровержение окончательных решений;

– документирование и юридическое закрепление решений администратора оценки.

Методология оценки предполагает, что оценка будет осуществляться в три этапа: предварительный, основной и заключительный.

На предварительном этапе основными действующими лицами являются заказчик оценки и эксперт Заказчик информирует все стороны относительно необходимости оценки профиля защиты или объекта оценки, обеспечивает эксперта необходимой документацией, материалами по профилю защиты и объекту оценки. Задачей эксперта является определение возможности успешного осуществления оценки на основе полученных материалов и по необходимости затребовать дополнительного обеспечения заказчика или разработчика.

Итогом подготовительного этапа является заключение между заказчиком и экспертом соглашения на осуществление работ по оценке объекта или профиля защиты.

Непосредственная оценка осуществляется на основном этапе. В процессе оценки эксперт рассматривает представленные ему материалы, профиль защиты или объект оценки. Эксперт может составлять ряд обзорных отчетов в которых могут содержаться его требования по предоставлению пояснений относительно применения требований контролирующего органа, запросы на дополнительную информацию по профилю защиты или объекту оценки у заказчика или разработчика, выявленные слабости недостатки и другая информация о ходе оценки.

Контролирующий орган осуществляет непрерывный мониторинг процесса оценки в соответствии со схемой оценки.

Результатом основного этапа является разработка и предоставление экспертом Технического отчета оценки (ТОО), который содержит обоснование принятого экспертом решения.

На заключительном этапе осуществляется рассмотрение и анализ ТОО всеми сторонами. Основным действующим лицом на этом этапе выступает контролирующий орган. Он осуществляет всесторонний анализ ТОО на предмет его соответствия общим критериям общей методологии и требования

схем оценки безопасности. Контролирующий орган принимает решение о согласии или несогласии с решением, изложенном в ТОО и готовит Итоговый отчет оценки на основе ТОО. При этом все стороны, вовлекаемые в процесс оценки, имеют право ознакомление с материалами Итогового отчета и могут требовать соответствующих пояснений.

5. Перспективы практической реализации положений ISO/IEC 15408 и CEM – 97/017

Выше рассматривались задачи, на решение которых направлено использование положений рассматриваемых нормативных документов. Эти задачи сформулированы и изложены в самих этих документах. Важно оценить перспективы применения положений документов на практике. При этом важно сделать эту оценку с позиций системного подхода к решению проблемы защиты информации.

В работе [11] изложены три основных задачи, решаемые в рамках системного подхода к решению сложной проблемы. В контексте рассмотренных документов эти принципы можно сформулировать следующим образом:

1) системный анализ сущности проблемы защиты информации;

2) разработка и обоснование полной и непротиворечивой концепции и методологии решения проблемы защиты информации, в рамках которой решение задачи защиты продукта или системы в конкретных условиях определяется в виде частного случая – разработкой профиля и проекта защиты;

3) системное использование методов и механизмов защиты информации при решении задачи синтеза (проектирования, разработки) безопасных продуктов и систем информационных технологий.

Видно, что предложенные документы направлены на решение первых двух задач. В стандарте ISO/IEC 15408 осуществлена полная декомпозиция проблемы защиты информации. Механизмы профиля и проекты защиты отражают суть концепции решения проблемы защиты информации.

Однако в документах нет методологии решения третьей задачи – задачи синтеза систем. Функциональные требования и требования адекватности, как и методология оценки безопасности, направлены в первую очередь на решение задачи оценки безопасности продукта или системы. Хотя их применение оказывает определенное регламентирующее влияние на проектирование, разработку и эксплуатацию систем. Здесь необходимо решать задачу установления соответствия целям защиты (которые выражаются через требования) и множеством средств и механизмов, которые имеются в нашем распоряжении для реализации этих целей.

Стандарт ISO/IEC 15408 предполагает создание электронного каталога профилей защиты, прошедших оценку и сертификацию, что позволит разработчикам использовать известные профили защиты при разработке новых продуктов и систем. Однако нужно сказать, что профиль (проект) защиты является не чем иным как сертифицированным и обоснованным решением задачи защиты информации в конкретных условиях эксплуатации продукта. Таким образом, можно сделать вывод, что последовательное применение положений стандарта при решении практических задач создает базу для разработки и создания экспертных систем в области защиты информации. А это в свою очередь позволяет перейти к разработке автоматизированных средств поддержки принятия решений в данной области и средств автоматизированного проектирования систем защиты информации. Одним из направлений использования результатов оценки разработанных профилей защиты, которые могут рассматриваться как управляющее воздействие на систёму при возникновении определенных угроз безопасности (ситуации), является применение модели ситуационного управления системами защиты информации. Это в перспективе может привести к созданию самомодифицирующихся систем защиты информации, которые с помощью администратора или автоматически будут модифицировать свою структуру и функции в зависимости от складывающихся условий эксплуатации и угроз.

Другим важным результатом разработки данных документов является отражение системности подхода к решению проблемы защиты информации и создание единой методологической базы решения задач защиты информации. Важно то, что не только схемы оценки безопасности должны разрабатываться в рамках единой методологии, но процессы проектирования и разработки новых продуктов и систем должны осуществляться с учетом норм и положений данных документов. Это является хорошим подспорьем и для отечественных специалистов.

Заключение

На наш взгляд, рассмотрение документов требует внимательного изучения и внедрения в отечественную практику разработки и оценки соответствующих изделий, продуктов и систем. Поскольку эти документы являются продуктом работы ряда организации и объектом международной стандартизации, то им можно доверять. Тем более Украина также является членом группы P Международной организации по стандартизации и участвует в голосовании решения по принятию этих документов.

Одной из особенностей стандартизации в области защиты информации является интернационализация стандартизации. Гиперскоростное развитие информационных технологий, создание всемирного единого информационного пространства, интеграция в это пространство нашего государства являются непреложными фактами. Создание адекватных и надежных систем защиты информации в таких условиях не под силу отдельному государству. И по этой причине необходимо осваивать и применять данный методологический аппарат в отечественной практике, адаптировать или разработать новые нормативные документы, которые будут учитывать положения международных стандартов.

Разработка отечественных уникальных схем оценки безопасности в рамках общих критериев и методологии позволит нам не только оценить собственные продукты и системы, но активно участвовать в сертификации изделий, продуктов и систем зарубежного производства, тем самым защитить свой рынок от низкопробной продукции.

На наш взгляд, реализация и применение норм и положений этих документов в отечественной и мировой практике даст новый толчок в развитии теории и методов защиты информации.

Список литературы: **1.** *Trusted* Computer Systems Evaluation criteria, US DoD 5200.28-STD, 1985. **2.** *Information* Tecnhology Security Evaluation Criteria, v. 1.2. –Office for Official ublications of the European Communities, 1991. **3.** *Canadian* Tusted Computer Product Evaluation Criteria, v. 3.0. Canadian System Security Centre, Communications Security Establishment, Government of Canada, 1993. **4.** *Federal* Criteria for Information Technology security. – NIST, NSA, US Government, 1993. **5.** *ISO/IEC 15408-1:1999* – Information technology – Security techniques – Evaluation criteria for IT security – Part1: Introduction and general model. **6.** *ISO/IEC 15408-2:1999* – Information technology – Security techniques – Evaluation criteria for IT security – Part 2: Security functional requirements. **7.** *ISO/IEC 15408-3:1999* – Information technology – Security techniques – Evaluation criteria for IT security – Part 3: Security assurance requirements. **8.** *CEM-97/017.* Common Evaluation Methodology for Information Technology Security – Part 1: Introduction and general model. **9.** *ISO/IEC 7498-2:1989.* – Information processing systems – Open Sistems Interconnection – Basic Reference Model – Part 2: Security Architecture. **10.** *Зегжда Д.П., Ивашко А.М.* Как построить защищенную информационную систему. – СПб: Мир и семья – 95, 1997. – 312 с. **11.** *Герасименко В.А.* Защита информации в автоматизированных системах обработки данных. Кн.1. – М.:Энергоатомиздат, 1994. – 400 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 21.03.2000

УДК 681.3.06: 519.248.681

В.И. ДОЛГОВ, И.В. ЛИСИЦКАЯ, С.А. ГОЛОВАШИЧ, А.С. БОНДАРЕНКО

ОБЕСПЕЧЕНИЕ СТОЙКОСТИ DES - ПОДОБНЫХ АЛГОРИТМОВ ШИФРОВАНИЯ К АТАКАМ ЛИНЕЙНОГО КРИПТОАНАЛИЗА ПРИ ИСПОЛЬЗОВАНИИ ТАБЛИЦ ПОДСТАНОВОК СЛУЧАЙНОГО ТИПА

В представленных ранее наших работах [1-3] рассматриваются пути и возможности обеспечения защищенности алгоритма шифрования DES от одной из опаснейших криптоаналитических атак – дифференциального криптоанализа. Этой работой мы продолжаем изучение вопросов безопасности шифра DES. Речь будет идти о другой криптоаналитической атаке, с помощью которой удалось поколебать уверенность в надежности американского стандарта – линейном криптоанализе. То, что таблицы стандарта не оптимизированы в отношении линейного криптоанализа, отмечается в ряде публикаций [3,4 и др.]. В [4] удалось найти краткое упоминание о работах по преодолению этого недостатка группы Кванджио Ким. Приводится даже пример построения таблиц S блоков, защищенных от атак линейного и дифференциального криптоанализа, однако не отмечается, насколько группе Кванджио Ким удалось продвинуться в этом направлении, как и не излагается сама методика отбора таблиц S блоков, а приведена лишь ссылка на критическое отношение Эли Бихама к результатам исследований отдельных этапов. Наши исследования показывают, что в отношении дифференциального криптоанализа приведенные в [4] таблицы нельзя считать надежными. Учитывая, что сама методика выполнения линейного криптоанализа, также как и возможности защиты от атак этого типа, остаются все еще мало изученными в Украине, в этой работе мы приводим краткое изложение принципов выполнения такой атаки для шифра DES и предлагаем свою версию решения задачи построения S блоков стандарта, защищенных от атак линейного криптоанализа.

Линейный криптоанализ – сравнительно новый тип криптонападения, предложенный Мацуи [3] в 1993 г. Этот метод использует линейную аппроксимацию для описания процедуры нападения на DES. Она заключается в нахождении ситуаций, когда сумма по модулю 2 некоторых битов открытого текста и некоторых битов соответствующего ему зашифрованного текста равна сумме по модулю 2 некоторых битов ключа. Если такая ситуация выполняется с некоторой вероятностью $p \neq 1/2$, то имеется возможность использовать собранные открытые тексты и соответствующие им шифрованные тексты для определения битов ключа.

Как и при дифференциальном криптоанализе алгоритма DES, сложность линейного криптоанализа определяется правилами построения S блоков [3], для описания свойств которых строятся специальные таблицы.

При построении таких таблиц просматриваются все возможные 4-битные выходы S блока, которые получаются при различных 6-битных значениях его входов. При этом вычисляются поразрядные произведения по модулю 2 входов S блока (6-битное число) и некоторого фиксированного 6-битного числа ("маски" по строкам) и соответствующие поразрядные произведения по модулю 2 выходов S блока и второго фиксированного теперь уже 4-битного числа ("маски" по столбцам). Эти фиксированные числа являются индексами входов в ячейку таблицы размера 64×16. Сама таблица, названная Мацуи линейной аппроксимационной таблицей, получается заполнением каждой из ячеек числом, соответствующим количеству линейных соотношений, выполняющихся для входных битов, прошедших маску по столбцам, и выходных битов, прошедших маску по строкам для этой ячейки, при вариации по всему множеству входов S блока. Сущность линейных соотношений заключается в равенстве нулю суммы по модулю 2 всех входных и выходных бит, прошедших обе маски. Математически отмеченные действия можно описать следующим образом.

Пусть, вектор $\bar{x}_i = \left(x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}\right)$ представляет собой 6-битное число, обозначающее один из $i = \overline{1,64}$ возможных входов S блока, а вектор $\bar{a}_l = \left(a_{l1}, a_{l2}, a_{l3}, a_{l4}, a_{l5}, a_{l6}\right)$, $l = \overline{1,64}$ – 6-битную входную «маску» (индекс таблицы по строкам).

Пусть, вектор $\bar{y}_p = \left(y_{p1}, y_{p2}, y_{p3}, y_{p4}\right)$ обозначает один $p = \overline{1,16}$ 4-битных выходов S блока, $\bar{b}_m = \left(b_{m1}, b_{m2}, b_{m3}, b_{m4}\right)$, $m = \overline{1,16}$ – 4-битную выходную «маску» (индекс таблицы по столбцам).

Обозначим $\bar{a}_l \cdot \bar{x}_i$ – двоичное скалярное произведение векторов $\bar{a}_l$ и $\bar{x}_i$, т.е. $\bar{a}_l \cdot \bar{x}_i = \bigoplus_{k=1}^{6} a_{lk} \cdot x_{ik}$.

Аналогично пусть $\bar{b}_m \cdot \bar{y}_p$ – двоичное скалярное произведение векторов $\bar{b}_m$ и $\bar{y}_p$, и, следовательно, $\bar{b}_m \cdot \bar{y}_p = \bigoplus_{k=1}^{4} b_{mk} \cdot y_{pk}$. В представленных выражениях символом $\oplus$ обозначена операция суммирования по модулю 2 (XOR).

Значения ячеек аппроксимационной таблицы можно представить в следующем аналитическом виде:

$$\theta_{lm} = \underset{\{\bar{x}_i\}}{\mu}\left(\bar{x}_i \cdot \bar{a}_l = \bar{y}_p \cdot \bar{b}_m\right) - 2^5 \Big|_{\,y_p = S(x_i);\ i=\overline{1,64},\ p=\overline{1,16}},$$

где функция $\mu(\cdot)$ обозначает количество случаев выполнения равенства в скобках при вариации по всем возможным значениям аргумента $\bar{x}_i$, $i = \overline{1, 64}$. В приведенной выше формуле использование различных индексов для входных и выходных значений S блока связано с тем, что размер множества входных значений $\{x_i\}$ в 4 раза больше размера множества выходных значений $\{y_p\}$. Слагаемое -2^5 используется для нормирования величин θ_{lm} относительно половинного (среднего) значения. В результате вероятность того, что аппроксимация в S блоке является правильной (соответствует значению линейной аппроксимационной таблицы θ_{lm}), дается выражением $p'_{lm} = (32 - \theta_{lm})/64 = 1/2 - \theta_{lm}/64$, или если ввести обозначение $p_{lm} = \theta_{lm}/64$, то $p'_{lm} = 1/2 - p_{lm}$. Вход со значением $p_{lm} = 0$, или что то же $\theta_{lm} = 0$ имеет вероятность $p' = 1/2$. Такой вход бесполезен для атаки на криптосистему. Любое ненулевое значение (положительное или отрицательное) может быть использовано для атаки.

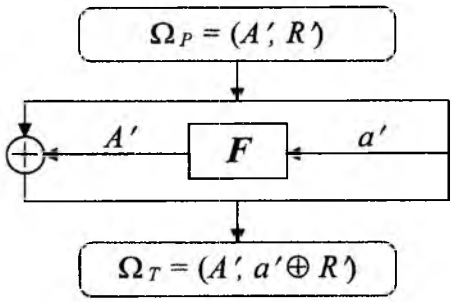
Здесь мы не будем сосредотачивать внимание на самой технике добывания ключей (основная атака Мацуи ориентирована на получение одного (единственного) бита ключа, а определение других битов требует уже дополнительных ухищрений [3]). Нас будет интересовать только то, что связано с оценкой стойкости алгоритма DES к рассматриваемой атаке. Это касается, прежде всего, построения аппроксимационных характеристик, под которыми понимаются системы взаимосвязанных линейных соотношений, распространенных на несколько циклов или S блоков процедуры шифрования, и оценки вероятностей одновременного выполнения всех линейных соотношений, попавших в цепочку (вероятностей аппроксимационных характеристик).

Приведем в связи с этим правила построения таких характеристик. Воспользуемся здесь обозначениями и определениями, предложенными в работе [3].

Определение 1. Одноцикловая характеристика есть форма $(\Omega_P, \Omega_T, \Omega_K, 1/2 + p)$, в которой $(\Omega_P)_L = (\Omega_T)_L = A$, $(\Omega_P)_R \oplus (\Omega_T)_R = a$, и для которой $1/2 + p$ есть вероятность того, что случайный входной блок P и его одноцикловое шифрование C с применением случайного подключа K удовлетворяет условию $P \cdot \Omega_P \oplus C \cdot \Omega_T \oplus K \cdot \Omega_K = 0$, где $(\cdot)$ обозначает двоичное скалярное произведение двух двоичных векторов, Ω_P – «маска», определяющая подмножество бит данных перед циклом, Ω_T – «маска», определяющая подмножество бит данных после цикла, и Ω_K – «маска», определяющая подмножество бит ключа, четность которых аппроксимируется; индексы L и R обозначают соответственно левую и правую половины блока данных.

Равенство $P \cdot \Omega_P \oplus C \cdot \Omega_T \oplus K \cdot \Omega_K = 0$ и есть линейная аппроксимация.

Примеры построения одноцикловых характеристик приведены на рис.1 и рис. 2 (здесь и далее мы пользуемся обозначениями и манерой изображения иллюстративного материала, предложенными в работе [3]).

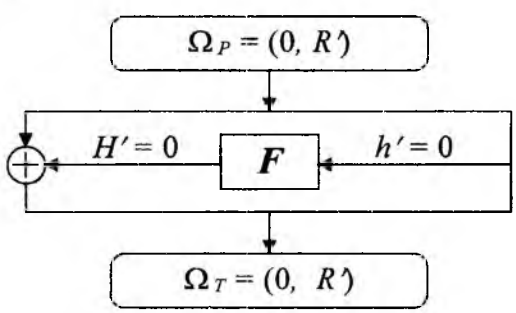


с вероятностью $1/2 - p_a$ задействован один ключевой бит

Рис. 1

Как и при дифференциальном криптоанализе аппроксимационные характеристики могут использовать более чем один S блок и распространяться на большее число циклов шифрования.

Определение 2. n – цикловая характеристика $\Omega^1 = \left(\Omega^1_P, \Omega^1_T, \Omega^1_K, 1/2 + p_1\right)$ может быть объединена с m – цикловой характеристикой $\Omega^2 = \left(\Omega^2_P, \Omega^2_T, \Omega^2_K, 1/2 + p_2\right)$, если Ω^1_T равно переставленному значению двух половинок Ω^2_P, т.е. $\left(\Omega^2_P\right)_L = \left(\Omega^1_O\right)_R$, $\left(\Omega^2_P\right)_R = \left(\Omega^1_O\right)_L$. Конкатенация характеристик Ω^1 и Ω^2 (если они могут быть объединены) есть $(n+m)$ – цикловая характеристика $\Omega = \left(\Omega^1_P, \Omega^2_T, \Omega^1_K \oplus \Omega^2_K, 1/2 + 2 \cdot p_1 \cdot p_2\right)$.



с вероятностью 1 нет задействованных ключевых бит

Рис. 2

Заметим здесь, что вероятность аппроксимации с двумя активными S блоками или вероятность двухцикловой аппроксимации есть $p'_1 p'_2 + \left(1 - p'_1\right)\left(1 - p'_2\right) = 1/2 + 2p_1 p_2$, так как сумма линейных соотношений будет линейным соотношением тогда, когда оба линейных соотношения равны нулю, и тогда, когда оба линейных соотношения равны единице (напомним, что $p'_i = 1/2 + p_i$).

Когда объединяется l характеристик с вероятностью p_i каждая (если это может быть выполнено), то вероятность результирующей характеристики определяется выражением

$$1/2 + p = 1/2 + 2^{l-1} \prod_{i=1}^{l} p_i \,. \qquad (1)$$

Заметим, наконец, что если линейная аппроксимация с вероятностью $1/2 + p$ известна, то атака на полный 16-цикловый DES требует около p^{-2} известных пар открытый – шифрованный текст, которые могут быть выбраны случайно [5].

В дальнейшем речь будет идти об атаках, использующих одноблочные характеристики. Введем понятие минимальной итеративной характеристики, под которой будем понимать характеристику, содержащую минимальное число циклов с задействованными ключевыми битами, среди которых имеется хотя бы один цикл тождественного типа (в котором нет задействованных ключевых битов), допускающую циклическое продолжение.

Для итеративной характеристики выполняется условие: выход характеристики является перестановкой левой и правой половинок входа, т.е. при входе $\Omega_P = (A', R')$ имеем $\Omega_T = (R', A')$, что и обеспечивает в соответствии с правилами построения характеристик (определения 1 и 2) ее циклическое продолжение. Заметим здесь сразу, что одной из особенностей линейного криптоанализа по сравнению с дифференциальным является то, что в линейной характеристике «свободной» является правая половина используемого множества входных бит, в то время как в дифференциальном криптоанализе это левая половина. Но тогда, если свободную часть взять равной нулю, то есть $R' = 0$, то тождественное одноцикловое преобразование, представленное на рис. 2, позволяет сразу ориентироваться на формирование тождественного нетривиального многоциклового преобразования $\Omega_P = (A', 0) \to \Omega_T = (A', 0)$, поскольку оно с помощью дополнительного тождественного (тривиального) преобразования приводится к требуемому виду: $\Omega_P = (A', 0) \to \Omega_T = (0, A')$. При этом удается еще один раз воспользоваться тождественным одноцикловым преобразованием, выполняющимся с вероятностью единица.

Заметим далее, что тождественное преобразование (нетривиального типа) можно реализовать только для нечетного числа циклов (в соответствии с правилом обмена левых и правых частей цепи Фейстеля). Наконец, можно сразу отметить, что нас интересуют возможности циклического продолжения характеристик, т.е. нас интересуют многоцикловые итеративные характеристики.

Теперь наша ближайшая задача построить такую характеристику для шифра DES. Будем исходить из принципа симметрии, который необходимо реализовать для входа и выхода итеративной характеристики с нечетным числом циклов. В качестве «центра симметрии» очевидно и должно выступить тождественное тривиальное преобразование (поскольку изначально рассматривается задача, когда такое преобразование имеется в единственном числе). В этой ситуации представляется естественным использование характеристики, представленной на рис. 3.

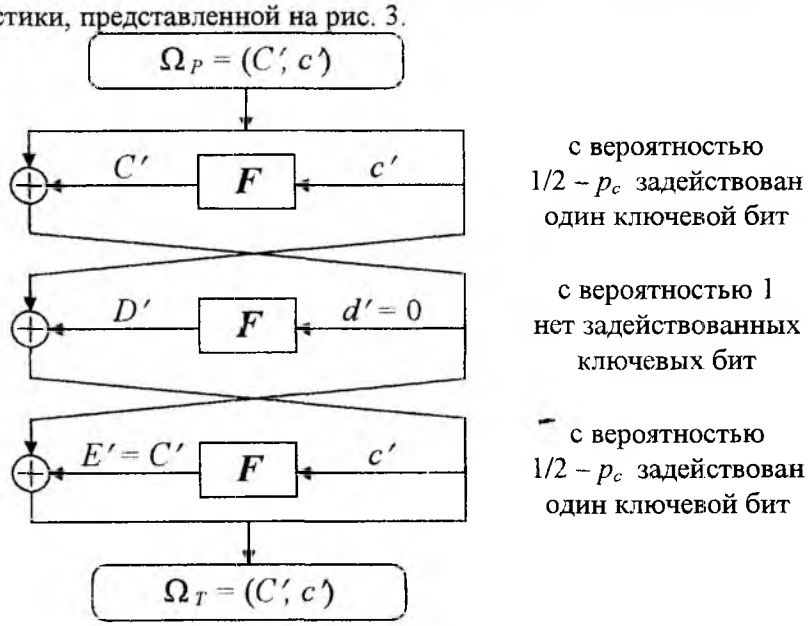


Рис.3

На этом рисунке представлена трехцикловая характеристика тождественного типа с ненулевыми значениями входов $\Omega_P = (C', c')$, причем $C' = F(c')$. Поскольку $E' = C'$, то, очевидно, что $e' = c'$. Очевидно также, что для этой характеристики симметричного типа c' и соответственно C' не могут быть равными нулю. Теперь нужно к этой характеристике подобрать начальную и конечную части, обеспечивающие получение итеративно продолжающейся характеристики $\Omega_P = (A', 0) \to \Omega_T = (A', 0)$. Это удается сделать, используя в обоих случаях двухцикловые характеристики (преобразования), представленные на рис. 4 и рис. 5.

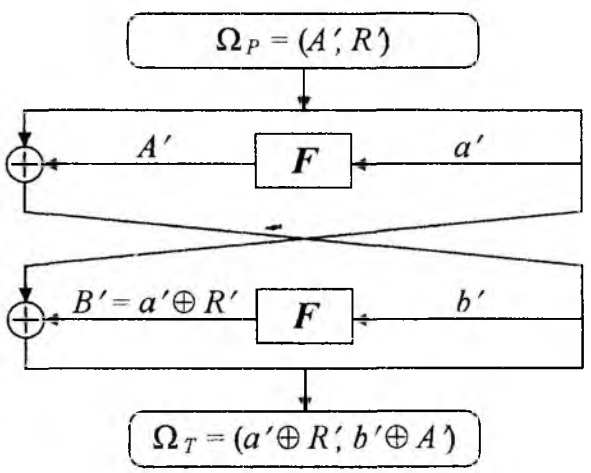
Условием сшивки характеристик, представленных на рис. 3, рис. 4 и рис. 5 выступают соотношения $C' = a' \oplus R'$, $\tilde{n}' = b' \oplus A'$.

Наша задача получить на выходе в качестве результата $\Omega_T = (A', R')$, что совпадает с исходным множеством битов $\Omega_P = (A', R')$ с точностью до порядка следования левой и правой половинок.

Для построенной семицикловой характеристики при $f' = a'$ и $g' = b'$ получаем $\Omega_T = (R', A')$. Остается сделать обмен левой и правой половинок множеств битов, участвующих в аппроксимации. Но как следует из рис. 3 такой обмен можно выполнить, если воспользоваться дополнительным тождественным циклом, который в свою очередь требует выполнения условия $R' = 0$.

В результате мы приходим к итеративной восьмицикловой характеристике, представленной на рис. 6. Именно характеристика такого типа использована в атаке Бихама [3]. Важно здесь подчеркнуть, что нельзя построить характеристики с большим числом тождественных преобразований, приходящихся на 8 циклов.

Любая другая характеристика с уменьшенным числом тождественных преобразований будет иметь результирующую вероятность не выше, чем вероятность характеристики минимального типа.



с вероятностью $1/2 - p_a$ задействован один ключевой бит

с вероятностью $1/2 - p_b$ задействован один ключевой бит

Рис.4

Для вероятности этой восьмицикловой характеристики в соответствии с (1) имеем:

$$P_8 = 2 \cdot \left(4 \cdot p_a p_b p_c\right)^2 = 2^5 \cdot \left(p_a p_b p_c\right)^2.$$

Для 16 цикловой характеристики, получающейся при итеративном продолжении минимальной 8-цикловой характеристики, соответственно получим:

$$P_{16} = 2 \cdot P_8^2 = 2^{11} \cdot \left(p_a p_b p_c\right)^4.$$

Следовательно, все показатели стойкости S блоков к атакам, использующим минимальные характеристики, определяются возможными значениями произведения трех вероятностей $p_a p_b p_c$, две из которых относятся к одному и тому же входу одного и того же S блока, а третья относится к другому S блоку, участвующему в формировании аппроксимационной характеристики. Действительно, из правил построения минимальной характеристики следует

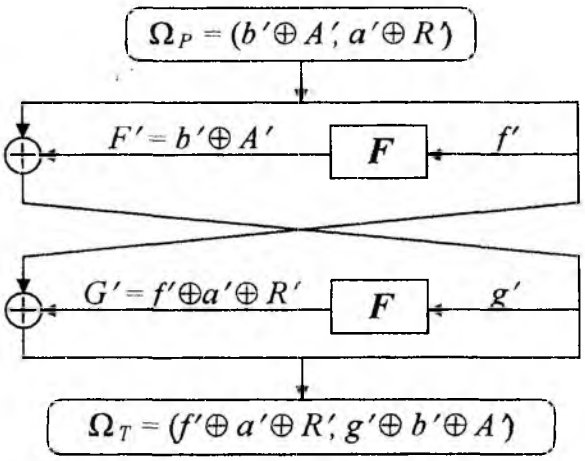
$$A' = F(a'),\ C' = F(a') \text{ и при этом } B' = F(b'), \text{ где } b' = A' \oplus C'. \qquad (2)$$

Но тогда, чтобы защититься от атаки, использующей минимальную характеристику, достаточно выбрать таблицы S блоков, исходя из условия, что максимально возможное значение произведения этих трех вероятностей удовлетворяет требованию

$$\left(P_{16}\right)^2 \le 2^{-55} \rightarrow p_a p_b p_c \le \sqrt[8]{2^{-77}} = 2^{-10}. \qquad (3)$$

Заметим здесь, что $2^{-10} = \dfrac{256}{64^3}$.

Теперь можно перейти к формированию критериев отбора S блоков устойчивых к атакам линейного криптоанализа. Его можно сформулировать как критерий для проверки линейных аппроксимационных таблиц для 3-цикловых аппроксимационных характеристик (являющихся первыми тремя циклами 8-цикловой минимальной характеристики).



с вероятностью $1/2 - p_f$ задействован один ключевой бит

с вероятностью $1/2 - p_g$ задействован один ключевой бит

Рис.5

Основная идея построения такой 3-цикловой аппроксимационной характеристики, как следует из соотношений (2), состоит в использовании двух наиболее вероятных одноблочных характеристик для одних и тех же однобитных входов (одного и того же S блока), разнесенных на один цикл. При этом сумма по модулю два выходов этих S блоков должна образовывать однобитный вход другого S блока на промежуточном цикле, также участвующего в формировании аппроксимационной характеристики. Выход этого S блока в свою очередь должен совпадать с входами разнесенных S блоков.

Тогда критерии для отбора таблиц S блоков, позволяющих защититься от атак, построенных на использовании таких 3-цикловых характеристик, можно сформулировать следующим образом.

Требование. Для обеспечения устойчивости шифра DES к известным атакам линейного криптоанализа необходимо и достаточно, чтобы максимальное значение произведения вероятностей $p_a p_b p_c$ одноцикловых характеристик, соответствующих формам $\left(\Omega_P^A, \Omega_T^A, \Omega_K^A, 1/2 + p_a\right)$, $\left(\Omega_P^B, \Omega_T^B, \Omega_K^B, 1/2 + p_b\right)$ и $\left(\Omega_P^C, \Omega_T^C, \Omega_K^C, 1/2 + p_c\right)$, где $\Omega_P^A = \left(A', 0\right) \rightarrow \Omega_T^A = \left(A', a'\right)$, $\Omega_P^B = \left(a', A'\right) \rightarrow \Omega_T^B = \left(a', A' \oplus b'\right)$, $\Omega_P^C = \left(A' \oplus b', a'\right) \rightarrow \Omega_T^C = \left(A' \oplus b', 0\right)$ (т.е. выполняются ограничения (2)), было меньше порогового значения 2^{-10}.

Фактически это трехцикловая характеристика вида $\left(\Omega_P^1, \Omega_T^3, \Omega_K^A \oplus \Omega_K^B \oplus \Omega_K^C, 1/2 + 4 p_a p_b p_c\right)$, для которой выполняется условие: при $\Omega_P^1 = \Omega_P^A = \left(A', 0\right)$ имеем $\Omega_T^3 = \Omega_T^C = \left(A' \oplus b', 0\right)$.

Приведем несколько замечаний относительно самой методики выполнения проверок.

Базовый метод, разработанный Мацуи и развитый Бихамом, использует одноблочные характеристики. Это значит, что вход b' является однобитным, так как в соответствии с (2) имеем $b' = F(a') \oplus F(c')$ – это сумма по модулю 2 выходных битов одного и того же S блока ($a' = c'$). А в соответствии с правилом завершающей цикловую функцию P подстановки выходы идентичных S блоков могут сформировать вход в один из S блоков только в случае, когда сумма $F(a') \oplus F(c')$ есть один единственный бит. Очевидно также, что должны быть однобитными и входы a' и соответственно $c' = a'$. В результате анализу подлежат все однобитные входные «маски» линейных аппроксимационных таблиц S блоков (индексы по строкам): 1_x, 2_x, 4_x, 8_x, 10_x и 20_x.

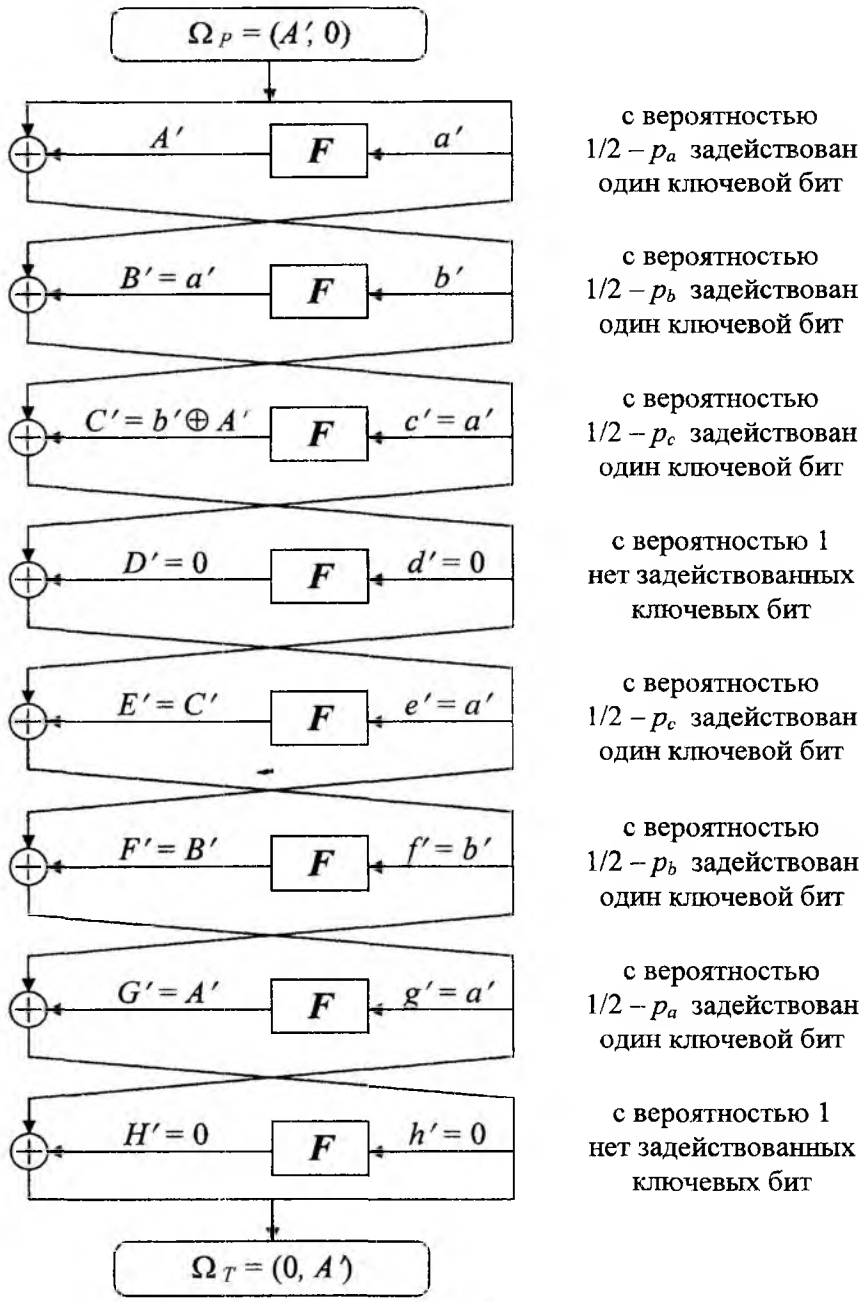


Рис. 6

Заметим, что в соответствии с принципами построения S блоков DES крайние пары битов 6-битных входов каждого из S блоков являются одновременно и входными битами соседних S блоков. Представленные варианты значений входов 1_x, 2_x, 4_x, 8_x, 10_x и 20_x (000001, 000010, 000100, 001000, 010000, 100000), тем не менее, всегда активизируют лишь один из S блоков, так как для входов (входных «масок») 1_x и 20_x линейных аппроксимационных таблиц S блоков все значения выходов (линейных аппроксимационных таблиц) являются нулевыми (для любой маски выходных значений одна и та же подстановка дает сбалансированный результат для числа одинаковых выходов), т.к. биты, соответствующие этим двум «маскам»,осуществляют выбор одной из четырёх перестановок, которые составляют S блок (уравновешенный результат для всех множеств битов, высекаемых из всех элементов перестановки).

При осуществлении проверки следует учитывать только те пары ячеек таблицы, для которых при одинаковой входной «маске» сумма по модулю два выходных «масок» даёт один единственный бит. Это пары входов по столбцам 1_x и 3_x, 1_x и 5_x, 1_x и 9_x, 2_x и 3_x, 2_x и 6_x, 2_x и A_x, 3_x и 7_x, 3_x и B_x, 4_x и 5_x, 4_x и 6_x, 4_x и C_x, 5_x и 7_x, 5_x и D_x, 6_x и 7_x, 6_x и E_x, 7_x и F_x, 8_x и 9_x, 8_x и A_x, 8_x и C_x, 9_x и

B_x, 9_x и D_x, A_x и B_x, A_x и E_x, B_x и F_x, C_x и D_x, C_x и E_x, D_x и F_x – всего 27 вариантов. Из этого числа сразу исключаются пары, для которых абсолютное значение хотя бы одной из ячеек аппроксимационной таблицы меньше или равно 2 (результирующая вероятность $p_a p_b p_c$ для этих случаев выходит за допустимые границы). Для отобранных пар определяются значения p_a и p_c, вычисляется соответствующее им значение $b' = F(a') \oplus F(c')$ и S блок для этого входа, затем по аппроксимационным таблицам находится значение p_b, при котором $F(b') = a'$.

Найденные значения p_a, p_b, p_c и проверяются на соответствие установленному критерию. При самом пессимистическом подходе всего потребуется выполнить $27 \cdot 4 \cdot 8 = 864$ проверок. На самом деле их будет на много меньше.

Разработанные критерии отбора таблиц подстановок устойчивых к атакам линейного криптоанализа были применены к таблицам S блоков группы Кванджио Ким, приведенным в [4]. Наши комментарии в отношении дифференциального криптоанализа представлены в начале статьи. Что касается линейного криптоанализа, то по результатам нашей проверки они полностью удовлетворяют выдвинутым в работе критериям отбора S боков.

Список литературы: **1.** *Лисицкая И.В., Головашич С.А., Олешко О.И., Олейников Р.В., Коряк А.С.* Построение таблиц подстановок для стандарта шифрования данных // Проблемы бионики. 1999. Вып.50. С. 185–194. **2.** *Лисицкая И.В., Олейников Р.В., Головашич С.А., Коряк А.С., Олешко О.И.* Анализ стойкости DES подобных алгоритмов шифрования при использовании таблиц подстановок случайного типа // Радиоэлектроника и информатика 1999. № 1. Стр 111–114. **3.** *Eli Biham* On Matsyi's Linear Cryptanalysis. Technion – Comput Science Department -Technic Report CS0813 - 1994. P 1-17. **4.** *Schneier B.* 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. **5.** *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.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06: 519.248.681

И.В. ЛИСИЦКАЯ, канд. техн. наук, А.С. КОРЯК

УТОЧНЕННЫЕ КРИТЕРИИ ОТБОРА ТАБЛИЦ ПОДСТАНОВОК С ЗАДАННЫМИ ХАРАКТЕРИСТИКАМИ СЛУЧАЙНОСТИ

В работах [1,2] предложены критерии отбора случайных таблиц подстановок, решающих задачу построения долговременных ключей для алгоритма шифрования ГОСТ 28147-89. Использование числовых конструкций типа подстановок характерно и для ряда других симметричных шифров. Конечно, при рассмотрении других шифров возникает необходимость вводить дополнительные ограничения и правила проверки [3]. Вместе с тем, по мере накопления опыта по применению, развиваемого в отмеченных работах подхода, появились дополнительные соображения и аргументы по обоснованию и уточнению некоторых параметров и правил отбора случайных таблиц подстановок, введенных ранее. В этой работе мы изложим ряд из сформулированных в [1] положений в новой редакции и с большей детализацией, а также приведем результаты статистического моделирования предлагаемых процедур отбора.

Напомним прежде всего, что под случайными здесь понимаются подстановки и таблицы подстановок, относящиеся к множеству наиболее вероятных случайных подстановок и случайных таблиц подстановок, и процедура проверки заключается в отбраковке подстановок и таблиц подстановок не входящих в это множество. Результаты анализа и статистических экспериментов по использованию критериев отбора, сформулированных в [1], говорят, однако, о том, что рассматриваемое в [1] множество допустимых подстановок и таблиц подстановок может быть расширено. Так в [4] показано (утверждается), что вполне приемлемыми характеристиками статистической безопасности обладают одно-цикловые подстановки и таблицы, составленные из одно-цикловых подстановок. Совершенно неоправданно исключать из рассмотрения и множество подстановок противоречивого типа [5]. Действительно, противоречивые подстановки, т.е. подстановки, не имеющие совпадающих элементов, должны считаться при составлении таблиц подстановок наиболее предпочтительными, так как они полностью исключают тождественные переходы (ситуации, когда подстановка как бы не участвует в криптографическом преобразовании). Более того, таблиц, составленных только из одно-цикловых или противоречивых подстановок, оказывается вполне достаточно, чтобы удовлетворить требованиям их использования в качестве секретных параметров шифров, как это требуется, например, для алгоритма ГОСТ 28147-89. Результаты проверки таких таблиц по критериям статистической безопасности [6] также полностью подтверждают эффективность их применения в алгоритмах симметричного шифрования. Отмеченное и побудило выполнить коррекцию введенных в [1] критериев отбора таким образом, чтобы расширить допустимое множество подстановок и таблиц подстановок за счет включения в него одно-цикловых, противоречивых и близких к ним подстановок. В этой работе излагается такая уточненная система критериев отбора случайных таблиц подстановок, рассматриваются расчетные соотношения, лежащие в их основе, и приводятся результаты статистической проверки случайных таблиц подстановок, полученных с помощью разработанного программного комплекса генерации долговременных ключей для шифра ГОСТ 28147-89.

Прежде всего, кратко напомним основные идеи подхода, развиваемого в [1].

Система (набор) из m различных подстановок n-ой степени $S_{n,m}$ (таблица подстановок) в дальнейшем записывается в виде расширения традиционного представления подстановки за счет добавления новых строк, т.е. в виде матрицы

$$S_{m,n} = \begin{bmatrix} 1 & 2 & 3 & \dots & n \\ i_{11} & i_{12} & i_{13} & \dots & i_{1n} \\ i_{21} & i_{22} & i_{23} & \dots & i_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ i_{m1} & i_{m2} & i_{m3} & \dots & i_{mn} \end{bmatrix} \qquad (1)$$

Верхняя строка называется нулевой, а остальные нумеруются от 1-ой до m-ой. Рассматривается задача построения некоторого заданного числа таких таблиц подстановок, которые удовлетворяют определенным критериям отбора (для шифра ГОСТ 28147-89 – $n = 16$, $m = 8$).

Предлагается использовать показатели и критерии проверки случайности трех уровней. На первом уровне проверяется соответствие показателей "случайности" отдельно взятой подстановки свойствам случайной равновероятной подстановки. Подстановки, прошедшие первый уровень проверки считаются уже подстановками случайного типа.

На втором уровне проверки оценивается соответствие характеристик случайности системы подстановок, попавших в таблицу, свойствам среднестатистической таблицы случайных подстановок. Таблицы подстановок, прошедших первые два уровня проверки, считаются случайными таблицами подстановок.

На третьем уровне осуществляется оценка степени подобия различных таблиц подстановок, из которых отбираются уже таблицы, выступающие в качестве системы (набора) долговременных ключей.

Соответствующие критерии отбора подстановок, таблиц подстановок и множеств таблиц подстановок названы критериями отбора первого, второго и третьего уровней.

Сущность уточненных правил, с помощью которых предлагается осуществлять отбор случайных таблиц подстановок, состоит в следующем.

Критерии отбора подстановок первого уровня, т.е. критерии отбора подстановок по индивидуальным характеристикам случайности формулируются, как и ранее в виде трех требований.

Требование 1.1. Число инверсий η_n в подстановке степени n должно удовлетворять условиям

$$\left|\eta_n - \frac{n(n-1)}{4}\right| \le a\sigma_\eta, \sigma_\eta = n^{3/2}/6 .$$

Требование 1.2. Число циклов ξ_n в подстановке степени n должно удовлетворять условиям

$$\xi_n \le \ln n + a\sigma_\xi, \sigma_\xi = \sqrt{\ln n} .$$

Требование 1.3. Число возрастаний θ_n в подстановке степени n должно удовлетворять условиям

$$\left|\theta_n - \frac{n}{2}\right| \le a\sigma_\theta, \sigma_\theta = \sqrt{n/12} .$$

Здесь изменения коснулись только требования, определяющего допустимое число циклов. Двухстороннее ограничение в этом требовании заменено односторонним, что позволило включить в допустимое множество подстановок и одно-цикловые подстановки.

При формировании критериев отбора подстановок на втором уровне рассматриваются таблицы, составленные из подстановок, которые прошли первый уровень проверки. Предлагаемая методика строится на основе понятия противоречивости подстановок (числа несовпадений элементов).

Практически для каждой таблицы из m подстановок n-й степени формируется двумерный метрический "портрет", т.е. таблица определяется двумя "векторами".

В первом случае определяется конфигурация $(t_0, t_1, t_3, \ldots, t_n)$ совпадений элементов в $N_k = m \cdot (m-1)/2$ попарных декомпозициях строк этой таблицы подстановок. Элемент конфигурации t_i, $i = 1, 2, \ldots, n$ представляет собой число пар строк среди общего их числа N_k, которые имеют совпадающих элементов, так что $\sum_{i=0}^{n} t_i = N_k$. На множестве возможных исходов $\{t_0, t_1, \ldots, t_n\}$ определяется закон распределения вероятностей $P(t = i)$ для числа i совпадений элементов в паре равновероятных подстановок n-ой степени, $i = 0, 1, 2, \ldots, n$.

Во втором случае определяется конфигурация $(\varsigma_0, \varsigma_1, \varsigma_2, \ldots, \varsigma_{\lfloor m/2 \rfloor})$ совпадений элементов в столбцах таблицы. Здесь элемент конфигурации ς_s, $s = 0, 2, \ldots, \lfloor m/2 \rfloor$ – это число столбцов с s повторениями (в том числе и многократными) элементов столбца, при этом $\sum_{s=0}^{\lfloor m/2 \rfloor} \varsigma_s = n$. Затем на множестве возможных исходов $\left\{\varsigma_0, \varsigma_1, \varsigma_2, \ldots, \varsigma_{\lfloor m/2 \rfloor}\right\}$ определяется закон распределения вероятностей $P(\varsigma = s)$

для s повторений (в том числе многократных) различных элементов в столбце таблицы подстановок, $s = 0, 1, 2, \ldots, \lfloor m/2 \rfloor$. Здесь и ранее $\lfloor x \rfloor$ обозначает наибольшее целое число x, не превосходящее x.

На основе полученных законов распределения вероятностей для числа совпадений элементов по строкам и столбцам таблиц подстановок строится эталонный портрет случайной среднестатистической таблицы подстановок, и в дальнейшем осуществляется отбор таблиц подстановок по степени их близости к эталону (для ГОСТ эталонный метрический портрет имеет вид (2, 7, 6, 1, 0), (9, 11, 6, 2, 0, . . . , 0); для DES – это конфигурация (11, 5, 0) по столбцам и конфигурация (2, 3, 1, 0, . . ., 0) по совпадениям в парах строк).

Изменения для критериев для отбора таблиц подстановок на втором уровне коснулись некоторого уточнения самих правил проверки. Остановимся в связи с этим более подробно на идеях построения решающих правил, использованных ранее.

Напомним, что в работах [1,2] при построении правил отбора на втором уровне для оценки степени близости конфигурации совпадений в столбцах и строках проверяемой таблицы к эталонной был взят двухсторонний критерий Пирсона, основанный на использовании таблиц χ^2 распределения Стьюдента, для применения которого необходимо выполнить требования в отношении значений ожидаемых частот (они должны быть большими 10). Анализ показывает, что такие условия строго не выполняются для рассматриваемых в этой работе шифров. В этом случае, как отмечено в [7], предельное распределение χ^2, приведенное в соответствующих таблицах, как правило, не дает надежных результатов т.е. таблицами пользоваться не следует. Тем не менее, в [3] были приведены аргументы в пользу возможности применения в рассматриваемом случае критерия Пирсона и требования 5 и 6 в работе [1] построены именно на использовании критерия χ^2 в чистом виде.

В рассматриваемом случае предлагается отойти от применения критерия χ^2. Во первых, предлагается воспользоваться вместо двухстороннего критерия односторонним. Во вторых, – оценку близости конфигураций совпадений в строках и столбцах проверяемой таблицы подстановок к эталонной предлагается проводить не путем вычисления и сравнения величины χ^2 с граничным значением, а на основе непосредственного сопоставления числа совпадений элементов по отдельным их типам (разновидностям) в проверяемой таблице подстановок с модифицированной эталонной. При этом для включения в число допустимых таблиц подстановок предельного случая – латинских прямоугольников (таблиц, составленных из противоречивых подстановок) в обоих случаях разрешается максимально возможное число несовпадений.

В итоге требования по отбору таблиц подстановок на втором уровне проверки предлагается сформулировать в следующем виде:

Требование 2.1. (не обязательное) В таблицу подстановок должны входить подстановки, не имеющие совпадений с нулевой строкой (не имеющие циклов нулевой длины).

Требование 2.2. Подстановки, вошедшие в таблицу, должны по конфигурации $(t_0, t_1, t_3, \ldots, t_n)$ совпадений элементов в N_k попарных декомпозициях строк таблиц подстановок удовлетворять критерию $t_0 \le N_k$, $t_1 \le t'_1$, $t_2 \le t'_2$, . . . , $t_n \le t'_n$, где элементы конфигурации t_i, $i = 1, 2, \ldots, n$ представляет собой число пар строк из общего их числа N_k, которые имеют i совпадающих элементов (для ГОСТ 28147-89 модифицированный эталон – $(t'_0, t'_1, t'_3, \ldots, t'_n) = (28, 11, 6, 2, 0, \ldots, 0)$), при этом $\sum_{i=0}^{n} t_i = N_k$.

Требование 2.3. Подстановки, вошедшие в таблицу, должны по конфигурации $(\varsigma_0, \varsigma_1, \varsigma_2, \ldots, \varsigma_{\lfloor m/2 \rfloor})$ совпадений элементов в столбцах таблицы подстановок, удовлетворять критерию $\varsigma_0 \le n$, $\varsigma_1 \le \varsigma'_1$, . . . , $\varsigma_{\lfloor m/2 \rfloor} \le \varsigma_{\lfloor m/2 \rfloor}$, где элементы конфигурации ς_s, $s = 0, 2, \ldots, \lfloor m/2 \rfloor$ – это числа столбцов с s повторениями (в том числе и многократными) элементов столбца (для ГОСТ 28147-89 модифицированный эталон – $(\varsigma'_0, \varsigma'_1, \varsigma'_2, \varsigma'_3, \varsigma'_4) = (16, 7, 6, 1, 0)$), удовлетворяющие ограничению $\sum_{s=0}^{\lfloor m/2 \rfloor} \varsigma_s = n$.

При формировании критериев отбора таблиц на третьем уровне проверки используется идея наложения таблиц и подсчета числа совпадающих элементов. Для этого случая также рассчитан

теоретический закон распределения вероятностей для числа совпадений элементов в паре наложенных таблиц подстановок и определены его числовые характеристики.

В результате требования по отбору таблиц подстановок на третьем уровне проверки сформулированы следующим образом.

Требование 3.1. Множество таблиц подстановок, используемых в качестве долговременных ключей, должно при всех попарных наложениях таблиц давать число совпадающих элементов q удовлетворяющее условиям $q \le m + \sqrt{m}$.

При формулировке этого требования в отличие от [1] также использовано одностороннее ограничение по максимуму.

Требование 3.2. Можно запретить и совпадение строк, стоящих на различных позициях таблиц если потребовать в записанном выше правиле, что для каждой пары таблиц процедура наложения выполняется со всеми циклическими подстановками строк одной из них.

Приведем основные расчетные соотношения, определяющие правила и показатели отбора случайных таблиц подстановок на втором и третьем уровнях проверки.

Закон распределения вероятностей $P(t = i)$ совпадений i элементов в паре равновероятных подстановок n-ой степени (в паре наложенных строк таблицы подстановок), используемый в формулировке требования 2.2, выражается через известное в комбинаторике [8] число беспорядков D_{ni} в виде

$$P(t = i) = 1 - \frac{D_{ni}}{n!} = 1 - \frac{1}{i!}\sum_{k=0}^{n-i}(-1)^k \frac{1}{k!}, i = 1,2,...,n \qquad (2$$

и оказывается близким к биномиальному.

Расчеты, выполненные по формуле (2), при $n = 16$, $m = 8$ иллюстрирует таблица 1.

Таблица 1

Число совпадений t	Вероятность $P(t = i)$
0	0,3316
1	$3{,}79 \cdot 10^{-1}$
2	$2 \cdot 10^{-1}$
3	$6{,}76 \cdot 10^{-2}$
4	$1{,}57 \cdot 10^{-2}$
5	$2{,}69 \cdot 10^{-3}$
6	$3{,}53 \cdot 10^{-4}$
7	$3{,}59 \cdot 10^{-5}$
8	$2{,}9 \cdot 10^{-6}$
9	$1{,}77 \cdot 10^{-7}$
⋮	⋮

Итоговое выражение для определения вероятности $P(\varsigma = s)$ числа s различных повторяющихся (в том числе и многократно) элементов в столбце таблицы (1) получено в [2]. Мы здесь его напомним

$$P(\varsigma = s) = \sum_{k_1=0}^{\left\lfloor \frac{m-2s}{s} \right\rfloor} \sum_{k_2=0}^{\left\lfloor \frac{m-s(k_1+2)}{s-1} \right\rfloor} \sum_{\cdots} \cdots \sum_{k_s=0}^{m-s(k_1+2)-(s-1)k_2-(s-2)k_3-\ldots-2k_{s-1}} P_{2+k_1,2+k_1+k_2,\ldots,2+k_1+k_2+\ldots+k_s} \qquad (3$$

В этом выражении $P_{i,j,\ldots,l}$ – вероятности композиций совпадений по $i,j,\ldots,l$ различных элементов. Например, расчетное соотношение для вероятности $P_{i,j,\ldots,l}$ в случае $i = 2 + k_1, j = 2 + k_1 + k_2$ $l = 2 + k_1 + k_2$, имеет вид

$$P_{2+k_1,2+k_1+k_2,2+k_1+k_2} = \frac{C_m^{2+k_1} C_{m-2-k_1}^{2+k_1+k_2} C_{m-4-2k_1-k_2}^{2+k_1+k_2} (n-1)_{m-3k_1-3-2k_2}}{2!\cdot(n-1)^m}, \tag{4}$$

$$k_1 = 0,1,\ldots,\left\lfloor \frac{m-6}{3} \right\rfloor k_2 = 1,2,3,\ldots,\left\lfloor \frac{m-3(2+k_1)}{2} \right\rfloor.$$

Здесь $(n-1)_m$ – m-размещение из $n-1$ элементов

$$(n-1)_m = \begin{cases} (n-1)\ (n-2)\cdots(n-m-2), m \le n-1 \\ 0, m > n-1 \end{cases} \tag{5}$$

В таблице 2 представлен закон распределения вероятностей $P(\varsigma = s)$ для значений $n = 16$, $m = 8$, рассчитанный по формулам (2)-(4).

Таблица 2

Число повторений s	Вероятность $P(\varsigma = s)$
0	0,1012
1	0,4433
2	0,3843
3	0,0698
4	0,0014

Этот закон, как показывает анализ, достаточно точно аппроксимируется в дискретных точках отсчетами функции нормального распределения.

Комбинаторные соображения позволяют представить записать и итоговое выражение для вероятности $P_k^{(n,m)}$ совпадения k элементов в m строках пары наложенных друг на друга таблиц типа (1)

$$P_k^{(n,m)} = \sum_{\substack{\sum_{i=1}^{m} k_i = k \\ \sum_{j=1}^{s} l_j = m}} C_m(l_1, l_2, \ldots, l_s) \prod_{i=1}^{m} P(t = k_i). \tag{6}$$

В этом выражении k_i, – число совпадающих элементов в i-той (i = 1, 2, . . . , m) паре строк рассматриваемой таблицы (композиции $(k_1, k_2, \ldots, k_8)$); l_{ij}, – число одинаковых k_i j-го типа, т.е. значений k_i, имеющих одно и то же число совпадений j = 1, . . . , s, s – число различающихся наборов совпадающих значений k_i, ($s \le m$).

В (6) использовано также обозначение $C_m(l_1, l_2, \ldots, l_s)$ – полиномиальный коэффициент, который определяется следующим образом [9]:

$$C_m(l_1, l_2, \ldots, l_s) = \frac{m!}{l_1! l_2! \ldots l_s!}.$$

Очевидно, что для вероятностей $P_k^{(n,m)}$, $k = 0,1,\ldots,nm$ должно выполняться условие

$$\sum_{k=0}^{mn} P_k^{(m)} = 1.$$

Расчеты, выполненные по формулам (6) и (2), иллюстрируют таблица 3 и таблица 4. Вторая таблица отличается от первой тем, что при ее построении учитывалось выполнение требования 2.1 (в таблицу подстановок не должны входить подстановки, элементы которых совпадают с соответствующими элементами нулевой строки).

Изложенные выше подходы к формированию методов и критериев отбора случайных подстановок были положены в основу разработки программного комплекса генерации и сертификации долговременных ключей для алгоритма ГОСТ-28147-89.

Таблица 3

Число совпадений k	Вероятность $P_k^{(16,8)}$
0	0,000258
1	0,002205
2	0,009335
3	0,026139
4	0,054455
5	0.090033
6	0,123045
7	0,142967
8	0,144158
9	0,128141
10	0,101658
11	0,072701
12	0,047256
13	0,0288111
14	0,015394
15	0,007800
16	0,003672
17	0,001613
18	0,000663
19	0,000256
20	0,000093
21	0,000032
22	0,000010
23	0,000003
24	0,000001
25	0,000000
⋮	⋮
сумма	1,000000

Таблица 4

Число совпадений k	Вероятность $P_k^{(16,8)}$
0	0,000146
1	0,001336
2	0,006059
3	0,018178
4	0,040577
5	0,071879
6	0,105251
7	0,131027
8	0,141556
9	0,134815
10	0,114593
11	0,087805
12	0,061150
13	0,038975
14	0,022868
15	0,012414
16	0,006262
17	0,002947
18	0,001298
19	0,000537
20	0,000209
21	0,000077
22	0,000027
23	0,000009
24	0,000003
25	0,000001
⋮	⋮
Сумма	1,000000

Некоторые результаты статистической проверки процедур отбора случайных подстановок и случайных таблиц подстановок с заданными характеристиками случайности представлены в таблицах 5÷8.

В таблице 5, приводится пример обработки файла статистики при отбраковке подстановок на первом уровне проверки. Численные значения показателей случайности подстановок установлены соответственно равными: для инверсий – 49-71, для возрастаний – 7-9, для циклов – 1-5. В правой колонке таблицы приводится количество подстановок в процентах от их общего числа, попавших в установленные границы.

Таблица 5

Проверяемый показатель случайности	Попало в интервал в %
инверсий	66%
возрастаний	77%
циклов	99%
инверсий и возрастаний	53%
инверсий и циклов	66%
возрастаний и циклов	76%
инверсий, возрастаний и циклов	53%

Теоретическая оценка числа случайных подстановок для шифра ГОСТ 28147-89, прошедших границы, установленные в [1], выполнялась для асимптотически нормальных законов распределения вероятностей соответствующих параметров.

Напомним, что для вероятности события, заключающегося в том, что случайная величина x, распределенная по нормальному закону $p(x)$ с параметрами $m_x = 0$ и $\sigma_x^2 = 1$, попадет в интервал $|x| \le a$, справедливо соотношение

$$P(|x| \le a) = \int_{-\infty}^{\infty} p(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 2\Phi(a) - 1.$$

Тогда, при $a = 1$ (для границ, установленных в эксперименте) имеем результат $P(|x| \le 1) = 0{,}6826$.

Если считать, что все три рассматриваемые случайные величины η_n, ξ_n и θ_n статистически независимы, то для вероятности порождения случайной подстановки, удовлетворяющей одновременно трем критериям случайности, можно получить оценку

$$P(\eta'_n \le 1,\ \xi'_n \le 1,\ \theta'_n \le 1) = (P(|x| < 1))^3 = 0{,}318.$$

Полученное расчетное значение достаточно близко повторяет результат таблицы 5. Заметим здесь, что более тщательные вычисления (уточнение пределов интегрирования, и др.), позволяют добиться еще более близкого подтверждения результатов эксперимента.

Что касается второго и третьего уровней проверки, то здесь представляет интерес задача определения потенциально возможного числа различных случайных таблиц подстановок (например, числа различных долговременных ключей для алгоритма ГОСТ 28147-89), которые можно сформировать при заданных значениях n и m.

Оценим сначала число таблиц подстановок, удовлетворяющих требованию 2.3.

Комбинаторные соображения позволяют для фиксированного "эталона" $(\varsigma'_0, \varsigma'_1, \varsigma'_2, \varsigma'_3, \varsigma'_4)$, записать формулу для общего числа различных допустимых по требованию 2.3. конфигураций совпадений элементов в таблицах подстановок по столбцам, в виде

$$N^{(n,\square)}\left(\varsigma'_0, \varsigma'_1, \ldots, \varsigma'_{\lfloor m/2 \rfloor}\right) = \sum_{\substack{\sum_{i=0}^{\lfloor m/2 \rfloor} k_i = n,\, k_i \le \varsigma'_i}} C_n\left(k_0, k_1, \ldots, k_{\lfloor m/2 \rfloor}\right).$$

Здесь уже k_i – число столбцов с i повторениями элементов в таблице подстановок, функция $C_n(k_0, k_1, \ldots, k_r)$ – это опять полиномиальный коэффициент, т.е. при $\sum_{i=1}^{r} k_i = n$

$$C_n(k_0, k_1, \ldots, k_r) = \frac{n!}{k_0! k_1! \cdots k_r!}. \qquad (7)$$

Если интересоваться только числом разрешенных конфигураций совпадений элементов в столбцах таблицы подстановок $\left(\varsigma_0, \varsigma_1, \varsigma_2, \ldots, \varsigma_{\lfloor m/2 \rfloor}\right)$, удовлетворяющих ограничениям $\varsigma_0 \le n$, $\varsigma_1 \le \varsigma'_1$, $\ldots$, $\varsigma_{\lfloor m/2 \rfloor} \le \varsigma'_{\lfloor m/2 \rfloor}$, то можно ввести вспомогательную числовую функцию

$$\psi_n(k_0, k_1, \ldots, k_r) = \begin{cases} 1, \text{ при } \sum_{i=0}^{r} k_i = n,\ k_i \le \varsigma'_i, \\ 0 \text{ в остальных случаях.} \end{cases}$$

Тогда формулу для определения размерности множества допустимых конфигураций совпадений элементов в столбцах таблицы подстановок можно представить в виде

$$N^{(n,m)}_{\left(\varsigma'_0, \varsigma'_1, \ldots, \varsigma'_{\lfloor m/2 \rfloor}\right)} = \sum_{\sum_{i=0}^{\lfloor m/2 \rfloor} k_i = n} \psi\left(k_0, k_1, \ldots, k_{\lfloor m/2 \rfloor}\right).$$

Для шифра ГОСТ 28147-89 (n = 16, m = 8) и модифицированного эталона $(\varsigma'_0,\varsigma'_1,\varsigma'_2,\varsigma'_3,\varsigma'_4)$ = (16 7, 6, 1, 0) вычисления с помощью ЭВМ приводят к результату

$$N^{(16,8)}_{(k_0,k_1,..,k_4)} = 112.$$

Заметим, что общее число возможных для заданных значений n и m различных вариантов совпадений элементов в столбцах таблицы подстановок можно получить, воспользовавшись свойствами полиномиальных коэффициентов [9],

$$\sum_{k_0+k_1+\ldots+k_r=n} C_n(k_0,k_1,\ldots,k_r) = (r+1)^n.$$

Так, при $r = \lfloor m/2 \rfloor = 4$ и $n = 16$ всего возможно $5^{16} \approx 1{,}53\cdot10^{11}$ различных вариантов конфигураций совпадений, из которых допустимыми являются только 112.

Для вероятности того, что произвольно взятая таблица удовлетворит оговоренным правилам проверки (требованию 2.3), соответственно можно записать выражение

$$P_c^{(m,n)} = \sum_{\substack{\sum_{i=0}^{\lfloor m/2 \rfloor} k_i = n,\, k_i \le \varsigma'_i}} C_n\left(k_0,k_1,\ldots k_{\lfloor m/2 \rfloor}\right) \prod_{s=0}^{\lfloor m/2 \rfloor} \left(P(\varsigma = s)\right)^{k_s}.$$

Расчеты, выполненные по этой формуле для параметров ГОСТ 28147-89, приводят к результату $P_c^{(m,n)} = 0{,}0781$, т.е. требованию 2.3 удовлетворяют около 10% всех таблиц подстановок (заметим, что в принципе можно получить более 10^{82} различных таблиц в пределах только одной конфигурации (2, 7, 6, 1, 0)).

Будем теперь интересоваться числом таблиц подстановок, удовлетворяющих требованию 2.2.

Выполним оценку ожидаемого числа таблиц подстановок, прошедших проверку по совпадениям элементов в парах строк. Как и в предыдущем случае, рассмотрим модифицированную эталонную конфигурацию $(t'_0,t'_1,t'_3,\ldots,t'_n)$ = (28, 11, 6, 2, 0, . . . , 0). Для общего числа возможных вариантов выбора таблицы с конфигурацией совпадений, удовлетворяющих требованию 2.2, можем записать выражение

$$N^{(n,m)}\left(t'_0,t'_1,\ldots t'_{\lfloor \frac{m}{2} \rfloor}\right) = \sum_{\substack{\sum_{i=0}^{n} q_i = \frac{m(m-1)}{2},\, q_i \le t'_i}} C_{\frac{m(m-1)}{2}}\left(q_0,q_1,\ldots,q_{\lfloor m/2 \rfloor}\right).$$

Расчеты общего числа различных конфигураций совпадений, выполненные по аналогии с предыдущим случаем, приводят к результату $N^{(16,8)}_{(t'_0,t'_1,..t'_4)} = 144$.

Для вероятности получения таблицы с конфигурацией совпадений в парах строк, удовлетворяющей требованию 2.2, здесь в рамках введенных выше обозначений можем записать выражение

$$P_s^{(m,n)} = \sum_{\substack{\sum_{i=0}^{\lfloor m/2 \rfloor} q_i = \frac{m(m-1)}{2},\, q_i \le t'_i}} C_{\frac{m(m-1)}{2}}\left(q_0,q_1,\ldots,q_{\lfloor m/2 \rfloor}\right) \prod_{i=0}^{\lfloor m/2 \rfloor} \left(P(t = i)\right)^{q_i}.$$

В этом случае для параметров ГОСТ 28147-89 приходим к результату $P_s^{(m,n)} = 0{,}0081$, т.е. требованию 2.2 удовлетворяют около 1% всех таблиц подстановок. Оценить реальные показатели отбраковки подстановок на втором уровне проверки позволяют также результаты статистического моделирования этого этапа проверки, представленные в таблице 6.

Таблица 6

Сгенерировано 132 таблицы, из них:	Число (%)
принято алгоритмом	10 (7.6%)
отброшено по совпадениям в столбцах	73 (55.3%)
отброшено по совпадениям в парах строк	111 (84.1%)

Из приведенных данных следует, что после второго уровня проверки остается $\sim 0(10^{82})$ различных таблиц, из которых на третьем уровне проверки будут уже отбираться варианты для реализации нужного количества долговременных ключей.

Результаты экспериментальных исследований по отбору таблиц подстановок на третьем уровне проверки приведены в таблице 7.

Сопоставление результатов таблицы 3 и таблицы 7 свидетельствуют о достаточно высоком совпадении экспериментальных результатов с расчетными.

Проверка статистической безопасности шифра ГОСТ 28147-89, выполненная по методике, изложенной в [6], практически повторила ранее полученные результаты и выводы.

Таблица 7

Количество k совпадений элементов в парах наложенных таблиц подстановок	Количество (%) пар таблиц из общего их числа 338182, имеющих k совпадений элементов	Эмпирический закон распределения вероятностей числа k совпадений элементов
0	101 (0%)	0,0003
1	770 (0%)	0,0023
2	3199 (1%)	0,0095
3	8826 (3%)	0,0261
4	17963 (5%)	0,0531
5	29237 (9%)	0,0864
6	39552 (12%)	0,1169
7	46215 (14%)	0,1455
8	46908 (14%)	0,1387
9	42706 (13%)	0,1263
10	34609 (10%)	0,1023
11	25928 (8%)	0,0767
12	17632 (5%)	0,0521
13	11134 (3%)	0,0329
14	6452 (2%)	0,0190
15	3568 (1%)	0,0105
...	...	...

В таблице 8 представлен пример долговременного ключа для шифра ГОСТ 28147-89, построенного с помощью предлагаемой методики. В нижних строках этой таблицы представлены значения числа совпадений в столбцах таблицы подстановок.

Таблица 8

№	Инв.	Возр.	Цикл.	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1	50	8	2	4	2	7	5	9	1	0	8	E	3	B	C	D	7	A	6
2	76	6	2	C	9	F	E	8	1	3	A	2	7	4	D	6	0	B	5
3	75	8	3	D	8	E	C	7	3	9	A	1	5	2	4	6	F	0	B
4	70	7	1	E	9	B	2	5	F	7	1	0	D	C	6	A	4	3	8
5	59	7	1	3	E	5	9	6	8	0	D	A	B	7	C	2	1	F	4
6	71	9	3	8	6	F	B	1	9	C	5	D	3	7	A	0	E	2	4
7	55	10	3	9	B	C	0	3	6	7	5	4	8	E	F	1	A	2	D
8	59	7	3	C	6	5	2	B	0	9	D	3	E	7	A	F	4	1	8
Ср.	64,37	7,75	2,25	1	2	2	1	0	1	3	3	0	1	1	2	1	1	1	2

Представляется, что изложенные результаты могут стать основой разработки общей методики построения таблиц подстановок, которые используются в симметричных шифрах.

Список литературы: **1.** *Горбенко И.Д., Лисицкая И.В.* Критерии отбора случайных таблиц подстановок для алгоритма шифрования по ГОСТ 2847-89 // Радиотехника. Всеукр. межвед. науч.-техн. сб. 1997. Вып. 103. С. 121–130. **2.** *Бильчук В.М., Лисицкая И.В.* Информационно-управляющие системы на железнодорожном транспорте. 1998. № 1. С. 10–17. **3.** *Лисицкая И.В., Головашич С.А., Олешко О.И., Олейников Р.В., Коряк А.С.* Построение таблиц подстановок для стандарта шифрования данных // Проблемы бионики. 1999. Вып. 50. С. 185-194. **4.** *Кононова И.В.* Оценка и анализ подмножества одно-цикловых подстановок // Информационные системы: Сб. научн. тр. - Харьков: НАНУ, ПАНУ, ХВУ, 1994. С. 37–46. **5.** *Кононова И.В.* Противоречивые подстановки в алгоритме ГОСТ 28147-89 // Информационные системы: Сб. научн. тр. Харьков: НАНУ, ПАНУ, ХВУ. 1995. С. 70–77. **6.** *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. **7.** *Крамер Г.* Математические методы статистики: Пер. с англ. - М.: ГИИЛ, 1948. - 631 с. **8.** *Математическая* энциклопедия: В 5 т. / Гл. ред. Виноградов И.М. - М.: Советская энциклопедия, 1979. - Т.2: Д-КОО. - 278 с. **9.** *Бронштейн И.Н. Семендяев К.А.* Справочник по математике для инженеров и учащихся Втузов, - М.: Наука, 1980. - 976 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06

С.А. ГОЛОВАШИЧ

КЛЮЧЕВЫЕ ГРУППЫ В АТАКАХ ДИФФЕРЕНЦИАЛЬНОГО КРИПТОАНАЛИЗА DES-ПОДОБНЫХ ШИФРОВ

В данной работе нас будут интересовать только принципы построения характеристик, используемых в атаках дифференциального криптоанализа (далее дифференциальных характеристик либо просто характеристик) DES-подобных шифров (основанных на «подстановке с расширением»), и оценка вероятности их осуществления. Вначале дадим базовые определения и рассмотрим классический подход к построению и оценке вероятностей дифференциальных характеристик, предложенный в открытой печати Эли Бихамом. Изложение материала будет выполняться на примере алгоритма DES, так как он наиболее хорошо изучен, и успешная атака дифференциального криптоанализа впервые была предложена именно для этого криптоалгоритма. Заметим, однако, что предлагаемый подход может быть применён и к другим шифрам, имеющим сходную структуру

Алгоритм DES построен на базе 16-цикловой цепи Фестеля. Основу алгоритма составляет цикловая функция (F-функция), которая включает в себя последовательность из 4-х базовых операций: расширения E ($32 \rightarrow 48$), сложение по модулю 2 с 48-битным подключом, ключенезависимое нелинейное преобразование (табличная подстановка $8 \times S(6 \rightarrow 4)$), перестановка бит P [1].

Далее входные и выходные воздействия любого функционального блока шифра будем рассматривать как бинарные вектора, а под разностью этих векторов (либо их изменением) будем понимать операцию сложение по модулю 2 (XOR).

Введём ряд понятий и определений из теории дифференциального криптоанализа [2].

Взаимосвязь входной и выходной разностей вида: определённое изменение данных на входе некоторого функционального блока либо фрагмента шифра, с некоторой вероятностью, вызывает фиксированное изменение на выходе ($\Delta X \rightarrow \Delta Y$), будем называть дифференциальной характеристикой.

В зависимости от охватываемого функционального блока будем различать одноблочные (состоящие из одного S-блока), многоблочные (состоящие из группы смежных S-блоков), одноцикловые (охватывающие один цикл) и многоцикловые (охватывающие несколько смежных циклов) характеристики.

Под вероятностью одноцикловой дифференциальной характеристики $p(\Delta Y \setminus \Delta X)$ будем понимать вероятность перехода входной разности ΔX в выходную разность ΔY, т.е. вероятность выполнения соотношения $F(X \oplus \Delta X) = Y \oplus \Delta Y$, где $Y = F(X)$.

Под вероятностью n-цикловой дифференциальной характеристики будем понимать вероятность последовательного выполнения цепочки из n определённых одноцикловых характеристик. Вероятность многоцикловой характеристики определяется произведением вероятностей одноцикловых характеристик её составляющих.

Цикл шифрования либо отдельный S-блок считается активным, если на его вход подаётся разность отличная от нуля ($\Delta X \neq 0$), в противном случае цикл (либо S-блок) считается пассивным ($\Delta X = 0$) и с вероятностью $p = 1$ сохраняет значение своего выхода ($\Delta Y = 0$).

Стойкость DES-подобных шифров к атакам дифференциального криптоанализа определяется свойствами используемых нелинейных преобразований (S-блоков). Авторы дифференциального криптоанализа Ади Шамир и Эли Бихам предложили для оценки свойств каждого из 8 S-блоков, содержащихся в алгоритме, воспользоваться так называемой таблицей распределения битовых разностей. Эта таблица имеет организацию 64×16, где первая координата (номер строки) соответствует входной разности ΔX (каждый S-блок имеет 6-ти разрядный вход), а вторая (номер столбца) соответствует выходной разности ΔY (каждый S-блок имеет 4-х разрядный выход). Значение каждой ячейки таблицы равно количеству входных векторов X, для которых $S(X \oplus \Delta X) = Y \oplus \Delta Y$, где $Y = S(X)$, при вариации по всем возможным X. Отношение этого значения к общему числу возможных входных векторов ($2^6 = 64$) соответствует вероятности выполнения некоторой одноблочной характеристики $S(\Delta X) \rightarrow \Delta Y$. В свою очередь вероятность одноцикловой характеристики вычисляется как произведение вероятностей всех одноблочных характеристик её составляющих.

Атака дифференциального криптоанализа эффективна, если её сложность (величина обратная вероятности соответствующей полноцикловой характеристики) меньше чем сложность «силовой атаки» (прямого перебора ключей).

Рассмотрим правило «сшивки» (объединения) нескольких одноблочных характеристик в одноцикловую характеристику. Правило, использованное авторами дифференциального криптоанализа, учитывает только «сшиваемость» входных разностей. В соответствии с этим правилом для «сшивки» одноблочных характеристик двух соседних S-блоков необходимо чтобы два правых бита разности ΔX_1 левого S-блока совпадали с двумя левыми битами разности ΔX_2 правого S-блока: $\Delta X_1 = \{\Delta x_0, \Delta x_1, \Delta x_2, \Delta x_3, \Delta x_4, \Delta x_5\}$ $\Delta X_2 = \{\Delta x_4, \Delta x_5, \Delta x_6, \Delta x_7, \Delta x_8, \Delta x_9\}$. Это следует из того, что после выполнения расширения E и сложения с ключом (рис. 1), на входы каждой пары соседних S-блоков попадают два общих входных бита разности. Назовём это требование правилом «сшивки» по разности (или правилом динамической «сшивки»).

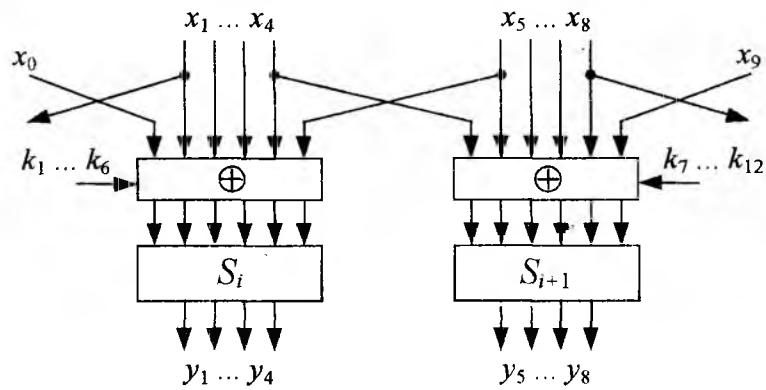


Рис. 1

Одноблочные характеристики, составляющие любую одноцикловую характеристику, всегда удовлетворяют указанному выше требованию. Вероятность результирующей одноцикловой характеристики авторы атаки вычисляют как произведение вероятностей всех одноблочных характеристик, участвующих в её построении, или как произведение вероятностей одноблочных характеристик активных S-блоков (на пассивных S-блоках происходит переход $0 \rightarrow 0$, вероятность которого всегда равна 1). Практически вероятность одноцикловой характеристики вычисляется как отношение произведения содержимого ячеек таблиц распределения дифференциальных разностей, соответствующих переходам (одноблочным характеристикам) активных S-блоков, к общему числу их возможных входных значений, т.е. 64^a, где a – количество активных S-блоков на данном цикле.

Приведенная методика вычисления вероятности одноцикловой характеристики, основана на допущении, что после сложения расширенного входного полублока данных с цикловым подключом (биты которого в пределах подключа независимы), входы любых двух S-блоков также становятся «статически» независимыми и поэтому на входах двух соседних S-блоков может возникнуть произвольное сочетание входных воздействий. Однако более тщательный анализ F-функции показывает, что использование в шифре DES сложения с подключом после E-расширения данных приводит к тому, что условия «сшивки» становятся зависимыми от ключевых битов. Далее будет показано, что в общем случае вероятность некоторой одноцикловой характеристики может принимать ряд дискретных значений, в зависимости от подключа шифрования, используемого в данном цикле, а величина, полученная по рассмотренной выше методике, соответствует среднему значению вероятности характеристики, при вариации по всем возможным вариантам подключа.

В первую очередь, отметим, что a соседних S-блоков имеют только $4 \times a + 2$ входных линий данных, и, следовательно, по ним можно подать только $2^{4 \times a + 2}$ (а не $2^{6 \times a}$) различных входных разностей. Далее нас будет интересовать характер взаимного влияния входов соседних S-блоков. Поэтому преобразуем классическую схему, приведенную на рис. 1 таким образом, чтобы сложение с ключом выполнялось до расширения E, как это показано на рис. 2, т.е. разделим преобразования, составляющие цикловую функцию F, на линейное (сложение входного блока с 32 битами подключа) и нелинейное (оставшаяся часть F-функции). Так как дифференциальные свойства алгоритма

пределяются соответствующими свойствами используемых нелинейных преобразований, то, ледовательно, нас будет интересовать «вторая» часть цикловой функции (перестановку P можно ключить из рассмотрения, так как её применение не влияет на ход дальнейших рассуждений).

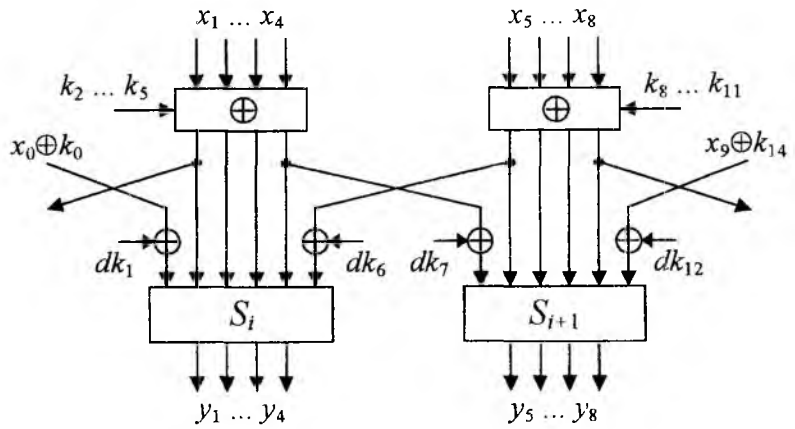


Рис. 2

В результате такого преобразования мы можем исключить из рассмотрения подключ в явном виде. Далее через x_i будем обозначать входные биты блока расширения E, полученные суммированием 32-х битного входного вектора с 32 битами подключа ($x_i := x_i \oplus k_j$). Оставшиеся 16 бит подключа будут участвовать в формировании 8 двухбитных векторов «ключевого смещения» dk – по одному вектору на каждую пару соседних S-блоков. Через dk_i обозначена сумма двух битов подключа, попадающих на входы соседних S-блоков, полученные (в результате расширения E) из одной линии данных, то есть:

$$dk_1 = k_1 \oplus k_0$$

$$dk_6 = k_6 \oplus k_8$$

$$dk_7 = k_7 \oplus k_5$$

$$dk_{12} = k_{12} \oplus k_{14}.$$

Из схемы, приведенной на рис. 2, видно, что если $X_1 = \{x_0, x_1, x_2, x_3, x_4, x_5\}$ и $X_2 = \{x_4', x_5', x_6, x_7, x_8, x_9\}$ – входные вектора соседних S-блоков, то

$$x_4' = x_4 \oplus dk_7$$

$$x_5' = x_5 \oplus dk_6$$

т.е. входные вектора двух соседних S-блоков по двум соответствующим разрядам всегда имеют фиксированную разность $\{dk_7, dk_6\} = \{k_7 \oplus k_5, k_6 \oplus k_8\}$, определяемую ключом шифрования. Назовём последнее свойство правилом «сшивки» по значению (или правилом статической «сшивки»).

Из последнего следует, что «сшивка» двух одноблочных характеристик возможна только в том случае, если пары одноблочных входных значений, участвующие в формировании одноблочных входных разностей, удовлетворяют правилу «сшивки» по значению, а, следовательно, и правилу «сшивки» по разности (правило динамической «сшивки» является частным случаем правила статической «сшивки»).

Исходя из всего выше сказанного, для вычисления вероятностей ключезависимых дифференциальных характеристик следует воспользоваться следующей методикой.

Значения в каждой ячейке «традиционной» таблицы дифференциальных разностей (описанной ранее), в общем случае, следует разбить на 2^4=16 ячеек соответствующих различным вариантам статической «сшивки», или, иначе говоря, к координатам входной ΔX и выходной ΔY разностей необходимо добавить координату X соответствующую фактическому значению 4 входных битов участвующих в «сшивке» (по 2 бита слева и справа).

Вероятность многоблочной характеристики, состоящей из a соседних S-блоков при некотором фиксированном ключе, может быть вычислена как сумма произведений вероятностей статически сшиваемых одноблочных характеристик, её составляющих:

$$p(k) = \sum_{x=0}^{4^{a-1}} \prod_{i=1}^{a} n[s_i, \Delta x, \Delta y, h(x, dk_i, i)] \Big/ 2^{4\times a+2}, \qquad (1)$$

где a – количество активных S-блоков в характеристике;

i – порядковый номер S-блока внутри многоблочной характеристики;

x – значение соответствующее двоичному представлению «статической сшивки»;

s_i – номер S-блока, имеющего в рассматриваемой характеристике индекс i;

$n[s_i, \Delta x, \Delta y, h]$ – количество входных значений S-блока s_i, для которых выполняется характеристика $\Delta x \rightarrow \Delta y$ и биты статической «сшивки» имеют значение h (т.е. элемент расширенной (новой) таблицы распределения дифференциальных разностей);

dk_i – вектор «ключевых смещений» между S-блоками s_i и s_{i-1} ($dk_0 = \{0,0\}$);

h – функция, возвращающая значения левого и правого «швов» для S-блока s_i, при котором его левый «шов» стыкуется с правым «швом» S-блока s_{i-1}.

Вероятность одноцикловой характеристики, состоящей из нескольких несвязанных (разделённых пассивным S-блоком) многоблочных характеристик будет равна произведению вероятностей многоблочных характеристик её составляющих.

В качестве примера рассмотрим две характеристики, предложенные Эли Бихамом для построения атаки с максимальной вероятностью, для стандартных таблиц DES [2]. В таблице 1 представлено разложение одноблочных обнуляющих характеристик, использованных в атаке Бихама по 16 «статическим швам» (верхняя строка). Каждый «шов» записан в 4-ричной системе счисления (старшая цифра соответствует левой паре входов S-блока, а младшая – правой паре); входные разности ΔX представлены в 16-ричном виде в виде индексов оригинальной таблицы, т.е. два старших бита определяют номер строки перестановки, а младшие четыре – номер элемента в этой строке.

Таблица 1

S	ΔX	00	01	02	03	10	11	12	13	20	21	22	23	30	31	32	33	Σ
S1	11	1	0	0	1	1	3	3	1	0	1	1	0	0	1	1	0	14
S2	29	0	0	0	2	0	1	0	1	0	1	0	1	0	2	0	0	8
	2B	2	0	0	0	0	0	2	0	2	0	0	0	0	0	2	0	8
S3	26	0	0	1	3	0	1	0	0	0	0	1	3	0	1	0	0	10

Для упрощения использования таблицы 1, её можно преобразовать следующим образом. Так как любая входная разность получается последовательным воздействием некоторой пары входных значений, то содержимое предыдущей таблицы можно представить в виде статистики распределения «переходов» по различным парам, в пределах некоторой входной дельты – получим таблицу 2. Значения в ячейках этой таблицы соответствуют количеству пар, для которых выполняется выбранная характеристика, в скобках в 4-ричной системе счисления указаны биты входной разности, соответствующие «швам».

Таблица 2

S1	00-03	01-02	10-13	11-12	20-23	21-22	30-33	31-32	Σ
ΔX = 11 (03)	1	0	1	3	0	1	0	1	7
S2	00-32	01-33	02-30	03-31	10-22	11-23	12-20	13-21	Σ
ΔX = 29 (32)	0	0	0	2	0	1	0	1	4
S2	00-32	01-33	02-30	03-31	10-22	11-23	12-20	13-21	Σ
ΔX = 2B (32)	2	0	0	0	0	0	2	0	4
S3	00-20	01-21	02-22	03-23	10-30	11-31	12-32	13-33	Σ
ΔX = 26 (20)	0	0	1	3	0	1	0	0	5

С целью дальнейшего сокращения количества значащих ячеек таблицы можно сгруппировать те из них, которые имеют идентичные наборы «статических швов», т.е. левая и правая двухбитные пары

которых совпадают. Полученная после такого преобразования таблица (см. строки блока S2 таблицы 3) может использоваться в качестве альтернативы таблицы 1. В такой таблице количество значащих ячеек для некоторого перехода (одноблочной характеристики) может принимать значения: 4 (входная разность по обоим «швам» отлична от 0), 8 (входная разность по одному из «швов» равна от 0), 16 (входная разность по обоим «швам» равна 0). Такое деление следует из того, что любая отличная от нуля 2-битная разность может быть получена из двух различных пар 2-битных значений. Таким образом, если разность по некоторому «шву» отлична от нуля, то по этому «шву» возможно только два варианта вероятности «сшивки», в противном случае - четыре.

При вычислении вероятности фиксированной многоблочной характеристики для крайних S-блоков можно объединить ячейки, отличающиеся только неиспользуемым «пассивным швом» (между активным и пассивным S-блоком). Получим таблицу 3.

Таблица 3

S1	00-03, 10-13, 20-23, 30-33		01-02, 11-12, 21-22, 31-32		Σ
ΔX = 11 (03)	1+1+0+0		0+3+1+1		7
S2	00-32, 02-30	01-33, 03-31	10-22, 12-20	11-23, 13-21	Σ
ΔX = 29 (32)	0+0	0+2	0+0	1+1	4
S2	00-32, 02-30	01-33, 03-31	10-22, 12-20	11-23, 13-21	Σ
ΔX = 2B (32)	2+0	0+0	0+2	0+0	4
S3	00-20, 01-21, 02-22, 03-23		10-30, 11-31, 12-32, 13-33		Σ
ΔX = 26 (20)	0+0+1+3		0+1+0+0		5

По таблице 3 легко рассчитать вероятность выполнения выбранной характеристики для произвольного подключа. В таблице 4 в столбце $p^{<1>}$ приведены вероятности выполнения двух рассматриваемых одноцикловых характеристик для различных вариантов вектора «ключевого смещения» (столбец *dk*), а в столбце keys указан процент подключей, для которых вероятность имеет указанное значение. В столбце $p^{<13>}$ приведены предельные значения вероятности осуществления атаки на полный вариант алгоритма (с помощью модифицированной 2R-атаки количество циклов понижается с 16 до 13 [2]). Предельное значение получается, если на всех активных циклах значения векторов «ключевого смещения» *dk* попадают в одну группу (принадлежат одной строке), т.е. если вероятность одноцикловой характеристики на всех активных циклах постоянна.

Таким образом, получаем, что вероятность лучшей дифференциальной криптоатаки на DES при стандартных таблицах подстановки для ряда ключей составит 2^{-43}, однако, для другой группы ключей (такой же размерности) вероятность снизится до 2^{-55}, т.е. будет соответствовать сложности «силовой атаки». В случае наиболее вероятной ситуации – когда для половины активных циклов (трёх) вероятность рассмотренной характеристики равна $112/2^{14}$, а для другой половины $28/2^{14}$ – вероятность полноцикловой характеристики будет равна 2^{-49}, что несколько ниже значения 2^{-47}, приведенного в работе Э. Бихама [2]. Значение 2^{-47} может быть получено, если вероятность одноцикловой характеристики принять равной среднему от двух фактически возможных значений: $112/2^{14} \times 8/16 + 28/2^{14} \times 8/16 = 70/2^{14}$.

Таблица 4

ΔX	$p^{<1>}$	$p^{<13>}$	keys	*dk*							
19600000	$28/2^{14}$	2^{-55}	8 / 16	00	02	10	12	20	22	30	32
	$112/2^{14}$	2^{-43}	8 / 16	01	03	11	13	21	23	31	33
1B600000	$112/2^{14}$	2^{-55}	8 / 16	00	02	10	12	20	22	30	32
	$28/2^{14}$	2^{-43}	8 / 16	01	03	11	13	21	23	31	33

Учитывая вид двух полученных групп «ключевого смещения», получим, что для обеих характеристик вероятность определяется одним битом вектора *dk*, т.е. только одной парой битов подключа (k_{12} и k_{14}). Анализ алгоритма развёртывания ключа показывает, что интересующие нас 6 пар битов ключа, соответствующих 6 активным циклам, формируются 11 битами ключа (разряды подключа k_{12} и k_{14} на 3 и 13 циклах соответственно, формируются одним 18-тым битом ключа,

остальные биты не повторяются), т.е. в каждой паре хотя бы один бит уникален. Таким образом, всего для анализируемой атаки возможно 7 вариантов вероятностей, в указанном выше диапазоне, а каждое из предельных значений вероятности возможно на 2^{50} ключей, т.е. на каждом 64-том ключе.

Следует отметить, что рассмотренное свойство было известно разработчикам стандарта (максимальная вероятность рассмотренных характеристик для 16 циклов равна $2^{-57,5}$), это подтверждает список требований к таблицам подстановки, приведенный в [3]. Рассмотрим два из них:

Для любых ненулевых 6-ти битовых различий входов не более чем 8 из 32 пар входов должны показывать одно и то же выходное различие.

Критерий, аналогичный вышеизложенному, но для случая трёх активных *S* блоков.

Первое требование вводит ограничение 8/32=16/64=1/4 на вероятность любой одноблочной характеристики, а второе - следует интерпретировать как ограничение равное $(1/4)^3$ на вероятность трехблочной характеристики для произвольного ключа шифрования, а, следовательно, и вектора «ключевого смещения», т.е. вероятность любой трёхблочной характеристики на произвольном ключе должна быть не хуже произведения предельных одноблочных вероятностей. Слово «аналогично» в последнем требовании можно интерпретировать как ограничение равное $2^{al/2}$ на количество пар, для которых выполняется выбранный переход, где $al = 4 \times a + 2$ – количество входных линий, способных активизировать a соседних S-блоков. Следовательно, предельная вероятность характеристики состоящей из a соседних S-блоков не должна превышать $2^{al/2} / 2^{al-1} = 2^{-2 \times a}$, что соответствует произведению ограничений a одноблочных характеристик.

Из всего выше изложенного следует, что традиционная методика оценки вероятностей дифференциальных атак не учитывает зависимость вероятностей характеристик от ключа шифрования, и поэтому не даёт возможность оценить реальную степень опасности отдельных характеристик, т.к. получаемая вероятность является усреднённым значением. Это связано с некоторым упрощением реальной схемы цикловой функции F, имевшем место при построении математической модели, удобной для описания дифференциального криптоанализа. Оригинальная методика игнорирует факт большей размерности циклового подключа по сравнению с размерностью входного блока данных и факт наличия «статической связи» между соседними S-блоками. Предложенная в статье методика свободна от этих недостатков и позволяет определить не только наиболее эффективную характеристику, но и множество ключей, на которых её вероятность будет максимальна.

Рассмотренное свойство дифференциальных характеристик (способность иметь вероятность, зависящую от применённого ключа шифрования) может проявляться и при других, отличных от DES, схемах цикловой функции. Это свойство может возникать в случае зависимости вида нелинейного преобразования от ключа шифрования. Также следует отметить другую важную особенность шифра DES – использование небиективных S-блоков (для которых пространство входных значений превышает пространство выходных значений), позволяет достигнуть меньших значений вероятностей дифференциальных характеристик, по сравнению с вариантом использования биективных S-блоков, т.к. в первом случае одноблочные характеристики отдельных S-блоков оказываются взаимосвязаны, и должны включаться в конечную характеристику одновременно.

Список литературы: **1.** *FIPS* PUB 46-2. Specifications for DATA ENCRYPTION STANDARD. U.S. DEPARTMENT OF COMMERCE/National Institute of Standards and Technology, 1993. **2.** *E. Biham, A. Shamir* Differential Cryptanalysis of the full 16-round DES. Technical Report - Computer Science Department, Technion, Israiel, 1993. **3.** *Schneier B.* 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.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06:519.248.681

В.И.ДОЛГОВ, д-р техн. наук, И.В. ЛИСИЦКАЯ, канд. техн. наук,
Р.В.ОЛЕЙНИКОВ, А.И.ШУМОВ

«СЛАБЫЕ» КЛЮЧИ В АЛГОРИТМЕ ШИФРОВАНИЯ ГОСТ 28147-89

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 проходит через все циклы алгоритма шифрования с вероятностью выше пороговой. 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 таблиц подстановок (так называемых S-блоков). 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 единственным 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 большего числа циклов, чем DES, ГОСТ обладает более высокой защищенностью от многих известных криптоаналитических атак. В нашей работе сделана попытка применить к ГОСТ 28147-89 элементы дифференциального криптоанализа и доказать существование и в этом алгоритме определённых слабостей.

Хотелось бы обратить внимание на некоторую аналогию между процессом поиска пар со специфическими различиями в дифференциальном криптоанализе и проверкой статистической безопасности симметричного шифра. И в дифференциальном криптоанализе, и при проверке статистической безопасности интересуются изменениями шифрованных текстов при небольших изменениях в соответствующих им открытых текстах. При оценке статистической безопасности изучается так называемый лавинный эффект, для которого требуется, чтобы изменение хотя бы одного бита открытого текста или ключа изменяло бы в шифртексте примерно половину битов. Результаты статистических испытаний и исследование лавинного эффекта в алгоритмах DES и ГОСТ 28147-89, приведенные в [2] для случайных открытых текстов, показывают, что оба шифра реализуют хорошие показатели статистической безопасности. В таблице 1 представлены данные из этой работы, характеризующие лавинный эффект на различных циклах шифрования для шифра ГОСТ. Использованы обозначения: m_w – математическое ожидание числа $W(\Delta_k)$ единичных бит в булевой побитной сумме Δ_k (сумме по модулю 2) для пары полученных на k-ом цикле шифртекстов, открытые тексты которых отличаются одним битом, σ^2_w – дисперсия числа единичных бит для этой же побитовой суммы.

Отсюда следует, что устойчивый лавинный эффект достигается уже на восьмом-девятом цикле шифрования. Поскольку стандарт использует 32 цикла шифрования, можно сделать предположение о хорошей статистической безопасности алгоритма.

Исследование лавинного эффекта позволяет оценить лишь среднестатистические характеристики процедуры шифрования и не исключает наличия отдельных пар открытых текстов, отличающихся малым числом битов, результат шифрования которых также отличается малым числом битов. Именно на этих особенностях и строится дифференциальный криптоанализ.

Таблица 1

Номер цикла	Изменен 1-й бит сообщения		Изменен 31-й бит сообщения		Изменен 63-й бит сообщения	
k	m_w	σ^2_w	m_w	σ^2_w	m_w	σ^2_w
1	3.33	0.67	2.33	0.55	1.00	0.00
2	7.30	4.32	5.39	3.15	3.00	0.78
3	12.33	15.88	10.02	9.81	5.99	3.21
4	18.14	23.50	15.50	19.20	9.98	10.03
5	24.28	29.44	21.24	28.03	15.38	20.46
6	28.97	23.86	26.45	25.36	21.60	26.91
7	31.32	18.75	30.41	20.93	26.23	32.13
8	31.80	16.14	31.65	16.26	29.64	24.04
9	32.17	16.07	32.10	16.16	31.61	16.51

Задачей этой работы и является изучение возможностей прохождения через циклы шифрования ГОСТ 28147-89 таких специфических различий. Для решения этой задачи воспользуемся элементами дифференциального криптоанализа.

Нас будет интересовать вероятность прохождения некоторой фиксированной разности через нелинейное преобразование (подстановку). Под разностью будем понимать, как и в классической атаке [1], сложение по модулю 2 пар открытых и зашифрованных текстов, а также промежуточных значений. Для вычисления интересующей нас вероятности используются так называемые таблицы распределения разностей S блоков, которые строятся следующим образом. Перебираются все возможные комбинации пар входов (все возможные пары из чисел от 0 до 15), а затем для каждой из них определяются значения результатов подстановки. Составляется таблица (размер которой 16x16), в которой индексом (входом) ячейки по строкам будет побитовая сумма по модулю 2 входных значений, а индексом (входом) по столбцам – сумма по модулю 2 выходных значений. Сами ячейки заполняются числами, соответствующими количеству попаданий в каждую из них при заданных входах. Для получения вероятностей нужно разделить эти числа на 16 (количество возможных пар входов).

Пример случайной подстановки и её таблицы распределения разностей приведен в табл.2 и 3. Отметим, что индексы в таблицах записаны в шестнадцатеричной системе счисления.

Таблица 2

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
7	4	2	13	9	3	1	8	0	6	14	10	15	5	11	12

Таблица 3

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	2	2	0	2	2	0	2	4	0	0	0	0	2
2	0	0	0	0	2	2	0	0	2	4	0	2	2	0	2	0
3	0	2	2	2	0	0	2	0	2	0	4	0	0	0	2	0
4	0	0	0	4	0	4	2	2	0	0	0	0	0	0	2	2
5	0	2	2	0	2	2	0	0	0	2	2	0	2	2	0	0
6	0	2	0	0	0	0	2	0	0	0	2	4	2	0	2	2

Окончание табл. 3

7	0	2	0	0	2	4	0	0	0	0	0	2	2	2	0	2
8	0	0	2	0	2	0	4	4	0	0	2	0	2	0	0	0
9	0	2	0	4	2	0	0	0	2	0	0	0	4	2	0	0
A	0	0	4	0	0	0	0	0	0	2	0	2	0	2	4	2
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
B	0	0	0	0	4	2	0	2	2	0	2	0	0	4	0	0
C	0	4	2	0	0	2	0	0	2	4	0	0	0	0	0	2
D	0	0	2	2	0	0	4	0	0	0	0	4	0	0	2	2
E	0	2	0	0	0	0	0	2	4	2	0	0	2	2	2	0
F	0	0	2	2	0	0	0	4	2	0	0	2	0	2	0	2

Из табл.3 следует, что с вероятностью $\frac{2}{16}$ входная разность 01h[1] (пары чисел 0 и 1, 2 и 3, 4 и 5 и т.д. – всего 16 комбинаций) переходит в выходную 03h. Аналогично можно определить вероятность перехода между произвольными заранее заданными входными и выходными разностями.

Отметим, что именно свойства таблиц распределения разностей в значительной мере определяют устойчивость шифра к дифференциальной атаке. Приведенная в качестве примера случайная таблица подстановки не имеет переходов с вероятностью, равной единице, и поэтому обладает неплохими свойствами устойчивости к дифференциальному криптоанализу. Однако всего существует $N_p = 16! \approx 2{,}09 \cdot 10^{13}$ подстановок типа «4 бита в 4», и некоторые из них гораздо менее устойчивы к дифференциальному криптоанализу, чем рассмотренная. Пример такой «слабой» подстановки и соответствующая таблица распределения разностей приведены в табл.4 и 5.

Таблица 4

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1	6	5	2	9	11	13	15	10	3	14	7	0	12	4	8

Таблица 5

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	4	0	4	0	0	0	8	0	0	0	0	0	0	0	0
3	0	4	0	4	0	8	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	4	0	12	0	0	0	0	0	0	0	0	0
5	0	0	0	0	12	0	4	0	0	0	0	0	0	0	0	0
6	0	8	0	0	0	4	0	4	0	0	0	0	0	0	0	0
7	0	0	0	8	0	4	0	4	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	8	0	0	0	4	0	4	0
9	0	0	0	0	0	0	0	0	0	0	8	0	4	0	4	0
A	0	0	0	0	0	0	0	0	0	4	0	4	0	0	0	8
B	0	0	0	0	0	0	0	0	0	4	0	4	0	8	0	0
C	0	0	0	0	0	0	0	0	4	0	4	0	0	0	8	0
D	0	0	0	0	0	0	0	0	4	0	4	0	8	0	0	0
E	0	0	0	0	0	0	0	0	0	8	0	0	0	4	0	4
F	0	0	0	0	0	0	0	0	0	0	0	8	0	4	0	4

Как видно из табл.5, входная разность 01h всегда (с вероятностью $p = \frac{16}{16} = 1$) переходит в выходную разность 02h. Отметим, что при случайной генерации перестановок заданному свойству удовле-

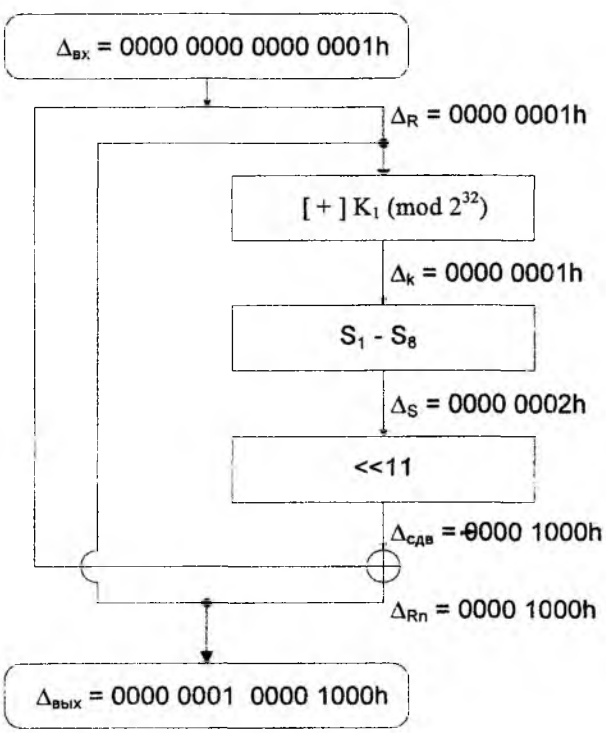
[1] Символ h после числа здесь и далее определяет запись числа в шестнадцатеричной системе счисления

творяет, в среднем, одна из 100 000, поэтому количество «слабых» подстановок можно оценить, приблизительно как $N_w = 10^{-5} \cdot N_p \approx 2{,}09 \cdot 10^8$.

Покажем, что при использовании таких «слабых» подстановок в ГОСТе для большого количества сеансовых ключей отмеченные выше показатели статистической безопасности не обеспечиваются.

Рассмотрим долговременный ключ, составленный из подстановок, обладающих заданным переходом, и будем проводить шифрование на одном и том же сеансовом ключе двух 64-битовых блоков, отличающихся младшим битом в правом полублоке. Входная разность перед первым циклом шифрования будет 0000 0000 0000 0001h. Левая половина (все нули в разности) будет гаммироваться с выходным значением цикловой функции, а правая без изменений поступит на следующий цикл и на вход цикловой функции в текущем цикле.

Первой операцией цикловой функции является сложение с ключом по модулю 2^{32}, при которой могут возникнуть переносы из младших разрядов в старшие. Поскольку для вычисления разности мы используем операцию сложения по модулю 2, то из-за влияния переносов существует некоторая вероятность, что после сложения с ключом разность изменится. Для исключения возникновения переноса в младшем разряде будем рассматривать ключи, у которых соответствующий бит равен нулю. Тогда разность 0000 0001h в младшей тетраде без изменения перейдет на вход подстановки и после неё трансформируется в значение 0000 0002h. После сдвига влево на 11 разрядов разность принимает вид 0000 1000h. Гаммирование с левой половиной двоичной разности на входе цикла даёт следующее значение разности на выходе первого цикла: 0000 0001 0000 1000h. Рассмотренные преобразования представлены на рисунке:



Описанные преобразования выполняются всегда для любых входных блоков, отличающихся младшим битом и чётным подключом. Вероятность прохождения разности через ключевое преобразование (сложение с ключом по модулю 2^{32}) зависит от значений битов ключа. В дальнейшем под «абсолютно слабым» ключом будем понимать такой, на котором вся 32-битная разность проходит через ключевой сумматор без изменений. Назовём ключ «слабым», если 32-битная разность проходит через ключевой сумматор без изменений с некоторой вероятностью, отличной от нуля.

Можно убедиться, что необходимым условием для прохождения разности через ключевое преобразование без изменения на «абсолютно слабом» ключе будет отсутствие переносов в младших 29 разрядах. Оно выполняется при любых входных значениях, если применяется подключ, у которого 29 младших битов нулевые. Если все подключи удовлетворяют этому условию, то прохождение разностей можно продлить на два цикла и более – вплоть до последнего цикла. Этапы этого преобразования в ходе шифрования иллюстрирует табл.6 (все значения разностей даны в шестнадцатеричной системе счисления).

Таблица 6

Номер	Разность (hex)	
цикла	левая половина	правая половина
0	0000 0000	0000 0001
1	0000 0001	0000 1000
2	0000 1000	0100 0001
3	0100 0001	0000 0010
4	0000 0010	0101 0001
5	0101 0001	1000 1000
6	1000 1000	0001 0101
7	0001 0101	0010 0000
8	0010 0000	0001 0100
9	0001 0100	1000 0000
10	1000 0000	0001 0000
11	0001 0000	0000 0000
12	0000 0000	0001 0000
13	0001 0000	1000 0000
14	1000 0000	0001 0100
15	0001 0100	0010 0000
16	0010 0000	0001 0101
17	0001 0101	1000 1000
18	1000 1000	0101 0001
19	0101 0001	0000 0010
20	0000 0010	0100 0001
21	0100 0001	0000 1000
22	0000 1000	0000 0001
23	0000 0001	0000 0000
24	0000 0000	0000 0001
25	0000 0001	0000 1000
26	0000 1000	0100 0001
27	0100 0001	0000 0010
28	0000 0010	0101 0001
29	0101 0001	1000 1000
30	1000 1000	0001 0101
31	0001 0101	0010 0000
32	0010 0000	0001 0100

«Абсолютно слабые» сеансовые ключи имеют нулевые младшие 29 битов (варьировать мы можем лишь 32 - 29 = 3 бита) в каждом из 8 подключей. Поэтому существует всего $\left(2^3\right)^8 = 2^{24}$ «абсолютно слабых» ключей, которые всегда дают заданное преобразование.

У «слабого» подключа в нуль установлены разряды, которым в двоичной разности соответствуют «1». Для описанного преобразования существует 2^{224} «слабых» ключей из 2^{256} всего возможных (соответственно вероятность генерации «слабого» ключа $p_w = \frac{2^{224}}{2^{256}} = 2^{-32}$). Для этих ключей единичные биты в разности без изменения будут проходить через ключевое преобразование, однако переносы могут с некоторой вероятностью возникнуть в предшествующих разрядах, что исказит разность. Оценим вероятность прохождения заданной разности через все 32 цикла с учетом возникновения переносов. Поскольку мы используем подстановки, всегда дающие нам заданный переход, вероятность будет зависеть только от сеансового ключа. На проявление этого свойства для каждого ключа будут влиять и шифруемые блоки. Назовём разряд, которому соответствует «1» в двоичной разности, актив-

ным. Разность будет искажена, если в шифруемых блоках возникнет перенос в активный разряд (поскольку двоичная разность до активного разряда равна нулю, то эта часть блоков совпадает, и перенос, если появится, то сразу в двух блоках). В активном разряде биты шифруемых блоков противоположны (один из блоков содержит 0, другой – 1). Ключ в разряде, соответствующем активному, содержит нуль. Поэтому в одном из блоков переноса из активного разряда не будет, а в другом блоке перенос обязательно возникнет. Практически всегда это приводит к искажению разности, но не исключено, что и искаженная разность будет преобразована в единичную на выходе сумматора. Однако вероятность этого события достаточно низкая, поэтому его рассматривать не будем. Итак, искажение разности произойдет при возникновении переносов в одном из блоков. Вероятность возникновения переноса в активный разряд – $\frac{1}{2}$. Это справедливо для любого активного разряда, исключая самые младшие. Отсюда вероятность искажения разности на первом цикле равна нулю (см. табл. 6, цикл 0), на втором и третьем – $\frac{1}{2}$ и т.д. Соответственно вероятность прохождения разности без искажений восьми циклов шифрования – $p_8 = 2^{-10}$, шестнадцати – $p_{16} = 2^{-19}$, и всего алгоритма – $p_{32} = 2^{-38}$. Можно дополнительно повысить вероятность прохождения заданной разности на выход алгоритма, подбирая значения левого полублока.

Соответственно для описанного преобразования разностей существует 2^{224} слабых ключей (при общем их числе 2^{256}), на каждом из которых заданная разность транслируется на выход алгоритма с вероятностью не менее 2^{-38}.

Здесь описан всего лишь один «потенциально опасный» вариант входной разности. В действительности для рассмотренной слабой подстановки существует 255 «потенциально опасных» входных значений разности только для правого полублока (от 0000 0001h до 1111 1111h), причем для восьми из них существует 2^{224} «слабых» ключей, для одного – 2^{192}, остальные располагают некоторым промежуточным количеством.

Можно рассматривать и сеансовые ключи, состоящие из комбинации «слабых» и «абсолютно слабых» подключей. Например, если второй подключ является «абсолютно слабым», а все остальные – «слабыми», то эффективная длина ключа составит $256 - 11 - 29 = 216$ битов, а вероятность распространения разности на выход – 2^{-32}. Использование сразу двух и более «абсолютно слабых» подключей ещё больше увеличивает вероятность успешной атаки.

Остались нерассмотренными и ряд других типов «слабых» долговременных ключей, для которых существуют «слабые» сеансовые ключи (существует много вариантов долговременных ключей, для которых в таблицах распределения разностей существует единичный переход).

Однако из уже представленных результатов следует, что для алгоритма ГОСТ 28147-89 существуют «слабые» подстановки и соответствующие им «слабые» сеансовые ключи, на которых не выполняются требования статистической безопасности. Более того, на «абсолютно слабых» сеансовых ключах заданная входная разность всегда преобразуется в определённую выходную и не зависит от шифруемых текстов.

Это ставит под сомнение вопрос о надёжности и безопасности алгоритма ГОСТ 28147-89 при использовании «слабых» долговременных ключей. На наш взгляд, представленные результаты свидетельствуют о том, что существующие методики генерации долговременных ключей, основанные на критериях проверки случайности[3],необходимо дополнить требованиями фильтрации «слабых» подстановок.

Список литературы: **1.** *E. Biham, A. Shamir.* Differential Cryptanalysis of DES-like cryptosystems. The Weizmann Institute of Science. Department of Applied Mathematics. Technical Report CS90-16. 1990. **2.** *Лисицкая И.В., Бондаренко А.С., Цепурит Т.В.* Сравнительный анализ механизмов образования лавинного эффекта в алгоритмах DES и ГОСТ 28147-89. 5-я Международная конференция «Теория и техника передачи, приёма и обработки информации». ХТУРЭ: Харьков, 1999. **3.** *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.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06

Е.Г.КАЧКО, канд. техн. наук, А.В.СВИНАРЕВ, канд. техн. наук, С.А.ГОЛОВАШИЧ

МЕТОДЫ И АЛГОРИТМЫ УСКОРЕНИЯ ВЫЧИСЛЕНИЙ В НЕСИММЕТРИЧНЫХ ПРЕОБРАЗОВАНИЯХ НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ

1. Введение

В настоящее время предложено множество алгоритмов несимметричных криптографических преобразований. Их стойкость основана на сложности решения некоторой математической задачи. Широко известны два класса задач криптоанализа: факторизация и нахождение дискретного логарифма в поле GF(p). В течение последних 5 лет проводятся исследования третьего класса несимметричных преобразований – преобразования на эллиптических кривых [1,2].

Эллиптическая кривая E над полем $\boldsymbol{Z}_p$, где p – простое, $p>3$, определяется уравнением

$$y^2 = x^3+ax+b, \qquad (1)$$

где $a, b \in Z_p$, $4a^3+27b^2 \not\equiv 0 \pmod p$. Множество точек на кривой $E(\boldsymbol{Z}_p)$ состоит из всех точек (x,y); $x, y \in \boldsymbol{Z}_p$, удовлетворяющих (1). В множество $E(\boldsymbol{Z}_p)$ также включается нулевая точка, обозначаемая как 0.

Для точек эллиптической кривой определим операцию сложения, удовлетворяющую следующим свойствам:

1. $P+0 = 0+P = P$ для $\forall P \in E(\mathrm{GF}(p))$

2. Для любой точки $P = (x,y)$, $P \in E(\mathrm{GF}(p))$, существует точка $Q = (x,-y)$, $Q \in E(\mathrm{GF}(p))$, такая, что $P+Q = 0$. Точка Q называется обратным элементом по отношению к P и обозначается как $(-P)$.

3. Для любых точек $P,Q \in E(\mathrm{GF}(p))$ существует точка $R=P+Q$; $R \in E(\mathrm{GF}(p))$.

Определим также операцию скалярного умножения точки на число. Пусть $P \in E(\mathrm{GF}(p))$, $c \in N$, тогда $c \times P = \underbrace{P + P + \ldots + P}_{c \text{ раз}}$.

Таким образом, множество $E(\mathrm{GF}(p))$, на котором определена операция сложения, образует абелеву группу.

2. Алгоритмы сложения на эллиптической кривой

Существует 2 способа внутреннего (машинного) представления точек эллиптической кривой – аффинные координаты и проективные координаты.

Аффинные координаты – пара чисел (x, y), удовлетворяющая уравнению кривой (1). Точка 0 не имеет аффинных координат, но для вычислений может быть представлена как пара чисел, не удовлетворяющих (1): (0, 0) при $b \neq 0$, (0,1) при $b=0$. Заметим, что при известном значении x можно из (1) вычислить y, поэтому точку можно однозначно представить как (x, y'), где $y' = y \bmod 2$.

Алгоритм 1. Сложение точек в аффинном представлении.
Вход: $P = (X_1,Y_1)$, $Q = (X_2,Y_2)$, $P \neq Q$.
Выход: $R = P+Q = (X_3,Y_3)$.
1. $L := \dfrac{X_2 - X_1}{Y_2 - Y_1}$.
2. $X_3 := L^2 - X_1 - X_2$.
3. $Y_3 := L(X_1 - X_3) - Y_1$.

Алгоритм 2. Удвоение точек в аффинном представлении.
Вход: $P = (X_1,Y_1)$.
Выход: $R = P+P = (X_3,Y_3)$.
1. $L := \dfrac{3X_1^2 + a_1}{2Y_1}$.
2. $X_3 := L^2 - 2X_1$.

3. $Y_3 := L(X_1 - X_3) - Y_1$.

Вычислительную сложность сложения точек можно оценить как

$$I_1(l) = 2I_{mul}(l) + I_{sqr}(l) + I_{inv}(l) + 6I_{add}(l),$$

где l – длина числа p в словах;
$I_{mul}(l)$ – вычислительная сложность умножения в поле GF(p);
$I_{sqr}(l)$ – вычислительная сложность возведения в квадрат в GF(p);
$I_{inv}(l)$ – вычислительная сложность нахождения обратного элемента в GF(p);
$I_{add}(l)$ – вычислительная сложность сложения и вычитания в GF(p).
Вычислительная сложность удвоения точки (т.е. сложения $P+P$) оценивается как

$$I_2(l) = 2I_{mul}(l) + 2I_{sqr}(l) + I_{inv}(l) + 5I_{add}(l).$$

Проективные координаты – числа (X, Y, Z) такие, что $x = X/Z^2$, $y = Y/Z^3$. Точка 0 имеет координаты $(\lambda^2, \lambda^3, 0)$, где λ – произвольное ненулевое число. Проективные координаты не являются однозначными, т.к. $(X, Y, Z) = (\lambda^2 X, \lambda^3 Y, \lambda Z)$.

Алгоритм 3. Сложение точек в проективном представлении.
Вход: $P = (X_0, Y_0, Z_0)$, $Q = (X_1, Y_1, Z_1)$, $P \neq Q$.
Выход: $R = P+Q = (X_2, Y_2, Z_2)$.
Если $P = Q$, алгоритм дает результат (0, 0, 0).

1. $L_1 := X_0 Z_1^2 - X_1 Z_0^2$
2. $L_2 := Y_0 Z_1^3 - Y_1 Z_0^3$
3. $L_3 := X_0 Z_1^2 + X_1 Z_0^2$
4. $L_4 := Y_0 Z_1^3 + Y_1 Z_0^3$
5. $Z_2 := Z_0 Z_1 L_1$
6. $X_2 := L_2^2 - L_3 L_1^2$
7. $Y_2 := ((L_3 L_1^2 - 2X_2)L_2 - L_4 L_1^3)/2$

Вычислительную сложность сложения точек можно оценить как

$$I_3(l) = 12I_{mul}(l) + 4I_{sqr}(l) + 10{,}5I_{add}(l).$$

Алгоритм 4. Удвоение точки в проективном представлении.
Вход: $P = (X_1, Y_1, Z_1)$.
Выход: $R = P+P = (X_2, Y_2, Z_2)$.

1. $L_1 := 3X_1^2 + aZ_1^4$
2. $Z_2 := 2Y_1 Z_1$
3. $L_2 := 4X_1 Y_1^2$
4. $X_2 := L_1^2 - 2L_2$
5. $Y_2 := L_1(L_2 - X_2) - 8Y_1^4$

Вычислительная сложность удвоения точки оценивается как

$$I_4(l) = 4I_{mul}(l) + 6I_{sqr}(l) + 12I_{add}(l).$$

Отметим, что для ускорения выполнения описанных алгоритмов целесообразно применить арифметику Монтгомери [3]. Используя численные значения сложности операций арифметики многократной разрядности, полученные в [4], определим оценки сложности операций на эллиптической кривой при программной реализации:

$I_1(l) = 5{,}3l^3 + 74l^2 + 124l$
$I_2(l) = 5{,}3l^3 + 88l^2 + 144l$
$I_3(l) = 260l^2 + 379l$
$I_4(l) = 152l^2 + 316l$

Для сравнения построим графики полученных зависимостей (рис.1, 2).

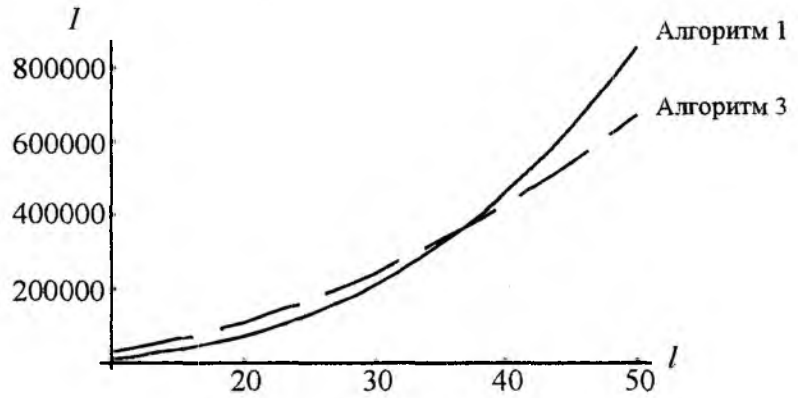


Рис.1

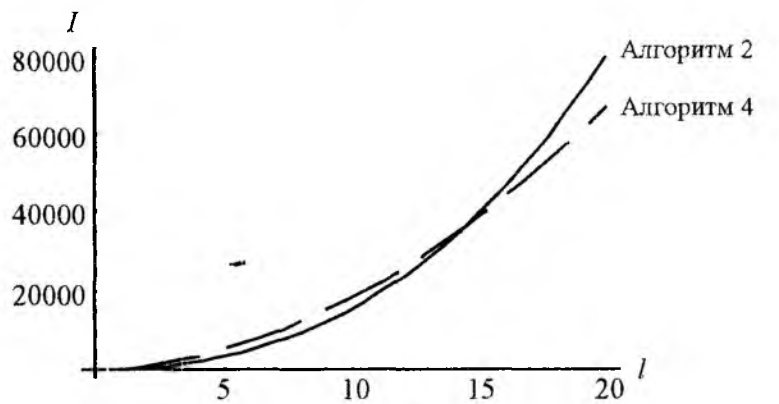


Рис.2

Очевидно, что для достаточно больших значений l $I_3<I_1$, $I_4<I_2$, т.е. в проективном представлении операции выполняются быстрее.

3. Алгоритмы скалярного умножения на эллиптической кривой

Для выполнения скалярного умножения $a\times P$ можно применить алгоритмы, аналогичные алгоритмам модульного возведения в степень – бинарного [4] и блочного [4,5].

Алгоритм 5. Бинарное скалярное умножение на эллиптической кривой.
Исходные данные: число $c\neq 0$, точка P, эллиптическая кривая $E = <a,b,p>$.
Результат: точка $Q = c\times P$.
1. Если $c=1$, то $Q:=P$; закончить работу алгоритма.
2. $k:=l_c-2$; $Q:=P$.
3. Для i, принимающего значения от k до 0, выполнить шаги 4-5.
4. $Q := Q+Q$.
5. Если i-й бит c равен 1, то $Q := Q+P$.
6. Закончить работу алгоритма.
Вычислительная сложность бинарного алгоритма составляет

$$I_{bin}(l_p, L_c) \approx L_c(I_{edbl}(l_p) + \rho I_{eadd}(l_p)), \qquad (2)$$

где L_c – длина числа c в битах;
ρ – доля единичных бит в числе c.
При $L_c = bl_p$, где b – размер слова в битах, выражение (2) примет вид

$$I_{bin} = (260\rho+152)bl_p^3 + (379\rho+316)bl_p^2 \qquad (3)$$

Алгоритм 6. Блочное скалярное умножение на эллиптической кривой.

Исходные данные: число $c \neq 0$, точка P, эллиптическая кривая $E = \langle a,b,p \rangle$.

Результат: точка $Q = c \times P$.

В алгоритме используется вспомогательная таблица $R[\,2^{d_{max}}\,]$.

В тексте алгоритма будут использоваться следующие обозначения [5]:

getblock(c, j) – операция выделения очередного блока из числа c, начиная с j-го бита;

end(C) – номер бита в числе c, на котором заканчивается блок C.

1. Если $c = 1$, то $Q:=P$; закончить работу алгоритма.
2. Вычислить $R[i]=i \times P$ для всех нечетных $i<2^m$.
3. $k := 1$; $Q := 1$.
4. Выделить блок: C:=**getblock**(c, 1).
5. Для k, принимающего значения от 1 до l_c, выполнить шаги 6-8.
6. $Q := Q+Q$
7. Если k = **end**(C) и $C \neq 0$, то: $Q := Q+R[C]$;
8. Если k = **end**(C), то C := **getblock**(c,k+1).
9. Закончить работу алгоритма.

Вычислительная сложность блочного алгоритма оценивается как:

$$I_{bl}(l_p, L_c) \approx I_{eadd}(l_p)(\,2^{d_{\max}-1} + \mu(L_c) + 1) + L_c I_{edbl}(l_p), \qquad (4)$$

где d_{max} – максимальная длина блока;

$\mu(L_c)$ – среднее количество блоков в числе c.

В [4] показано, что $\mu(L_c) = \dfrac{L_c}{2^{1-d_{\max}} + d_{\max}}$. Определим, при каком значении d_{max} функция (4) достигает минимума. Функцию (4) можно представить в виде суперпозиции двух функций: $I_{bl}=f(g(d_{max}))$, где $f(x)= x\, I_{eadd}(l_p) + (L_c-1) I_{edbl}(l_p)$; $g(d_{max})= 2^{d_{\max}-1} + \dfrac{L_c}{2^{1-d_{\max}} + d_{\max}} + 1$.

Т.к. $f(x)$ – монотонно возрастающая функция, то значение d_{max}, при котором функция $g(x)$ достигнет минимума, будет являться искомым значением d_0. Построив таблицу значений $g(x)$ для некоторых часто используемых длин показателя, найдем d_0 (см. табл. 1).

Таблица 1

x	L_c			
	160	192	224	256
3	54,2	64,1	73,9	83,8
4	47,8	55,5	63,3	71,1
5	48,6	54,9	61,2	67,6
6	59,5	64,8	70,1	75,4
7	87,8	92,4	96,9	101
d_0	4	5	5	5

Далее определим, при какой длине показателя степени блочный алгоритм скалярного умножения получает преимущество по быстродействию перед бинарным.

Используя (2) и (4), оценим выигрыш алгоритма 6 по сравнению с алгоритмом 5:

$$\Delta I = I_{bin}(l_p, L_c) - I_{bl}(l_p, L_c) =$$

$$= I_{eadd}(l_p)(\rho(L_c-1) - 2^{d_{\max}-1} - \frac{L_c}{2^{1-d_{\max}} + d_{\max}} - 1). \qquad (5)$$

Требуется найти значение L_c, при котором (5) обращается в 0.

Пусть $f(x)=\rho(x-1) - 2^{d_{\max}-1} - \dfrac{x}{2^{1-d_{\max}} + d_{\max}} - 1$. Тогда искомое значение $L(Y)$ совпадает с корнем уравнения $f(x) = 0$.

При ρ = 0,5 и d_{max} = 5 (см. табл. 1) это уравнение имеет корень $x_0 \approx 57{,}8$, следовательно, при $L_c \geq 58$ бит блочный алгоритм быстрее бинарного.

В алгоритме проверки цифровой подписи Эль-Гамаля на эллиптических кривых (например, ECDSA), встречается выражение вида $Q = a_1 \times P_1 + a_2 \times P_2$. Для ускоренного вычисления таких выражений можно использовать параллельный блочный алгоритм [4, 5].

Алгоритм 7. Параллельное блочное умножение на эллиптической кривой.

Исходные данные: числа c_i, точки P_i; $i=\overline{1,k}$; эллиптическая кривая E = <*a*,*b*,*p*>.

Результат: точка $Q = c_1 \times P_1 + c_2 \times P_2 + \ldots + c_k \times P_k$.

В алгоритме используется вспомогательная таблица $R[2^{d_{max}}, k]$.

1. $m := (\max L(Y_i)) - 1$; $Q := 0$.
2. Дополнить все c_i слева нулями до длины m бит.
3. Заполнение таблицы R:

$$R[i,j] := \begin{cases} 0; & j \bmod 2 = 0 \\ j \times P_i; & j \bmod 2 = 1 \end{cases}; \; i=\overline{1,k}; \; j=\overline{1, 2^{d_{max}} - 1}$$

4. Выделить блоки: C_i:=**getblock**(c_i, 1), $i=\overline{1,k}$.
5. Для j, принимающего значения от 1 до m, выполнить шаги 6-7.
6. Q := $Q+Q$.
7. Для i, принимающего значения от 1 до k, выполнить шаги 8-9.
8. Если j=end(C_i) и $C_i \neq 0$, то: Q:=Q+R[C_i,i].
9. Если j=end(C_i), то C_i:=**getblock**(c_i,j+1).
10. Закончить работу алгоритма. ▪

Вычислительная сложность алгоритма 7 оценивается как

$$I_{pm}(l_p, m, k) = mI_{edbl}(l_p) + \left(k\left(2^{d_{\max}-1}\right) + \sum_{i=1}^{k} \mu(c_i) \right) I_{eadd}(l_p),$$

а выигрыш по сравнению с алгоритмом 6 –

$$\Delta I = \left(\sum_{i=1}^{k} l_{c_i} - m \right) I_{edbl}(l_p).$$

Полученные оценки вычислительной сложности показывают, что арифметические операции на эллиптической кривой выполняются медленнее, чем соответствующие операции на GF(p) при одинаковой длине p. В то же время алгоритмы на эллиптических кривых обладают большей стойкостью: 160-битовое простое число в ECDSA соответствует по стойкости 1024-битовому числу в DSA, а 256-битовое - 2048-битовому [2].

4. Результаты экспериментальных измерений вычислительной сложности

В табл. 2 приведены временные характеристики несимметричных преобразований на эллиптических кривых, измеренные на Pentium-200 MMX.

Таблица 2

Длина ключа, бит	Время, мс		
	ECDSA, формирование	ECDSA, проверка	ECDH
160	22,7	28,0	22,7
192	31,5	40,3	31,5
224	42,6	55,0	42,6
256	56,5	73,0	56,5

Для сравнения приведем характеристики несимметричных преобразований в поле GF(p) (табл.3).

Таблица 3

Длина ключа и модуля, бит	Время, мс		
	DSA, формирование	DSA, проверка	DH
160 / 512	9,7	11,1	9,7
160 / 1024	31,2	37,2	31,2
256 / 2048	170,3	210,4	170,3

Приведенные результаты показывают, что с возрастанием длины ключевых параметров применение эллиптических алгоритмов является все более эффективным с точки зрения производительности: для ECDSA со 160-битовым ключом выигрыш составляет 37%, а для 256-битового ключа – 3 раза.

5. Заключение

Полученные результаты показывают, что применение цифровых подписей на эллиптических кривых над полем GF(p) может дать значительное повышение производительности систем защиты информации. В то же время интерес представляют исследования вычислительной сложности операций на эллиптических кривых над полем GF(2^m).

Список литературы: **1.** *ANSI X9.62.* Public Key Cryptography for the Financial Services Industry: The Elliptic Curve Digital Signature Algorithm (ECDSA). Draft. ANSI, 1998. **2.** *IEEE P1363.* Standard Specifications for Public-Key Cryptography. Draft. IEEE, 1999. **3.** *Горбенко И.Д., Качко Е.Г., Свинарев А.В.* Повышение быстродействия алгоритмов арифметики многократной точности // Безопасность информации, 1997. №1. С.6-12. **4.** *Свинарев А.В.* Методы и средства комбинированных несимметричных криптографических преобразований информации с уменьшенной вычислительной сложностью. Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.06. ХТУРЭ, Харьков, 1998. **5.** *Свинарев А.В.* Методы ускорения процедур цифровой подписи класса Эль-Гамаля // Радиотехника. Всеукр. межвед. науч.-техн. сб. 1997. Вып.104. С.173-178.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06

Д.И. ЛАВРИНЕНКО

ПРИМЕНЕНИЕ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ В КРИПТОГРАФИЧЕСКИХ ПРЕОБРАЗОВАНИЯХ

Введение

Криптографические системы, использующие несимметричные алгоритмы шифрования [1], базируются на специфическом математическом аппарате, который оперирует с целыми числами, разрядность которых превышает размер машинного слова. К таким системам часто предъявляются жесткие требования по производительности, которые напрямую зависят от вычислительной сложности математических алгоритмов. Поэтому решение задач минимизации вычислительной сложности этих алгоритмов и как следствие повышение скорости преобразований является важной задачей. При несимметричных преобразованиях наибольшую вычислительную сложность имеет операция модульного возведения в степень, которая базируется на операции модульного умножения. В данной статье производится краткий обзор существующих методов модульного умножения и вопрос выполнения модульного умножения «больших» целых чисел с использованием быстрого преобразования Фурье.

Для проведения сравнительной оценки вычислительной сложности введем следующие обозначения:

X, Y,... (большие буквы) — целые неотрицательные числа многократной точности. Причём числом **многократной точности** называется целое число, разрядность которого превышает размер машинного слова.

x, y,... (маленькие буквы) — числа однократной точности.

b — длина машинного слова в битах.

$B = 2^b$.

$L(X)$ — длина числа X в битах, т.е. $2^{L(X)-1} \le X < 2^{L(X)}$.

$l(X)$ — длина числа X в словах, т.е. $B^{l(X)-1} \le X < B^{l(X)}$.

X_i — i-й блок числа X; $i=0,1,\ldots, (L(X)-1)/b$, тогда $X=(X_{l(X)-1}\ldots X_1X_0)$

I_{a0} — вычислительная сложность одной машинной команды сложения, вычитания или пересылки.

I_{l0} — вычислительная сложность одной машинной команды цикла или перехода.

I_{m0} — вычислительная сложность одной машинной команды умножения или деления.

$x \bmod y$ — остаток от деления x на y.

$x \equiv y \pmod m$ — x сравнимо с y по модулю m.

$\lfloor x \rfloor$ – целая часть x (наибольшее целое число, не превосходящее x).

Наиболее простым и естественным методом умножения является выполнение операции «в столбик». Эти алгоритмы известны сейчас как арифметика Кнута [2]. Приведем схему алгоритма умножения в арифметике Кнута.

Исходные данные: числа X и Y.

Результат: число $Z = X \cdot Y$, $l(Z) = l(X)\, l(Y)$.

1. $Z:=0$.
2. Для i, принимающего значения от 0 до $l(X)-1$, выполнить шаг 3.
3. Для j, принимающего значения от 0 до $l(Y)-1$, выполнить шаги 4-5.
4. $T:=X_iY_j$; $l(T)=2$.
5. $(Z_{l(Z)-1}\ldots Z_{i+j}) := (Z_{l(Z)-1}\ldots Z_{i+j}) + T$.

Исходя из набора операций в каждом шаге и количества выполнений шага, вычислительную сложность алгоритма оценим как:

$$I_m(l(X), l(Y)) = (3I_{a0}+I_{l0}+I_{m0})\, l(X)\, l(Y)+I_{l0}\, l(X)+(l(X)+l(Y))I_{a0}.$$

При $l(X)=l(Y)=l$ получаем:

$$I_m(l) = I_{m0}\, l^2+I_{l0}(l^2+l)+I_{a0}\,(3l^2+2l).$$

1. Особенности применения БПФ для вычисления произведения

Вычислять произведение двух чисел многократной точности можно также исходя из полиномиального представления числа. Если представить исходные числа в виде полиномов и применить к ним операцию свертки, то получим полиномиальное представление произведения исходных чисел. Здесь мы можем применить для выполнения операции свертки спектральные преобразования, в частности преобразование Фурье. Кратко изложим сущность и свойства преобразования Фурье [3].

Преобразование Фурье связано с вычислением полиномов и их интерполяцией. Пусть

$$p(x) = \sum_{i=0}^{n-1} a_i x^i$$

– полином $(n-1)$-й степени. Его можно однозначно представить двумя способами: списком его коэффициентов $a_0, a_1 \ldots a_{n-1}$ и списком его значений в n различных точках $x_0, x_1 \ldots x_{n-1}$. Вычисление преобразования Фурье вектора $[a_0, a_1 \ldots a_{n-1}]^T$ эквивалентно превращению представления полинома $\sum_{i=0}^{n-1} a_i x^i$ списком его коэффициентов в представление его списком значений в точках $\omega^0, \omega^1 \ldots \omega^{n-1}$ (физический аналог операции – получение спектра сигнала). Вычисление обратного преобразования Фурье эквивалентно интерполяции полинома по его значениям в точках $\omega^0, \omega^1 \ldots \omega^{n-1}$ (восстановление сигнала по его спектру). Поясним смысл обозначений $\omega^0, \omega^1 \ldots \omega^{n-1}$.

Обычно преобразование Фурье определяется над кольцом комплексных чисел. Обобщая, можно определить преобразование Фурье над произвольным коммутативным кольцом $(R, +, \cdot, 0, 1)$. Элемент ω из R, такой что

1. $\omega \neq 1$,
2. $\omega^n = 1$,
3. $\sum_{j=0}^{n-1} \omega^{jp} = 0$ для $1 \le p < n$,

называется *примитивным корнем n-й степени из единицы*. Элементы $\omega^0, \omega^1 \ldots \omega^{n-1}$ называются *корнями n-й степени из единицы*.

Можно осуществлять преобразование Фурье на множестве точек, отличных от корней из единицы. Но выбирая степени корня ω, мы делаем вычисления преобразований особенно простыми.

Прямое или обратное преобразование Фурье вектора a из R^n можно вычислить за $O(n^2)$ операций. Но если n – степень числа 2, то можно сократить число операций до $O(n \log n)$. Эта модификация известна как быстрое преобразование Фурье (БПФ) [3].

Одно из основных приложений преобразования Фурье – вычисление свёртки двух векторов [4]. Свёртка двух векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ равна обратному преобразованию, применённому к покомпонентному произведению их образов прямого преобразования. Формально это записывается так: $a \otimes b = F^{-1}(F(a) \cdot F(b))$. Заметим что коэффициенты произведения полиномов – это в точности компоненты свёртки векторов коэффициентов этих полиномов. Если теперь представить числа в полиномиальном виде, то станет очевидным использование преобразований Фурье для вычисления произведения двух чисел.

Обычно преобразование Фурье выполняется в кольце комплексных или вещественных чисел. Но для вычисления произведения нужен точный результат. Избежать погрешностей при работе с вещественными числами можно, если производить вычисления в конечном кольце – кольце R_m целых чисел по модулю m, где m будет таким, чтобы в R_m существовал примитивный корень n-й степени из единицы. Если n – степень числа 2, то подходящее m существует всегда. В частности, если ω и n – степени числа 2, то m можно выбрать как $\omega^{n/2} + 1$. Числа вида $2^p + 1$ известны как числа Ферма, а вычисления по такому модулю известны как теоретико-числовые преобразования [3]. Модульные вычисления по такому модулю заметно проще, чем по произвольному модулю.

Приведем схему алгоритма БПФ:

Исходные данные: число A.

Промежуточные данные: массив S, $l(S) =$ n.

Результат: массив B = БПФ(A).

1. Для i, принимающего значения от 0 до n, выполнить шаг 2.
2. S[i] = A_i.
3. Для l, принимающего значения от 0 до $\log_2 n$-1, выполнить шаг 4.
4. Для i, принимающего значения от 0 до n-1, выполнить шаг 5.
5. S[i] = S[i & ~(2^l)] + $\omega^{inv(i)>>l<<l} \cdot$S'[i | ($2^l$)] mod m.
6. Для i, принимающего значения от 0 до l(S)–1, выполнить шаг 7.
7. B[i] = S[inv(i)].

Здесь inv(i) обозначает битово-инверсное представление числа i в разрядной сетке из $\log_2 n$ разрядов. Строка 5 алгоритма представляет собой не что иное, как базовую (двухточечную) операцию БПФ [5], известную также под названием операции «бабочка» [3].

Обратное преобразование Фурье аналогично прямому, если в шаге 5 ω^p заменить на ω^{-p} и на шаге 7 добавить умножение на n^{-1}.

Если рассматривать машинную реализацию БПФ, то можно заметить, что если ω и n – степени числа 2, то все умножения на ω^p и n^{-1} можно заменить операциями двоичного сдвига, которая обычно более эффективна, чем операция умножения. Все значения степеней ω можно вычислить заранее.

2. Анализ вычислительной сложности БПФ

Для оценки вычислительной сложности БПФ введем обозначения

I_{Fs} — вычислительная сложность одной операции сдвига.

I_{Fa} — вычислительная сложность одной операции модульного сложения или вычитания.

I_{Fm} — вычислительная сложность одной операции модульного умножения.

I_{Fv} — вычислительная сложность одной операции пересылки.

Разрядность операндов этих операций, в отличие от операций в классическом умножении, зависит от разрядности исходных данных и от параметров ω и n. Кроме того, операции выполняются по специальному модулю поэтому они обозначаются отдельно.

Оценим вычислительную сложность прямого БПФ как:

$$I_{FFTD}(n) = I_{Fs}(n \cdot \log_2 n) + I_{Fa}(n \cdot \log_2 n) + I_{Fv}(n \cdot (\log_2 n + 2)) + I_a(n \cdot \log_2 n).$$

Вычислительная сложность обратного преобразования есть

$$I_{FFTR}(n) = I_{Fs}(n \cdot (\log_2 n + 1)) + I_{Fa}(n \cdot \log_2 n) + I_{Fv}(n \cdot (\log_2 n + 2)) + I_a(n \cdot \log_2 n).$$

Для того, чтобы умножить два числа с использованием БПФ, необходимо получить прямые БПФ-преобразования исходных чисел, осуществить их покомпонентное умножение и получить обратное преобразование результата умножения. Но если рассматривать умножение как составную часть алгоритма возведения в степень, то можно опустить одно прямое преобразование, т.к. второй аргумент можно вычислять заранее и оставлять в базисе БПФ. Оценим теперь общую вычислительную сложность умножения с применением БПФ.

$$I_{FFTR}(n) = I_{Fs}(n \cdot (2 \cdot \log_2 n + 1)) + I_{Fa}(2 \cdot n \cdot \log_2 n) + I_{Fv}(2 \cdot n \cdot (\log_2 n + 2)) + I_a(n \cdot \log_2 n) + I_{Fm} n.$$

Анализируя применение БПФ для умножения чисел многократной точности, нельзя не отметить несколько особенностей и ограничений метода.

Во-первых, БПФ требует, чтобы оно осуществлялось на количестве точек $n = 2^p$. Это также требует, чтобы полиномиальное представление исходных чисел тоже имело 2^p коэффициентов. Если $n \neq 2^p$, то нужно преобразовать схему БПФ, как это указано в [3], что повлечет за собой дополнительные вычислительные затраты. Другой выход из положения – добавить к полиномиальному представлению столько нулевых коэффициентов, чтобы их общее количество стало 2^p, что влечет за собой увеличение размерности БПФ, что также увеличивает вычислительную сложность преобразования. Из всего сказанного выше следует, что применение БПФ наиболее эффективно при умножении чисел определенной длины.

Во-вторых, при выборе количества точек БПФ также имеются ограничения. Это следует из $m = \omega^{n/2} + 1$, $\omega = 2^p$, и $L(X) = n \cdot L(b)$, где b – блок, на которые разбивается исходное число (по своей математической сути блоки – это коэффициенты полиномиального представления числа). Покажем это.

m должно быть не меньше величины максимального элемента свертки, чтобы не вносить искажений, значит $L(m) \geq 2 \cdot L(b) \cdot \log_2 n$. Но для реализации удобнее, если m как можно меньше и кратно разрядной сетке ЭВМ, поскольку в противном случае цена отдельной операции по модулю m значительно (квадратично) возрастает. Это предопределяет $L(m) = 2 \cdot L(b)$, а как исправить возникающие искажения, связанные с компонентом $\log_2 n$, будет сказано ниже.

$m = \omega^{n/2} + 1$, $\omega = 2^p$, причём $\omega \geq 2$, отсюда следует $L(m) \geq n/2$. Из $n = L(X) / L(b)$ и $L(m) = 2 \cdot L(b)$ следует $n = 2L(X) / L(m)$. В результате получаем

$$L(m) \geq L(X) / L(m) \Rightarrow L(m) \geq \sqrt{L(X)},$$

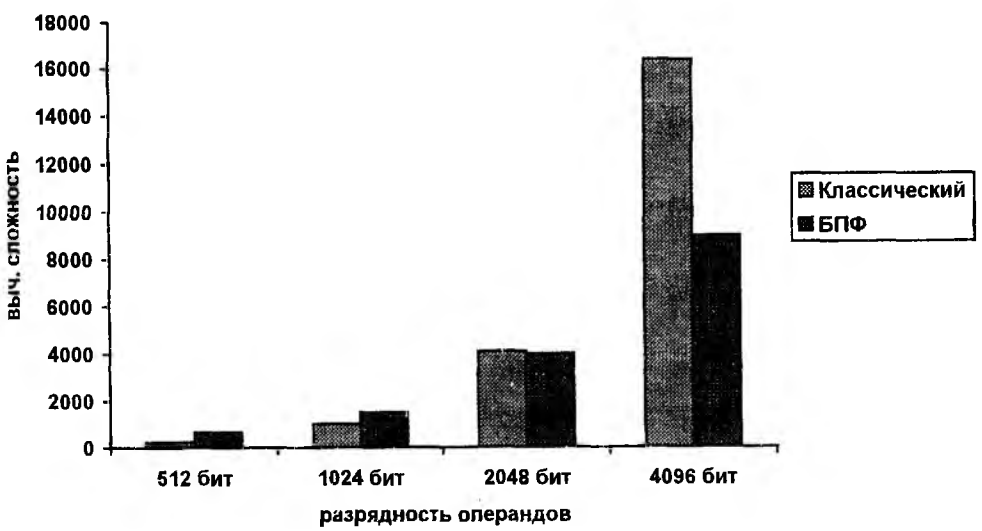
$$n \leq 2 \cdot L(m) \Rightarrow n \leq 2 \cdot \sqrt{L(X)}$$

Теперь для того, чтобы исправить искажения, возникшие в результате применения модуля меньшей длины, воспользуемся китайской теоремой об остатках. Здесь теорема будет применима, если мы вычисляем компоненты свёртки дважды: по модулю m и по модулю n (эти числа будут взаимно просты, так как n – степень 2, а m – нечётное). Вычисления компонентов свёртки по модулю мы вычисляем с помощью преобразования Фурье, как было объяснено выше, а про вычисления свёрток по модулю 2^p объясняются в [4].

3. Экспериментальное исследование вычислительной сложности умножения на БПФ

Известно, что каждая из элементарных операций на разных типах ЭВМ имеет различную вычислительную сложность. Сравнивая формулы для вычислительной сложности умножения многократной точности классическим методом и методом с преобразованиями Фурье, можно заметить, что метод с преобразованиями Фурье будет тем выгоднее, чем меньшую вычислительную сложность имеют операции по модулю $\omega^{n/2} + 1$ (операции сложения, вычитания, сдвига, умножения), которые на большинстве ЭВМ имеют вычислительную сложность примерно в 3-4 раза большую, чем по модулю $\omega^{n/2}$. Поэтому применение БПФ в умножении многократной точности будет тем выгоднее, чем меньше будет вычислительная сложность этих преобразований. Исходя из этого можно дать следующие рекомендации:

- следует добиться минимальной вычислительной сложности для вычислений по модулю $\omega^{n/2} + 1$. Возможно, такие вычисления будет выгодно производить на специализированных ЭВМ;
- поскольку при умножении классическим методом количество операций умножения значительно больше, чем при умножении с БПФ, то последний метод будет выгоднее на ЭВМ, где операции умножения будут иметь большую вычислительную сложность по сравнению с операциями сложения, вычитания и сдвигов.



Сравнивая формулы оценки вычислительной сложности для алгоритмов умножения многократной точности классическим методом и с БПФ , можно заметить, что зависимость вычислительной сложности алгоритма от длины исходных операндов для обоих алгоритмов носит полиномиальный характер, но порядок полинома для классического алгоритма будет выше, чем для алгоритма с БПФ. Поэтому эффект от применения БПФ будет присутствовать всегда, начиная с определённой длины исходных операндов, даже если вычислительная сложность операций в БПФ будет сравнительно велика.

Проведённый вычислительный эксперимент показал (рисунок), что на процессорах Intel Pentium эффект от применения БПФ в умножении будет появляться для исходных операндов длиной не менее 2048 бит. Это объясняется в первую очередь относительно большой вычислительной сложностью команд сдвигов и операций по модулю $\omega^{n/2}+1$ на процессорах этого типа.

Для сравнения возьмём сигнальный процессор ADSP 21063, архитектура которого значительно отличается от процессоров Intel. Операции умножения на этом процессоре имеют такую же вычислительную сложность, как и операции сложения и сдвига, поэтому эффект от применения БПФ здесь будет иметь место при ещё большей длине исходных операндов (порядка 16284 бит и более).

Заключение

Так как в современных криптосистемах в основном применяются операнды разрядностью в 512 и 1024 бит, то имея в виду реализацию на процессорах архитектуры Intel, можно сказать, что БПФ может найти применение в криптосистемах с повышенными требованиями безопасности, т.е. при длинах модулей 2048 и более бит. Что касается реализации на ЭВМ другой архитектуры, здесь перспективность БПФ определяется вышеупомянутыми рекомендациями.

Список литературы **1.** *Diffie W., Hellman M.E.* New directions in cryptography // IEEE trans. on Information Theory.– 1976.– V.IT-22.– №6.– pp. 644-654. **2.** *Кнут Д.* Искусство программирования для ЭВМ. В 3-х т. — Т.2. Получисленные алгоритмы. — М.: Мир, 1977. — 387 с. **3.** *Цифровая* обработка сигналов. Ч.1. Учебное пособие. И.Д. Горбенко, А.С. Трощило, К.В. Бессарабенко. МО СССР, 1988. **4.** *Ахо А., Хопкрофт Дж., Ульман Дж.* Построение и анализ вычислительных алгоритмов – М.: Мир,1979. – 536 с. **5.** *Рабинер Л. Гоулд Б.* Теория и применение цифровой обработки сигналов – М.: Мир, 1978. – 848 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 27.03.2000

УДК 681.3.06

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

РАСШИРЕННОЕ ПОЛЕ ГАЛУА $GF(2^M)$. ВЫЧИСЛИТЕЛЬНАЯ СЛОЖНОСТЬ ПРОСТЕЙШИХ ОПЕРАЦИЙ НАД РАСШИРЕННЫМ ПОЛЕМ $GF(2^M)$

Введение

В последние годы в несимметричной криптографии достаточно широкое распространение нашел математический аппарат теории групп, полей и колец. Так в, ставших уже классическими, криптоалгоритмах класса Эль-Гамаля [1] и в криптопротоколах типа Диффи-Хелмана [2] преобразования осуществляются в простом поле Галуа. В недавно появившихся модификациях этих алгоритмов преобразования осуществляются на эллиптических кривых над простым или расширенным полем Галуа. Применение перечисленных криптоалгоритмов позволяет реализовать ряд преимуществ по сравнению с классическими, к примеру, симметричными криптоалгоритмами. В то же время преобразования больших чисел требуют значительных вычислительных затрат, при этом наибольшей вычислительной сложностью обладает операция умножения по модулю, деление по модулю и возведение в степень по модулю. Рядом преимуществ обладают криптографические преобразования осуществляемые над полем Галуа $GF(2^m)$. Для этого поля также остается актуальной задача вычислительной сложности выполнения базовых операций над большими числами.

Расширенное поле Галуа $GF(2^m)$ имеет два основных базисных представления - полиномиальное и нормальное. Целью настоящей статьи является изучение влияния используемого базиса представления на вычислительную сложность выполнения основных арифметических операций , обоснование и выбора наиболее предпочтительного из них. Определим эти базисы.

При использовании полиномиального базиса каждый элемент поля $GF(2^m)$ представляется двоичным полиномом степени не выше чем $(m-1)$ или битовой строкой его коэффициентов $(a_{m-1}...a_2a_1a_0)$. В наиболее общем случае элемент представляется в виде полинома, имеющего вид

$$b_i=a_{m-1}\,t^{m-1} + \ldots + a_2t^2 + a_1t + a_0 .$$

Для него полиномиальный базис имеет вид

$$B = \{t^{m-1}, \ldots, t^2, t, 1\} .$$

В рассмотренном поле преобразования осуществляются по двойному модулю $p(t),2$, где $p(t)$ является примитивным полиномом. Анализ показывает, что вычислительная сложность выполняемых операций по модулю $p(t)$ зависит от количества ненулевых коэффициентов в примитивном полиноме. Наименьшая вычислительная сложность требуется для трехчлена $p(t)=t^m + t^k + 1$. Однако примитивный трехчлен существует не для всех значений m, для остальных значений m существует пятичлен $p(t)= t^m + t^a + t^b + t^c + 1$, где $m>a>b>c$.

Нормальный базис задается множеством вида

$$b_i =a_0\theta + a_1\theta^2 + a_2\theta^{2^2} + \ldots + a_{m-1}\theta^{2^{m-1}}$$

Наиболее предпочтительным является Гауссовский нормальный базис. Его отличие от нормального базиса заключается в том, что 2^m не делится на 8.

Рассмотрим сущность и алгоритмы основных выполняемых арифметических операций в полиномиальном и нормальном Гауссовском базисе.

1. Алгоритмы операций с полиномиальным базисом

1.1 Сложение по модулю 2

Вход: $a=(a_{m-1}...a_2a_1a_0)$, $b=(b_{m-1}...b_2b_1b_0)$

Выход: $c=(c_{m-1}..c_2c_1c_0)=a+b\ (mod\ 2)$

1. *Для i:=0 до m-1 с шагом 1 выполнить*

1.1 $c_i:=a_i \oplus b_i$

2. *$c=(c_{m-1}...c_2c_1c_0)$, конец*

1.2 Вычитание по модулю 2

Операция вычитания имеет идентичный характер вычисления с операцией сложения.

Вход: $a=(a_{m-1}...a_2a_1a_0)$, $b=(b_{m-1}...b_2b_1b_0)$

Выход: $c=(c_{m-1}...c_2c_1c_0)=a-b(mod\ 2)$

1. *Для i:=0 до m-1 с шагом 1 выполнить*

1.1 $c_i:=a_i \oplus b_i$

2. *$c=(c_{m-1}...c_2c_1c_0)$, конец*

1.3 Приведение *c* по двойному модулю *p(t),2*

Вход: $a=(a_k...a_2a_1a_0)$, $p(t)=(p_m...p_0)$, *где k>m*

Выход: $c=(c_{m-1}...c_2c_1c_0)=a(mod\ p(t),2)$

1. c:=a

2. i:=l(a)-1

3. Если i=m-1 то конец алгоритма

4. Если a_i=0 то перейти на 10

5. j:=m

6. Если j <0 то перейти на 10

7. $c_{i-(m-j)}=c_{i-(m-j)} \oplus p_j$

8. j:=j-1

9. Перейти на 6

10. i:=i-1

11. Перейти на 3

1.4 Умножение по двойному модулю *p(t),2*

Вход: $a=(a_{m-1}...a_2a_1a_0)$, $b=(b_{m-1}...b_2b_1b_0)$, $p(t)=(p_m...p_0)$

Выход: $c=(c_{m-1}...c_2c_1c_0)=a^*\ b(mod\ p(t),2)$

1. c:=0

2. *i:=0*

3. *Если i>l(b)-1 то перейти на 12*

4. *Если b_i=0 то перейти на 10*

5. *j:=0*

6. *Если j>l(a)-1 то перейти на 10*

7. *$c_{i+j}:=c_{i+j} \oplus a_j$*

8. *j:=j+1*

9. *перейти на 6*

10. *i:=i+1*

11. *перейти на 3*

12. *c:=c (mod p(t),2)// по алгоритму 2.3*

13. *$c=(c_{m-1}...c_2c_1c_0)$, конец*

1.5 Возведение в квадрат по двойному модулю *p(t),2*

Вход: $a=(a_{m-1}...a_2a_1a_0)$, $p(t)=(p_m...p_0)$

Выход: $c=a^2(mod\ p(t),2)$

1. c:=0

2. *i:=0*

3. *Если i>l(a)-1 то перейти на 7*

4. $c_{i+i}:=c_{i+i}\oplus a_i$

5. *i:=i+1*

6. *перейти на 3*

7. *c:=c (mod p(t),2)// по алгоритму 2.3*

8. $c=(c_{m-1}\ldots c_2c_1c_0)$, *конец*

1.6 Возведение в квадрат с предвычислениями

Вход: $a=(a_{m-1}\ldots a_2a_1a_0)$, матрица *S // вычисление матрицы описано в* [3]

Выход: $c=a^2$

1. *c:=0*

2. *c:=a*S*

3. $c=(c_{m-1}\ldots c_2c_1c_0)$, *конец*

1.7 Возведение в степень по двойному модулю *p(t),2*

Вход: положительное число *k*, поле $GF(2^m)$ и элемент α.

Выход: $x=\alpha^k(mod\ p(t),2)$.

1.*Пусть* $k=k_r\,k_{r-1}\ldots k_1\,k_0$ *бинарное представление k, где большее количество* k_r *равняется 1.*

2.*Установим* $x:=\alpha$.

3.*Для* $i:=r-1$ *до 0 с шагом -1 выполнить*

3.1 *Установим* $x:=x^2$.

3.2 *Если* $k_i=1$ *то* $x:=\alpha x$.

4.*Выход x, конец.*

Возведение в квадрат и умножение в алгоритме выполняются с помощью 1.4 и 1.5.

1.8 Вычисление квадратного корня по двойному модулю *p(t),2*

Вход: $a=(a_{m-1}\ldots a_2a_1a_0)$, *p(t)*

Выход: $x=\sqrt{\alpha}\ (mod\ p(t),2)$

По [3] $\sqrt{\alpha}=a^{m-1}mod\ p(t)$, поэтому извлечение квадратного корня сводиться к возведению в степень по методике алгоритма 1.7

1.*Пусть* $m-1=k_r\,k_{r-1}\ldots k_1\,k_0$ *бинарное представление k, где большее количество* k_r *равняется 1.*

2.*Установим* $x:=\alpha$.

3.*Для* $i:=r-1$ *до 0 с шагом -1 выполнить*

3.1 *Установим* $x:=x^2$.

3.2 *Если* $k_i=1$ *то* $x:=\alpha x$.

4. *Выход x, конец.*

1.9 Деление по двойному модулю *p(t),2*

Вход: $a=(a_{l-1}\ldots a_2a_1a_0)$, $b=(b_{k-1}\ldots b_2b_1b_0)$, $l<m$, $k<m$

Выход: $c=(c_{m-1}\ldots c_2c_1c_0)=a*b^{-1}(mod\ p(t),2)$

1. *Пусть* $m-1=b_rb_{r-1}\ldots b_0$ *бинарное представление m-1*

2. $\eta:=b$, $k:=1$

3. *Для* $i:=r-1$ *до 0 с шагом 1 выполнить*

3.1 $\mu:=\eta$

3.2 *Для* $j:=1$ *до k с шагом 1 выполнить*

3.2.1 $\mu:=\mu^2$

3.3 $\eta:=\eta*\mu$, $k:=2*k$

3.4 *Если* $b_i=1$, *то* $\eta:=\eta^2*b$, $k:=k+1$

4. *Выход* $a*\eta^2$, *конец.*

Мы рассматриваем общий случай, существуют более эффективные способы для вычисления деления по двойному модулю $p(t),2$.

2. Алгоритмы операций с нормальным базисом

2.1 Сложение

Аналогично алгоритму сложению 1.1 в полиномиальном базисе.

Вход: $a = (a_0\ a_1\ \ldots\ a_{m-1}), b = (b_0\ b_1\ \ldots\ b_{m-1})$

Выход: $c = (c_0\ c_1\ \ldots\ c_{m-1}) = c + b$

1. *Для* $i := m-1$ *до* 0 *с шагом* -1 *выполнить*

1.1 $c_i := a_i \oplus b_i$

2. $c = (c_0\ c_1\ \ldots\ c_{m-1})$, *конец*

2.2 Вычитание

Аналогично алгоритму вычитания 1.2 в полиномиальном базисе.

Вход: $a = (a_0\ a_1\ \ldots\ a_{m-1}), b = (b_0\ b_1\ \ldots\ b_{m-1})$

Выход: $c = (c_0\ c_1\ \ldots\ c_{m-1}) = c + b$

1. *Для* $i := m-1$ *до* 0 *с шагом* -1 *выполнить*

1.1 $c_i := a_i \oplus b_i$

2. $c = (c_0\ c_1\ \ldots\ c_{m-1})$, *конец*

2.3 Возведение в квадрат

Вход: $a = (a_0\ a_1\ \ldots\ a_{m-1})$,

Выход: $c = a^2$

1. $c := a$

2. $c :=$ *ПравыйЦиклическийСдвиг*$(c,1)$

3. $c = (c_{m-1} \ldots c_2 c_1 c_0)$, *конец*

В записи *ПравыйЦиклическийСдвиг*$(c,1)$ выполняется сдвиг на 1 бит.

2.4 Умножение

Вход: $a = (a_0\ a_1\ \ldots\ a_{m-1}), b = (b_0\ b_1\ \ldots\ b_{m-1})$, матрица M, вычисление которой приведено в [3]

Выход: $c = (c_0\ c_1\ \ldots\ c_{m-1})$

1. $x := a$.

2. $y := b$.

3. *Для* $k := 0$ *до* $m-1$ *с шагом* 1 *выполнить*

3.1 *Вычисление произведение при помощи матрицы*

$c_k := x\ M\ y^{tr}$

где y^{tr} *транспонированная матрица вектора* y.

3.2 $x :=$ *ЛевыйЦиклическийСдвиг*$(x,1)$ *и* $y :=$ *ЛевыйЦиклическийСдвиг*$(y,1)$

4. $c = (c_0\ c_1\ \ldots\ c_{m-1})$, *конец*

2.5 Возведение в степень

Алгоритм аналогичен последовательности операций возведения в степень в полиномиальном базисе, но со своими алгоритмами умножения и возведения в степень.

Вход: положительное число k, поле $GF(2^m)$ и элемент α.

Выход: $x = \alpha^k$.

1. *Пусть* $k = k_r\ k_{r-1}\ \ldots\ k_1\ k_0$ *бинарное представление* k, *где большее количество* k_r *равняется 1.*

2. *Установим* $x := \alpha$.

3. *Для* i *от* $r-1$ *до* 0 *с шагом* -1 *выполнить*

3.1 *Установим* $x := x^2$.

3.2 *Если* $k_i = 1$ *то* $x := \alpha x$.

4. *Выход x, конец.*

2.6 Вычисление квадратного корня

Вход: $a = (a_0\ a_1\ \ldots\ a_{m-1})$

Выход: $c = \sqrt{\alpha}$

1. c:=a

2. c:=ЛевыйЦиклическийСдвиг(c,1)

3. $c=(c_{m-1}\ldots c_2 c_1 c_0)$, конец

2.7 Деление

Аналогичен 1.9, но со своими операциями возведение в квадрат и умножения.

Вход: $a = (a_0\ a_1\ \ldots\ a_{m-1}), b = (b_0\ b_1\ \ldots\ b_{m-1})$

Выход: $c=(c_{m-1}\ldots c_2 c_1 c_0)=a*b^{-1}$

4. Пусть $m-1=b_r b_{r-1}\ldots b_0$ бинарное представление m-1

5. $\eta:=b$, $k:=1$

6. For i от r-1 до 0 делать

3.5 $\mu:=\eta$

3.6 For j:=1 до k делать

3.2.1 $\mu:=\mu^2$

3.7 $\eta:=\eta*\mu$, $k:=2*k$

3.8 Если $b_i=1$, то $\eta:=\eta^2*b$, $k:=k+1$

5. Выход $a*\eta^2$

3. Расчет вычислительной сложности

Для оценки вычислительной сложности разобьем команды на классы:

1-я группа – команды сложения, вычитания, пересылки, умножение битов;

2-я группа – команды цикла, перехода, команды сдвигов;

3-я группа – команды умножения и деления.

Используя подходы, изложенные в [4], проведем анализ вычислительной сложности приведенных арифметических операций в полиномиальном и нормальном Гауссовском базисах. Полученные результаты приведены в таблице 1.

Ниже *l(a), l(b), l(m-1)* длина чисел в битах.

Таблица 1

№ алгоритма	№ шага	Среднее количество повторений *j*-го шага в зависимости от входных параметров	Сложность команды
1	2	3	4
1.3	1	*l(a)*	*1*
	2	*1*	*1*
	3	*l(a)-m-1*	*1*
	4	*l(a)-m-1*	*2*
	5	*n(l(a)-m-1)*	*1*
	6	*n(l(a)-m-1)*	*2*
	7	*n(l(a)-m-1)*m*	*1*
	8	*n(l(a)-m-1)*m*	*1*
	9	*n(l(a)-m-1)*m*	*2*
	10	*l(a)-m-1*	*1*
	11	*l(a)-m-1*	*2*

1	2	3	4
1.4	1	*m*	*1*
	2	*1*	*1*
	3	*l(b)*	*2*
	4	*l(b)*	*2*
	5	*l(b)*	*1*
	6	*l(b)*	*2*
	7	*l(b)*l(a)*	*1*
	8	*l(b)*l(a)*	*1*
	9	*l(b)*l(a)*	*2*
	10	*l(b)*	*1*
	11	*l(b)*	*2*
	12	*1*	*Сложн(1.3)*
	13	*1*	*1*
1.5	1	*l(a)*	*1*
	2	*1*	*1*
	3	*l(a)*	*2*
	4	*l(a)*	*1*
	5	*l(a)*	*1*
	6	*l(a)*	*2*
	7	*1*	*Сложн(1.3)*
	8	*1*	*1*
1.6	1	*l(a)*	*1*
	2	*1*	*m(2m-1)*
	3	*1*	*1*
1.7	1	*1*	*0*
	2	*1*	*1*
	3	*l(k)*	*2*
	3.1	*l(k)*	*Сложн(1.6)*
	3.2	*n*l(k)*	*Сложн(1.4)*
	4	*1*	*1*
1.8	1	*1*	*0*
	2	*1*	*1*
	3	*l(m-1)*	*2*
	3.1	*l(m-1)*	*Сложн(1.6)*
	3.2	*n*l(m-1)*	*Сложн(1.4)*
	4	*1*	*1*
1.9	1	*1*	*0*
	2	*l(b)*	*1*
	3	*l(m-1)*	*2*
	3.1	*l(m-1)*	*1*
	3.2	$n*(2^{l(m)+1}+2^{l(m)}-l(m)-2)$	*2*
	3.2.1	$n*(2^{l(m)+1}+2^{l(m)}-l(m)-2)$	*Сложн(1.6)*
	3.3	*l(m-1)*	*Сложн(1.4)*
	3.4	*n* l(m-1)*	*Сложн(1.6)+ Сложн(1.4)*
	4	*1*	*Сложн(1.6) +Сложн(1.4)*

Продолжение таблицы 1

1	2	3	4
2.3	1	$l(a)$	1
	2	1	2
	3	1	1
2.4	1	$l(a)$	1
	2	$l(b)$	1
	3	m	2
	3.1	m	m^2+m
	3.2	m	4
	4	1	1
2.5	1	1	0
	2	1	1
	3	$l(k)$	2
	3.1	$l(k)$	*Сложн(2.3)*
	3.2	$n*l(k)$	*Сложн(2.4)*
	4	1	1
2.6	1	$l(a)$	1
	2	1	2
	3	1	1
2.7	1	1	0
	2	$l(b)$	1
	3	$l(m-1)$	2
	3.1	$l(m-1)$	1
	3.2	$n*(2^{l(m)+1}+2^{l(m)}-l(m)-2)$	2
	3.2.1	$n*(2^{l(m)+1}+2^{l(m)}-l(m)-2)$	*Сложн(2.3)*
	3.3	$l(m-1)$	*Сложн(2.4)*
	3.4	$n*l(m-1)$	*Сложн(2.3)+ Сложн(2.4)*
	4	1	*Сложн(2.3) +Сложн(2.4)*

Заметим, что в алгоритме 2.4 на шаге 2.1 рассчитана сложность через подсчет простейших операций при умножении матрицы размерностью m на m на вектор размерностью m, но при программной реализации эта сложность значительно снизится за счет архитектурно зависимых оптимизаций.

В таблице 2 приведена суммарная вычислительная сложность приведенных алгоритмов, полученная сложением вычислительных сложностей каждого шага соответствующих алгоритмов.

Таблица 2

№ алгоритма	Вычислительная сложность	Примечание
1	2	3
1.3	$l(a)+1+(7.5+2m)(l(a)-m-1)=m^2+3.25m-6.5$	*Для расчета зависимости вычислительной сложности от «m» выбираем l(a) равной 1.5m*
1.4	$m+2+10l(b)+4l(a)l(b)+$ *Сложн(1.3)*$=5m^2+14.25m-4.5$	$l(a)=l(b)=m$
1.5	$7l(a)+2+$ *Сложн(1.3)*$=m^2+10.25m-4.5$	$l(a)=m$
1.6	$l(a)+m(2m-1)+1=2m^2+1$	$l(a)=m$

1	2	3
1.7	$2+l(k)(2+ Сложн(1.6)+0.5 Сложн(1.4))=2+l(k)(4.5\ m^2+7.125m+0.75)$	$l(a)=m$
1.8	$2+l(m-1)(2+ Сложн(1.6)+0.5 Сложн(1.4))=40.5\ m^2+64.125m+8.75$	$n=0.5,\ l(a)=l(b)=m,\ l(m-1)=9,$ так как m ограничено в диапазоне $160 \le m \le 512$ и большая часть чисел представлены девятью битами
1.9	$l(b)+l(m-1)(3+ Сложн(1.4))+ n*(2^{l(m)+1}+ 2^{l(m)}-l(m)-2)(2+ Сложн(1.6))+(Сложн(1.4)+ Сложн(1.6))(0.5l(m-1)+1)=2027.5m^2+207.625m+2883.25$	$n=0.5,\ l(a)=l(b)=m.\ l(m-1)=l(m)=9$
2.3	$l(a)+3=m+3$	$l(a)=m$
2.4	$l(a)+l(b)+m(m^2+m+4)+1=m^3+m^2+8m+1$	$l(a)=l(b)=m$
2.5	$2+l(k)(2+ Сложн(2.3)+0.5 Сложн(2.4))=2+l(k)(0.5\ m^3+0.5m^2+5m+5.5)$	$n=0.5,\ l(a)=m$
2.6	$l(a)+3=m+3$	$l(a)=m$
2.7	$l(b)+l(m-1)(3+ Сложн(2.4))+ n*(2^{l(m)+1}+ 2^{l(m)}-l(m)-2)(2+ Сложн(2.3))+ (Сложн(2.4)+ Сложн(2.3))(0.5l(m-1)+1)=14.5m^3+14.5m^2+1094.5m+4918$	$n=0.5,\ l(a)=l(b)=m,\ l(m-1)=9$

При расчете суммарной сложности все вероятности появления символов 1 и 0 брались равными 50%.

4. Сравнение вычислительной сложности алгоритмов на эллиптической кривой над полем $GF(2^M)$

Проведем сравнение вычислительной сложности основных операций выполняемых при криптографических преобразованиях на эллиптической кривой над полем $GF(2^M)$. При преобразовании на эллиптической кривой основными операциями является: сложение двух точек и удвоение точки. Проведем анализ сложности при выполнении операции удвоение - основной операции при скалярном произведении большого числа на точку, принадлежащей эллиптической кривой.

Удвоение точки вычисляется по формуле $2*(x_1,y_1)=(x_3,y_3)$, где

$$x_3 = \lambda^2 + \lambda + a,$$
$$y_3 = x_1^2 + (\lambda + 1)\, x_3$$
$$\lambda = x_1 + \frac{y_1}{x_1}.$$

1. Полиномиальный базис

$Сложн(\lambda)=2027.5m^2+208.625m+2883.25$

$Сложн(x_3)= 2\ m^2+2m+1$

$Сложн(y_3)= 7\ m^2+16.25m-3.5$

Сложн(Суммарная1)$= Сложн(\lambda)+Сложн(x_3)+ Сложн(y_3)= 2036.5m^2 + 226.875m + 2880.75$

2. Нормальный базис

$Сложн(\lambda)=14.5m^3+14.5m^2+1095.5m+4918$

$Сложн(x_3)=3m+3$

Сложн(y_3) = $m^3+m^2+11m+4$

Сложн(Суммарная2)= *Сложн(λ)* + *Сложн(x_3)* + *Сложн(y_3)* = $15.5m^3 + 15.5m^2 + 1109.5m + 4925$

Теперь мы можем проанализировать зависимость вычислительной сложности алгоритмов от значения m. Полученные результаты приведены в таблице 3.

Таблица 3

Значение m	**Полиномиальный базис**	**Нормальный базис**
1	2	3
15	14515	2417
30	57579	14708
45	129281	46833
60	229621	108602
75	358600	209824
90	516217	360307
105	702472	569859
120	917365	848289
135	1160897	1205406
150	1433067	1651018
165	1733875	2194934
180	2063322	2846963
195	2421407	3616913
256	4172657	8167237
512	16686728	65156593
1024	66739382	520637257

Полученные результаты представлены на рисунке 1.

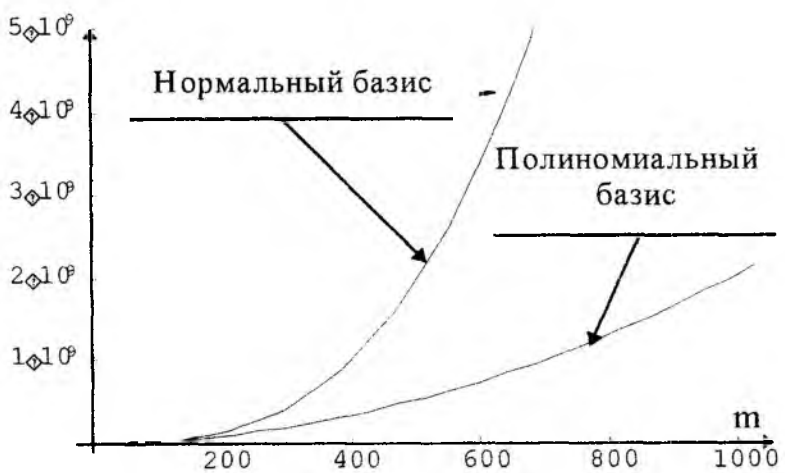


Рис. 1

Кривые, показанные на рисунке 1, имеют две точки пересечения: $m \approx 1.254$, $m \approx 129.941$. Но масштаб рисунка 1 не дает возможность этого увидеть, поэтому приведем рисунок 2 с более крупным масштабом.

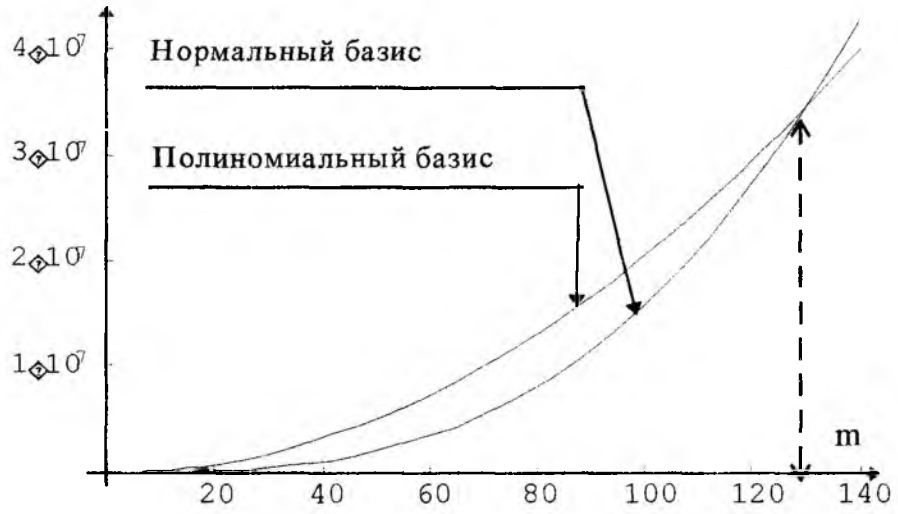


Рис. 2

Заключение

Таким образом, более предпочтительным является нормальный Гауссовский базис, сточки зрения минимизации вычислительной сложности при $m<129$, а при $m>129$ - полиномиальный базис. Учитывая то, что преобразования на эллиптической кривой используют $m>160$, в целом более предпочтительным является полиномиальный базис.

Список литературы: **1.** T. *ELGAMAL*, A public key cryptosystem and a signature scheme based on discrete logarithms, IEEE Transactions on Information Theory, 31 (1985), 469-472. **2.** *M. Blaze, W. Diffie, R. Rivest, B. Schneier, T. Shimomura, E. Thompson, AND M. Wiener*, Minimal key lengths for symmetric ciphers to provide adequate commercial security, January 1996. **3.** *IEEE P1363* / D11 (Draft Version 11). Standard Specifications for Public Key Cryptography. Annex A (Informative). Number-Theoretic Background. **4.** *Кнут Д.* Основы программирования для ЭВМ: В 3 т. – М.: Мир, 1977. Т.2. Получисленные алгоритмы.– 724 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.324.067

О.В. ЖИЛИН, А.И. БАЗИЛЕВСКИЙ

ИСПОЛЬЗОВАНИЕ ОДНОСТОРОННЕГО ПРЕОБРАЗОВАНИЯ, ОСНОВАННОГО НА ФУНКЦИЯХ ЛЮКА В НЕСИММЕТРИЧНЫХ КРИПТОСИСТЕМАХ

Введение

В настоящее время вопросы защиты информации в компьютерных системах приобрели особую актуальность. Это связано с более широким применением компьютерной техники в государственных и коммерческих структурах. В связи с все возрастающими требованиями к целостности, конфиденциальности и аутентичности сообщений и их источников все большее внимание уделяется несимметричным алгоритмам шифрования.

Самым первым, и в то же время по-прежнему наиболее популярным несимметричным алгоритмом является алгоритм RSA [1], разработанный Ривестом, Шамиром и Адлманом в 1978 году. На данный момент существует множество реализаций этого алгоритма, он стандартизирован и широко применяется. Разумеется, что хорошо изучены и слабые стороны этого алгоритма. В связи с этим постоянно происходит поиск новых криптоалгоритмов, которые бы превзошли RSA по показателю Быстродействие/Стойкость. Одной из таких попыток являются системы класса Эль-гамаля [2]. Эти системы более стойкие, чем RSA, однако работает на порядок медленнее.

В данной работе рассматривается криптосистема, основанная на математическом аппарате функций Люка. Эти системы имеют перед RSA ряд преимуществ в стойкости, но, к сожалению, их быстродействие уступает RSA. Реально реализуемые на ЭВМ методы построения таких систем были предложены П. Смитом и М. Ленном в [3].

1. Анализ несимметричных криптосистем

Впервые принципы построения несимметричных криптосистем были рассмотрены Диффи и Хелманом в [4]. Они предложили использовать в криптографии односторонние функции. Эти функции характерны тем, что вычисление обратной функции за разумное время вычислительно сложно или невозможно, если неизвестна дополнительная информация, обычно называемая ключом. Вид функции и ключ не зависят от посылаемого сообщения. Обычно используется числовая функция, в качестве ключа используется число определенной длины, вычисление функции называется преобразованием по закону открытого ключа, а вычисление обратной функции – преобразованием по закону секретного ключа. Наибольшая безопасность в криптосистеме достигается, если имеется возможность производить секретные преобразования только у одного из абонентов. Это позволяет хранить секретный ключ в строгой тайне. Открытый ключ, предназначенный для шифрования сообщений, может быть широко распространен, однако зашифрованные сообщения будет иметь возможность читать только хозяин секретного ключа.

Диффи и Хелман указали на то, что такая же методология может быть использована для цифровой подписи сообщений, то есть обеспечения их аутентичности. В этом случае подписывающий сообщение абонент выполняет над открытой подписью преобразование по закону секретного ключа, а все владельцы открытого ключа могут проверить подпись, осуществив преобразование по закону открытого ключа и проанализировав полученную открытую подпись.

После публикации статьи Диффи и Хелмана было предложено множество односторонних функций. Наиболее широко известной и применяемой является функция, используемая в криптоалгоритмах RSA и Эль-Гамаля. Эти алгоритмы основаны на возведении сообщения, рассматриваемого как число, в степень по модулю большого числа.

Преобразование по закону открытого ключа в системе RSA представляется следующим соотношением:

$$f_{RSA} = M^e \pmod N. \qquad (1)$$

Если M – шифруемое сообщение, то $f_{RSA}(M)$ – криптограмма M'. Для того чтобы сделать f_{RSA} односторонней функцией, N выбирается как произведение 2-х больших простых чисел p, q. Для 1преобразования по закону секретного ключа необходимо число d, такое что:

$$e \cdot d \equiv 1 (\mathrm{mod}\,(p-1)\cdot(q-1)) \quad (2)$$

Если e взаимнопросто с (p-1)(q-1), то d всегда может быть вычислено при известных p, q. Преобразование по закону секретного ключа выглядит следующим образом:

$$M = (M')^d \,(\mathrm{mod}\, N). \quad (3)$$

При этом необходимо соблюдение условия M<N. Как видно, преобразования (1) и (3) отличаются только используемым ключом e или d.

Первой трудностью при реализации RSA является то, что в качестве N используются длинные числа (минимум 512 бит). В связи с этим возникает проблема реализации арифметики многократной точности. В настоящее время существуют алгоритмы возведения в степень, позволяющие возводить длинные числа в степень за время $O(\log_2 N)$, где N – показатель степени.

Функция f_{RSA} является односторонней, так как не существует способа вычисления M из M' по формуле (3) без знания d, а также способа вычислить d из e и N, кроме факторизации N. Было доказано, что любой метод получения d приблизительно эквивалентен или сложнее факторизации. В настоящее время методом решета общего числового поля было факторизовано число длиной 428 бит, однако, выбирая N достаточно большим (например, 1024 бита) можно обеспечить определенную стойкость системы. Сегодня RSA считается системой с доказуемо стойкостью (вероятно стойкой). Стойкость RSA обеспечивается сложностью задачи факторизации модуля.

Одной из слабых сторон RSA системы является то, что для цифровой подписи (ЦП) по RSA возможен адаптивный криптоанализ с выбранным текстом. При этом возможно получить подпись ложного сообщения после того, как абонент-жертва атаки подпишет несколько сообщений, вид которых специально подобран злоумышленником. Эта атака возможна из-за мультипликативности ЦП по RSA, то есть

$$M^d \cdot L^d = (M \cdot L)^d. \quad (4)$$

Если злоумышленнику известны M, M^d, L, L^d, то ML и $(ML)^d$ могут быть вычислены без знания d.

Такая ситуация, возможно, трудно реализуема, так как сообщения перед подписью обычно хешируют, однако применение хеш-функций (ХФ) не может гарантировать невозможность подделки подписи. Возможна следующая последовательность действий злоумышленника:

Выбирается ложное сообщение, хешируется, если необходимо и раскладывается на простые сомножители.

Генерируется множество "невинных" сообщений, они, если необходимо, хешируются, факторизуются и среди них выбираются те сообщения, которые содержат множители, полученные на шаге 1 или множители других сообщений, полученных на данном шаге.

Добываются подписи сообщений, полученных на шаге 2.

В процессе элиминации получают подписи простых чисел-сомножителей сообщений.

Путем перемножения подписей сомножителей ложного сообщения в соответствующей степени получаем подпись ложного сообщения.

Для реализации такой атаки необходимо большое количество сообщений, однако существуют и другие способы добычи подписей набора простых чисел, и данная атака может оказаться реально осуществимой. Это означает, что хеширование не может гарантированно защитить подпись от подделки. В отличие от RSA подписи в описываемой здесь системе, надежно защищены от таких атак, так как они не мультипликативны.

2. Определение и свойства функции Люка

Данная система основана на односторонней функции, отличной от тех, что используются в системах RSA и Эль-Гамаля, определяемой функциями Люка. Поскольку эти функции можно считать обобщением степеней, то преобразования по закону открытого и секретного ключей производится

аналогичным образом и все атаки на эту систему могут быть использованы и для RSA, но не наоборот что доказывает, что LUC является более стойкой системой.

Функции Люка являются частным случаем линейных рекуррентных соотношений высшег порядка. Если $P_1, P_2, ..., P_m$ - целые числа то можно определить последовательность целых чисел $\{T_n\}$ следующим образом:

$$T_n = P_1 \cdot T_{n-1} + P_2 \cdot T_{n-2} + .. + P_m \cdot T_{n-m} \quad (5)$$

$T_1, T_2, .. T_{m-1}$ определяются независимо. Выражение (5) называется линейным рекуррентным соотношением (ЛРС) m-го порядка. Можно доказать, что последовательность, определяемая ЛРС первого порядка состоит из степеней P_1, умноженных на константу T_0. Следовательно, ЛРС более высоких порядков являются обобщением степеней и не удивительно, что RSA система может быть реализована с использованием ЛРС порядка, большего, чем первый.

В данном случае мы обсудим только ЛРС 2-го порядка, которую можно представить в виде:

$$T_n = P \cdot T_{n-1} - Q \cdot T_{n-2} \quad (6)$$

В виде (5) выражение (6) можно представить следующим образом:

Пусть a и b - корни следующего полинома 2-го порядка:

$$x^2 - P \cdot x + Q = 0 \quad (7)$$

Пусть c_1 и c_2 – целые числа, тогда последовательность $\{c_1 a^n + c_2 b^n\}$ обладает следующим свойством:

$$P \cdot (c_1 \cdot a^{n-1} + c_2 \cdot b^{n-1}) - Q \cdot (c_1 \cdot a^{n-2} + c_2 \cdot b^{n-2}) = c_1 \cdot a^n + c_2 \cdot b^n \quad (8)$$

Таким образом, любая последовательность вида (6) может быть представлена в виде $\{c_1 a^n + c_2 b^n\}$, где $T_0 = c_1 + c_2$, $T_1 = c_1 a + c_2 b$. Следует отметить, что если T_0 и T_1 – целые, то последовательность будет состоять из целых чисел, хотя c_1, c_2, a, b могут быть комплексными.

Нас интересует два подмножества последовательностей вида (6). Они обозначаются $\{U_n\}$ и $\{V_n\}$ и определяются:

$$U_n = \frac{a^n - b^n}{a - b} \quad \left(c_1 = \frac{1}{a - b} = -c_2\right), \quad (9)$$

$$V_n = a^n + b^n \quad (c_1 = 1 = c_2). \quad (10)$$

Это всегда будут последовательности целых чисел, первые два члена которых

$$U_0 = 0, U_1 = 1, \; V_0 = 1, V_1 = P. \quad (11)$$

Эти последовательности зависят только от P и Q и называются функциями Люка от P и Q. Иногда они обозначаются Un(P,Q) и Vn(P,Q). Расширенная теория этих функций обсуждается Леммером в работе [5].

Следует отметить, что для любого целого N выполняется соотношение:

$$U_n(P \bmod N, Q \bmod N) = U_n(P, Q) \pmod N \quad (12)$$

Доказательство этого соотношения можно получить методом математической индукции. Аналогично

$$V_n(P \bmod N, Q \bmod N) = V_n(P, Q) \pmod N \quad (13)$$

Поскольку корни полинома (7) a и b удовлетворяют соотношениям:

$$a + b = P, \; a \cdot b = Q, \quad (14)$$

то нетрудно вывести множество соотношений между функциями Люка Un и Vn и коэффициентами ЛРС (6), P и Q. Дискриминант полинома (7), $D = P^2 - 4Q$, может быть выражен через a и b:

$$D = (a - b)^2. \quad (15)$$

Приведем некоторые из этих соотношений:

$$V_{2n} = V_n^2 - 2 \cdot Q^n, \tag{16}$$

$$V_{2n-1} = V_n \cdot V_{n-1} - P \cdot Q^n, \tag{17}$$

$$V_{2n+1} = P \cdot V_n^2 \cdot V_{n-1} - Q \cdot V_n \cdot V_{n-1} - P \cdot Q^n, \tag{18}$$

$$V_n^2 = D \cdot U_n^2 + 4 \cdot Q^n, \tag{19}$$

$$2 \cdot V_{n+m} = V_n \cdot V_m + D \cdot U_n \cdot U_m, \tag{20}$$

$$2 \cdot Q^m \cdot V_{n-m} = V_n \cdot V_m - D \cdot U_n \cdot U_m. \tag{21}$$

Еще одним важным свойством функций Люка является:

$$V_n(V_k(P,Q),Q^k) = V_{nk}(P,Q). \tag{22}$$

Из данного выражения следует, что мы можем сформулировать правило умножения функций Люка, аналогичное правилу умножения степеней, где индекс функции Люка будет выступать аналогом показателя степени. Приняв Q=1 получим более простое соотношение:

$$V_{nk}(P,1) = V_n(V_k(P,1),1). \tag{23}$$

Существует связь между индексом функции Люка и делителями ее значения. Для пояснения необходимо ввести определение символа Лежандра:

Пусть a – целое, p – простое. Символ Лежандра $L(a,p)$ равен 1, если существует b, такое что $b^2 \equiv a \pmod p$, иначе он равен -1.

Если p - простое, не делитель Q или D, и e=L(D,p), то, как Леммер показал в [5],

$$U_{k(p-e)}(P,Q) \equiv 0 \pmod p \text{ для любого целого к.} \tag{24}$$

а также

$$V_{k(p-e)}(P,Q) \equiv 2 \cdot Q^{k(1-e)/2} \pmod p \text{ для любого целого к.} \tag{25}$$

Для функций Люка, зависящих от взаимнопростых целых P и Q, существует обобщение функции Эйлера, называемой функцией Леммера [5]. В данном случае нам необходимо применить ее для числа N вида N=pq, где p и q различные нечетные простые числа. Тогда функция Леммера определяется следующим образом:

$$T(N) = [p - L(D,p)] \cdot [q - L(D,q)]. \tag{26}$$

Как и в случае функции Эйлера полное произведение необязательно для получения нужного результата, достаточно найти наименьшее общее кратное (НОК) сомножителей (26). Тогда обобщенная функция Леммера определяется следующим образом:

$$S(N) = НОК([p - L(D,p)] \cdot [q - L(D,q)]). \tag{27}$$

Из выражения (27) следует, что выражения (24) и (25) можно записать в виде:

$$U_{kS(N)}(P,1) \equiv 0 \pmod N \text{ для любого целого к.} \tag{28}$$

$$V_{kS(N)}(P,Q) \equiv 2 \pmod N \text{. Для любого целого к.} \tag{29}$$

Тогда если N=pq – произведение двух различных нечетных простых чисел, P<N и взаимнопростое с N, e – любое число взаимнопростое с S(N), а d вычислено по расширенному алгоритму Евклида из соотношения ed=kS(N)+1, тогда:

$$V_d(V_e(P,1),1) \equiv P \pmod N. \tag{30}$$

Этот результат позволяет определить одностороннюю функцию, аналогичную задаваемой формулами (1) и (3), основанную на функциях Люка. Однако явно видно отличие – отсутствие

симметрии между $V_e(P,1)$ и инверсной функцией $V_d(R,1)$. d и e связаны через S(N), то есть через квадратичные вычеты дискриминанта D, который зависит от P.

Для функций Люка справедливо также соотношение

$$L(D,p) = L\left(P^2 - 4, p\right) = L(V_e^2(P,1) - 4, p) \quad (31)$$

Это означает, что значение S(N) одно и то же для P и функции Люка $V_e(P,1)$.

Используя полученные результаты можно построить систему с открытыми ключами, аналогичную RSA. Пусть выбраны числа N и e, такие что N - произведение двух различных нечетных простых чисел p и q, e взаимнопросто с (p-1)(q-1)(p+1)(q+1). Пусть M - сообщение, взаимнопростое с N. Определим преобразование по закону открытого ключа как

$$f_{LUC}(M) = V_e(M,1) \pmod N. \quad (32)$$

Для определения соответствующего преобразования по закону открытого ключа необходимо число d, вычисляемое из соотношения

$$de = 1 \pmod{S(N)}. \quad (33)$$

Преобразование по закону секретного ключа определяется заменой ключа e на d:

$$M = V_d(M',1) \pmod N. \quad (34)$$

Из соотношения (30) следует, что соотношение (34) будет всегда верно при соблюдении всех условий.

3. Возможности реализации криптографических систем с использованием функций Люка

При реализации системы LUC возникают две главные проблемы. Первая из них - вычисление значений функций V_e и V_d, для больших значений e и d, а также то, что значение ключа d зависит от сообщения. Первая проблема решаема, так как из соотношений (16)-(19) следует, что для вычисления функций Люка можно использовать обычный двоичный алгоритм для возведения длинных чисел в степень по модулю. При этом могут быть использованы любые эвристические процедуры, применяемые для данного алгоритма (например, блочный метод) [6]. Такой метод вычисления гарантирует, что вычисление функции Люка по временным затратам сравнимо с возведением числа в соответствующую степень и превосходит его всего на 50%.

Что касается второй трудности, то на самом деле существует всего 4 возможных значения ключа d, так как существует всего 4 возможных значения S(N): НОК[(p+1)(q+1)], НОК[(p+1) (q−1)], НОК[(p−1) (q+1), НОК[(p−1) (q−1). Все эти значения известны когда известен модуль N, поэтому соответствующие 4 значения ключа d могут быть вычислены заранее. Для выбора нужного ключа при произведении преобразования по закону секретного ключа вычисляется дискриминант D сообщения или криптограммы, затем вычисляются символы Лежандра L(D,p), L(D,q) и выбирается соответствующий ключ. Это также означает, что, в отличие от RSA, делители модуля p и q должны храниться вместе с секретными ключами и, разумеется, в строгом секрете.

Наличие 4 ключей приводит к возрастанию времени преобразования по закону секретного ключа. Вычисление символов Лежандра L(D,p), L(D,q) по вычислительной сложности равно $O(\log_2 p)+O(\log_2 q)$. Таким образом, данный процесс занимает приблизительно на 80% больше времени, чем аналогичный процесс в RSA.

Существует способ избежать вычисления и значения четырех секретных ключей и составляющих модуля. Для этого вместо функции S(N) в выражение (33) подставляется функция R(N), определяемая следующим образом:

$$R(N) = (p-1)(q-1)(p+1)(q+1). \quad (35)$$

При этом размер ключа d возрастет в два раза, и, соответственно, возрастет время вычисления функции V_d, причем практически в два раза.

Аналогично RSA, преобразование по закону секретного ключа может быть осуществлено злоумышленником только если он найдет способ вычисления V_d без знания d, или вычисления d из e и

N. Решение первой проблемы, как и в RSA, - полный перебор значений. Отличие в решении второй проблемы заключается в том, что для каждой пары e и N существует 4 ключа d, применяемые в зависимости от сообщения.

Поскольку функции Люка - обобщение степеней, то удачная криптографическая атака на LUC является успешной и для RSA. Однако, так как в LUC применяются дополнительные усложнения, то обратное утверждение неверно. Например, поскольку подпись LUC не мультипликативна, то адаптивный криптоанализ для такой системы не применим. Из этого следует, что LUC криптографически сильнее RSA.

Фактически, любая криптографическая система, основанная на возведении в степень, может быть преобразована для использования функций Люка для определения односторонней функции. Например, можно разработать систему класса Эль-Гамаля, основанную на функциях Люка.

По классическому алгоритму шифрования по Эль-Гамалю, в системе генерируется простое число p и первообразный элемент a. Затем каждый абонент генерирует себе секретный ключ x, затем рассылает открытый ключ $y=a^x$. Для того, чтобы зашифровать сообщение генерируется случайный сеансовый ключ k и вычисляется $L=y^k \bmod p$. Вычисляются две части криптограммы: $c1=a^k$, $c2=L\cdot M$, которые отправляются абоненту.

На приемной стороне вычисляется $L=(ak)^x=(c1)^x$, с использованием секретного ключа x. Затем вычисляется обратный элемент L, $L'L=1 \bmod p$, и восстанавливается исходное сообщение $M=c2\cdot L'$.

Для шифрования с использованием функций Люка процесс происходит аналогично. Открытый ключ y в этом случае $y=V_x(g,1) \bmod p$. Нам также необходим сеансовый ключ k. Затем мы вычисляем $G=V_k(y,1) \bmod p$. После этого вычисляются две части криптограммы $d1=V_k(g,1) \bmod p$, $d2=G\cdot M \bmod p$.

На приемной стороне вычисляется $G=V_x(d1,1) \bmod p$, с использованием секретного ключа x. Затем вычисляется обратный элемент G, $G'\cdot G=1 \bmod p$, и восстанавливается исходное сообщение $M=d2\cdot G'$.

Следует заметить, что, как и в классическом алгоритме, криптограмма в два раза больше модуля.

Такое же преобразование к функциям Люка можно провести для прикладных алгоритмов класса Эль-Гамаля, в которых длина криптограммы гораздо меньше за счет использования двухмодульного преобразования.

На основании теоретических сведений, изложенных выше, были алгоритмически реализованы подпрограммы для вычисления функций Люка, а также производящие генерацию ключей и производящие преобразования по закону этих ключей для системы LUC, а также для шифрования и цифровой подписи, классической и прикладной по алгоритмам класса Эль-Гамаля.

Разработка производилась на алгоритмическом языке C. Результатом разработки является модули с подпрограммами, которые могут быть откомпилированы для подключения и использования в 16-битных и 32-битных приложениях. Компиляция для 16-битных приложений производилась с помощью компилятора Borland C++ 3.1, а для 32-битных – с помощью Borland C++ Builder.

Далее рассматриваются временные характеристики разработанного программного обеспечения.

В таблице 1 приводятся временные характеристики вычисления функций Люка для 512 битного модуля под управлением ОС MS-DOS. Данные в таблице 1 показывают, что вычисление V-функции Люка приблизительно в два раза дольше возведения в степень по модулю, что соответствует теоретическим оценкам.

Таблица 1

Функция	Время выполнения, с
Блочное возведение в степень по Монтгомери	0.143900
Вычисление V- функции Люка	0.301000
Вычисление UV - функции Люка	0.472400
Вычисление UV - функции Люка по блочному алгоритму	0.420200

В таблице 2 приведены временные характеристики системы LUCRSA.

Таблица 2

Действие	Время выполнения, с
Подпись (Расшифровка)	0.362500
Проверка подписи (Шифрование)	0.316900

В таблице 3 приведены временные характеристики криптосистемы шифрования по классическому Эль-Гамалю с использованием функций Люка.

Таблица 3

Действие	Время выполнения, с
Шифрование	0.580100
Расшифровка	0.294400

В таблице 4 приведены временные характеристики криптосистемы ЦП по классическому Эль-Гамалю с использованием функций Люка.

Таблица 4

Действие	Время выполнения, с
Подпись	0.431700
Проверка	1.132600

В таблице 5 приведены временные характеристики криптосистемы ЦП DSS с использованием функций Люка.

Таблица 5

Действие	Время выполнения, с
Подпись	0.075800
Проверка	0.288900

Полученные результаты показывают, что криптосистемы с использованием функций Люка уступают традиционным аналогам по быстродействию приблизительно в два раза.

Заключение

В данной работе проведено исследование о применимости функций Люка в качестве односторонних функций в несимметричных алгоритмах, а также приведены практические результаты такого использования. Результаты показывают возможность использования разработанных алгоритмов в криптографических системах. Они показывают, что по времени работы эти алгоритмы сравнимы с традиционными, однако теоретические исследования показывают более высокую стойкость этих алгоритмов. В частности показана невозможность применения адаптивного криптоанализа в системах с использованием алгоритмов, основанных на функциях Люка.

Как показал Шеннон, применение в криптосистеме разнообразных алгоритмов позволяет существенно увеличить стойкость системы. В настоящее время в мире в большинстве применяются традиционные криптосистемы, основанные на возведении в степень по модулю, следовательно, хорошо изучены сильные и слабые стороны таких систем. Изменение алгоритма при остающейся неизменной математической базе редко позволяет получить алгоритм, стойкость которого на порядок выше предшественников. Поэтому необходим постоянный поиск и обоснование нового математического аппарата для построения криптосистем. Одним из примеров такого математического аппарата являются криптосистемы, основанные на функциях Люка.

Список литературы: **1.** *R.L. Rivest, A. Shamir, L.M. Adleman.*– A method for obtaining digital signatures and public key cryptosystems. Comm. ACM, 1978. pp120–126. **2.** *T. El Gamal.* A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE Transactions of Information Theory (July 1985) pp100–140. **3.** *P. J. Smith, M. J. Lennon* – LUC: A New Public Key System. AsiaCrypt'93. pp 230–250. **4.** *W. Diffie, M. Hellman* - New directions in cryptography, IEEE Transaction of Information Theory, 22 (1976) pp644-654. **5.** *H. Lehmer* - An extended theory of Lucas' functions, Annals of Math., 31 (1930) pp 419-448.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06

М.А. КРИВОШЛЫК, Е.Г. КАЧКО, канд. техн. наук

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ ММХ ДЛЯ ВЫПОЛНЕНИЯ ЦЕЛОЧИСЛЕННЫХ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ НАД ЧИСЛАМИ МНОГОКРАТНОЙ ТОЧНОСТИ

Объем и сложность данных, обрабатываемых современными компьютерами, стремительно увеличиваются. Постоянно возрастающая мощность аппаратных средств, используемых злоумышленником для реализации угроз нарушения конфиденциальности, целостности, аутентичности, доступности информации приводит к необходимости создания все более совершенных систем защиты. Такие системы могут быть созданы как за счет применения более совершенных алгоритмов, так и за счет увеличения размера используемых ключевых параметров. Последнее приводит к усложнению вычислений и возрастанию времени криптопреобразований, однако сложность криптоанализа увеличивается при этом по экспоненциальному закону.

Решение задачи криптографической защиты информации требует выполнения большого объема целочисленных вычислений, таких как арифметические действия над числами многократной точности. Данные операции приводят к сильной загрузке центрального процессора. В связи с этим целесообразно было бы попытаться переложить часть вычислений с CPU на другие вычислительные устройства компьютера. Такими устройствами являются блок с плавающей точкой (FPU) и устройство для выполнения ММХ команд.

Технология ММХ была разработана фирмой Intel для ускорения мультимедиа и коммуникационных программ. Она включает в себя новые команды и типы данных и основана на их параллельной обработке. При этом сохраняется полная совместимость с существующими операционными системами и программным обеспечением. ММХ-технология - одно из самых значительных усовершенствований со времени создания 32-разрядной архитектуры. В основе ММХ лежит принцип SIMD (Single Instruction Multiple Data), т.е. одной командой можно обработать сразу несколько единиц информации.

Эффективное использование возможностей FPU и ММХ-технологии при выполнении целочисленных вычислений над числами многократной точности может привести к повышению производительности компьютера и значительно уменьшить время выполнения криптопреобразований.

ММХ-регистры отображены на поля мантиссы в FPU-регистрах. Значение, записываемое в ММХ-регистр, автоматически появляется в младших битах (биты 63-0) соответствующих FPU-регистров. При этом в поле порядка (биты 78-64) и в знаковый бит (бит 79) заносятся единицы. Значение поля TOS (Top Of Stack) слова состояния FPU устанавливается в 0 после выполнения каждой ММХ-команды. Значение мантиссы, записываемое в FPU-регистр с помощью FPU-команды, автоматически появляется в соответствующем ММХ-регистре [1].

Отображение ММХ-регистров фиксировано и не зависит от значения поля TOS (биты 11-13 в регистре состояния FPU). В обозначении MMn n указывает на физический номер регистра, в отличие от регистров FPU, где в обозначении Stn n указывает на относительный номер регистра (относительно поля TOS).

После выполнения любой ММХ-команды (кроме EMMS) значения всех полей регистра тэгов устанавливаются в 00. Команда EMMS устанавливает значения всех полей регистра тэгов в 11. Значение регистра тэгов не оказывает никакого влияния на ММХ-регистры или выполнение ММХ-команд.

Так как ММХ и FPU используют физически одни и те же регистры, для сохранения и восстановления контекста ММХ используются команды FSAVE (Store FP state) и FRSTOR(Restore FP state).

ММХ-технология вводит 4 новых 64-разрядных типа данных :

- упакованные байты;
- упакованные слова;
- упакованные двойные слова;
- учетверенное слово.

В памяти новые типы данных располагаются так, как это принято в Intel-архитектуре, то есть по принципу младший байт первым .

MMX поддерживает новую арифметику, называемую арифметикой с насыщением (Saturation arithmetic). Сравним ее с привычной арифметикой с циклическим переносом (Wraparound arithmetic).

В режиме циклического переполнения в случае переполнения результат обрезается и возвращаются только младшие биты результата. Данный метод хорошо известен, он используется при операциях над целочисленными регистрами. В режиме насыщения результат операции, который выходит за границу размера данных, насыщается до предельно возможного значения для используемого типа данных.

Большую проблему при использовании MMX для целочисленных вычислений представляет то, что MMX-команды не сообщают о переполнении с помощью генерации исключений или установки каких-либо флагов.

Эмуляция MMX-команд невозможна.

Система команд MMX состоит из 57 инструкций, которые можно подразделить на следующие классы:

- арифметические команды;
- команды преобразования;
- команды сравнения;
- команды логических операций;
- команды переноса данных;
- команды сдвига;
- команда EMMS.

Задача использования технологии MMX для целочисленных арифметических вычислений разбивается на 2 подзадачи:

- использование MMX совместно с целочисленными командами процессора;
- использование MMX и команд FPU параллельно с целочисленными командами.

Первая подзадача состоит в попытке использования 64-разрядной архитектуры и нового набора арифметических команд MMX.

Вторая подзадача - в разбиении кода приложения на 2 части. Первая часть при этом выполняется в центральном процессоре, вторая параллельно в MMX и FPU . Это возможно благодаря независимости конвейеров CPU и FPU.

Рассмотрим возможности решения первой задачи. MMX позволяет выполнить такие целочисленные операции, как сложение, вычитание и умножение. При их выполнении возникают следующие проблемы.

1. Неприменима арифметика с насыщением, поскольку в случае переполнения или заема результат автоматически округляется, что недопустимо для целочисленных вычислений. В результате система команд MMX сокращается для нас почти в 2 раза.

2. Арифметика с циклическим переносом может применяться, однако в случае переноса или заема не формируются никакие флаги, а результат обрезается. Это приводит к тому, что, например, в случае сложения 2-х чисел максимальной для MMX длины мы не можем быть уверены в старшем бите результата.

3. Большое число MMX команд предназначено для работы со знаковыми числами, то есть старший бит операндов недоступен для вычисления этими командами.

Из 2-й и 3-й проблем можно найти следующий выход: необходимо использовать числа не максимальной длины. Однако разрядность MMX-операндов выровнена на границе слова, а это значит, что уменьшение длины значащей части слагаемых (множителей) на 1 бит сделает ее не кратной степени двойки и приведет с одной стороны к получению правильного результата, а с другой стороны к следующему:

- большие затраты времени, связанные с переводом исходных чисел в такой формат;
- значительно усложняется синтез результата из промежуточных значений в случае работы с числами многократной точности.

Поскольку мы работаем исключительно с такими числами, то данный метод неприменим.

Реальным выходом в данной ситуации оказывается применение операндов с разрядностью в 2 раза меньшей, нежели исходная. Это требует в 2 раза больше выполнений команд, но при этом операнды выровнены на границе слова, что облегчает анализ результата (то есть данный метод требует меньше пред и поствычислений).

Далее приведен пример программы сложения чисел многократной точности длиной 1024 бита с использованием команд MMX.

```
.586
.MMX
model       flat
extrn       ExitProcess:proc
data
x dd    32      dup (0ffffffffh)
y dd    32      dup (0ffffffffh)
z dd    33      dup (?)
.code
begin:  lea         ebx,offset x
lea         esi,offset y
lea         edi,offset z

  mov       cx,32
  mov       eax,1
movd        mm3,eax
movd    mm7,eax
psllq       mm3,32
mov     edx,0ffffh
movd        mm4,edx

for1:mov        eax,[ebx]               ;Загрузка операндов
mov     edx,eax
and     eax,0ffffh
shr         edx,16
movd        mm0,edx
psllq   mm0,32
movd        mm1,eax
por     mm0,mm1
mov     eax,[esi]
mov     edx,eax
and     eax,0ffffh
shr         edx,16
movd        mm1,edx
psllq   mm1,32
movd        mm2,eax
por     mm1,mm2

paddd   mm0,mm1             ;Выполнение сложения
movq    mm2,mm0
movq    mm1,mm0
psrlq   mm1,16
pand        mm1,mm3
```

```
paddd   mm1,mm2
pand    mm0,mm4
psrlq   mm1,16
pxor    mm1,mm7
por     mm0,mm1

movd    edx,mm0             ;Запись результата
  add   [edi],edx
  xor   eax,eax
l2:   jnc   l1
  lea   eax,[eax+4]
  adc   [edi+eax],0
  jmp   l2
l1:   psrlq   mm0,32
movd    eax,mm0
  xor   edx,edx
  add   [edi+4],eax
l3:   jnc   l4
  lea   edx,[edx+4]
  adc   [edi+edx+4],0
  jmp   l3

l4:   add   ebx,4
  add   esi,4
  add   edi,4
  loop  for1

call    ExitProcess
```

Анализ временных параметров программ, выполняющих сложение и умножение целых чисел многократной точности показал их неэффективность. Так сложение MMX выполняется на 75% медленнее целочисленного, а умножение MMX - почти в 4 раза медленнее. Исходя из этого можно сделать вывод о неприменимости 64-разрядной архитектуры MMX и нового набора команд вместо целочисленных арифметических вычислений многократной точности.

Рассмотрим теперь возможности решения задачи разбиения кода приложения на две части с целью параллельного выполнения целочисленных команд и команд MMX и FPU.

К сожалению в данном случае также возникает ряд сложностей. Основной проблемой является то, что при разработке MMX-технологии фирма Intel не предусматривала необходимость совместной работы с данными в FPU и MMX. Сама Intel в технической документации по своим процессорам неоднократно подчеркивала необходимость обнулять регистры FPU при переходе к командам MMX и регистры MMX с помощью команды EMMS при переходе к FPU. Это связано с тем, что хотя по сути дела данные находятся в одних и тех же физических регистрах, реально они имеют разный формат. Как уже было описано выше, при занесении значения в регистр MMX в поле порядка соответствующего ему регистра FPU (биты 78-64) и в знаковый бит (бит 79) заносятся единицы, а значение поля TOS (Top Of Stack) слова состояния FPU устанавливается в 0 после выполнения каждой MMX-команды. Таким образом сложно рассчитать, какой FPU-регистр соответствует конкретному MMX-регистру. Кроме того, для команд FPU данные MMX в FPU-регистрах видны как вещественные знаковые. И наоборот: после выполнения операций над целыми числами в регистрах FPU в MMX-регистрах окажется бесполезный набор цифр.

Конечно, из данной проблемы можно найти выход путем преобразования данных MMX перед выполнением действий в FPU в формат, понятный командам FPU (и наоборот), однако для этого необходимо использовать целочисленные команды процессора, что неприменимо по условию поставленной задачи.

Итак, приходится констатировать невозможность совместного использования данных FPU и MMX. Это фактически означает, что все вычисления придется проводить только силами MMX. А это, к сожалению, невозможно ввиду ограниченности набора команд MMX. Таким образом, данную задачу решить невозможно.

Кроме того, в ходе исследований было определено, что время переключения процессора из режима FPU в режим MMX (и наоборот) очень велико. Это значит, что выигрыш в производительности может быть получен только в том случае, когда используются большие непрерывные участки MMX-кода. Однако при реализации алгоритмов алгоритмов многократной точности этого достичь не удалось.

Мы рассмотрели 2 попытки использования технологии MMX для целочисленных арифметических вычислений, однако они оказались неудачными. Так как других возможностей использования MMX для данной цели не видно, можно сделать вывод о неприменимости MMX для снижения нагрузки на целочисленный процессор.

Некоторые проблемы, описанные в данной статье, на момент ее написания уже удачно решены в технологии Streaming SIMD Extention, реализованной фирмой Intel в процессорах Pentium III и являющейся логичным продолжением технологии MMX. Анализ возможностей использования этой технологии для выполнения арифметических операций над числами многократной точности будет произведен в последующих исследованиях.

Список литературы: **1.** *Бердышев Е.* Технология MMX. Новые возможности процессоров P5 и P6. М.: Диалог-МИФИ, 1998. – 234 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3

Е.Г. КАЧКО, канд. техн. наук, М.В. БЛАГАЙ

ИСПОЛЬЗОВАНИЕ DLL ПРИ РАЗРАБОТКЕ ЗАЩИЩЕННЫХ ПРОГРАММ

В настоящее время широкое распространение получила технология динамического связывания библиотек (Dynamic Link Libraries), которая позволяет оптимально использовать и хранить готовые процедуры и функции, а также ускоряет разработку больших проектов. Практически любая современная операционная система имеет возможности динамического подключения библиотек. Эта статья описывает некоторые структурные особенности реализации динамических библиотек в ОС WINDOWS, снижающие защищенность проектов в целом. Предлагаются методы повышения защищенности таких проектов.

1. Динамическое связывание библиотек

Существует два основных формата файлов динамических библиотек: более старый 16 битный NE (New Executable) и относительно новый 32 битный, наиболее распространенный, PE (Portable Executable) или COFF. NE впервые появился в Windows 3.0, его внутренняя структура довольно сложна и в настоящее время такие DLL мало используется. Формат PE распространен очень широко, так как используется практически везде – в исполняемых файлах приложений под Windows, в динамических библиотеках, в объектных файлах. Структура PE или COFF описывается в документации MICROSOFT.

Рассмотрим простейший пример. Пусть нам необходимо написать программу для кодирования файлов, для входа в которую необходимо ввести персональный ключ-пароль. Имея представление об алгоритмах кодирования, мы выбираем наиболее приемлемый из них, и либо реализуем его сами в виде DLL, либо покупаем готовую динамическую библиотеку. Защиту паролем также реализуем при помощи DLL, в которой производится проверка допустимости введенного пароля. Такая структурная организация позволяет оперативно вносить изменения в систему кодирования и проверки пароля заменой или модификацией DLL. Имея несколько библиотек с различными алгоритмами кодирования можно гибко менять возможности готового приложения, заменой библиотек.

Допустим, программа написана и функционирует без отказов. Рассмотрим возможные пути взлома, которые возникают в связи с использованием динамически подключаемых библиотек.

Пускай, в этом гипотетическом приложении мы используем две динамические библиотеки CODER.DLL и PASSWORD.DLL. На самом деле, имена DLL несут довольно небольшую смысловую нагрузку. В CODER.DLL имеется две функции Coding и Decoding, а в PASSWORD – функция PasswordCheck. Использование длинных символьных имен функций и переменных, описывающих назначение функции или переменной, позволяет лучше ориентироваться в исходном тексте программы, и является хорошим тоном в программировании. Обычно, при компиляции все имена заменяются адресами, но не в случае с DLL. Если используются динамические библиотеки, существует вероятность того, что имена функций останутся в коде программы.

Рассмотрим сам процесс динамического связывания. При компиляции программы, использующей функции из DLL, компилятор заменяет вызов функции из DLL пустышкой следующего вида:

```
Call Dword Ptr [ Function_Stub_Address ] ,
```

либо второй вариант

```
Call near Function_Stub
.....
Function_Stub: Jmp Dword Ptr [Function_Stub_Address]
```

Этот фрагмент кода в мнемониках ассемблера означает *Переход по адресу, находящемуся по адресу Fuction_Stub_Address* . Таким образом, в коде программы не указывается адрес функции, а только

адрес, по которому хранится указатель на функцию. Во время выполнения программы, процессор перейдет по адресу, который указан в двойном слове Function_Stub_Adderss. Вся работа по установке правильных адресов во все такие заглушки выполняется операционной системой на этапе загрузки программы или библиотеки в память. Загрузчик ОС получает полную информацию, необходимую для связывания, из файла программы, причем существует два способа связывания – импортирование по имени функции, либо по номеру функции. Рассмотрим формат PE файла.

2. Формат PE файла

Файлы формата PE [1]состоят из секций, каждая из которых имеет свое назначение. Из основных можно назвать – Text, Data, BSS, Relocs, TLS, idata, edata. В первой собран весь код программы, во второй все инициализированные данные, в третьей – не инициализированные данные. Секция Relocs предназначена для хранения информации о необходимых привязках, она используется если образ программы (Program Image) загружается по другому адресу, отличному от стандартных 4Mb для Windows 95 или 1Mb для Windows NT. В секции TLS описаны все нити или цепочки (Threads) программы. В последних двух секциях хранится информация об экспорте – импорте. На основе именно этой информации и происходит связывание динамических библиотек.

Object table:

#	Name	VirtSize	RVA	PhysSize	Phys off	Flags
01	.text	00026000	00001000	00025800	00000600	[CER]
02	.data	0001C000	00027000	00005400	00025E00	[IRW]
03	.tls	00001000	00043000	00000200	0002B200	[IRW]
04	.idata	00001000	00044000	00000A00	0002B400	[IR]
05	.edata	00001000	00045000	00000200	0002BE00	[IR]
06	.relocs	00002000	00047000	00001A00	0002CC00	[ISR]

Key to section flags:
C - contains code
E - executable
I - contains initialized data
R - readable
S - shareable
W - writeable

Exports from Coder.dll
11 exported name(s), 11 export addresse(s). Ordinal base is 1.

Ordinal	RVA	Name
0000	00028ea8	_DebugHook
.....	.....	.....
0003	000035b8	_gostCoding
0004	00003646	_gostDecoding
0005	000013dc	_readCodeKeys
.....	.....	.....
0011	00001468	_shifrFile

Рис. 1

[1] Питрек Мэтт. СЕКРЕТЫ системного программирования в Windows 95; Киев: Диалектика 1996

Существует множество программ, позволяющих получить информацию о PE файле и его секциях Практически все языки, позволяющие писать под Windows, имеют утилиты для получения подобной информации. Рассмотрим нашу гипотетическую программу при помощи утилиты TDUMP из Турбо Ассемблера пятой версии.

Файл, полученный при помощи TDUMP, содержит много информации о программе, но нас прежде всего, интересует таблица секций и секции импорта-экспорта. В таблице секций подробно описано, где в файле и в памяти располагаются секции программы (рисунок). Рассмотрим таблицу секций на примере простейшей DLL:

Как видно из рисунка, таблица секций содержит полное описание всех секций файла. Нас наиболее интересует импорт, в главный исполняемый файл, функций из DLL. В EXE файле эти функции будут находиться в секции idata, в DLL они попадут в секцию edata.

Просмотрев секцию экспорта в DLL, можно найти символьное имя функции, определить язык, на котором написана эта функция, и получить точку входа в функцию. Например, функция gostCoding скорее всего занимается кодированием, написана на языке C (признаком применения C может служить символ подчеркивания '_', который автоматически вставляется компилятором) и применяет правила передачи параметров C.

Термин RVA, применяемый в таблицах, расшифровывается как относительный виртуальный адрес (Relative Virtual Address). Это смещение в памяти по отношению к адресу, с которого начинается отображение файла в памяти. Phys off это смещение в файле программы Для вычисления точки входа в функцию необходимо вычислить смещение точки входа по отношению к началу секции кода и скорректировать полученное значение с учетом смещения секции кода в файле. Для этого возьмем значение RVA функции из таблицы экспорта, вычтем из этого значения RVA секции .text из таблицы секций и прибавим значение из поля Phys off для секции .text. В файле CODER.DLL по адресу 0x2BB8 (0x35B8-0x1000+0x600) находится следующий код:

```
00002BB8:    55        Push      Ebp
00002BB9:    8BEC      Mov       Ebp,esp
00002BBB:    83C4F0    Add       Esp,FFFFFFF0
```

Это стандартный код пролога, создаваемый компилятором для всех функций. Данный фрагмент выделяет 16 байт в стеке для локальных переменных функции. Замена первого байта 0x55 на 0xC3, код команды возврата из подпрограммы, даст очень интересный результат:

```
00002BB8:    C3        ret
00002BB9:    8BEC      mov       Ebp,esp
00002BBB:    83C4F0    add       esp,FFFFFFF0
```

Теперь процедура не выполнится ни разу – процессор сразу же выйдет из функции. Таким образом, можно очень легко и быстро отключить любую функцию из любой DLL. Получив точку входа в функцию можно внести любые изменения, получить код функции, а так же перехватить передаваемые функции параметры.

Для перехвата параметров используется немного другой механизм, основанный на тех же принципах. Пишется фрагмент кода, сохраняющий параметры в буфере и выполняющий переход на точку входа в функцию. В этом случае корректируется таблица экспорта DLL, либо таблица импорта EXE файла – в зависимости от того к какому файлу приписывается этот фрагмент. При вызове функции, сначала выполняется этот фрагмент затем сама функция. При желании, можно перехватить и результат – модифицировав адрес возврата на точку входа своего фрагмента кода. Учитывая, что обычно PE файлы имеют файловое выравнивание по 512 байтам (размер сектора на диске), добавление небольшого фрагмента даже не изменит длины файла. При этом автоматически решается проблема раздельных адресных пространств Windows NT. При желании, можно добавить новые секции к PE файлу так как обычно после нее в файле имеется около 400 байт свободного пространства. При длине описания сек-

ции 40 байт получаем место для 10 дополнительных секций. В данный момент существуют вирусы, которые используют свободное пространство в конце секций для хранения своего кода, при этом длина инфицированного файла не изменяется. Одним из таких вирусов является CIH.

В нашем случае функция PasswordCheck прерывается таким образом чтобы любой пароль считался корректным, функция Coding просто прерывается. Вся процедура внесения изменений не займет много времени и может выполнятся вручную, без применения утилит типа TDUMP с помощью простого редактора.

3. Анализ возможной угрозы и варианты её предотвращения

Рассмотренный метод взлома основан на использовании динамических библиотек с импортированием по имени. Нетрудно догадаться, что использование импортирования по номеру существенно затруднит взлом. Именно так поступают программисты MICROSOFT, если необходимо скрыть какую либо функцию. Найти нужную функцию, имея только точки входа теоретически реально, но практически почти невозможно. Однако использование импортирования по номеру функции не рекомендовано, так как при совпадении номеров могут возникнуть проблемы совместимости, по - этому большинство приложений используют импортирование по имени функции И конечно этот метод взлома не работает в случае, когда функция не вынесена в DLL.

Существует множество методов противодействия, опишем вкратце основные из них.

Избежать всего выше приведенного можно очень простым и старым методом - просто не выносить особо критичные секции кода из основной программы. Однако при этом теряется гибкость, о которой говорилось в начале статьи.

Другая возможность состоит в использовании импортирования по номеру функции. Однако этот метод имеет следующий недостаток – если кто-либо напишет DLL с функцией, импортируемой по такому же номеру, как и у вашей функции, то могут возникнуть проблемы.

Третий вариант состоит в использовании оверлеев. Можно попробовать использовать и стандартные DLL, но загружать их в процессе работы программы, а не на этапе загрузки. При этом перед загрузкой проверяется целостность файла.

Последний, наиболее простой и результативный метод состоит в том, чтобы скрыть имена функций в DLL и проверить их сигнатуру перед вызовом. Этот метод устраняет возможность определения предназначения функции по её имени и дополнительно проконтролировать код функции на точке входа. Сокрытие имен может производиться либо на этапе разработки, либо по окончании отладки всего приложения.

В первом случае функции, выносимые из основной программы в DLL, называют абсолютно нейтральными именами – например буквенно-цифровым кодом. В главном файле на каждую внешнюю функцию пишется либо функция-пустышка, либо макрос с «интуитивно» понятным именем. В этом случае в секциях импорта/экспорта останется только цифровой код, как имя внешней функции. Если код будет достаточно длинен (5 – 10 символов) возможность совпадения с именем уже существующей функции будет исключена.

Во втором случае сокрытие имен производится в уже готовом приложении, после окончания отладки. Процесс сокрытия состоит в замене имен функций как в DLL так и в EXE файлах буквенно-цифровым кодом. Такой подход позволяет вести разработку и отладку приложения без дополнительных трудностей с именами, а после её окончания при помощи небольшой утилиты провести сокрытие имен.

Обычно приложения под Windows имеют в составе кроме DLL, написанных программистом, несколько стандартных DLL, поставляемых с языком программирования. Файлы DLL также имеют имена, по которым можно определить их принадлежность. Сокрытие может выполняться и с именами DLL. Для этого файлы переименовывают и поправляют их имена в секции импорта EXE файла.

Для контроля сигнатуры функции необходимо сохранить в основном файле приложения небольшой фрагмент начала кода функции, а перед её вызовом сравнить её с кодом вызываемой функции.

Адрес функции может быть легко получен в любом языке программирования, однако необходимо уточнить для конкретного компилятора, как создаются заглушки импортируемых функций. В качестве адреса функции может передаваться адрес двойного слова Function_Stub_Address, либо адрес заглушки Function_Stub, в этом случае реальный адрес функции будет храниться по этому адресу. Данный метод позволяет проконтролировать точку входа, обнаружить вышеописанную модификацию кода и попытку перехвата параметров функции.

Описанный метод позволяет эффективно предотвратить взлом приложения, использующего динамическое связывание.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.142.35

М.П. КАРПІНСЬКИЙ, д-р техн. наук, Я.І. КІНАХ

ВИКОРИСТАННЯ МЕТОДІВ ФАКТОРИЗАЦІЇ ДЛЯ ОЦІНКИ НАДІЙНОСТІ СИСТЕМИ ШИФРУВАННЯ RSA

Бурхливий розвиток комп'ютеризації всіх сфер життя надає нові можливості для національних економік. Поширення інформаційних технологій має і свій негативний аспект: це відкриває шлях до антисоціальної і злочинної поведінок. Крім того, що комп'ютерні злочини наносять значні економічні збитки, суспільство стає все залежнішим від роботи комп'ютеризованих систем у різноманітних сферах життя — від керування рухом літаків і поїздів до медичного обслуговування та національної безпеки. Будь-який збій у функціонуванні таких систем може привести до реальної загрози життю людей. Стрімке зростання глобальних комп'ютерних мереж, а також можливість під'єднання до них через звичайні телефонні лінії посилюють можливості їх використання для несанкціонованого доступу.

У порівнянні з високорозвиненими країнами інформаційна безпека України поки що залежить від комп'ютерних мереж значно менше. На сьогодні в нашій державі основна маса несанкціонованих доступів спостерігається у фінансово-кредитній сфері. Але у недалекому майбутньому такі несанкціоновані доступи можуть викликати глобальні катастрофи. Введення сучасної системи управління повітряним рухом, поширення телекомунікаційної мережі, впровадження системи електронних платежів, використання комп'ютерів у діяльності правоохоронних органів та керуванні військами значно збільшили інтерес до несанкціонованого доступу серед користувачів і програмістів [1].

Система шифрування інформації RSA на сьогодні стала де-факто світовим стандартом, що реалізується в якості самостійних програмних продуктів і у складі продуктів прикладного програмного забезпечення. Систему RSA використовують у світових банківських мережах, зокрема для роботи з кредитними картками. Вона зустрічається в таких стандартах: SSL, S-HHTP, S-MIME, S/WAN, STT і PCT.

Система RSA належить до асиметричних криптографічних систем. Їх суть полягає в тому, що кожний користувач генерує два ключі, які пов'язані деяким співвідношенням. Один ключ функціонує відкрито, інший є таємним. Текст шифрується відкритим ключем адресата. Процес дешифрування можна здійснити тоді і тільки тоді, якщо відомий таємний ключ. Дану систему можна використовувати як самостійний засіб захисту, так і при розподілі ключів, а також як засіб аутентифікації.

Перед шифруванням текст кодують у зручну для роботи систему числення. Закодований текст розбивають на блоки $B_i \in Z_n$ і перетворюють блоки згідно з правилом:

$$E(B_i) = B_i^e \pmod n, \tag{1}$$

де e – відкритий ключ, такий що $e < \phi(n)$, $(e, \phi(n)) = 1$ - найбільший спільний дільник, $\phi(n)$ – функція Ейлера, n – модуль перетворення, що є добутком двох, бажано «сильних», простих чисел p і q достатньо великої розрядності (p і q не розголошуються).

В результаті отримаємо криптотекст, що також формується з блоків $P_i = E(B_i)$. Очевидно, що $P_i \in Z_n$. Процес дешифрування відбувається за правилом:

$$D(P_i) = P_i^d \pmod n. \tag{2}$$

Тут d – таємний ключ, що пов'язаний з відкритим ключем таким співвідношенням:

$$ed \equiv 1 (\mathrm{mod}\, \phi(n)). \tag{3}$$

Під час практичної роботи з криптотекстами виникає задача дешифрування, коли таємний ключ d є невідомим. Оскільки таємний і відкритий ключі пов'язані відомим співвідношенням, то обчисливши

значення функції $\phi(n)$, можна взнати таємний ключ за відкритим. Відомо, що $\phi(n) = (p - 1)(q - 1)$ і поставлена задача зводиться до обчислення p і q, де $pq = n$. Тоді постає задача факторизації.

Згідно з твердженням 4.4 [2], обчисливши два квадратні корені y та y' з деякого числа x за модулем n, можна твердити: найбільший спільний дільник $(y + y', n)$ є одним з дільників p або q числа n, за умови, що

$$y \neq \pm y' (\mathrm{mod}\, n). \tag{4}$$

Таким чином, поставлена задача зводиться до розв'язання конгруенції $y^2 = (y')^2 (\mathrm{mod}\, n)$. Серед ефективних методів розв'язку останньої конгруенції заслуговують на увагу метод квадратичного решета і метод решета числового поля. Останній метод вважається найперспективнішим на сьогоднішній день. Розглянемо його детальніше.

Подамо число n у формі

$$n = r^e - s, \tag{5}$$

де $r > 0$, $s \neq 0$. При цьому r і $|s|$ є достатньо малими.

Виберемо мінімальні $d \in Z_{>0}$ і $k \in Z_{>0}$, такі, що $kd \geq e$. Звідси випливає, що

$$r^{kd} \equiv s r^{kd-e} (\mathrm{mod}\, n). \tag{6}$$

Нехай $m = r^k$, $c = s r^{kd-e}$. Тоді

$$m^d \equiv c (\mathrm{mod}\, n). \tag{7}$$

Сформуємо многочлен

$$f(x) = x^d - c \in Z[x], \tag{8}$$

де α – корінь многочлена.

Побудуємо гомоморфізм φ такий, що відображає $Z[\alpha]$ в Z / nZ. Метод решета поля дозволяє відшукати пару цілих алгебраїчних чисел a і b, які зустрічаються в співвідношенні

$$\varphi(a + \alpha b) = (a + mb \bmod n). \tag{9}$$

Отримані числа a і b використовують для знаходження розв'язку конгруенції

$$y^2 = (y')^2 (\mathrm{mod}\, n). \tag{10}$$

Під час пошуку можна обмежитися головними ідеалами Z[α] простої норми, бо вони єдині містять алгебраїчні цілі числа форми $a + \alpha b$, де a і b - взаємопрості числа.

Множину головних ідеалів Z[α] простої норми визначають пари чисел p і c_p, де p - просте число і $c_p \in \{0, 1, \ldots, p - 1\}$. Число c_p повинно задовольняти умову $f(c_p) \equiv 0 \bmod p$. Під час пошуку пар можна використовувати головні ідеали норми p, породжені p і $\alpha - c_p$, або еквівалентні головні ідеали Z[α], що перетворюються гомоморфізмом в Z / pZ, і α відображається в c_p. Зокрема, число $a + \alpha b$ знаходиться в головному ідеалі, що відповідає парі чисел p і c_p, якщо тільки $a + c_p b \equiv 0 \bmod p$. Головний ідеал $a + \alpha b$ характеризується нормою $N(a + \alpha b) = a^d - c(-b)^d \in Z$.

Зафіксуємо межу розкладу $B \in R_{>0}$. Значення B визначаємо експериментально. Нехай a і b цілі числа і $b > 1$. Припустимо, що:

$$|N(a + \alpha b)| = \prod_{p \leq B} p^{v_p} \tag{11}$$

$$|a+mb| = \prod_{p \le B} p^{w_p}, \tag{12}$$

де v_p, $w_p \in Z$.

Для продовження процесу факторизації представимо числа $a + \alpha b$ у вигляді:

$$a + \alpha b = (\prod_{u \in U} u^{t_u})(\prod_{g \in G} g^{v_g}), \tag{13}$$

де v_g, $t_u \in Z$. U - множина деяких груп. G - множина головних ідеалів Z[α] простих норм. Звідси випливає:

$$((\prod_{u \in U} \varphi(u)^{t_u})(\prod_{g \in G} \varphi(g)^{v_g}) = \prod p^{w_s} \bmod n. \tag{14}$$

При наявності достатньої кількості пар чисел *a, b* формуємо вирази (15) і (16), обчислюємо значення функції *x(a, b)* за допомогою методу Гауса, додавши вектори за модулем 2 в рівності (14). Одержуємо:

$$\prod_{a,b} (a + \alpha b)^{x(a,b)} = \left((\prod_{d \in U} u^{\overline{v_u}}) \cdot (\prod_{g \in G} g^{\overline{v_g}}) \right)^2, \tag{15}$$

$$\prod_{a,b} (a + mb)^{x(a,b)} = \left(\prod_{p \le B} p^{\overline{w_p}} \right)^2. \tag{16}$$

Прирівнявши праві частини виразів (15) і (16), отримаємо:

$$\left((\prod_{u \in U} \varphi(u)^{\overline{t_u}})(\prod_{g \in G} \varphi(g)^{\overline{v_g}}) \right)^2 = \left((\prod_{p \le B} p^{\overline{w_p}})^2 \bmod n \right). \tag{17}$$

Тут $\overline{t_u}, \overline{v_g}, \overline{w_p} \in Z$. З рівності (17) знаходимо цілі числа y і y', що задовільняють конгруенцію $y^2 = (y')^2 (\bmod n)$.

Одержані результати розв'язку дозволяють достеменно визначити таємний ключ системи RSA.

Час роботи алгоритму що реалізує метод решета числового поля оцінюється виразом $\exp\left((c + o(1))(\ln n)^{1/3}(\ln \ln n)^{2/3}\right)$, c=1,526. Під час факторизації методом решета числового поля числа довжиною в сто десяткових цифр необхідно проаналізувати принаймні 900 ідеалів простої норми. Матриця (14) має розмір порядку 10^5 на 10^5. Для її обробки необхідно виконати кількість арифметичних операцій порядку 10^{13}. Для порівняння наведемо вираз, що дозволяє оцінити час роботи алгоритму квадратичного решета: $\exp\left((1 + o(1))(\ln n)^{1/2}(\ln \ln n)^{1/2}\right)$ [3]. Рисунок 1 дозволяє наочно спостерігати перевагу методу решета числового поля над методом квадратичного решета, бо графічна ілюстрація виразу, що дозволяє оцінити складність методу квадратичного решета - 2, мажорує відповідний вираз для решета числового поля - 1 для будь-якого допустимого n.

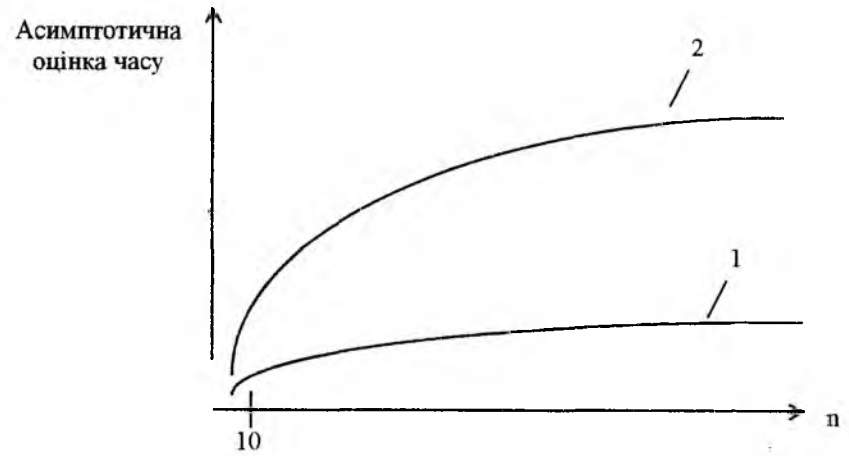


Рис. 1

Список літератури: **1.** *Криміналістика* / П.Д.Біленчук, О.П.Дубовий, М.В.Салтевський П.Ю.Тимошенко / За редакцією П.Д. Біленчука. – К.: АТІКА, 1998. – 416 с. **2.** *Вербіцький О.В.* Вступ д криптології. – Львів: ВНТЛ, 1998. – 247 с. **3.** *Lenstra A.K., Lenstra H.W., Manasse M.S., Pollard J.M.* Th number field sieve. Online access through WWW: http://www.rsasecurity.com/rsalabs/faq/.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 10.02.2000

УДК 681.3

И.Д. ГОРБЕНКО, д-р техн. наук, Е.Г. КАЧКО, канд. техн. наук, С.А. КОВАЛЕВ, асп.

СЕРВИС АУТЕНТИФИКАЦИИ KERBEROS

Важное место при построении систем защиты информации занимает сервис аутентификации Kerberos. В данной работе выполнен обзор возможных протоколов и предлагается программная модель такого сервиса.

Рассмотрим назначение и функции Kerberos [1-7].

Имеется открытая, незащищенная сеть. В ее узлах находятся субъекты: сервера и клиенты. (По определению: сервер предоставляет услуги клиентам; клиент – программное обеспечение, пользователи и др.)

Необходимо подтвердить подлинность (аутентичность) клиента С серверу S и наоборот.

При этом используется следующая предпосылка: у каждого субъекта есть собственный секретный ключ. Для пользователей (под английским термином – User – понимается физическое лицо, в отличие от термина Client, под которым понимается субъект, потребляющий услуги) специфицируется алгоритм преобразования пароля в секретный ключ. Подлинность субъекта отождествляется со знанием его секретного ключа.

Субъект А может подтвердить свою подлинность субъекту В, просто переслав ему свой секретный ключ КА. Но сеть незащищена – ключ могут перехватить, исказить. Требуется менее прямолинейный способ демонстрации знания своего секретного ключа.

Эту задачу выполняет Kerberos – третья сторона, которой доверяют все субъекты, и, соответственно, которая хранит секретные личные ключи всех субъектов. Kerberos не полагается на:

- средства аутентификации, реализованные в операционных системах сетевых компьютеров;
- на подлинность сетевых адресов;
- на физическую защищенность всех сетевых компьютеров.

Kerberos производное от Cerberus. Цербер – трехголовый пес, охраняющий вход в царство мертвых. Сервер аутентификации Kerberos основан на протоколе аутентификации третьей стороны Нидхама (Needham) и Шредера (Schroeder) (1978) с изменениями Дэнинга (Denning) и Сако (Sacco) (1981). Проектирование и реализация с 1-й по 4-ю версии Kerberos'а осуществлялась в рамках проекта Athena Стивом Миллером (Steve Miller) из Digital Equipment Corporation и Клиффордом Ньюменом (Clifford Neuman) (сейчас в Information Sciences Institute of University of Southern California). А также принимали участие и многие другие, в том числе: Джером Салтцер (Jerome Saltzer) - технический директор проекта Athena и Джефри Шилер (Jeffrey Sciller) - менеджер сети университета MIT (Massachusetts Institute of Technology). Над пятой версией работает команда разработчиков MIT во главе с Теодором Тсо (Theodore Ts'o).

Кроме доказательства аутентичности субъектов Цербер позволяет распространить сеансовый и субсеансовый ключи.

Основная идея Kerberos такова: очень сложно из зашифрованного текста извлечь ключ шифрования, также сложно расшифровать закрытое сообщение (без знания секретного ключа). Отсюда идея: послать нужную субъекту информацию, зашифровав ее на его секретном ключе.

1. $K \rightarrow C$: $\{M\}Kc$; $\{T\}Ks \subset M$
2. $C \rightarrow S$: $\{T\}Ks$,

где К - Kerberos;
С - клиент;
S - сервер;
Кх - ключ шифрования принадлежащий субъекту X;
{Y}Кх - информация Y зашифрована секретным ключом субъекта X.

Достоинства:

- Конфиденциально;
- Воспользоваться информацией может только целевой субъект;
- В открытом виде нигде не появляется.

Рассмотрим несколько вариантов протокола обмена при аутентификации [1,5,6]. При этом пока не рассматриваются следующие вопросы:

- первоначальный обмен ключами между Kerberos'ом и субъектами;
- как субъекты хранят свои ключи;
- как осуществляется администрирование Kerberos.

Первый вариант.

1. $C \rightarrow K$: c, s, ...
2. $K \rightarrow C$: $\{d1\}_{Kc}$; $\{T_{c\text{-}s}\}_{Ks}$
3. $C \rightarrow S$: d2, $\{T_{c\text{-}s}\}_{Ks}$

где $T_{c\text{-}s}$ - *билет клиента С к серверу S;*

d1 - некоторая информация, которую Kerberos передает клиенту и серверу (в билете к серверу);

d2 - часть информации из d1, ее сервер сравнивает с такой же, находящейся в билете.

Чтобы с помощью Kerberos получить доступ к серверу S, клиент С посылает запрос, содержащий сведения о себе и о требуемой услуге.

В ответ Kerberos возвращает две порции информации: так называемый билет, зашифрованный секретным ключом сервера, и копию части информации из билета, зашифрованную секретным ключом клиента.

Клиент должен расшифровать вторую порцию данных и переслать ее вместе с билетом серверу. Сервер, расшифровав билет, может сравнить его содержимое с дополнительной информацией, переданной клиентом.

Совпадение свидетельствует о том, что клиент смог расшифровать предназначенные ему данные - ведь содержимое билета никому, кроме сервера и Kerberos, недоступно – и продемонстрировал знание своего секретного ключа. Исходя из предпосылки считаем, что субъект тот, за кого себя выдает.

Это очень упрощенный вариант протокола. Он не обеспечивает полностью защищенный процесс аутентификации. Например, злоумышленник, перехватив диалог, может послать серверу ту же информацию и выдать себя за второго С. В более практичных вариантах процедур проверки аутентичности добавляют:

- разделяемый сеансовый ключ для обеспечения конфиденциальности взаимодействия;
- срок годности этого ключа, билета, по истечении которого необходимо заново пройти процедуру аутентификации (стандартный срок – 8 часов – может быть изменен в зависимости от политики безопасности данной организации);
- для подтверждения подлинности сервера S клиенту С – клиент шифрует передаваемую серверу дополнительную информацию с помощью сеансового ключа, который Kerberos поместил в билет, зашифрованный на секретном ключе сервера. Сервер, расшифровав дополнительную информацию клиента, возвращает часть ее клиенту (тоже соответственно зашифрованную на сеансовом ключе);
- для защиты от воспроизведения (выдачи злоумышленником себя в качестве второго клиента С):
- временной штамп – удостовериться в "свежести" сообщения;
- случайное число – неповторяемость сообщения и подтверждение целостности цепочки сообщений (помещая это число или закономерно модифицированное (например, +1)).

Второй вариант.

1. $C \rightarrow K$: {c, s1, ...}
2. $K \rightarrow C$: {n, s,timeexp $K_{c\text{-}s}$, ...}$_{Kc}$, {T}$_{Ks}$
3. $C \rightarrow S$: {s1, ts, ck ...}$_{Kc\text{-}s}$, {T}$_{Ks}$
4. $S \rightarrow C$: {ts, c, timeexp,...}$_{Kc\text{-}s}$,

где $K_{c\text{-}s}$ - разделяемый клиентом С и сервером S сеансовый ключ;
s1 - необходимый клиенту сервис;
n - одноразовое число;
timeexp - срок годности билета;
ts - временной штамп;
ck - контрольная сумма.

В билет помещается следующая информация:

имя С - сервер определяет клиента, запрашивающего определенный сервис;

имя самого сервиса – подтверждение правильной расшифровки и, соответственно, правильной генерации сообщения к серверу со стороны клиента, а также знания разделяемого сеансового ключа (сеансовый ключ помещается в ответ Kerberos'a клиенту и шифруется на секретном ключе клиента);

сетевой адрес – для частичной (сетевой адрес тоже можно подделать) защиты от кражи использования билета, а также для идентификации клиента;

Порции информации, позволяющие убедиться в подлинности предъявившего их субъекта, назовем аутентификатором (А).

Если клиенту понадобятся услуги нескольких серверов, то придется, соответственно, несколько раз доказывать свою подлинность, взаимодействовать, использовать секретный ключ. Для минимизации времени активного использования секретного ключа разделяют Kerberos на сервер начальной аутентификации (AS – Authentication Server) и сервер выдачи билетов (TGS – Ticket Granting Server).

Третий вариант.

1. $C \rightarrow AS$: c, tgs1, ...
2. $AS \rightarrow C$: {$K_{c\text{-}tgs}$, tgs, ...}$_{Kc}$, {Tc-tgs}$_{Ktgs}$
3. $C \rightarrow TGS$: {A_c}$_{Kc\text{-}tgs}$, {$T_{c\text{-}tgs}$}$_{Ktgs}$, s1,...
4. $TGS \rightarrow C$: {$K_{c\text{-}s}$, s,...}$_{Kc\text{-}tgs}$, {$T_{c\text{-}s}$}$_{Ks}$
5. $C \rightarrow S$: {Ac}$_{Kc\text{-}s}$, {$T_{c\text{-}s}$}$_{Ks}$
6. $S \rightarrow C$: {As}$_{Kc\text{-}s}$,

где AS - Authentication Server – сервер начальной аутентификации;
TGS - Ticket Granting Server - сервер выдачи билетов;
$T_{c\text{-}tgs}$ - билет клиента С к серверу TGS;
Ac - аутентификатор клиента С;
As - аутентификатор сервера S.

AS выдает TGT (Ticket Granting Ticket – "билет на билеты"), на основании которого у TGS получает билеты к другим серверам.

В том числе, при использовании глобальных сетей, субъекты из одной области управления могут взаимодействовать с субъектами из другой области управления. (Естественно, при наличии соглашения между областями управления и после обмена ключами между Kerberos'ами этих областей. Возможны как иерархические, так и специальные отношения и соглашения между областями управления).

В протоколе это отражается в разделении TGS на локальный locTGS и удаленный remTGS. При наличии договорных отношений локальный TGS выдает билеты к удаленному TGS, а тот – билеты к удаленным серверам.

Четвертый вариант.

1. $C \rightarrow$ locAS: c, tgs,...
2. locAS $\rightarrow C$: {$K_{c\text{-}tgs}$, tgs,...}$_{Kc}$, {$T_{c\text{-}tgs}$}$_{Ktgs}$

3. $C \rightarrow locTGS$: $\{A_c\}_{Kc\text{-}tgs}$, $\{T_{c\text{-}tgs}\}_{Ktgs}$, tgs-rem,...
4. $locTGS \rightarrow C$: $\{K_{c\text{-}remTGS}$, remTGS,...$\}_{Kc\text{-}tgs}$, $\{T_{c\text{-}remTGS}\}_{KremTGS}$
5. $C \rightarrow remTGS$: $\{A_c\}_{Kc\text{-}remTGS}$, $\{T_{c\text{-}remTGS}\}_{KremTGS}$, remS,...
6. $RemTGS \rightarrow C$: $\{K_{c\text{-}remS}$, remS,...$\}K_{c\text{-}remTGS}$, $\{T_{c\text{-}remS}\}_{KremS}$
7. $C \rightarrow remS$: $\{A_c\}_{Kc\text{-}remS}$, $\{T_{c\text{-}remS}\}_{KremS}$
8. $RemS \rightarrow C$: $\{A_{remS}\}_{Kc\text{-}remS}$

Помимо описанных возможностей существует множество дополнительных. Например, вводится срок годности – устанавливается не только окончание, но и начало срока (так называемые "билеты на потом"). Вводится возможность продления билета и максимальный срок такого продления; передача прав одного субъекта на действия от имени другого (сервер печати к файлам пользователя). Каждому билету приписываются разнообразные флаги.

Для реализации выбран третий вариант протокола. Он соединяет в себе такие противоположные качества, необходимые в экспериментальной версии, как простота и функциональность. Он самый простой из функциональных и самый функциональный из простых. Этот протокол сравнительно легко расширить до наиболее сложного известного протокола и внести пока не изобретенные усовершенствования.

Проанализируем возможность применения средств симметричного шифрования (СШ). В оригинальных разработках системы Kerberos на всех этапах применяется симметричное шифрование. После размышлений над этим фактом авторы пришли к выводу о том, что в данной ситуации (эта версия, этот протокол, опытный характер разработки) использование несимметричных систем криптопреобразований является неконструктивным. Увеличивается время преобразований, усложняется генерация ключей (например, удваивается база ключей).

Существующие реализации Kerberos основаны на использовании симметричных алгоритмов, принятых в качестве стандартных в странах-разработчиках. Для использования системы в отечественных средствах, необходима разработка системы на основе стандартов шифрования, принятых на Украине. Так как время жизни стандарта кратковременно, одним из важных требований к системе является возможность простой замены используемых алгоритмов, длин ключей, способов их хранения.

Кратко опишем результаты выполнения этапов анализа и проектирования.

Было принято решение произвести пробную реализацию сервиса Kerberos и выполнить ее в объектно-ориентированной парадигме (как проектирования, так и программирования).

В результате анализа проблемной области были выделены классы и составлена иерархия (упрощенное изображение иерархии уровня приложений показано на рисунке 1), которая, в зависимости от уровня абстракции, раскладывается на три плоскости:

- классы абстракций уровня приложений;
- классы абстракций уровня сообщений
- классы абстракций уровня базовых структур данных.

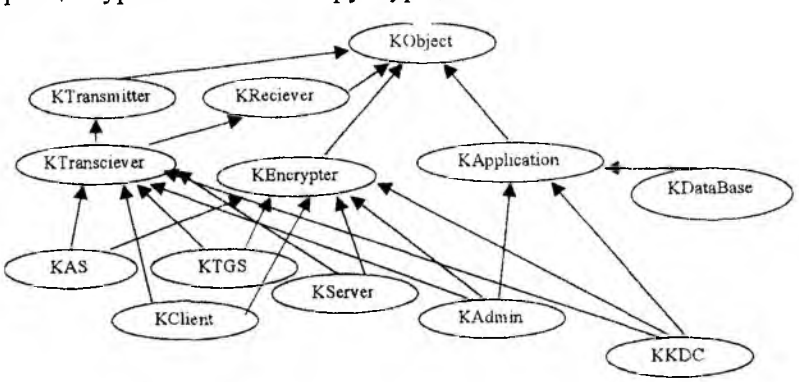


Рис. 1

На данный момент в иерархии 52 класса, что вынуждает более внимательно подойти к вопросам управления библиотекой классов [8,9]. При реализации мы придерживаемся следующих соглашений:

- класс ↔ файл – Установление однозначной связи. Повышает гибкость библиотеки и обеспечивает раздельную компиляцию.
- файл ↔ переменная препроцессора – Позволяет на этапе компиляции отслеживать какие файлы уже компилировались.
- условная компиляция – Использование препроцессорной обработки позволяет полностью исключить отладочный код из коммерческой версии.
- идентификация классов – каждый класс должен «уметь» сообщить о состоянии своего объекта, его выполненных действиях, размещении в памяти.
- расширяемость – необходимость явно спроектировать возможность расширения путем предоставления свободно модифицируемого каркаса [9].
- настраиваемость – необходимость предусматривать пути использования и приспособления данной библиотеки [9].

Вследствие постоянно увеличивающегося количества классов в иерархии автором в настоящее время рассматривается возможность использования CASE – средств (например, Rational Rose) [10,11].

Список литературы: **1.** *Вьюкова Н.* Сервер аутентификации Kerberos //Открытые системы. 1996. №1. с 44-50. **2.** *John Kohl and B. Clifford Neuman.* The Kerberos Network Authentication Service (Version 5). Internet Request for Comments RFC 1510. September 1993. **3.** *John Linn.* The Kerberos Version 5 GSS-API Mechanism. Internet Request for Comments RFC 1964. **4.** *B. Clifford Neuman and Theodore Ts'o. Kerberos:* An Authentication Service for Computer Networks, IEEE Communications, 32(9):33-38. September 1994. **5.** *Brian Tung.* The Moron's Guide to Kerberos. http://gost.isi.edu/brian/security/kerberos.html **6.** *Bill* Bryant. Designing an Authentication System: a Dialogue in Four Scenes. 1988. Afterword by Theodore Ts'o, 1997 http://web.mit.edu/kerberos/www/dialogue.html **7.** *S.M. Bellovin and M. Merritt.* Limitations of the Kerberos Authentication System. In *Proceedings of the Winter 1991 Usenix Conference*. January 1991. ftp://research.att.com/dist/internet_security/kerblimit.usenix.ps **8.** *Дьюхаст С., Старк К.* Программирование на C++. Пер с англ. Киев: «ДиаСофт», 1993. 272 с. **9.** *Страуструп Б.* Язык программирования C++, Часть вторая. Пер. с англ. Киев: «ДиаСофт», 1993. 296 с. **10.** *Вендоров А.М.* Современные методы и средства проектирования информационных систем. Москва: Финансы и статистика, 1998. 176 с. **11.** *Калянов Г.Н.* Консалтинг при автоматизации предприятий: Научно-практическое издание. Серия «Информатизация России на пороге XXI века». Москва: СИНТЕГ, 1997. 316 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.3.06

Е.Г.КАЧКО, канд. техн. наук, В.А. ЖЕЛЕЗНЯК

РАЗРАБОТКА КРИПТОПРОВАЙДЕРА. ИНТЕГРАЦИЯ КРИПТОПРОВАЙДЕРА В СИСТЕМУ

1. Введение

Человечество перешло к новой фазе своего развития – информационной, характеризующейся широким внедрением информационных технологий во все сферы его деятельности. Основной особенностью этой фазы является создание и коллективное использование массивов и баз данных, баз знаний, интенсивный обмен сообщениями, передача команд управления, широкое использование и распространение различного программного обеспечения. Одной из важнейших задач, возникающих при обмене информацией, является обеспечение ее целостности и конфиденциальности.

В связи со всеобщей стандартизацией все большее количество функций переходит от приложений к операционным системам (ОС). Например, до появления Windows, каждому приложению, работающему с принтером, было необходимо иметь в своем составе библиотеки для работы со всеми возможными видами принтеров. Аналогичная ситуация наблюдалась и для видеокарт, звуковых карт, а также коммуникационного оборудования – вообще любого оборудования. При стремительном росте ассортимента оборудования, начавшемся в 90х годах, такой способ перестал быть приемлемым. Появление сложных ОС позволило свести все драйвера оборудования под централизованное управление ОС и избежать ненужного дублирования. Однако процесс стандартизации не ограничился драйверами устройств – все большее количество функций, например, интерфейса с пользователем, ранее считавшихся прерогативой приложений, переходят к ОС.

В современных ОС, разработанных фирмой Microsoft был сделан еще один важный шаг к стандартизации – в ОС Windows теперь входят поддержка CryptoAPI – API, предоставляющего стандартизованный интерфейс для выполнения разнообразных криптографических операций, реализованная через криптопровайдеры (Cryptographic Service Provider, CSP) – особые DLL, интегрируемые в операционную систему. Интерфейс между приложениями и криптопровайдерами осуществляется через функции, реализованные в стандартном модуле AdvAPI32.DLL.

2. Постановка задачи

Многие, а в последнее время и большинство, программ требуют перед началом работы выполнить некоторую последовательность действий:

- копирование файлов на винчестер (возможно, с разархивированием);
- создание конфигурационного файла (файла настроек) или записей в реестре;

Обычно, все подготовительные операции выполняет специальная программа – ***инсталлятор***. В случае инсталляции криптопровайдера (Crypto Service Provider) рекомендованная Microsoft процедура состоит из следующих пунктов:

- копирование файлов провайдера в системную папку;
- создание записей в реестре;
- запись в реестр цифровой подписи (ЦП) провайдера.

Последний пункт является самым сложным, т.к. для получения цифровой подписи разрабатываемого провайдера необходимо обратиться в Государственный Департамент США, что невозможно по нескольким причинам.

Целью работы является исследование криптопровайдеров ОС WINDOWS, разработка пользовательского криптопровайдера и его инсталляция.

В данной работе решается проблема генерации ключа цифровой подписи для инсталлятора, встраивание открытого ключа в систему, генерация цифровой подписи длиной 1024 бита для метода RSA, инсталляция криптопровайдера таким образом, чтобы программы проверки целостности и подлинности «признали» эту программу, как свою.

3. Программная модель CryptoAPI

Microsoft Cryptographic Application Program Interface (CryptoAPI) – это набор функций для цифровой подписи и шифрования.

Все криптооперации реализованы с помощью независимых модулей, называемых криптопровайдерами (Cryptographic Service Provider, CSPs). По тем же причинам, по которым правильно разработанные приложения не вызывают напрямую драйвера графических устройств и аппаратуру, они не работают напрямую с криптопровайдерами и криптоаппаратурой.

Криптосистема Microsoft состоит из большого числа модулей. Три исполнимых части – непосредственно приложение, операционная система, и криптопровайдер.

Приложение связывается с ОС через функции, называемые программный интерфейс криптоприложения (Cryptographic Application Program Interface, CryptoAPI). ОС связывается с провайдерами через множество функций (Cryptographic Service Provider Interface, CryptoSPI). Как показали исследования, эти два набора функций практически идентичны: функция CryptoAPI CryptAcquireContext вызывает функцию CryptoSPI CPAcquireContext и т. д. Существующая разница, например, в списке параметров, обусловлена необходимостью изолировать область данных провайдера от приложений.

Еще раз отметим, что приложение не связывается с криптопровайдерами напрямую. Вместо этого, все обращения к криптографическим функциям выполняются через операционную систему. Параметр для каждой CryptoAPI функции указывает для операционной системы который CSP использовать, чтобы выполнить фактическую криптографическую операцию. Подробнее этот вопрос обсуждается в разделе 3.

Примечание: При попытке одновременного вызова двух криптофункций из разных потоков одного приложения вызов одной из них будет задержан до выхода из другой.

Криптопровайдеры - это независимые модули, которые и выполняют реальную криптоработу. В идеале они написаны полностью независимо от любого приложения, так что любое приложение выполняется с множеством провайдеров. На самом деле, некоторые приложения могут иметь очень специфические требования, которые потребуют заказного CSP.

CSP состоит, как минимум, из DLL и подписи файла. Подпись файла необходима, чтобы ОС распознала CSP. Операционная система проверяет правильность этой сигнатуры периодически, чтобы гарантировать, что CSP не модифицирован. Подробнее этот вопрос рассматривается в 1.4.

В Windows NT операционная система[1)] может просмотреть область данных провайдера. Можно, конечно, не хранить ключи в открытом виде и раскрывать их непосредственно перед использованием, но поскольку в любой момент может произойти передача управления другому процессу, то такой метод сам по себе не применим. Возможно, решением будет разработка своего драйвера – работа драйверов не может быть прервана;

В Windows 95/98, не только система, но и любой процесс, применив специальные методы, может просмотреть любую область памяти. В данной работе проблема защиты от несанкционированного доступа не решалась.

При реализации криптопровайдера следует иметь в виду, что изоляция ключей и централизация криптографических операций в аппаратных средствах безопасней программной реализации.

Проверка целостности и подлинности провайдера. Проверка осуществляется по следующему алгоритму [1]:

1) Вычисляется сжатый образ проверяемого провайдера по алгоритму MD5;

2) из системной библиотеки AdvAPI32.Dll берется значение основного открытого ключа и модуля криптопреобразований. Т. к. эти данные хранятся в зашифрованном виде, то предварительно производится декодирование (сложение по модулю 2 с константой);

3) сигнатура из реестра дешифруется на открытом ключе;

4) сверяются значения, полученные в 1) и 3). В случае равенства целостность и подлинность файла провайдера считаются доказанными;

5) в случае, если проверка не удалась, то берется дополнительный ключ и повторяются 3) и 4).

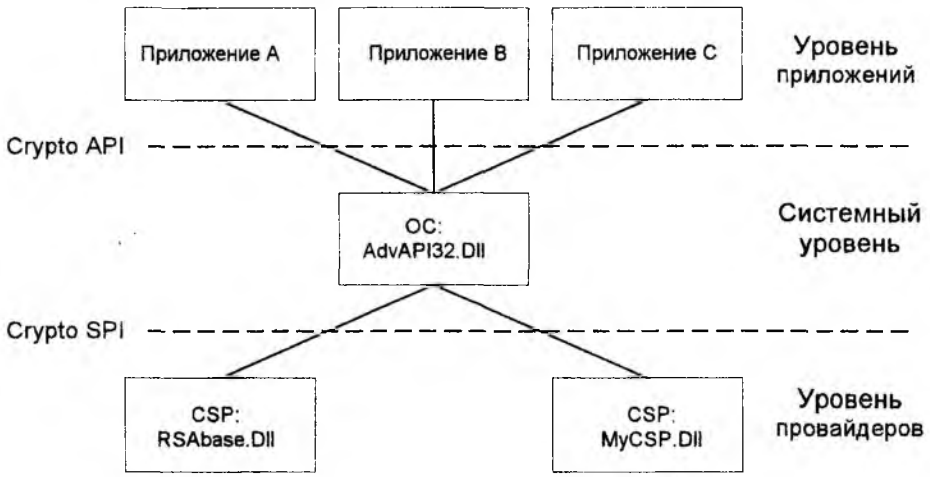
Необходимо добавить, что заменив один байт[2)] в теле AdvAPI32, целостность которой нигде не проверяется, можно добиться полного отключения проверки.

4. Интерфейс между приложениями, ОС и криптопровайдерами

Всякая работа приложений с функциями криптопровайдеров осуществляется не непосредственно, а через вызовы системных функций (рисунок)

[1)] В т.ч. и драйвера, а до установки ServicePack 3, и любой процесс

[2)] Код условного перехода на код безусловного



Для начала работы приложение вызывает функцию CryptAcquireContext (AdvAPI32.Dll) дл получения уникального контекста провайдера (аналогично вызову GetDC в функциях графическо подсистемы). Переданные параметры содержат имя требуемого криптопровайдера, идентификато алгоритма, а также параметры этого алгоритма.

При вызове CryptAcquireContext Windows предпринимает следующие действия:

- По переданному имени провайдера и записям в реестре (HKLM\ SOFTWARE\ Microsoft Cryptography\ Defaults\ Provider Types\ ..., HKLM\ SOFTWARE\ Microsoft\ Cryptography\ Defaults Provider\...) определяет путь и имя DLL провайдера;
- Проверяет целостность этой DLL – по коду DLL, сигнатуре в реестре (HKLM\ SOFTWARE Microsoft\ Cryptography\ Defaults\ Provider\ Signature) по алгоритмам MD5 и RSA. Заменой одного байт в AdvAPI32.DLL, целостность которой нигде не проверяется, можно добиться отключения проверки целостности;
- Загружает DLL провайдера и создает массив указателей на экспортируемые функции. Есл хотя бы одна из этих функций реализована не будет – выход с ошибкой;
- Вызывает CPAcquireContext из DLL провайдера

5. Интеграция нового провайдера в ОС

При установке нового провайдера должны выполняться следующие действия:

- Копирование файла (DLL) провайдера в системный каталог Windows;
- запись в реестр имени провайдера и полного имени соответствующей DLL;
- запись в реестр цифровой подписи (ЦП) провайдера.

На последнем пункте следует остановиться более подробно: для получения подпис провайдера необходимо обратиться в Microsoft, что неприемлемо по причине нарушени суверенных прав Украины.

В ходе исследований было найдено два способа решения этой проблемы – можно ил отключить ее вышеописанным методом или сгенерировать сигнатуру для нового провайдера Способ отключения проверки, однако, неприемлем, т.к. приводит к отключению проверк целостности одновременно всех провайдеров, в т.ч. и стандартных, следовательно, дл разработанного инсталлятора был выбран второй метод.

Алгоритм генерации ЦП:

1) по алгоритму RSA генерируются ключи для ЦП;

2) осуществляется подпись провайдера;

3) открытый ключ и модуль шифруются сложением по модулю 2 с константой;

4) клиенту передаются: зашифрованные открытый ключ и модуль ЦП, а также ЦП и D провайдера;

5) в системе клиента записывается в реестр ЦП и заменяется резервный ключ в AdvAPI32.

Заключение

Несмотря на то, что на данный момент функции CryptoAPI используются сравнительно редко, в будущем они, судя по всему, станут более распространены. Например, в Windows 2000 на базе CryptoAPI будет реализована новая файловая система - Encrypted File System (EFS), обеспечивающая, в отличие от FAT и NTFS, реальное шифрование данных на винчестере. Т.к. за пределы США и Канады будет экспортироваться только версия, поддерживающая ключи ограниченной длины, то разработка провайдеров приобретает вдвойне большее значение.

В рамках дальнейшей работы необходимо более тщательно исследовать механизмы проверки целостности и подлинности в ОС Windows 98 и Windows NT, а также разработать и реализовать механизмы более надежной защиты целостности и подлинности провайдеров и ключей.

Список литературы. 1. *Железняк В.А.* Разработка криптопровайдера. Интеграция криптопровайдера в систему. Материалы 3-го международного молодежного форума «Радиоэлектроника и молодежь в XX1 веке», 20-23 апреля 1999 г. Харьков, 1999, с. 489 – 490

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.327.8

П.В. КОЛЕСНИКОВ

ОБЗОР ПРОТОКОЛОВ ЗАЩИТЫ ИНФОРМАЦИИ В ОТКРЫТЫХ СЕТЯХ

В настоящее время банковская и другая информация, требующая защиты, в лучшем случае передается по корпоративным сетям или защищенным каналам связи. Построение таких сетей требует дополнительных затрат на их создание и использование, поэтому является не эффективным. Альтернатива корпоративным сетям- это объединение локальных сетей или отдельных машин, используя уже существующие сети, созданные на базе арендуемых и коммутируемых каналов связи сетей общего пользования (таких как Интернет).

Для организации такой сети необходима, там где требуется, защита сети от воздействия вирусов, злоумышленников, результатов ошибок в администрировании сети, а также от других угроз. Сейчас в мире компании-производители программного обеспечения и оборудования предпринимают усилия по разработке открытых (свободных для распространения и реализации) протоколов и стандартов в области защиты информации. Целью данной работы является исследование существующих протоколов с точки зрения обеспечения безопасности.

Эти протоколы предусматривают организацию защиты данных на различных уровнях Модели Взаимосвязи Открытых Систем (ВОС):

Таблица 1

Протоколы защиты	**Уровень ВОС**
SHTTP, S/MIME, PGP	Прикладной
SOCKS, SSL/TLS	Сеансовый
IPSec, SKIP	Сетевой
PPTP, L2TP	Канальный

Можно выделить следующие закономерности в реализации протоколов.

Чем ниже уровень ВОС, на котором организуется защита, тем она прозрачнее для приложений и незаметнее для пользователей; однако, тем меньше набор реализуемых услуг безопасности, и тем сложнее организация управления

Чем выше уровень ВОС, на котором реализуется защита, тем шире набор услуг безопасности, надежнее контроль доступа и проще конфигурирование правил доступа; однако, тем сложнее становится защита для приложений и пользователей. Применение одновременно протоколов защиты на различных уровнях модели усиливает действие каждого из протоколов

Протоколы сетевого уровня, как правило, являются фильтрами сообщений.

Протокол SOCKS описывает взаимодействие клиентов через proxy-серверы по протоколу TCP/IP Общая схема взаимодействия по протоколу SOCKS v4 сводится к следующему:

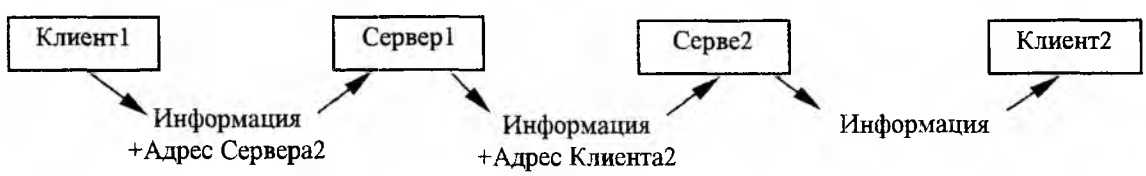


Рис. 1

Пользователь и удаленный сервер взаимодействуют друг с другом по цепочке соединений. Пользователь, желающий установить соединение с каким-либо сервером в сети, соединяется вместо этого с SOCKS-сервером и сообщает ему адрес удаленного сервера, тем самым криптоаналитик, перехвативший сообщение не может знать конечного адресата сообщения, т.к. оно передается в зашифрованном виде. Далее сам SOCKS-сервер соединяется с удаленным сервером-адресатом и передает ему сообщение. В протоколе предусмотрена аутентификация серверов и клиентов.

Одним из широко используемых протоколов в Интернет является протокол TLS. Это дальнейшая реализация протокола SSL. Главная цель TLS протокола- обеспечить сохранность и аутентичность данных при обмене сообщениями между двумя приложениями. Свойства TLS протокола:

- Создание соединения между двумя сторонами, закрытого алгоритмами шифрования.
- Способность расширения.

Протокол обеспечивает основной подход, с возможностью настроить нужные алгоритмы шифрования. Это дает возможность достичь еще две подцели. Во-первых, не нужно создавать новый протокол, который, возможно, будет обладать более слабой защитой. Во-вторых, созданная библиотека функций протокола не будет столь громоздкой. Одно из достоинств TLS протокола- это его независимость от протокола приложения. Для протоколов уровня высшего чем TLS он является совершенно прозрачным. Кроме того, TLS стандарт не определяет, как именно протоколы ведут защиту по TLS; решение как реализовать TLS протокол в конкретной ситуации и интерпретировать сертификаты аутентификации зависит от разработчика протоколов, реализованных над TLS.

<table>
<tr><td>TLS Протокол приветствия</td><td>TLS Протокол смены шифра</td><td>TLS Протокол сообщений</td><td>HTTP</td><td>FTP</td><td>SMTP</td><td>...</td></tr>
<tr><td colspan="3" rowspan="3">TLS Протокол обмена записями</td><td colspan="4">Сжатие</td></tr>
<tr><td colspan="4">Шифрование</td></tr>
<tr><td colspan="4">Аутентификация</td></tr>
<tr><td colspan="7">TCP</td></tr>
<tr><td colspan="7">IP</td></tr>
</table>

Рис. 2

Протокол разбит на два уровня: TLS протокол обмена записями и TLS—клиенты. TLS протокол обмена записями на вход получает запись, разбивает данные на блоки, сжимает их, накладывает подпись, шифрует и отправляет получателю. Информация от протокола обмена записями передается выше по стеку протоколов, либо к приложениям клиента, либо к другим составляющим протокола TLS. В последнем случае это служебная информация, необходимая для организации защищенного обмена. Существует 4 вида служебных протоколов (TLS-клиентов): протокол приветствия, сигнальный протокол, протокол смены шифра, протокол обмена данными приложения.

Протокол обмена записями находится на нижнем уровне протокола TLS и на вершине транспортного протокола TCP. TLS протокол обмена записями обеспечивает защиту соединения, которая основана на свойствах.

Соединение защищено симметричными алгоритмами шифрования (такими как DES, RC4). Ключи для этих алгоритмов генерируются уникальными для каждого соединения и основаны на общем секрете, выработанном на другом протоколе (таком как TLS приветствие). TLS протокол обмена записями может также использоваться без шифрования.

Поддерживается аутентичность соединения. Передача сообщений содержит проверку аутентичности, для которой используются алгоритмы хеш—функций такие как SHA, MD5. Протокол обмена записями может выполняться без алгоритма аутентификации, но обычно он используется.

Состояние соединения содержит информацию о том, как осуществляется обмен информацией в текущем соединении. Каждое состояние соединение определяет свои алгоритмы сжатия и шифрования. Всегда при обмене существует 4 состояния соединения: для записи, для чтения (от клиента к серверу и наоборот) и два дежурных состояния для записи и чтения.

Каждое состояние соединения включает следующую информацию:

- Состояние алгоритма шифрования: ключ шифрования; для блочных алгоритмов в режиме поточного шифрования- синхромаркер, для поточного шифра- состояние ключа.
- Закрытый ключ аутентификации
- Счетчик соединения: число размером 64 бита, которое увеличивается после каждой посланной записи и сбрасывается при установке нового текущего состояния соединения.

Протокол приветствия определяет параметры для дежурных состояний, и момент, с которого эти состояния станут активными. В первоначальном текущем состоянии, после активизации соединения, использование алгоритмов шифрования и сжатия заблокировано, пока нет договоренности о параметрах. **Протокол смены ключевых параметров** получает всего один байт на свой вход, который указывает на необходимость сменить текущие состояния для протокола обмена записями. Это сообщение требует подтверждения со стороны получателя для синхронной смены состояний.

Во время работы протокола приветствия производятся следующие действия и устанавливаются параметры защиты для соединений чтения и записи в протоколе TLS:

- Проводится аутентификация клиента и сервера с использованием X509v3 сертификатов.
- Выбирается основной алгоритм шифрования (null, «Поточное шифрование», «RC4 с ключом 40 бит», «RC4 с ключом 128 бит», «Блочные шифры в поточном режиме», «RC2 с ключом 40 бит», «DES с ключом 40 бит», «DES с ключом 54 бит», «тройной-DES с ключом 168 бит», «Idea (128 бит)», «Fortezza (96 бит)»).

Этот параметр включает размер ключа в алгоритме, сколько ключей является закрытыми, размер блока шифра, тип алгоритма (stream, block).

- Выбирается алгоритм аутентификации (null, «MD5» 128-бит, «SHA-1 160-бит»), включая размер данных, возвращаемых алгоритмом аутентификации.
- Выбирается метод сжатия и дополнительная информация, которая необходима для этого алгоритма.
- Выбирается главный секрет: 48 бит, общий для обоих сторон соединения.
- Выбирается случайное значение клиента: 32 битное значение, предоставляемое клиентом.
- Выбирается случайное значение сервера: 32 битное значение, предоставляемое сервером.

Главный секрет и случайные значения клиента и сервера (всего 102 бита) хешируются в последовательность байт, которая разбивается и присваивается следующим ключевым переменным: секрет подписи клиента, секрет подписи сервера, ключ шифрования клиента, ключ шифрования сервера, синхромаркер клиента (при использовании DES для поточного шифрования), синхромаркер сервера (при использовании DES для поточного шифрования). Таким образом информации в общем секрете должно быть достаточно для получения ключевых параметров с допустимой степенью надежности. Как видно из соответствующих значений для реализации протокола с экспортными ограничениями, для генерации ключей мы имеем всего 102 бита, что является очень маленьким числом.

Параметры клиента используются сервером, когда он принимает и обрабатывает сообщения, а параметры сервера- клиентом. После того как параметры защиты выбраны и сгенерированы ключи, могут быть назначены текущие состояния соединения.

Передаваемое сообщение перед передачей по сети разбивается на блоки длины 2^{14} байт или меньше, над которыми производятся операции в следующим порядке. Первая операция- сжатие, с точки зрения криптографии не представляет особого интереса.

Шифрование и аутентификация. Аутентификация приводится перед шифрованием информации по следующей формуле:

$$\text{HMAC\_hash(MAC\_Секрет, счетчик, сжатый текст)} \qquad (1)$$

Шифрование происходит по формуле

$$\text{Chipher(сжатый текст, подпись, \{возможное выравнивание длины\})} \qquad (2)$$

Добавление выравнивания длины до размера блока шифра происходит при использовании блочных шифров.

Рассмотрим **протокол приветствия**. Уровень, находящийся выше TLS, не может быть всегда уверен, что установлено максимально надежное соединение между двумя сторонами: есть ряд способов для криптоаналитика заставить протокол установить минимально защищенное соединение. Протокол организован так, чтобы уменьшить этот риск, но возможно, например, блокировать доступ к порту, на котором реализован сервер защиты, чтобы заставить вести неавторизованное соединение. Пользователь высшего уровня должен для себя решить, какая защита ему необходима и никогда не передавать данные по каналу меньшей защищенности, чем необходимо. Защищенным является канал, использующий 3DES с 1024 битным ключом RSA с узлом, чей сертификат был верифицирован, в отличие от соединения с ключом 40 бит. Эти цели достигаются при использовании протокола приветствия. Приветствие проходит в два этапа. На первом этапе производится аутентификация клиента и сервера. Клиент отсылает запрос на сертификат сервера. Возможны варианты, когда соединение устанавливается без проверки сертификатов, с проверкой сертификата сервера, и с сертификатами сервера и клиента. На втором этапе происходит выбор алгоритмов шифрования и аутентификации. Возможно, что в ходе приветствия не будет выбрано применение никаких алгоритмов защиты.

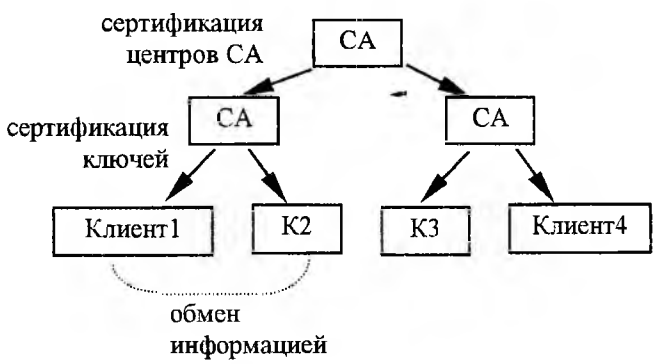
Аутентификация клиента и сервера происходит по их сертификатам, которыми обмениваются во время протокола приветствия. Сертификат содержит открытый ключ клиента, для

иесимметричного алгоритма. В сертификате содержится информация о пользователе, с которым будет установлена связь. Достоверность этой информации гарантируется подписью третьей стороны (службы сертификации), которой доверяют оба участника обмена сообщениями. Кроме того, указывается дата действия сертификата.

Таблица 2

Название поля	**Содержимое**
Информация о пользователе	Идентификационное имя, открытый ключ
Информация о подписи	Имя службы сертификации, цифровая подпись
Период действия	Дата начала действия, дата конца действия
Административная информация	Версия, серийный номер

После истечения срока действия сертификата или после его компрометации, информация о нем заносится в список компрометации (Revocation List), который может быть доступен всем пользователям. Подпись службы сертификации можно проверить на основании ее открытого ключа. Возникает вопрос: нельзя ли как-нибудь проверить подлинность самого этого ключа и как можно гарантировать его аутентичность. Этот вопрос решен. На самом деле существует иерархия центров аутентификации, открытый ключ каждого следующего из них сертифицирован центром, более старшим в этой иерархии. Таким образом, увеличивается надежность всей системы.



CA- (Center Authority) -центр сертификации

Рис. 3

Как показывает анализ TLS протокола, он обеспечивает необходимые функции целостности, причастности и скрытости содержания, но длины ключей, используемых в прототипе, не обеспечивают требуемого уровня криптостойкости. Например, длина ключа для симметричных шифров (общий секрет) составляет 48 бит, поэтому предлагается использование этих протоколов, но с ключами, отвечающими современному состоянию вычислительной техники.

В заключение следует заметить, что на смену арифметике, используемой сейчас в протоколе TLS, приходит арифметика с использованием эллиптических кривых. Существует возможность настройки протокола на использование таких криптосхем. Можно построить протокол, реализованный на шифрах: схема шифрования с использованием эллиптических кривых, цифровая подпись на эллиптических кривых, схема Диффи-Хелмана на эллиптических кривых. Следующая реализация протокола TLS предположительно должна содержать эти алгоритмы.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 681.32

В.А. ГОРБАЧЕВ, канд. техн. наук, М.А. ВОЛК, канд. техн. наук, С.Н. САРАНЧА

МЕТОДЫ СЕРТИФИКАЦИИ МИКРОПРОЦЕССОРНЫХ КОМПОНЕНТОВ

На современном этапе развития электронных систем все большую роль играет информация. Именно поэтому вопросам защиты информации сейчас уделяется все больше внимания.

Однако достаточно большой класс потенциальных угроз безопасности остался вне рассмотрения специалистов по защите информации – это класс аппаратно реализуемых угроз. Интенсивно развиваются методы и способы электронного терроризма – несанкционированного доступа к информации, расположенной в компьютере, с целью ее кражи, разрушения, взлома защиты и использования не по назначению.

Особенности, способствующие данному феномену:

- Компьютер, как средство сберегания, обработки и передачи информации, сделался объектом электронного терроризма;
- Стремительное развитие технологий производства ИС сделало возможным использование аппаратных закладок с большими возможностями;
- Бесконтрольные закупки импортной вычислительной и оргтехники.

Исследования показали, что наиболее вероятным местом расположения источников таких угроз (ИУ) является центральный процессор, поскольку он имеет доступ практически ко всем ресурсам персонального компьютера и может проникать под любые слои защиты систем безопасности. Способом реализации угрозы в данном случае выступают не специфицированные для данного компонента функции.

Для защиты от угроз данного типа предлагается перед установкой компонентов в ПЭВМ производить процедуру *сертификации*. Под сертификацией понимается комплекс организационно-технических мероприятий, в результате которых подтверждаются показатели и характеристики образца.

При сертификации аппаратных средств решаются такие задачи:

- оценка соответствия технических показателей аппаратуры установленным техническим нормам;
- определение уровня физических полей, возникающих при работе устройств, и степени их опасности с точки зрения появления КНСД к информации;
- наличие закладных устройств, предназначенных для снимания и передачи информации;
- обеспечение заданного уровня защищенности информации, циркулирующей в программно-аппаратной среде.

Задача сертификации средств вычислительной техники решается в рамках технической диагностики, основные функции которой как науки определяются обеспечением качества средств и эффективности их использования по назначению на этапах проектирования, производства и эксплуатации. Процесс диагностирования включает в себя два этапа:

а) проверка того, что устройство работает исправно;

б) поиск дефектов, если проверка дала негативный результат.

С точки зрения сертификации электронных средств на наличие источников угрозы необходимо выполнить такие процедуры:

а) сертификация на соответствие специфицированным функциям;

б) сертификация на наличие не специфицированных функций;

в) анализ этих функций при позитивном результате второго этапа.

С точки зрения нормального функционирования микропроцессора не специфицированная функция может рассматриваться как неисправное поведение чипа. В таком случае для поиска не специфицированных функций может быть применен весь имеющийся на сегодняшний момент аппарат диагностирования неисправностей в микропроцессорах.

Трудность диагностирования неисправностей в микропроцессорных структурах определяется следующими факторами:

–все современные микропроцессоры представляют собой интегральные схемы сверхбольшой степени интеграции, в которых нельзя наблюдать сигналы на внутренних точках системы. В то же время введение многочисленных контрольных выводов явно нерационально. Следовательно, необходимы такие тест-процедуры, при использовании которых для информации о наличии неисправностей требовались бы только нормальные вводы и выводы схемы;

–отсутствие детальной информации об объекте диагностики;

–большое число выводов и высокие частоты функционирования компонентов требуют использования специального технического обеспечения в виде комплексов диагностики.

Микропроцессор с точки зрения теории автоматов может быть рассмотрен как последовательностный автомат. Однако для последовательностных автоматов не разработано единых и универсальных методов построения и оценки качества тест-последовательностей. Это связано с тем, что, во-первых, каждый возможный проверочный вход, как правило, можно вычислять для каждого возможного состояния схемы. Таким образом каждый элемент памяти, содержащийся в схеме, удваивает объем вычислений, необходимых для отыскания теста. Во-вторых, существует проблема начальной установки. Прежде, чем мы сможем применить некоторый тест к последовательностной схеме, мы должны суметь установить ее в известное фиксированное состояние или в крайнем случае мы должны знать, в каком состоянии она находится. Это может быть сделано, если первому тесту предшествует *установочная последовательность*. Необходимо также чтобы эти процедуры были в значительной мере автоматизированными.

Рассмотрим микропроцессор как объект диагностики.

С точки зрения теории автоматов микропроцессор может быть описан как

$$Y = A(X;S;\delta;\gamma) \qquad (1)$$

где Y - множество выходов

X - множество входов

S - множество состояний

$\delta : S \times X \to S$ функция переходов

$\lambda : S \times X \to Y$ функция фыходов

Однако при таком представлении возникает проблема в представлении данных о функционировании микропроцессора в связи с высокой сложностью выполняемых им функций.

В настоящее время существует два различных подхода к процедуре построения диагностирующих последовательностей – детерминированный и стохастический. Детерминированный метод опирается на детальное знание структуры исследуемого устройства и определенный класс обнаруживаемых неисправностей. Стохастический метод может рассматривать объект диагностики как "черный ящик" с заданными функциями (в том или ином виде).

При детерминированном способе построения диагностирующей последовательности [1] предлагается модель МП — это совокупность взаимосвязанных функций, реализуемых компонентами оборудования (называемых механизмами), каждый из которых представлен частью оборудования МП с неизвестной или частично известной структурой для реализации определенной функции. Оборудование МП дифференцируется на механизмы хранения и передачи данных, управления передачей данных, обработки данных, управления обработкой. Тест МП определяется как совокупность тестов отдельных механизмов.

1. Механизм обработки данных выполняет арифметические и логические операции, модификацию операндов и результата, выработку признаков результата, операции адресной арифметики.

2. Механизм управления обработкой данных выполняет дешифрацию операций, модификацию операций, операндов, результата, активизацию операций и их модификаций.

3. Механизм хранения и передачи данных представлен совокупностью регистров и шин.

4. Механизм управления передачей данных осуществляет адресацию и выборку регистров, отрабатывает реакции на внутреннее состояние и переходы.

5. Механизм отработки реакций на внешние сигналы и управления вводом-выводом осуществляет взаимодействие с внешней средой.

Проектирование теста для МП есть процесс построения теста для каждого механизма. При этом будем считается, что неисправности одного механизма не влияют на проверку исправности .другого.

Преимуществом данного метода является его простота и универсальность, хорошая применимость для микропроцессоров с малой разрядностью регистров, малым числом регистров и небольшим набором операций.

В рамках стохастического подхода к проблеме генерации диагностирующих последовательностей [2] можно выделить метод построения тестов на основе аппарата цепей Маркова. При этом микропроцессор может представляться функциональной моделью – графом информационной связанности, структурно-функциональным графом или алгоритмическим описанием на основе регистровых передач. Такая методика позволяет обнаруживать неисправности типа "чувствительность к определенным последовательностям команд". Преимуществом такого подхода является отсутствие модели неисправности, недостатком – большая длина тест-последовательности (и, следовательно, большое время тестирования) и высокая сложность оценки результатов.

Общим недостатком вышеперечисленных методов является отсутствие в модели объекта временных характеристик и неприменимость к микропроцессорам с суперскалярной архитектурой. В то же время все современный микропроцессоры широкого применения построены именно по суперскалярной архитектуре. Рассмотренные выше методы не могут обнаруживать ошибки в устройствах, функционирующих на основе алгоритмов с нечеткой логикой (например устройство предсказания ветвлений Branch Prediction Unit или кэш-память). Однако ошибки в кэш-памяти или блоке предсказания ветвлений никак не сказываются на правильности работы микропроцессора с точки зрения выполняемых функций, поскольку данные устройства являются программно прозрачными, однако приводят к замедлению работы микропроцессора и понижению реального iCOMP индекса по сравнению с заданным.

Ввиду отсутствия априорных знаний о структуре и функционировании ИУ мы можем производить тестирование только с точки стохастического подхода. Вопросы детерминированного тестирования без модели неисправности в настоящее время только начинают рассматриваться [3], однако объектом исследования выступают комбинационные схемы, что является абсолютно неприменимо в данной ситуации.

Основными вопросами стохастического подхода являются методы моделирования объектов диагностики, методы построения тест-последовательности и методы оценки результатов тестирования.

В качестве метода описания объекта диагностики выбираем процессный метод. Преимущество данного метода заключается в том, что фактически любая современная микросхема обладает моделью, написанной на языке VHDL, где процесс является одной из основных структурных единиц. Поскольку единого определения такого понятия как процесс нет, мы под процессом понимаем аппаратно реализуемое преобразование цифровой информации. Таким образом любое цифровое устройство можно представить как совокупность параллельно протекающих процессов. Микропроцессор можно представлять на структурно-процессном уровне, когда каждому структурному блоку ставится в соответствие процесс и на функционально-процессном уровне, когда выполнению каждой команды ставится в соответствие процесс. В первом случае мы можем говорить о процессе конвейеризации, чтения-записи кэш-памяти и т.д. Во втором случае мы можем рассматривать процессы выполнения команд пересылки данных, обработки данных и управления. Таким образом процессное представление легко интегрирует в себя методы представления объекта диагностики, использующиеся как в детерминированном, так и в стохастическом методах генерации тестов.

Каждый процесс П характеризуется набором входных и выходных линий и набором внутренних состояний и фактически является автоматом (1). В множестве входных линий выделяется подмножество линий $X_a \in X$, изменение уровня сигнала на которых активизирует процесс. Данное подмножество сигналов называется списком чувствительности.

Информация передается от процесса к процессу с помощью сигналов. При этом один сигнал является выходным для одного процесса и входным для другого. Таким образом процессы формируют между собой информационные отношения типа предок-потомок. Условимся называть процесс-предок *управляющим*, если его выходные линии входят в список чувствительности процесса-потомка. В ином случае будем называть процесс-предок *информационным*.

Исходя из вышесказанного возможно построить граф информационной связанности процессов ГИСП, где вершинам будут являться процессы, а дугами – сигналы. В качестве основной вершины графа выберем процесс взаимодействия с внешней средой. Далее строим покрытие графа, активизирующее все управляющие дуги графа.

Модель для проверки процесса хранения и передачи данных представлена совокупностью регистров и связей между ними. Это определяет граф регистровых передач (ГРП), который имеет регистры-вершины $R=(R_0,..R_m)$ и две дополнительные вершины IN, OUT, отображающие внешнюю среду, $J^A=\{I_1,\ldots,I_r\}$ — множество команд пересылок и ветвлений в МП.

Модель неисправностей процесса хранения и передачи данных представлена следующими допущениями.

1. Любой разряд регистра или любая линия передачи данных может принимать константные значения 0 или 1.

2. Две любые линии передачи данных могут быть замкнуты.

3. Допускается наличие указанных неисправностей с любым числом разрядов регистров и линий передачи данных,

4. Пути передачи данных, источники и приемники информации выбираются правильно.

Таким образом допускается наличие кратных константных неисправностей и неисправностей типа "короткое замыкание" в информационных связях процессов. Построение проверяющего теста переноса должно удовлетворять условиям: для каждой пары разрядов должен существовать набор с различными значениями сигналов в этих разрядах; для каждого разряда должны существовать по крайней мере два набора с различными значениями сигналов в атом разряде. Таким образом, минимальной длины тест переноса содержит $\log_2 n+1$ наборов, где n — число разрядов.

Процедура 1 построения минимизированного теста путем составления неизбыточной совокупности путей ГРП, покрывающей все дуги, представляет собой задачу покрытия: для каждого пути выбирается кратчайшая последовательность команд, активизирующая этот путь; при этом каждому ребру пути ставится в соответствие дизъюнкция команд, помечающих это ребро; составляется некоммутативная конъюнкция дизъюнкций всех ребер пути от IN до OUT; выполняется переход от КНФ к ДНФ, раскрывая скобки с сохранением порядка конъюнкций, применяя, где это возможно, операцию поглощения; выбирается терм полученного выражения, состоящий из минимального числа команд.

Минимальные термы, активизирующие все участки выбранных путей представляют собой искомое решение, состоящее из набора команд, выполнение которых с операндами и адресами теста переноса обеспечит проверку процесса хранения и передачи данных тестируемого микропроцессора. Процедура 1 может быть адаптирована к проверке любых информационных связей между процессами.

В качестве модели для проверки процесса управления передачей данных рассматривается некоторое множество n-разрядных регистров $R=\{R_0,..,R_{m-1}\}$. Выборка регистра приемника или источника информации обеспечивается дешифрацией номера этого регистра, явно или неявно адресуемого в команде МП. Для этого применяются дешифраторы регистровых файлов, мультиплексоры, демультиплексоры.

Модель дефектов представлена неисправностью записи в регистры $R\rightarrow R_i/R\rightarrow fl(R_i)$, вместо $R\rightarrow R_i$ осуществляется запись $R\rightarrow fl(R_i)$, где $fl(R_i)$ — произвольное подмножество регистров, в том числе и пустое множество. Наличие таких неисправностей допускается для нескольких регистров R_i. Неисправность чтения $R_i\rightarrow R/f2(R_i)\rightarrow R$ определяется чтением информации в приемник R из регистров $f2(R_i)$ вместо R.

Процедура 2 позволяет находить все неисправности данного типа. Она состоит из двух фаз – прямой и обратной. В прямой фазе в каждый регистр записывается последовательно, в порядке возрастания номеров, двоичный код его номера, а затем в таком же порядке читается содержимое каждого регистра. В обратной фазе, в порядке возрастания номеров, в каждый регистр R_i (i=0,m-1) записать двоичный код m-1-i и в том же порядке выполнить чтение их содержимого.

Таким образом применяются процедура 1, процедура 2 и процедура 3 детерминированного метода. Формируются входные воздействия для процесса взаимодействия с окружающей

средой и подаются на выводы микропроцессора. Полученные реакции объекта сравниваются с результатами имитационной модели. Во многом первый этап диагностирования перекликается с задачами, выполняемыми самим микропроцессором при выполнении процедуры самотестирования BIST. Однако поскольку содержание процедуры BIST различна для каждого типа микропроцессора, то необходимость проводить детерминированный этап диагностирования остается. Во время этого этапа при сравнительно небольшом количестве поданных тест-векторов проверяется весьма большое количество неисправностей.

На втором этапе производится диагностирование стохастическим методом. При этом единицей входных воздействий служит *цикл шины*. Такой подход позволяет экономить память входных воздействий (при поддержке таких функций аппаратной частью комплекса диагностирования). Для построения входной последовательности используется аппарат многосвязанных марковских цепей. Под многосвязанной марковской цепью понимается последовательность, в которой вероятность перехода в следующее состояние зависит не только от текущего состояния, но и от n предыдущих состояний (n – глубина связанности). Под состоянием будем понимать процедуру формирования того или иного цикла шины. Таким образом можно значительно уменьшить размер матрицы переходных вероятностей по сравнению с [2] ввиду того, что число разнообразных циклов шины для современных процессоров не превышает 10.

Процедура тестирования начинается с подачи сигнала сброса и приведения микропроцессора в заранее заданное состояние. Далее алгоритм представляет собой последовательное выполнение следующих шагов:

1. Вычисление нового состояния марковской цепи.
2. Доопределение необходимых параметров цикла (адрес, данные).
3. Изменение матрицы вероятностей – уменьшение вероятности перехода в наиболее часто встречающееся состояние и увеличение остальных вероятностей.
4. Повторение шагов 1-3.
5. Подача сформированных таким образом тест-векторов на реальный объект и на имитационную модель, сбор реакций и их анализ.

Данный алгоритм, полагая память входных воздействий и реакций достаточно большой для хранения сформированной таким образом последовательности, не рассматривает способы установки микропроцессора в заранее определенное состояние и методы приостановки-запуска процедуры диагностирования.

Для применения данного подхода к диагностированию микропроцессорных компонентов необходимо обладать знаниями о структурной схеме компонента, путях передачи данных, реализуемых шинных циклах и иметь процессную модель данного компонента. Аппаратной поддержкой служит комплекс диагностики, содержащий в своем составе средства подачи тестовых воздействий на объект диагностирования и средства сбора реакций объекта на данные воздействия.

Предложенный метод диагностирования является применимым для задач поиска источников аппаратно реализуемых угроз безопасности информации в микропроцессорных системах.

Список литературы: **1.** *Хаханов В.И.* Техническая диагностика элементов и узлов персональных компьютеров. К.:ИСМО, 1997. 308с. **2.** *Клисторин И.Ф., Гремальский А.А.* Функциональный контроль микропроцессорных устройств. Минск: Знание, 1990г. 90с. **3.** *Raimund Ubar, Dominique Borrione.* Design Error Diagnosis in Digital Circuits without Error Model. TIMA research report. 1999 june, pp 1-5.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 09.02.200

УДК 681.394.74

В.К. СТЕКЛОВ, д-р техн. наук, И.А. ТАРАСЕНКО

ОПТИМАЛЬНЫЙ АЛГОРИТМ КОГЕРЕНТНОЙ ОБРАБОТКИ РАЗЛИЧНЫХ ВАРИАНТОВ МНОГОЧАСТОТНЫХ ГРУППОВЫХ СИГНАЛОВ В СОВРЕМЕННЫХ УСТРОЙСТВАХ ДОСТУПА К СЕТЯМ ПЕРЕДАЧИ ДАННЫХ

Развитие сети Internet и других сетевых технологий и протоколов предопределило необходимость повышения скорости доступа к сетям передачи данных. В настоящее время основными сетевыми приложениями являются передача текстовой и графической информации, работа с базами данных, как доступ к ним, так и постоянные их обновления и репликации, передача подвижного изображения и голоса и другие приложения, требующие для своей реализации больших сетевых ресурсов и высокой скорости доступа. Решением данной проблемы можно считать использование волоконно-оптических соединений, однако экономически это оправдано только для крупных клиентов, так как существующая кабельная инфраструктура полностью состоит из медных кабелей. Поэтому был предложен новый метод передачи с использованием обычной витой медной пары – DSL (Digital Subscriber Line). Этот метод основывался на том, что пропускная полоса частот витой медной пары не ограничивается частотой 3,4 кГц, а имеет более реальное ограничение в диапазоне 1,1 МГц (при использовании витой пары на сравнительно дальние расстояния), хотя с увеличением частоты увеличивается и затухание, соответственно снижается дальность и надежность работы таких устройств. Использование достаточно широкой полосы частот для передачи, предопределяет повышенную чувствительность данных устройств шумам с различными частотными характеристиками. Поэтому, для этих устройств важно определение оптимального приема многопозиционных сигналов и использование более помехозащищённой системы многопозиционных сигналов.

Остановимся более подробно на технологии ADSL, которая на данный момент является наиболее прогрессирующей из-за оптимального соотношения скорости работы и дальности, разности скоростей передачи и приема (1:10 – практически идеально для доступа к сети Internet), возможности одновременной передачи данных и голоса, а также количества поддерживаемых приложений. Эта технология интересна для рассмотрения нашего алгоритма когерентной обработки многочастотных групповых сигналов тем, что она использует амплитудно-фазовую модуляцию и манипуляцию различной кратности как тип линейного кодирования сигналов.

Принципиально существует 2 основных линейных кода – это CAP (Carrierless Amplitude/Phase) и DMT (Discrete Multitone).

Рассмотрим технологию DMT, которая принципиально представляет собой многоканальный модем (или устройство, работающее по принципу частотного разделения каналов). Данная технология использует принцип разделения спектра на большое количество узкополосных каналов и параллельную передачу маленьких фрагментов данных в каждом канале одновременно. Каждый из каналов модулируется отдельно (используется амплитудно-фазовая модуляция), при этом частота модуляции равна средней частоте каждого канала. Согласно стандарта ANSI T1.413 спектр делится на подканалы по 4,3125 кГц каждый. Нижняя частота выделяется для возможности передачи голоса одновременно с данными. Выделяется 256 подканалов, нумерация - левосторонняя, из них 32-250 (136 KHz - 1.1 MHz) подканалы для передачи в одном направлении (относительно пользователя это приём данных или downstream) и 6-31 (24 KHz - 136 KHz) подканалы для передачи данных в обратном направлении (относительно пользователя – передача или upstream), при этом подканалы 16 и 64 выделяются под пилот-каналы для синхронизации.

Таким образом, каждый подканал представляет собой канал передачи данных со скоростью 4 Кбода и модуляцией QAM и при использовании 15-16 кратной манипуляции скорость передачи может достигать 60-64 Кбит/с. Однако, кратность манипуляции для каждого из подканалов определяется независимо друг от друга и полностью зависит от соотношения сигнал шум в данном подканале и может быть от 0 до 16 бит/сек/Гц, в данном случае зависимость прямо пропорциональная. Таким образом, данная система линейного кодирования является наиболее адаптирующейся к условиям линии.

Далее предлагается рассмотреть оптимальный алгоритм когерентной обработки многочастотных групповых сигналов. Этот алгоритм даст возможность использовать более высокий уровень кратности манипуляции по сравнению с классическим методом приёма при одинаковом соотношении сигнал/шум, тем самым он может дать значительный прирост скорости передачи данных и возможность использования модемов даже на линиях сравнительно низкого качества с незначительными изменениями в отношении сигнал/шум с течением времени, а также с фазовыми дрожаниями.

Итак, в чем же заключается данный алгоритм? Принцип оптимального алгоритма когерентной обработки многочастотных групповых сигналов состоит в том, что на определённом «скользящем» интервале принятых информационных посылок после принятия решения о варианте полученного сигнала происходит приведение и усреднение проекций принятых вариантов сигнала по отношению к какому либо из определенных заранее вариантов сигнала. Затем, из полученного нового варианта по заранее известным отношениям между вариантами сигналов системы строится новая система сигналов, которая будет адаптирована под состояние данного канала или подканала передачи и будет использоваться приёмником для принятия решений о вариантах полученных сигналов в дальнейшем.

При современной цифровой обработке принятых многочастотных сигналов удобно переходить от высокочастотного сигнала к его отображению через координаты в двумерном пространстве, что на практике соответствует, например, операциям переноса спектра или разделения ортогональных сигналов в многоканальной системе. Тогда оптимальный алгоритм приёма можно представить в следующем виде:

$$i = \arg\min\left[(x_0 - x_j)^2 + (y_0 - y_j)^2\right], \quad j = 1,2,...,m. \tag{1}$$

Входящие величины x_0 и y_0 определяются в результате обработки принятой посылки сигнала, а величины x_j и y_j, число которых равно 2m, должны быть известны априорно или вычислены (оценены) в процессе приёма предыдущих посылок сигнала. Для вычисления оценок проекций вариантов сигнала x_j и y_j воспользуемся методом приведения и усреднения проекций принятого сигнала.

Пусть, $\Delta\tilde{\varphi}_n$ - разность фаз между вариантом сигнала, в пользу которого в демодуляторе принято решение на n-ой посылке и первым вариантом сигнала (в данном случае приведение и усреднение будем проводить на основе 1-го варианта какой-то системы сигналов), $\tilde{a}_n$ - амплитуда сигнала, пользу которого принято решение на n-ой посылке. Естественно, $\Delta\tilde{\varphi}_n$ принимает значения из дискретного множества разрешенных фаз, определяемых заранее определенной системой сигналов. Что касается $\tilde{a}_n$, то эта величина равна фактической амплитуде, принятой на n-ой посылке смеси сигнала с шумом, однако в дальнейшем она отождествляется с амплитудой того варианта сигнала, в пользу которого вынесено решение на n-ой посылке. Волнистой чертой в обоих случаях подчеркивается, что эти оценки могут быть ошибочными. Тогда приведенные проекции x_{1n} и y_{1n} принятого сигнала на n-ой посылке вычисляется через принятые проекции x_{0n} и y_{0n} по формулам:

$$x_{1n} = \frac{a_1}{\tilde{a}_n}(x_{0n}\cos\Delta\tilde{\varphi}_n + y_{0n}\sin\Delta\tilde{\varphi}_n), \quad y_{1n} = \frac{a_1}{\tilde{a}_n}(y_{0n}\cos\Delta\tilde{\varphi}_n - x_{0n}\sin\Delta\tilde{\varphi}_n). \tag{2}$$

Подчеркнём, что $\tilde{a}_n$ и $\Delta\tilde{\varphi}_n$ определяются решением о переданном на n-ой посылке вариантом сигнала, принятом по результатам обработки величин x_{0n} и y_{0n}. Далее, величины, полученные путём приведения, усредняются на "скользящем" интервале в M посылок, предшествующих обрабатываемой в данный момент:

$$\tilde{x}_1 = \frac{1}{M}\sum_{n=1}^{M}\frac{a_1}{\tilde{a}_n}(x_{0n}\cos\Delta\tilde{\varphi}_n + y_{0n}\sin\Delta\tilde{\varphi}_n), \quad \tilde{y}_1 = \frac{1}{M}\sum_{n=1}^{M}\frac{a_1}{\tilde{a}_n}(y_{0n}\cos\Delta\tilde{\varphi}_n - x_{0n}\sin\Delta\tilde{\varphi}_n). \tag{3}$$

Волнистой чертой над $\tilde{x}_1$ и $\tilde{y}_1$ отмечается, что это оценки. Оценки являются несмещёнными и эффективными. Отметим, что при вычислении оценок нет необходимости в априорных сведениях о средней мощности приходящего сигнала, т.к. в этот алгоритм входят только отношения амплитуд. Из оценок следующим образом можно сформировать несмещенные и эффективные оценки проекций всех остальных вариантов сигнала, входящих в данную систему сигналов.

$$\tilde{x}_j = \frac{a_j}{a_1}(\tilde{x}_1\cos\Delta\varphi_j - \tilde{y}_1\sin\Delta\varphi_j), \quad \tilde{y}_j = \frac{a_j}{a_1}(\tilde{x}_1\sin\Delta\varphi_j + \tilde{y}_1\cos\Delta\varphi_j), \tag{4}$$

где $\Delta\varphi_j$ - известная разность фаз между j и 1-м вариантами сигналов системы, а a_j и a_1 их амплитуды соответственно.

Таким образом, мы получаем, адаптированную под соотношение сигнал/шум в канале, новую систему сигналов, на основе которой будет приниматься решение о варианте поступающих посылок

сигнала. Данный алгоритм достаточно сложен и его реализация стала практически возможна только благодаря развитию микропроцессорной техники.

Большим преимуществом использования данного алгоритма приёма является то, что он универсален, т.е. он может быть использован практически для любой системы многопозиционных сигналов. Поэтому предлагается сравнить рассчитанные теоретически помехоустойчивости двух систем многопозиционных сигналов с использованием оптимального алгоритма когерентного приёма многочастотных групповых сигналов. В качестве основы для сравнения возьмем 4-х кратную 16-ти позиционную квадратурную амплитудную модуляцию (рис. 1), а в качестве оппонента 16-ти позиционную систему сигналов, представленную на рис. 2.

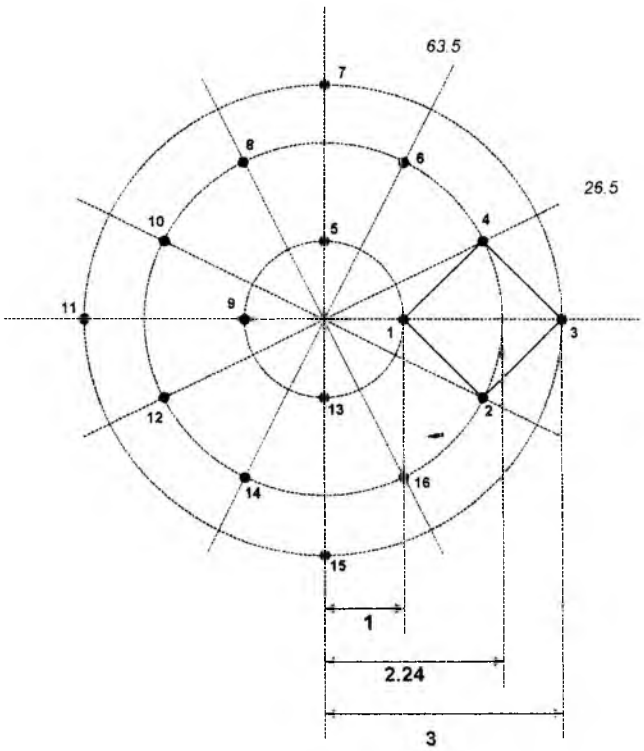


Рис. 1

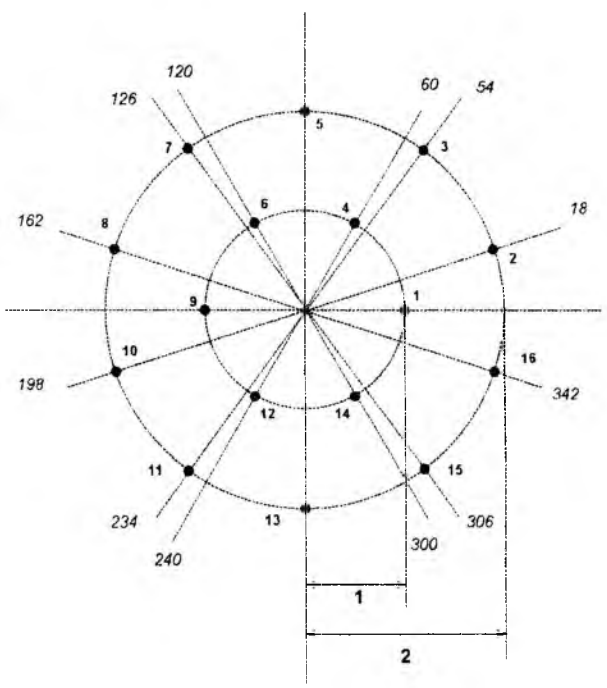


Рис. 2

Для теоретического расчета помехоустойчивости данных систем использовалась программа на языке Mathcad с использованием метода статического моделирования. В качестве источника дискретной информации использовалась сформированная псевдослучайная последовательность чисел от 0 до 15. Моделирование воздействия шума рассчитывалось на основе средней энергии сигналов системы и заданного значения отношения сигнал/шум. При определении новой системы сигналов согласно алгоритма оптимального когерентного приема многочастотных групповых сигналов использовался интервал усреднения равный 20. Полученные результаты можно представить в виде графиков зависимости вероятности ошибок от отношения сигнал/шум для различных систем сигналов, приведенных на рис. 3.

Кривые 4 и 2 соответственно характеризуют теоретическую помехоустойчивость для шестнадцати позиционной квадратурной амплитудной модуляции и для системы сигналов, представленной на рис. 2. Как видно из графиков выигрыш от применения системы сигналов, представленной на рис. 2, составляет около 2-х дБ. Графики 1 и 3 получены при теоретических расчётах с использованием оптимального алгоритма когерентного приема многочастотных групповых сигналов с интервалом усреднения равным 20. Таким образом, можно сделать вывод о том, что выигрыш при применении оптимального алгоритма когерентного приёма многочастотных групповых сигналов составляет около 2-х дБ. Поскольку данный метод универсален и может быть использован для приема любой системы многочастотных групповых сигналов, то использование оптимального метода приёма совместно с более помехоустойчивой системой сигналов может дать суммарный выигрыш в отношении сигнал/шум около 4 дБ. Тем самым можно будет повысить дальность, надежность и скорость работы высокоскоростных устройств доступа к сетям передачи данных.

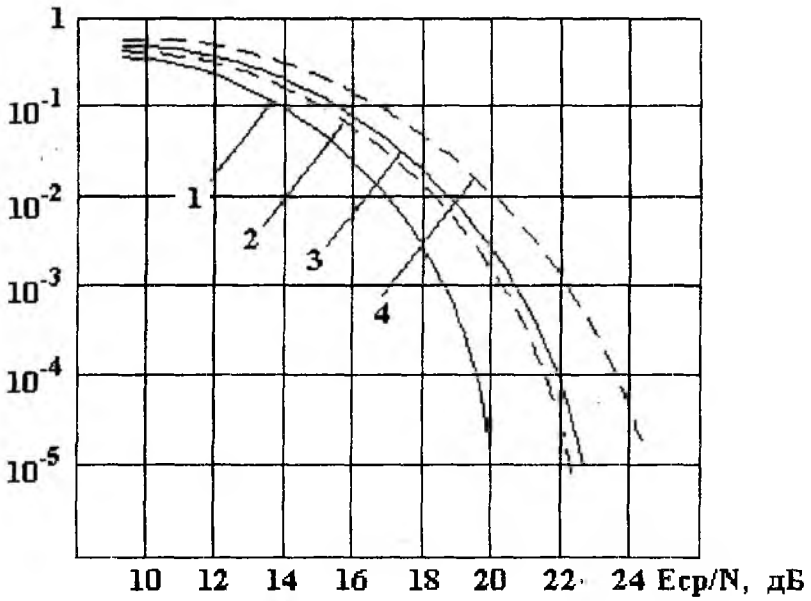


Рис. 3

Далее рассмотрим возможное применение оптимального алгоритма когерентного приёма многочастотных групповых сигналов в современных устройствах доступа к сетям передачи данных, а именно в ADSL модемах, использующих линейное кодирование DMT. Данные модемы при установлении соединения и определения начальной частотной зависимости отношения сигнал/шум могут использовать различные методы передачи и приёма синхросигналов, а также последующей подстройки приёмника или передатчика. При определении подканалов с более высоким уровнем шума возможно использование другой помехоустойчивой системы многопозиционных групповых сигналов. После установления соединения зависимость отношения сигнал/шум от частоты может незначительно изменяться под влиянием различного рода помех и наводок. Поэтому имеет смысл дополнительно, время от времени, использовать метод подстройки системы сигналов на основе принимаемых информационных посылок. Как один из вариантов такой подстройки можно использовать описанный ранее оптимальный алгоритм когерентного приёма многочастотных групповых сигналов. Обработку принимаемых сигналов по данному алгоритму можно производить постоянно для всех каналов либо циклично поканально на интервале в 20 посылок (теоретически найденный оптимальный интервал), определяя при этом разность между начальными значениями проекций сигнала и приведенными и усредненными проекциями, с последующим обновлением начальной системы сигналов при превышении данной разностью определенного значения. При этом все вышеуказанные действия будут производиться с информационными посылками сигнала без разрыва связи. Использование данной реализации даёт максимальную адаптированность модемов по отношению к изменяющимся линейным характеристикам, но это приведёт к значительному усложнению приёмных устройств и соответственно к повышению их стоимости. Таких возможных вариантов и механизмов инициализации оптимального алгоритма когерентной обработки многопозиционных групповых сигналов можно реализовать достаточно много с учётом различных характеристик системы приёма.

Таким образом, с учётом всего вышеописанного, приёмник современного высокоскоростного устройства доступа к сети, использующего оптимальный алгоритм когерентного приёма многочастотных групповых сигналов, схематично можно представить в следующем виде (рис. 4):

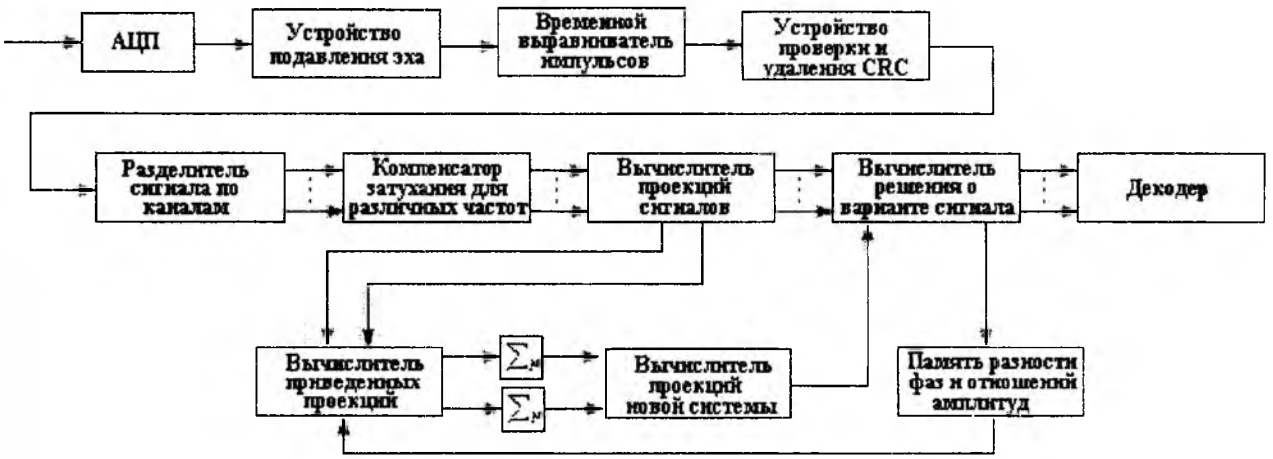


Рис. 4

В статье проанализированы методы повышения скорости в сетях доступа. Представлен универсальный алгоритм оптимального приёма, позволяющий анализировать модем для различных систем многопозиционных сигналов. Рассчитаны кривые помехоустойчивости для двух систем шестнадцатипозиционных сигналов.

Список литературы: **1.** *Шварцман В.О., Емельянов Г.А.* Теория передачи дискретной информации. - М.: Связь, 1986. **2.** *Мизин И. А., Богатырёв В. А., Кулешов А. П.* Сети коммутации пакетов. - М.: Радио и связь, 1986. **3.** *Бомштейн Б. Д., Киселев А.К., Моргачев Б.Т.* Методы борьбы с помехами в каналах проводной связи. - М.: Связь, 1975. **4.** *American* National Standards Institute, ANSI T1.413-95, Asymmetric Digital Subscriber Line (ADSL) Metallic Interface, 1995.

Киевский институт связи
Украинской государственной
академии связи им. А.С. Попова

Поступила в редколлегию 22.03.2000

УДК 681.394.74

С.Н. СКЛЯРЕНКО, канд. техн. наук, С.И. МНИЩЕНКО

ИНТЕГРАЦИЯ ИНТЕЛЛЕКТУАЛЬНОЙ И МОБИЛЬНЫХ СЕТЕЙ ПРИ СОЗДАНИИ ГЛОБАЛЬНОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

Основным стратегическим направлением развития и совершенствования средств связи в Европе считается создание телекоммуникационной базы для информационной базы в составе EII – Европейской Информационной Инфраструктуры – с перспективой её внедрения в Глобальную Информационную инфраструктуру [1]. Решение проблем EII предполагает опережающее развитие телекоммуникационной среды как основы взаимодействия информационных систем в соответствии с моделью взаимодействия открытых систем. Аналогичные задачи ставятся и в России [2,3].

Экспоненциальный рост числа пользователей, как услугами мобильной связи, так и услугами мультимедиа и, как следствие, возрастающая конкуренция, заставит операторов телекоммуникационных сетей уже в ближайшее время:

- улучшить качественные показатели сетей;
- значительно расширить перечень предоставляемых услуг (с учётом запросов пользователя, но не ограничений той или иной технологии);
- расширить взаимодействие между системами, базирующимися на разных технологиях.

Сведения о текущем и прогнозируемом, на ближайшие 5-10 лет, состоянии рынка телекоммуникационных услуг обуславливают следующие требования к сетевым подсистемам:

- поддержка локальной и глобальной мобильности в широком смысле (мобильность терминала, мобильность пользователя, мобильность услуг и т.д.);
- предоставление услуг с использованием технологии мультимедиа;
- поддержка различных технологий (роуминг между системами различных стандартов);
- поэтапность и преемственность (эволюционность) развития.

Разрабатываемая в настоящее время система UMTS-Универсальная Мобильная Телекоммуникационная Система должна удовлетворить всем указанным требованиям. Она призвана интегрировать в своём составе фиксированные и мобильные цифровые сети связи.

Современные телекоммуникационные сети представляют собой исключительно сложные объекты. Естественно, что анализ и синтез алгоритма функционирования столь сложного объекта, как телекоммуникационная сеть информационной инфраструктуры, немыслима без использования правил функциональной декомпозиции [4]. При рассмотрении вопросов, связанных со структурой телекоммуникационной сети, как правило, используется метод функциональной декомпозиции этого объекта в двух направлениях (по горизонтали и вертикали). На рис.1 показана объективная декомпозиция базовой телекоммуникационной сети. Каждый из элементов сети должен иметь стандартные интерфейсы как по горизонтали (между одноименными уровнями разнородных сетей), так и по вертикали (между смежными функциональными уровнями).

При рассмотрении вопросов, связанных с интеграцией базовых составляющих фиксированных и мобильных сетей, наибольший интерес представляет интеграция соответствующих компьютерных платформ (платформы IN), ведающих оперативным управлением процедурами предоставления услуг [5]. От успешного решения подобных задач зависят темпы расширения перечня услуг связи, предоставляемых на множестве разнородных сетей.

В настоящее время вопросы интеграции компьютерных платформ фиксированных (ISDN) и мобильных (GSM) сетей интенсивно развиваются в контексте разработки стандартов на систему UMTS, которая в свою очередь призвана играть роль системы обобщенного доступа к услугам EII.

Эволюционность является основным лейтмотивом работ, связанных с развитием и стандартизацией основных положений концепции UMTS. В настоящее время интенсивно проводятся работы по сопряжению сетевых подсистем двух главных систем (GSM и ISDN/DECT) к общей конечной цели – сетевой подсистеме системы UMTS.

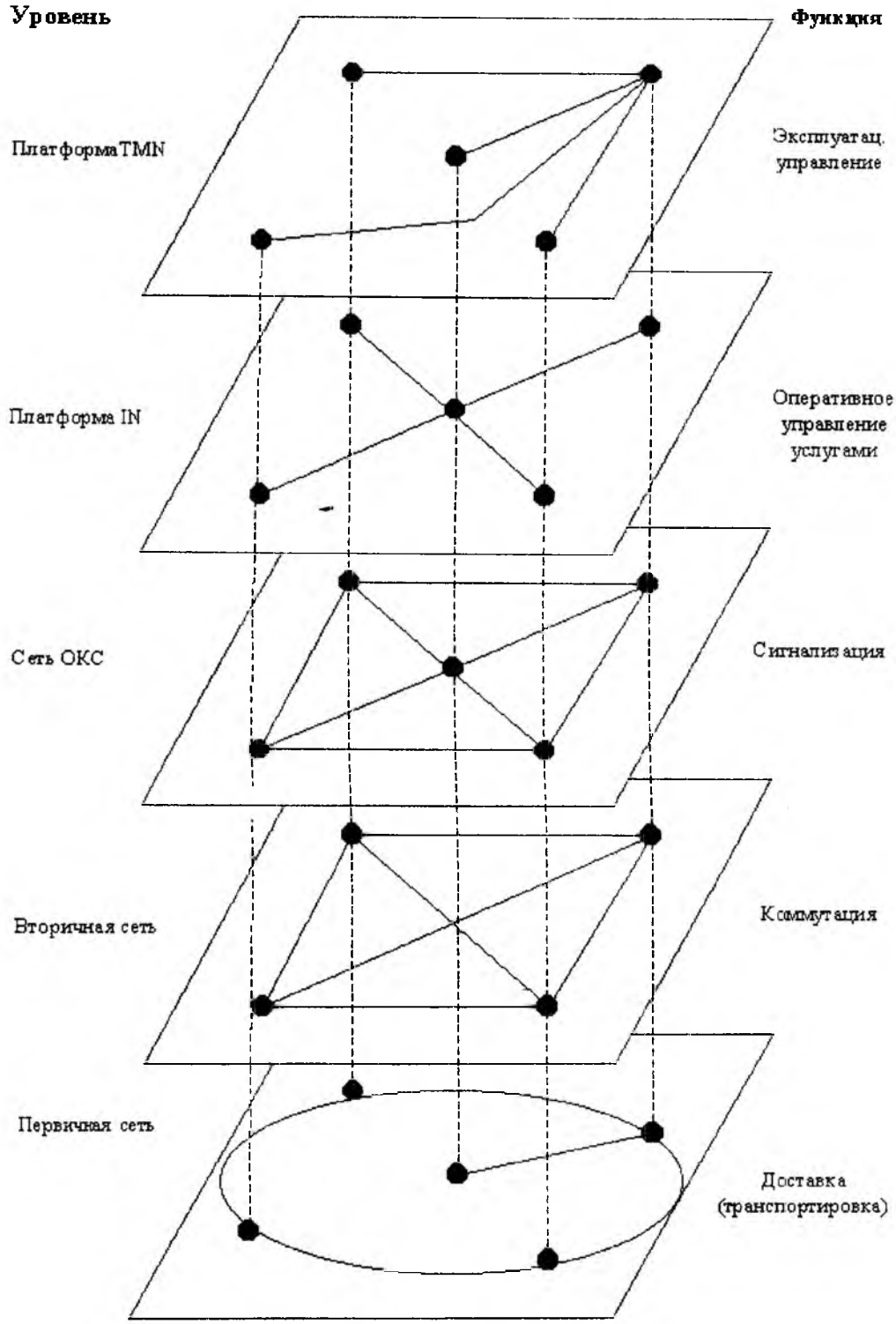

Уровень
Функция
Платформа TMN
Эксплуатац. управление
Платформа IN
Оперативное управление услугами
Сеть ОКС
Сигнализация
Вторичная сеть
Коммутация
Первичная сеть
Доставка (транспортировка)

Рис. 1

Сегодня сетевые подсистемы стационарных (ISDN, Internet) и мобильных (GSM) сетей функционируют по отдельности. Однако согласно концепции ближайшей задачей является решение задачи их интеграции на базе совместной интеллектуальной компьютерной платформы, обеспечивающей персонализацию услуг и автоматизацию их предоставления использованием сетевых ресурсов мобильных систем (GSM), спутниковых систем, ISDN, Интернет. Персонализация и автоматизация означает:

- предоставление услуги – персональный профиль где угодно и когда угодно;
- требования к среде передачи – радиоэфир, металлические и оптические кабели;
- координированность – вызовы и сообщения спецслужбам;
- замкнутость – комплексность транзакции (безшлюзовая технология).

Первый этап интеграции мобильных и стационарных сетей будет происходить путем интеграции соответствующих интеллектуальных платформ с ориентацией на управление процедурами предоставления дополнительных услуг в сочетании с услугами мобильности. В связи с этим существенное изменение претерпят как платформа IN фиксированных сетей, так и компьютерная платформа сетей GSM на пути их конвергенции и последующей интеграции в составе интеллектуальной платформы UMTS. Первым шагом в данном направлении является создание платформы CAMEL в сетях GSM.

Концепция CAMEL определяет принципы интеграции структурных свойств Интеллектуальной сети и сетевой подсистемы GSM, а также определяет механизм поддержки в сочетании с услугой роуминга ряда специфических услуг, не стандартизированных в спецификациях GSM.

Концепция CAMEL разрабатывается поэтапно. В настоящее время обсуждается ближайшее и долгосрочное её развитие в контексте разработки концепции VHE - Виртуальное Домашнее Окружение, которая в свою очередь разрабатывается для UMTS.

CAMEL задает принципы поддержки услуг IN в сетях GSM, а соответствующие спецификации первой фазы могут рассматриваться как:

часть спецификаций ITU IN CS-1 (например, в части операций, задаваемых протоколом взаимодействия объектов платформы IN, и используемых при этом точек DP);

расширение возможностей CS-1 в части услуги мобильности терминала.

Для поддержки услуг мобильности разработаны новые механизмы, обеспечивающие:

обмен информацией между базами данных GSM и CAMEL;

поддержку дополнительных услуг GSM, подобных услуге "Перенаправление вызова", при межсетевом взаимодействии.

Кроме того, определены процедуры блокирования передачи сообщений на иностранных языках и запросов HLR со стороны SCP. Функциональная структура платформы GSM/CAMEL на первой фазе развития показана на рис.2.

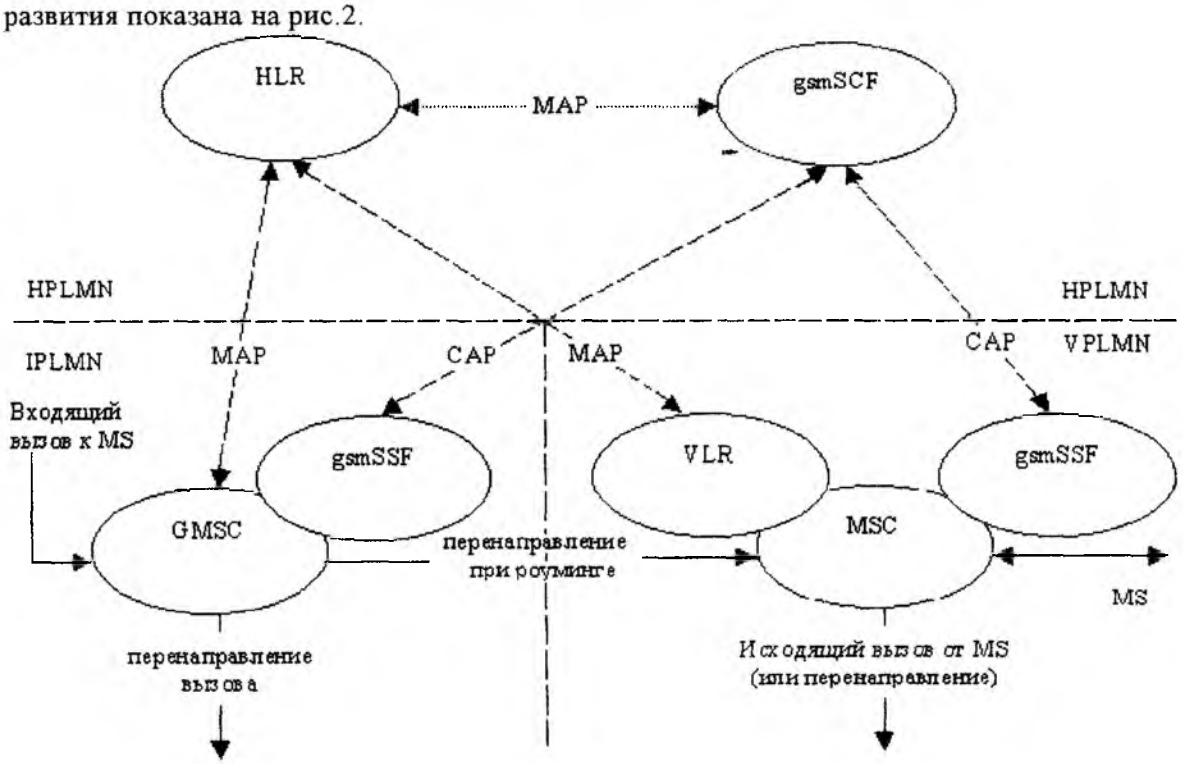


Рис. 2

Приведенные на рис.2 сокращения обозначают:

HPLMN - Home Public Land Mobile Network - Сеть подвижной радиосвязи общего пользования, в которой прописан вызываемый абонент;

IPLMN - Interrogating PLMN - Сеть подвижной радиосвязи общего пользования, в которой находится вызывающий абонент;

VPLMN - Visited PLMN - Сеть подвижной радиосвязи общего пользования, в которой находится вызываемый абонент;

MSC - Mobile Switch Center - Центр коммутации сети подвижной радиосвязи;

GMSC - Gateway MSC - Шлюзовой центр коммутации сети подвижной радиосвязи;

HLR (VLR) - Home (Visited) Location Register - Серверы "интеллектуальной" платформы GSM;

SCF/SSF - Service Control/Switch Function - Элементы функциональной плоскости Концептуальной Модели Интеллектуальной Сети;

MS - Mobile Station - Мобильный терминал;

MAP - Mobile Application Protocol - Прикладной протокол поддержки услуги мобильности;

CAP - CAMEL Application Protocol - Прикладной протокол поддержки услуг CAMEL;

В отличие от платформы IN здесь введен интерфейс между gsmSSP и gsmSCP разных сетей.

В настоящее время основные усилия по стандартизации направлены на поддержку первой фазы платформы CAMEL, ориентированной на обработку запросов услуг IN в HPLMN, VPLMN и IPLMN.

Реализация решений, связанных с поддержкой конвергенции ресурсов компьютерных платформ IN и GSM, позволит интегрировать эти платформы в рамках платформы CAMEL. В свою очередь это позволит достичь положения, при котором пользователь сможет получать привычные для него телекоммуникационные услуги независимо от того, в какой сети он находится на момент их запроса. Избавление пользователя от необходимости знать границы и тип сети является решающим фактором в части маркетинга будущих телекоммуникационных услуг.

Список литературы: **1.** *ETSI/TA*22(95)5, SRC6REP, JMM /al, 31 May,1995. **2.** *Голубков А.С.* Перспективы развития российской Информационной Инфраструктуры и вхождение в мировое информационное сообщество. Вестник-3, Комитет при президенте РФ по политике информатизации, 1996. **3.** *Булгак В.Б.* Перспективы развития электросвязи в России и ее вхождение в Глобальную Информационную Инфраструктуру. Электросвязь, №8, 1995, с.2-3. **4.** *Буч Г.* Объективно-ориентированное проектирование с примерами применения: Пер. С англ. - М.:Конкорд, 1992, 519 с. **5.** *ETSI ETR* 271 Universal Mobile Telecommunications System (UMTS); Objectives and overview (UMTS 01.01), February, 1996.

Киевский институт связи
Украинской государственной
академии связи им. А.С. Попова

Поступила в редколлегию 22.03.2000

УДК 621.396

А.А. РУДЕНКО, А.А. ГРИНЬ

КОРРЕКТИРУЮЩИЕ АЛГОРИТМЫ СИСТЕМ ФАП

Системы фазовой автоподстройки (ФАП) предназначены для согласования (идентификации) фаз переменных напряжений и их можно использовать в радиолокации, связи, электротехнике, телемеханике и других областях, где требуется обеспечить синфазность напряжений переменного тока [1-3].

Рассмотрим возможность построения системы ФАП с принципом управления по отклонению (с обратной связью) (рис. 1,а). На входы 1 и 2 поступает задающее $U_1(t)=U_{m1}\cos[\omega t+\varphi_1(t)]$ и управляемое $U_1(t)=U_{m2}\cos[\omega t+\varphi_2(t)]$ напряжения одинаковой частоты, сдвинутые по фазе на угол $\alpha(t)=\varphi_1(t)-\varphi_2(t)$, в общем случае изменяющийся во времени. Задача состоит в обеспечении равенства фаз этих напряжений. В состав системы входят фазовой дискриминатор ФД, сглаживающий фильтр (фильтр нижних частот) Ф, управляемый фазовращатель ФВ. Если ФД имеет косинусную статистическую характеристику, то при сдвиге фаз между поступающими на его вход напряжениями, равном 90^0, на его выходе – нулевое напряжение.

Для цифровых систем ФАП с управлением по отклонению управляющая микро-ЭВМ (УМЭВМ) входит в замкнутый контур управления и может быть использована для программной реализации корректирующих алгоритмов управления [1]. В общем случае цифровые системы ФАП обладают рядом преимуществ: дискретные (цифровые) элементы обеспечивают более высокую точность передачи и преобразования информации, и во многих случаях оказываются проще в конструктивном отношении аналогичных непрерывных систем ФАП. В настоящей работе рассматриваются корректирующие алгоритмы цифровых систем ФАП с УМЭВМ в замкнутом контуре.

Программная реализация корректирующего алгоритма может быть записана в виде дискретной передаточной функции

$$K_k(z)=\frac{x(z)}{\Delta\varphi(z)}=\frac{\sum_{i=1}^{k}b_j z^{-j}}{\sum_{j=1}^{s}Q_j z^{j}+1}=\frac{D(z)}{F(z)}, \qquad (1)$$

где $\Delta\varphi(z)=Z\{\Delta\varphi[n]\}$; $x(z)=Z\{x[n]\}$.

Дискретной передаточной функции (1) соответствует разностное уравнение:

$$x(n)=b_0\Delta\varphi[n]+b_1\Delta\varphi[n-1]+\ldots+b_k\Delta\varphi[n-k]-(a_1x[n-1]+a_2x[n-2]+\ldots+a_sx[n-s]). \qquad (2)$$

Структурная схема программной реализации корректирующего алгоритма (2), соответствующая прямому программированию, показана на рис. 1,б.

Рассмотрим возможные корректирующие алгоритмы в цифровых системах ФАП.

ПИД – пропорционально-интегральный дифференциальный корректирующий алгоритм.

Если в выражении (1) $k=2$ и $s=1$, то дискретная передаточная функция корректирующего алгоритма имеет вид

$$K_k(z)=\frac{b_0+b_1z^{-1}+b_2z^{-2}}{1-z}. \qquad (3)$$

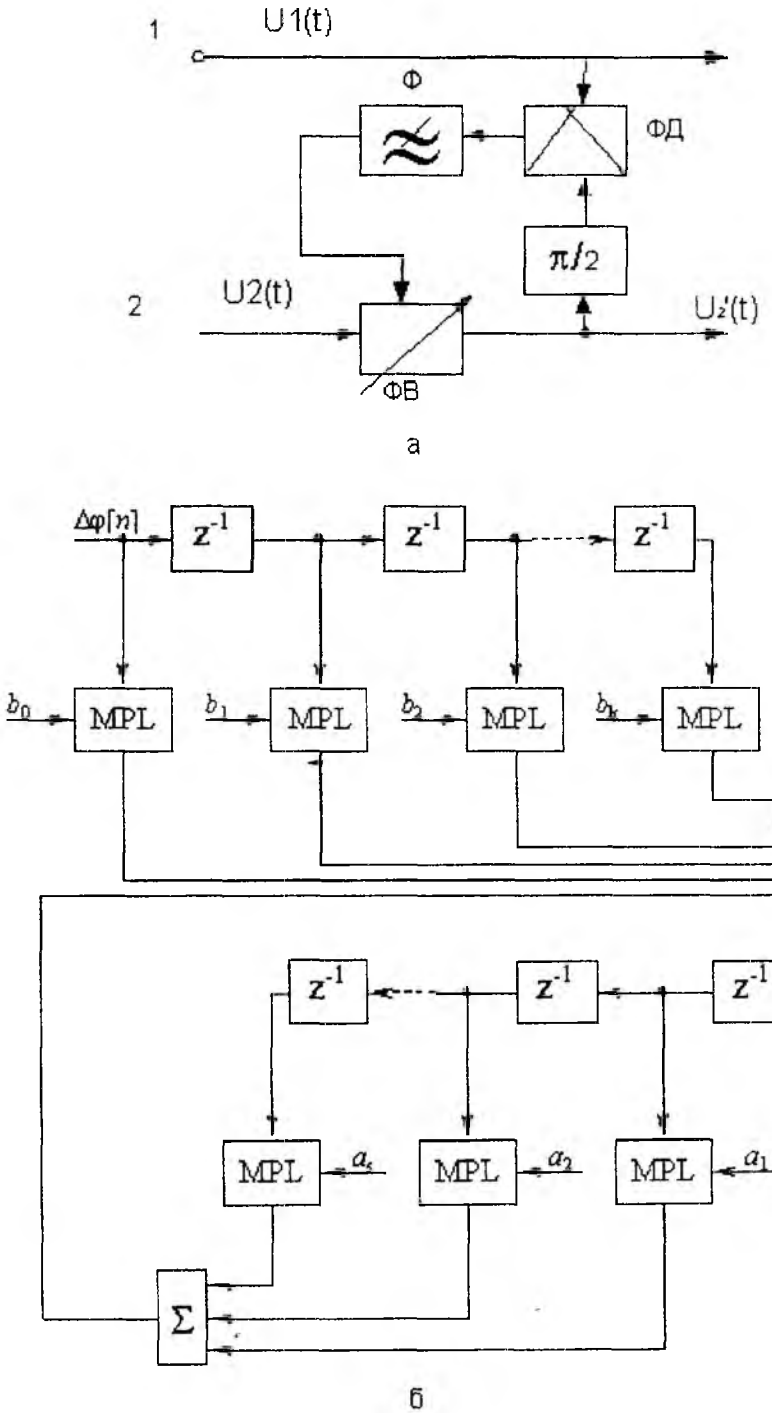


Рис. 1

Структурная схема алгоритма программной реализации передаточной функции (3) изображена на рис. 2,а.

ПИ – пропорционально-интегральный изодромный корректирующий алгоритм управления.

Если в выражении (1) k=1 и s=1, то дискретная передаточная функция корректирующего алгоритма имеет вид

$$K_k(z) = \frac{b_0 + b_1 z^{-1}}{1 - z^{-1}}. \qquad (4)$$

Структурная схема программной реализации ПИ – корректирующего алгоритма, соответствующая выражению (4) изображена на рис. 2,б.

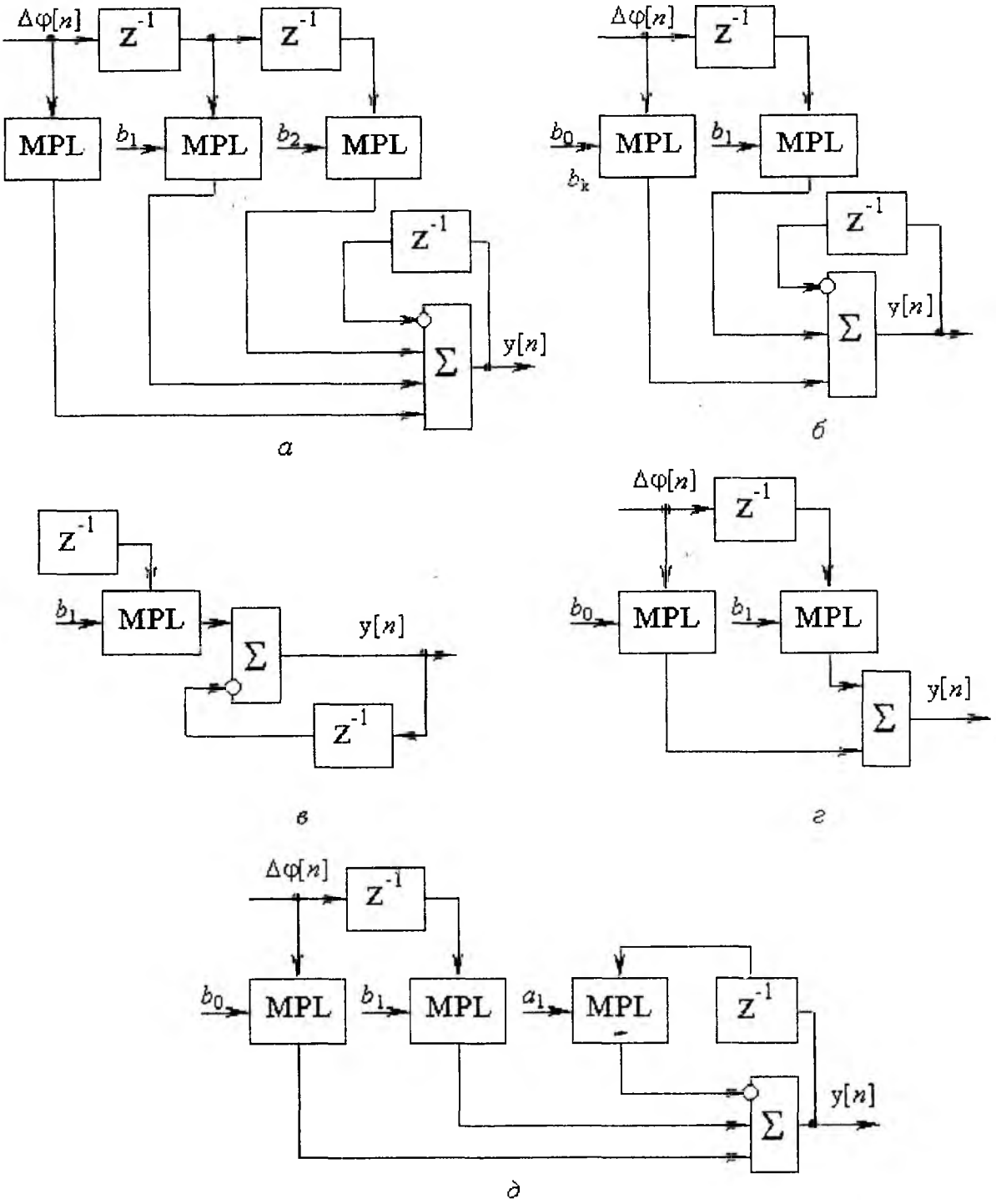


Рис. 2

И – интегральный корректирующий алгоритм управления.

Если в выражении (1) k=1 и s=1, b_0=0, то дискретная передаточная функция корректирующего алгоритма имеет вид:

$$K_k(z) = \frac{b_1 z^{-1}}{1 - z^{-1}}, \tag{5}$$

а структурная схема его программной реализации изображена на рис. 2,в.

П – пропорциональный изодромный корректирующий алгоритм управления.

Если в выражении (1) $b_0 \neq 0$, a_j=0 (j=1,s), то

$$K_k(z) = b_0 .$$

Значение b_0=1 соответствует единичному коэффициенту передачи УМЭВМ. При b_0>1 сигнал управления усиливается, при b_0<1 – ослабляется.

ПД – пропорционально-дифференциальный корректирующий алгоритм управления.

Если в выражении (1) k=1 и s=0, то дискретная передаточная функция корректирующего алгоритма имеет вид:

$$K_k(z) = b_0 + b_1 z^{-1},$$

а структурная схема его программной реализации изображена на рис. 2,г.

ИД – интегро-дифференциальный корректирующий алгоритм.

Если в выражении (1) k=1 и s=1, то дискретная передаточная функция корректирующего алгоритма имеет вид:

$$K_k(z) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1}},$$

а структурная схема его программной реализации изображена на рис. 2,д.

Рассмотренные дискретные передаточные функции корректирующихся алгоритмов позволяют их использовать в цифровых системах ФАП для повышения точности и быстродействия.

Предложены корректирующие алгоритмы управления в цифровых системах фазовой автоподстройки и приведены структурные схемы программной реализации корректирующих алгоритмов.

Список литературы: **1.** *Бойко Н.П.* Системы автоматического управления на базе микро-ЭВМ. – К.: Техника, 1989. – 182с. **2.** *Зайцев Г.Ф., Стеклов В. К.* Радиотехническая система автоматического управления высокого точности. – К.: Техника, 1988 - 208 с. **3.** *Автоматическая* подстройка фазового набега / Под. ред. *М. В. Капронова.* – М.: Сов. радио, 1972. - 175 с.

Укртелеком, г.Киев,
Харьковский электротехникум связи

Поступила в редколлегию 22.03.2000

УДК 621.396.2

В.И.АНТЮФЕЕВ, д-р техн. наук, В.Н.БЫКОВ, канд. техн. наук, А.С. ВИЛЬЧИНСКИЙ, А.М. ГРИЧАНЮК, М.Г. ШОКИН

ОЦЕНКА ТОЧНОСТИ ИЗМЕРЕНИЯ КООРДИНАТ ОБЪЕКТОВ МАТРИЧНЫМИ КОРРЕЛЯЦИОННО-ЭКСТРЕМАЛЬНЫМИ СИСТЕМАМИ НАВИГАЦИИ

В работе [1] получены аналитические соотношения для оценки потенциальной точности определения координат объектов простой геометрической формы (полосы, прямоугольника на однородном фоне) матричными системами землеобзора, в частности, матричными корреляционно-экстремальными системами навигации (КЭСН) летательных аппаратов (ЛА) по наземным ориентирам.

Количественная оценка потенциальной точности местоопределения КЭСН согласно положений работы [1] осуществляется по следующей методике.

1. Строится модель текущего изображения (ТИ), формируемого матричной КЭСН в процессе обзора земной поверхности, осуществляемого по ходу движения ЛА с определенной скоростью и под определенным углом к визируемой поверхности. Предъявляются требования к информационным датчикам матричной КЭСН.

2. Определяются размеры визируемого многолучевой антенной "кадра" (формы растра), параметры разрешаемых элементов на земной поверхности, обусловленные размерами проекций сечения диаграмм направленности антенны (ДНА) парциальных каналов (датчиков), законы движения разрешаемых элементов и законы изменения выходных сигналов матричного датчика.

3. Задается представление визируемого объекта - ориентира навигации и окружающего фона. Делаются допущения о наличии или отсутствии априорной информации об оцениваемом параметре и оцениваемых мешающих (ограничивающих) параметрах. На основе этих допущений выбирается критерий оптимальности. Находится система уравнений, описывающая алгоритм оптимальной обработки сигналов матричной КЭСН. Решение системы уравнений позволяет найти глобальный экстремум критерия оптимальности и дисперсии оценок анализируемых параметров. Дисперсии или среднеквадратические ошибки (СКО) оцениваемых параметров являются искомой мерой точности.

Данная методика может быть применена для оценки точности определения координат объектов различной формы: точечных (занимающих малую часть кадра), протяженных и площадных (занимающих значительную часть кадра), а также радиолокационных и радиометрических КЭСН, работающих в разных участках радио- и видимого диапазонов электромагнитных волн. Различия в оценке для разных систем состоят в первоначальных высотах и углах визирования кадра, параметрах сигналов, излучаемых и/или переотражаемых объектом и фоном.

В данной работе представилось целесообразным оценить точность измерения координат объектов сложной формы матричными радиометрическими (РМ) КЭСН миллиметрового диапазона волн (ММД).

В качестве модели анализируемого изображения принята зонная модель, в которой объект с одним "яркостным" наполнением (радио-яркостной температурой) расположен на однородном фоне с другим яркостным наполнением.

Обзор земной поверхности осуществляется параллельно с помощью многолучевой антенны матричной РМ КЭСН, движущейся в плоскости xy под определенным углом к оси z системы координат x,y,z, связанной с поверхностью земли, из точки (x_0,y_0,z_0). Многолучевая антенна формирует матрицу из N_1 строк и N_2 столбцов, наклон плоскости, в которой лежат оси строчных ДНА, задан углом β_i $(i \in \overline{1,N_1})$ относительно вектора скорости ЛА, а положение оси ДНА в строке характеризуется углом α_{ij} $(i \in \overline{1,N_1}, j \in \overline{1,N_2})$. Ширина каждой ДНА по уровню 3 дБ равна θ_x в угломестной и θ_y – в азимутальной плоскости. Каждая парциальная ДНА аппроксимируется гауссовской поверхностью, пересчитанной к координатам x,y.

Объект сложной формы представляет собой набор N прямоугольников в плоскости xy с вершинами (a_i,c_i), размерами $\Delta l_{xi},\Delta l_{yi}$ вдоль соответствующих осей и радиояркостными

температурами T_i. Обозначим через $\varepsilon_x = a_1$, $\varepsilon_y = c_1$ координаты точки прицеливания, $\Delta T_i - T_{\text{тм}}$ – радиояркостный контраст i-го объекта относительно фона. Будем полагать, что датчики РМ изображения являются безынерционными, их собственные шумы взаимно независимы и представляют собой нормально распределенные случайные величины с нулевым средним значением и дисперсией σ^2, при этом взаимодействие шума с полезным сигналом имеет аддитивный характер.

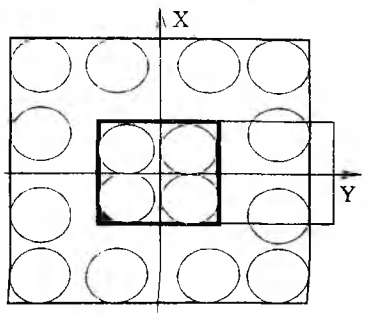


Рис. 1

На рис. 1 приведена геометрия визирования объекта прямоугольной формы, в частном случае квадрата 70×70 м (аналог крыши дома с битумным покрытием) с возможностью его расширения по одной из сторон (по оси у). Большой квадрат обозначает визируемый на земной поверхности кадр. Размеры кадра обусловлены величиной СКО основной инерциальной системы навигации ЛА, которая устраняется с помощью КЭСН [2]. С высоты полета ЛА 1 км размеры кадра составляют 180×180 м. Кружками в кадре обозначено примерное расположение проекций парциальных ДНА многолучевой антенны.

На рис. 2а, б представлены два не соприкасающихся и два соприкасающихся объекта прямоугольной формы с размерами 70×70 м каждый, с возможностью перемещения одного из объектов (верхнего) вдоль оси у (аналог двух крыш либо одной крыши сложной формы) и 20 м и соединяющей их вертикальной дорожки длиной 60 м и шириной 40м.

На рис. 3 показано визирование объекта сложной формы, в кадр попадает пересечение трех бетонных дорожек: двух горизонтальных дорожек шириной 40 м и 20 м и соединяющей их вертикальной дорожки длиной 60 м и шириной 40 м. Полная длина горизонтальных дорожек гораздо больше размеров кадра. Фоном для объектов является травяной покров на земной поверхности.

Оцениваемыми параметрами принимаемого и анализируемого матричной РМ КЭСН сигнала являются координаты $(\varepsilon_x, \varepsilon_y)$. В связи с отсутствием априорной информации об оцениваемых параметрах в качестве критерия оптимальности выбран критерий максимума правдоподобия. Пользуясь аналогичными описанными в работе [1] приемами, для СКО оценок параметров $(\varepsilon_x, \varepsilon_y)$ можно получить соотношения

$$\sigma_x = \|\mathrm{f}_y\| / F(\varepsilon_x, \varepsilon_y), \tag{1}$$

$$\sigma_y = \|\mathrm{f}_x\| / F(\varepsilon_x, \varepsilon_y), \tag{2}$$

позволяющие оценить потенциальную точность местоопределения ЛА, обеспечиваемую матричной КЭСН, в частности матричной РМ КЭСН.

В формулах (1), (2) $\mathrm{f} = [f_{ij}], i \in \overline{1, N_1}, j \in \overline{1, N_2}$;

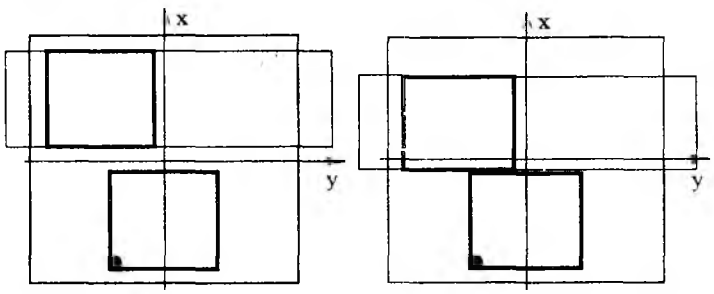
$$F(\varepsilon_x, \varepsilon_y) = \|\mathrm{f}_x\|^2 \|\mathrm{f}_y\|^2 - (\mathrm{f}_x, \mathrm{f}_y), \quad \mathrm{f}_{x(y)} = \left[\frac{\partial f_{ij}}{\partial \varepsilon_x(\varepsilon_y)}\right];$$

$$f_{ij}(\varepsilon_x, \varepsilon_y) = \sum_{k=1}^{N} \Delta T_k \left[\Phi\left(\frac{a_k + \Delta l_{xk} - x_{ij}^0}{\delta_x}\right) - \Phi\left(\frac{a_k - x_{ij}^0}{\delta_x}\right)\right] \times$$

$$\times \left[\Phi\left(\frac{c_k + \Delta l_k - y_{ij}^0}{\delta_y}\right) - \Phi\left(\frac{c_k - y_{ij}^0}{\delta_y}\right)\right];$$

$\|\mathrm{f}\| = \left(\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} f_{ij}\right)$ – норма матрицы; δ_x, δ_y – описанные в [1] параметры аппроксимирующей ДНА гауссоиды.

На рис. 4а,б приведены зависимости СКО определения координат объекта прямоугольной формы (соответственно, y-я и x-я составляющие) от величины Δl_y при отношении сигнал-шум в ТИ $q = \Delta T_1 / \sigma = 3$. При этом значение радиояркостной температуры фона T_{TM}=270K, объекта T_1=267K, СКО шума σ=1K.



а) б)

Рис. 2

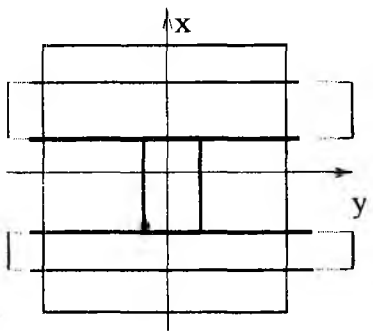
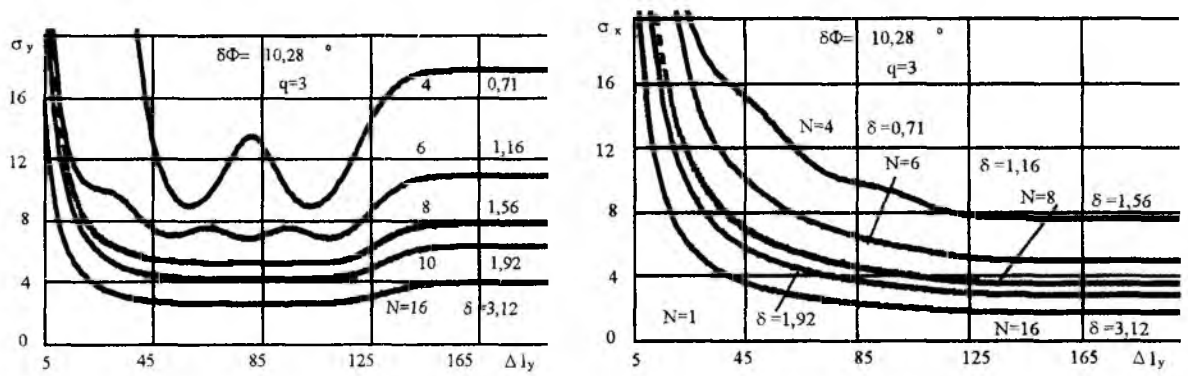


Рис. 3

Полный угол обзора кадра с размерами 180×180 м с высоты 1 км составляет $\delta\Phi = 10{,}28^{\circ}$, необходимым условием является размещение лучей всех парциальных ДНА матричной антенны в указанном угле (кадре). Цифрами на рисунках обозначено число лучей ДНА в квадратной матрице, а также соответствующие значения коэффициентов пересечения соседних ДНА δ. Ширина луча



а) б)

Рис. 4

парциальных ДНА $\theta \approx 2^{o}$ при диаметре параболической или линзовой антенны 30 см (рабочая длина волны $\lambda \approx 8$ мм).

На графиках можно выделить три области. Первая соответствует низкой точности местоопределения вследствие того, что объект при фиксированных размерах по оси x (70 м) имеет незначительные угловые размеры по оси y относительно θ_y. Затем, по мере роста σ_x монотонно уменьшается, что обусловлено накоплением сигнала, принимаемого по строчным лучам, а σ_y колеблется относительно некоторого среднего значения, при этом амплитуда колебаний снижается с увеличением коэффициента пересечения δ. Информативность изображения в направлении оси y при этом падает, точность σ_y уменьшается до значения, которое она принимает при бесконечном удлинении стороны прямоугольника.

Результаты, приведенные на рис. 4, показывают, что применение матричных РМ КЭСН с количеством лучей 6×6 и более позволяет обеспечить точность местоопределения координат площадного объекта до единиц метров. Увеличение размерности матрицы (более 8×8) при неизменном $\delta\Phi$ нецелесообразно (неоправданно), во-первых, из-за снижения отношения сигнал-шум в изображении вследствие возрастания межканальных помех и, во-вторых, в многолучевых антеннах

оптического типа величина δ ограничена значением ≈ 2 из-за минимального углового разрешения, определяемого по критерию Рэлея [3].

Существует несколько причин, которые способны привести к существенному уменьшению отношения сигнал-шум. К ним, в первую очередь, относится ослабление радиоволн ММД в дожде большой интенсивности (затухание радиоволн 8-мм диапазона в дожде с интенсивностью 16 мм/ч равно 4-5 дБ/км). Умеренные дожди интенсивностью порядка 4 мм/ч, а также мокрый снег с таким же содержанием воды способны сгладить (уменьшить) контраст пары объект-фон. Поскольку составляющие точности обратно пропорциональны отношению сигнал-шум, необходимо обеспечить запас по величине q как за счет работы по высококонтрастным объектам (парам, сообществам объектов), так и благодаря применению РМ датчиков матричных КЭСН, обладающих высокой температурной чувствительностью.

Увеличение угла визирования по одной из координат (по координате x) до 45° при неизменных размерах кадра (180×180 м) приводит к уменьшению угла визирования кадра до 5,16°, при этом точность ухудшается на 8...10%, коэффициент пересечения лучей парциальных ДНА возрастает. Так, для матрицы 8×8 коэффициент пересечения равен 1,56, для матрицы 16×16 – 3,12. В этой связи целесообразно осуществлять визирование кадра на земной поверхности под углами 0°...30° от вертикали.

На рис. 5...7 представлены зависимости СКО $\sigma_{y,x}$ от величины сдвига L перемещаемого прямоугольного объекта относительно фиксированного (при этом объекты не пересекаются). Расчеты произведены в случае вертикального визирования для матрицы 8×8.

На рис. 7 осуществлено сравнение оценок точности для двух матриц 8×8 и 16×16. Рис. 6 дополняет

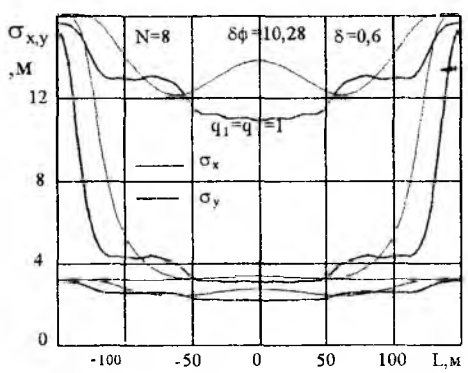


Рис. 5

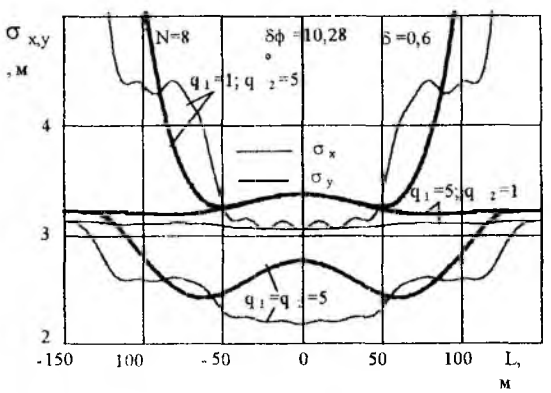


Рис. 6

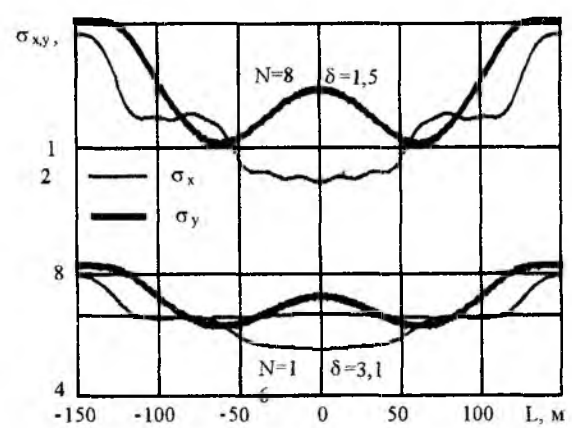


Рис. 7

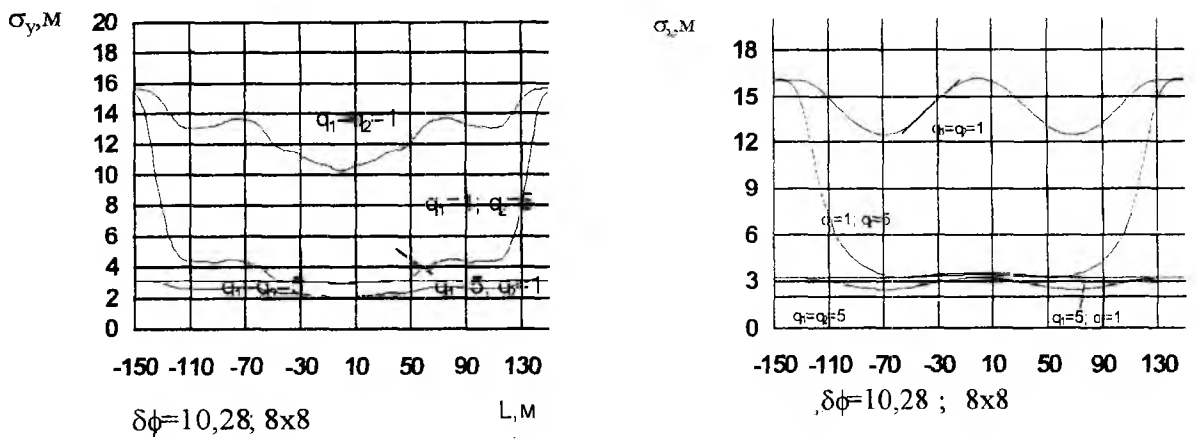
рис. 5 путем изменения масштаба в области малых значений ошибки.

На рис. 8а, б приведены аналогичные зависимости величины СКО для двух касающихся объектов.

Рис. 9 иллюстрирует результаты сравнения величин СКО для касающихся (Δx =70м) и не касающихся объектов (Δx =90м). Результаты, приведенные на указанных рисунках, получены для случая, когда точка прицеливания находится в левом нижнем углу основного объекта

Анализ полученных результатов позволяет сделать следующие выводы. Наблюдается строгая симметрия зависимости величины ошибки до появления в кадре верхнего объекта и после его выхода из кадра.

При совпадении положения объектов по оси x наблюдается минимум ошибки σ_y в связи с накоплением сигнала по большому количеству лучей парциальных ДНА и некоторое увеличение



а) б)

Рис.8

ошибки σ_x. Наилучшая точность наблюдается при большом отношении сигнал-шум у двух объектов.

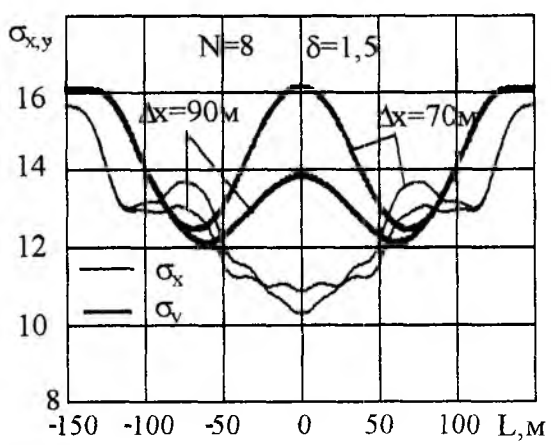
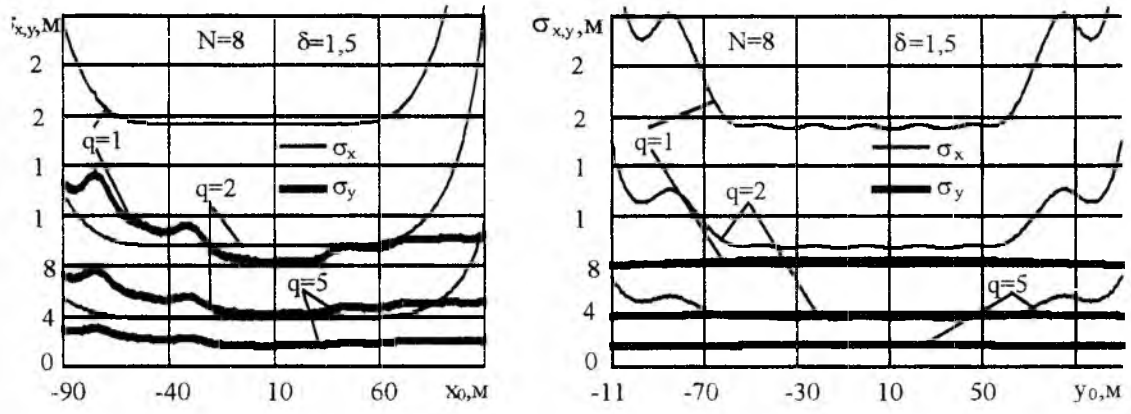


Рис. 9

Уменьшение данного отношения до единицы у вспомогательного объекта незначительно ухудшает точность "привязки" к основному объекту. При малом отношении сигнал-шум у основного объекта величина точности определяется точностью вспомогательного объекта до того момента, пока он не вышел из кадра, а далее точностью (низкой) основного объекта. Определение координат несоприкасающихся объектов осуществляется с меньшей СКО, чем соприкасающихся, так как в этом случае привязка КЭСН идет по двум объектам, окруженным однородным фоном. На рис. 10а,б приведены зависимости СКО определения координат сложного объекта (рис. 3) при перемещении визируемого кадра по оси x (рис. 10а) и оси y (рис. 10б).

Точка прицеливания находится на пересечении нижней горизонтальной и вертикальной дорожек. Следует отметить, что существенное ухудшение точности местоопределения данного объекта наблюдается в случае выхода из кадра вертикальной соединяющей дорожки с расположенной на ней точкой прицеливания. При этом объект воспринимается матричной КЭСН в виде полосы бесконечной длины с отличительными признаками только по одной координате.



а) б)

Рис. 10

В качестве выводов необходимо отметить следующее. Разработанная методика позволяет производить количественные оценки потенциальной точности измерения координат точечных, протяженных и площадных объектов с помощью матричных КЭСН. Анализируемые объекты могут быть использованы в качестве навигационных ориентиров, координаты которых при поэтапном визировании могут быть определены матричной КЭСН с высокой точностью.

Список литературы: **1.** *Агафонов Ю.Н., Антюфеев В.И., Быков В.Н., Гричанюк А.М.* Точность измерения координат объектов матричными системами землеобзора // Радиотехника. 1998. Вып.108. С. 63-68. **2.** *Быков В.Н., Гричанюк А.М.* Анализ влияния размеров эталонного и текущего изображений на функционирование корреляционно-экстремальных систем навигации летательных аппаратов // Радиотехника. 1998. Вып.105. С.122-125. **3.** *Ingvessonm K.S.* Imaging front-end systems for millimeter waves and submillimeter waves // SPIE. Instrumentation for submillimeter spectroscopy. 1985. Vol. 598. P.104-113.

Харьковский военный университет

Поступила в редколлегию 17.01.2000

УДК 621.396.62

Г.В. АЛЁШИН, д-р техн. наук, А.А. ТРУБЛИН

ОБ ОПТИМАЛЬНОСТИ ЧАСТОТНО-СЕЛЕКТИВНЫХ СРЕДСТВ АВИАЦИОННОЙ РАДИОСВЯЗИ, РАБОТАЮЩИХ В РАВНОМЕРНО ЗАГРУЖЕННОМ ЧАСТОТНОМ ДИАПАЗОНЕ

В работе решена задача об оптимальном, по критерию электромагнитной совместимости (ЭМС), распределении усилий селективных устройств радиоэлектронных средств (РЭС), работающих в равномерно загруженном диапазоне волн, и проверена оптимальность параметров избирательности современных средств радиосвязи.

Статистика роста количества РЭС в Украине, а также участие ВС Украины в современных учениях показали, что проблема ЭМС РЭС становится со временем всё острее. Отсутствие единого адекватного математического аппарата до 1992 года не позволяло хотя бы оценить качество боевой совместной работы радиосредств, в том числе средств авиационной радиосвязи в произвольной электромагнитной обстановке (ЭМО). Результаты работ [1, 2] принципиально позволяют это сделать, для известной ЭМО.

Однако производство современных серийных РЭС авиационной радиосвязи, использующее интуитивные инженерные решения, не гарантирует оптимального распределения усилий устройств селекции сигналов. Поэтому целью данной работы является постановка и решение задачи оптимизации, по критерию ЭМС, параметров избирательности РЭС по основному и побочным каналам приема сигналов при ограниченных ассигнованиях на создание соответствующих устройств, а также проверка оптимальности параметров существующих средств. Такая работа проведена впервые. Она полезна как при расчете эффективности существующих, так и для создания перспективных радиоэлектронных средств.

В работах [1, 2] предложен общий критерий качества ЭМС η_i для i-х РЭС различных типов, удовлетворяющий требованиям простоты, адекватности учёта ЭМО и общим условиям его применимости. В данной задаче используется лишь частотная селекция.

$$\eta_i^{-1} = \frac{\Delta f_{1i}}{\Delta f_{2i} + \Delta f_{3i}}, \tag{1}$$

где $\Delta f_{1i} = \int_{-\infty}^{\infty} \left|\dot{k}_i(f)\right|^2 \cdot \left|\dot{s}_i(f)\right|^2 \partial f$,

$$\Delta f_{2i} = \sum_{j=1}^{k} \Delta f_{2j}, \quad \Delta f_{2j} = \int_{-\infty}^{\infty} \left[\left|\dot{k}_i(f)\right|^2 \cdot \left(\dot{s}_i(f) \cdot \dot{s}_j^*(f) + \dot{s}_i^*(f) \cdot \dot{s}_j(f)\right)\right] \partial f,$$

$$\Delta f_{3i} = \int_{-\infty}^{\infty} \left|\dot{k}_i(f)\right|^2 \cdot \left|\sum_{j=1}^{n} \dot{s}_j(f)\right|^2 \partial f,$$

$\dot{k}_i(f), \dot{s}_i(f), \dot{s}_j(f)$ - соответственно нормированные частотные характеристики и спектры i-го РЭС и спектры мешающих сигналов j-х РЭС.

Использовать соотношения Δf_k в формуле (1) не всегда приемлемо для произвольной ЭМО. Для типичного случая полной и равномерной загруженности частотного диапазона, когда

$$\dot{S}_j(f_{0j}) \approx \dot{S}_{j-1}(f_{0(j-1)}),$$

критерий качества ЭМС РЭС η_i [1] можно существенно упростить:

$$\eta_i^{-1} = \sum_{k=1}^{m} \sigma_k^{-2}, \tag{2}$$

где σ_k - избирательность по k-му каналу приема сигналов.

Соотношение (2) представляет собой по существу отношение суммы мощностей помех по паразитным каналам приема к мощности "своего" i-го сигнала в точке приема.

Ответственным этапом оптимизации распределения избирательности по побочным каналам приема при ограничениях на стоимость является получение зависимостей стоимости от избирательности. Путем анализа принципиальных схем базовых образцов техники авиационной радиосвязи Р-155П, Р-863, "Рябина-М1" были получены статистические данные: обеспечиваемая избирательность по побочным каналам приема (зеркальному, соседнему, каналу прямого прохождения промежуточной частоты), стоимость узлов и деталей, израсходованная на обеспечение соответствующего значения избирательности. Эти статистические данные приведены в таблице.

Зеркальный канал		Соседний канал		Канал ПЧ	
Избирательность $\sigma_з$,Дб	Стоимость $C_з$, грн.	Избирательность $\sigma_с$,Дб	Стоимость $C_с$, грн.	Избирательность $\sigma_{пч}$,Дб	Стоимость $C_{пч}$, грн.
60	45,87	60	61,47	70	64,78
70	57,68	70	72,30	80	72,92
90	84,62	80	109,56	100	124,94

Для получения функциональных зависимостей $C(\sigma_{i(k)})$ была составлена программа для ЭВМ на алгоритмическом языке Pascal, которая проводит аппроксимацию функций методом наименьших квадратов с приближением функции к квадратичной зависимости. Для побочных каналов приема, приведенных в таблице, были полученные следующие результаты:

$$C(\sigma_з) = 0{,}01\sigma_з^2 + 0{,}41\sigma_з\ ,$$
$$C(\sigma_с) = 0{,}02\sigma_с^2 - 0{,}18\sigma_с + 0{,}16\ , \qquad (3)$$
$$C(\sigma_{пч}) = 0{,}01\sigma_{пч}^2 - 0{,}04\sigma_{пч} + 0{,}19.$$

В общем виде задачу оптимизации параметров избирательности РПУ РЭС $\sigma_k \quad \forall\, k \in [1,n]$ по критерию минимума η_i^{-1} (1) при ограничениях на суммарную стоимость $C_{доп}$ соответствующих устройств, можно представить в виде:

При $$\min_{\{\bar{\sigma}\}} \eta_i^{-1}(\bar{\sigma}) = \min_{\{\bar{\sigma}\}} \sum_{k=1}^{n} \sigma_k^{-2}, \qquad \sum_{k=1}^{n} C_k(\sigma_k) \le C_{доп}. \qquad (4)$$

С учетом линеаризации ограничений задача оптимизации примет вид:

При $$\min_{\{\bar{\sigma}\}} \eta_i^{-1}(\bar{\sigma}) = \min_{\{\bar{\sigma}\}} \sum_{k=1}^{n} \sigma_k^{-2}, \qquad \sum_{k=1}^{n} C'_k(\sigma_k)\cdot\sigma_k \le \Delta C_{доп}\ , \qquad (5)$$

где $\Delta C_{доп} = C_{доп} - C(\bar{\sigma}_0) - \sum_{k=1}^{n} C'_k(\sigma_{k0})\cdot\sigma_{k0}$.

Оптимальное решение $\sigma_{j\,opt}$ задачи (5), полученное методом множителей Лагранжа, и оптимум η_i^{-1} запишутся следующим образом

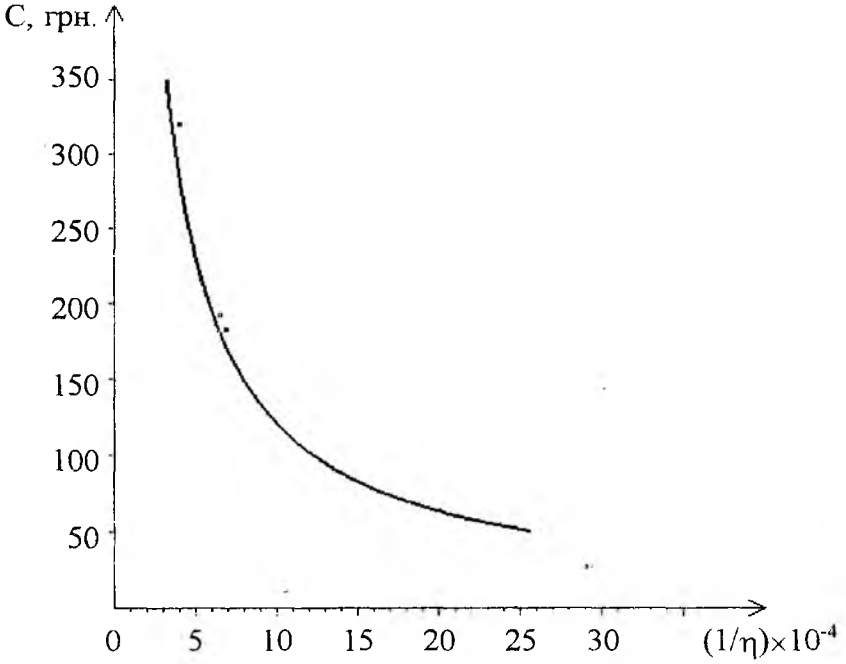
$$\sigma_{j\,opt} = \frac{\Delta C_{доп}}{\sqrt[3]{C'_j}\cdot\sum_{k=1}^{n}(C'_k)^{2/3}}, \qquad \forall\ j,k \in [1,n], \qquad (6)$$

и

$$\eta_{i\,\min}^{-1} = \frac{\left[\sum_{k=1}^{n}(C'_k)^{2/3}\right]^3}{\Delta C_{доп}^2}. \qquad (7)$$

Зависимости между оптимальными значениями показателей качества иногда называют кривыми обмена [3]. Такое название обусловлено их противоречивостью, что очевидно из зависимости $\eta_{i\,\min}^{-1}$ от

$C_{доп}$ (7) и рисунка, на котором изображена кривая обмена для зеркального, соседнего каналов и канала промежуточной частоты.



Точками на рисунке обозначены значения величины η_i^{-1}, которые обеспечиваются соответствующими селекторами приемных устройств. Вид кривых обмена и положение точек реальных радиоприемных устройств примечателен. Близость точек реальных радиоприемных устройств к кривой обмена прогнозировалась, поскольку исследовавшиеся радиоприемные устройства уже давно находятся в эксплуатации и их параметры интуитивно приведены инженерами к значениям, близким к кривой обмена. Кривая обмена позволяет оценить, насколько технологичны и технически оптимальны устройства РЭС, обеспечивающие ЭМС.

Кривая обмена, показанная на рисунке позволяет также прогнозировать показатели ЭМС перспективных РЭС, оптимально их выбирать при разработке РЭС при заданных уровнях взаимных помех и для заданного энергетического потенциала РЭС и, кроме того, может служить основой для оценки эффективности РЭС и для разработки соответствующих оптимальных стандартов Украины.

Список литературы: **1.** *Алёшин Г.В.* Основы построения оптимальных информационно-измерительных радиотехнических систем. - Харьков: ХВУ, 1994. 252 с. **2.** *Алёшин Г.В.* Эффективность радиотехнических устройств оценивания параметров сигнала. - Харьков. ХВВКИУ РВ, 1992. 103 с **3.** *Гуткин Л.С.* Проектирование радиосистем и радиоустройств. - М.: Радио и связь, 1986. 288 с.

Харьковский Институт Лётчиков
Военно-Воздушных Сил Украины

Поступила в редколлегию 14.03.2000

УДК 621.391

С.С. МАРТЫНЕНКО

ОБНАРУЖЕНИЕ ИМПУЛЬСНОГО СИГНАЛА НА ФОНЕНЕГАУССОВСКИХ ПОМЕХ

В работе делается попытка отойти от традиционного подхода в задаче обнаружения радиоимпульса, считая, что сигнал принимается на фоне аддитивной негауссовской помехи.

Пусть имеется независимая выборка случайных величин $\vec{\xi} = \{\xi_1, \xi_2, ..., \xi_n\}$ объемом n. Причем при осуществлении гипотезы H_1 (наличие сигнала) для выборки справедливо представление:

$$\xi_v = S_v + n_v,$$

а при осуществлении гипотезы H_0 (отсутствие сигнала) - имеет место:

$$\xi_v = n_v.$$

Случайная помеха n_v является негауссовской случайной величиной, описываемой последовательностью кумулянт $\{\chi_1, \chi_2, ..., \chi_s\}$ в общем случае с отличными от нуля кумулянтами 3-го и высших порядков, что говорит о негауссовости помехи[1]. Значения сигнала $S_v = S(t_v)$ представляют собой выборку из видео- или радиоимпульсов, взятых в моменты времени t_v, $v = \overline{1,n}$. Необходимо по выборке определить, какая осуществляется гипотеза H_1 или H_0.

В качестве решающего правила используется полиномиальное степени S решающее правило вида:

$$\gamma(\vec{\xi}) - h_0 \underset{H_0}{\overset{H_1}{\gtrless}} 0, \tag{1}$$

где

$$\gamma(\vec{\xi}) = \sum_{i=1}^{s} \sum_{v=1}^{n} h_{iv} \xi_v^i, \tag{2}$$

$$h_0 = \frac{1}{2}(E_0 + E_1). \tag{3}$$

В последнем выражении E_0 и E_1 - математическое ожидание решающей функции $\gamma(\vec{\xi})$ при гипотезе H_1 и H_0 соответственно.

Неопределенные коэффициенты h_{iv} находятся оптимальными по критерию КУ [2] из минимума функционала:

$$Q_s = \frac{G_0[\gamma] + G_1[\gamma]}{(E_1 + E_0)^2}, \tag{4}$$

где

$$G_0[\gamma] = M[(\gamma - M\gamma)^2 / H_0] = \sum_{i=1}^{s} \sum_{j=1}^{s} \sum_{v=1}^{n} h_{iv} h_{jv} F_{(i,j)v}(H_0), \tag{5}$$

$$G_1[\gamma] = M[(\gamma - M\gamma)^2 / H_1] = \sum_{i=1}^{s} \sum_{j=1}^{s} \sum_{v=1}^{n} h_{iv} h_{jv} F_{(i,j)v}(H_1). \tag{6}$$

Функции $F_{(i,j)v}$ выражаются через моменты случайных величин ξ_v порядка i, j и $(i+j)$ при гипотезах H_1 и H_0 следующим образом:

$$F_{(i,j)v}(H_1) = m_{(i+j)v} - m_{iv} m_{jv}, \quad F_{(i,j)v}(H_0) = u_{(i+j)v} - u_{iv} u_{jv}, \tag{7}$$

$$F_{(i,j)v} = F_{(i,j)v}(H_0) + F_{(i,j)v}(H_1). \tag{8}$$

Показано, что оптимальные значения коэффициентов h_{iv}, минимизирующие правую часть Q_s, находятся из системы линейных алгебраических уравнений:

$$\sum_{j=1}^{s} h_{iv}[F_{(i,j)v}(H_0)+F_{(i,j)v}(H_0)]=m_{iv}-u_{iv}, \quad i=\overline{1,s};\ v=\overline{1,n}. \tag{9}$$

При этих значениях коэффициентов выполняется равенство:

$$G_o[\gamma]+G_1[\gamma]=(E_1-E_0).$$

Поэтому

$$Q_S=(E_1-E_0)^{-1}=[\sum_{i=1}^{s}\sum_{v=1}^{n}h_{iv}(m_{iv}-u_{iv})]^{-1}. \tag{10}$$

Количество извлекаемой информации J_S из выборочных значений ξ_v о различии гипотез H_1 и H_0 с критерием качества Q_S связано обратно - пропорциональным соотношением:

$$Q_S=(E_1-E_0)^{-1}=[\sum_{i=1}^{s}\sum_{v=1}^{n}h_{iv}(m_{iv}-u_{iv})]^{-1}. \tag{11}$$

Применим общие выражения для получения алгоритмов обнаружения импульсного сигнала на фоне негауссовских помех и исследуем характеристики решающих правил для $S=1,2,3$.

При $S=1$ решающее правило имеет вид:

$$\sum_{v=1}^{n}h_{1v}(\xi_v-0.5S_v)\underset{H_0}{\overset{H_1}{\gtrless}}0, \tag{12}$$

где оптимальное значение h_{1v} равно:

$$h_{1v}=\frac{S_v}{2\chi_2}.$$

Для этого коэффициента значение критерия качества равно:

$$Q_1=(\sum_{v=1}^{n}\frac{l_v^2}{2\chi_2})^{-1}=(\frac{q}{2}\sum_{v=1}^{n}l_v^2)^{-1}, \tag{13}$$

где q - отношение сигнал/шум по мощности,

l_v - значение импульса с единичной амплитудой.

Отметим, что полученное решающее правило является линейным и совпадает с оптимальным решающим правилом, полученным при гауссовской помехе.

При степени полинома $S=2$ решающее правило имеет вид:

$$\sum_{v=1}^{n}h_{1v}(\xi_v-\frac{1}{2}S_v)+\sum_{v=1}^{n}h_{2v}[(\xi_v^2-\frac{1}{2}(S_v^2+2\chi_2^2)]\underset{H_0}{\overset{H_1}{\gtrless}}0, \tag{14}$$

где оптимальные коэффициенты h_{1v} і h_{2v} находятся из решения системы линейных алгебраических уравнений:

$$\begin{cases}h_{1v}F_{(1,1)v}+h_{2v}F_{(1,2)v}=S_v,\\ h_{1v}F_{(2,1)v}+h_{2v}F_{(2,2)v}=S_v{}^2,\end{cases} \tag{15}$$

где $F_{(i,j)v}$ - совместные корреляционные моменты, которые равны суме корреляционных моментов $F_{(i,j)v}(H_0)$ и $F_{(i,j(v}(H_1)$ для гипотез H_0 та H_1 соответственно.

В данном случае:

$$F_{(1,1)v}=2x_2 \quad F_{(1,2)v}=F_{(2,1)v}=2x_3+2S_vx_2,$$

$$F_{(2,2)v}=2(x_4+2x_2^2)+4S_vx_3+4S_vx_2^2+4S_v{}^2x_2.$$

Из решения системы (15) получим выражения для оптимальных коэффициентов:

$$h_{1\nu} = \frac{\Delta_{1\nu}}{\Delta_\nu}, \quad h_{2\nu} = \frac{\Delta_{2\nu}}{\Delta_\nu}, \tag{16}$$

$$\Delta_{1\nu} = 2\chi_2^{5/2} l_\nu q^{1/2} (\gamma_4 + 2 + q l_\nu^{\ 2} + q^{1/\ 2} l_\nu \gamma_3),$$

$$\Delta_{2\nu} = -2\chi_2^2 q^{1/\ 2} l_\nu \gamma_3 \qquad \Delta_\nu = 2\chi_2^3 (\gamma_4 + 2 + q l_\nu^2 - \gamma_3^2).$$

В последних выражениях γ_3, γ_4 – коэффициенты асимметрии и эксцесса соответственно. Очевидно, что при $S = 2$ решающее правило является нелинейным, так как выборочные значения дополнительно возводятся в квадрат и суммируются с определенными коэффициентами $h_{2\nu}$. Выигрыш в точностных характеристиках решающего правила при $S = 2$ по сравнению с $S = 1$ равен:

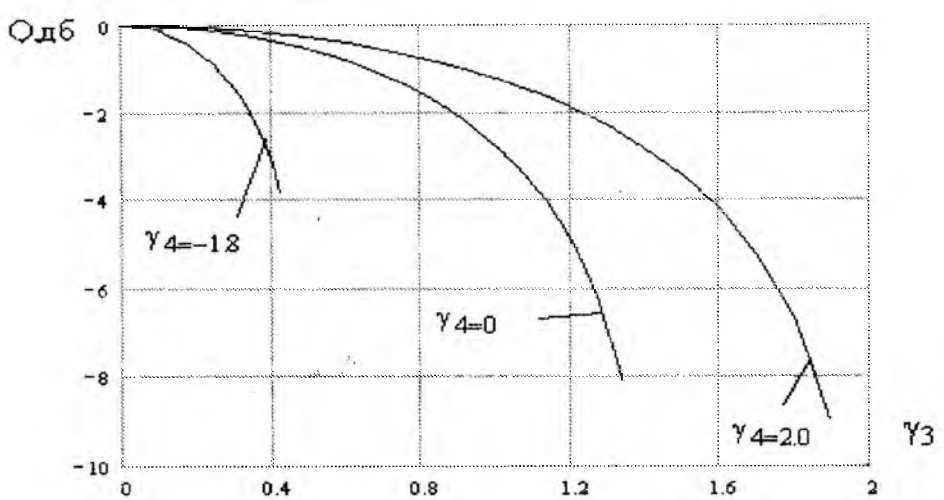
$$\frac{Q_2}{Q_1} = \frac{\sum_{\nu=1}^{n} l_\nu^2}{\sum_{\nu=1}^{n} l_\nu^2 \left(1 + \frac{\gamma_3^2}{\gamma_4 + 2 - \gamma_3^2 + q l_\nu}\right)}. \tag{17}$$

Из этого выражения видно, что решающее правило при $S = 2$ в общем случае более предпочтительно при негауссовской помехе, когда $\gamma_3 \neq 0$. График зависимости отношения (17) от коэффициента экцесса γ_3 при отношении сигал/шум $q = 0.2$ для различных значений $\gamma_4 (\gamma_4 = -1.8;\ 0;\ 2.0)$ приведен на рисунке. Из приведенного графика видим, что чем больше негауссовость помехи, т.е чем больше γ_3 отлична от нуля, тем большее количество информации о различии гипотез можно извлечь из выборочных значений.

Выигрыш имеет место как для видео-, так и радиоимпульсов и может достигать до -9 dB при $q = 0.2$. С увеличением q отношение Q_2/Q_1 уменьшается. Если $\gamma_3 = 0$, то никакого выигрыша нет.

При степени полинома $S = 3$ решающее правило имеет вид:

$$\sum_{\nu=1}^{n} h_{1\nu}(\xi_\nu - 0.5 S_\nu) + \sum_{\nu=1}^{n} h_{2\nu}[(\xi_\nu^2 - \frac{1}{2}(S_\nu^2 + 2\chi_2)] + \sum_{\nu=1}^{n} h_{3\nu}[(\xi_\nu^2 - \frac{1}{2}(S_\nu^3 + 3 S_\nu \chi_2 + 2\chi_3)] \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{18}$$



В данном случае решающее правило также будет нелинейным и выборочные значения подвергаются не только квадратичному преобразованию, но и возведению в куб и суммированию с определенными коэффициентами. Аналитические выражения для коэффициентов $h_{1\nu}, h_{2\nu}, h_{3\nu}$ имеют громоздкий вид поэтому приводить их не будем.

Рассмотрим случай, когда распределение помехи симметричное, т.е. $\gamma_3 = \gamma_5 = 0$. Для обнаружителя степени $S = 2$ количество извлекаемой информации равно количеству извлекаемой информации при $S = 1$. Для полиномиального обнаружителя степени $S = 3$ количество извлекаемой информации определяется выражением:

$$J_3 = \frac{q}{2}\sum_{\nu=1}^{n}\left[1 + \frac{(\gamma_4 + 0.5qS_\nu^{\,2})^2}{\gamma_6 - \gamma_4^{\,2} + 9\gamma_4 + 6 + \frac{9}{4}qS_\nu^{\,2}(\gamma_4 + 2)}\right]. \qquad (19)$$

Из этого выражения видно, что при $S = 3$ количество извлекаемой информации о различии гипотез в общем случае увеличивается, по сравнению со случаем $S = 1$ и $S = 2$. Это увеличение зависит от коэффициента эксцесса γ_4 и от отношения сигнал/шум q.

Анализ точностных характеристик показал, что с увеличением коэффициента эксцесса γ_4, т.е. чем больше негауссовость помехи, качество степенного обнаружителя при степени $S = 3$ по сравнению с линейным обнаружителем также увеличивается, но при малых отношениях сигал/шум ($q < 1$). Так при значении $q = 0.2$ выигрыш достигает значения 4÷5 дб.

При больших значениях q выигрыша почти никакого нет.

Список литературы: 1. *Малахов А.Н.* Кумулянтный анализ негауссовых случайных процессов и преобразований. М.: Сов. радио. 1979, 376 с. 2. *Кунченко Ю.П., Мельяновский П.А., Слюсаренко В.М.* Применение функциональных полиномов для обнаружения радиосигналов на фоне негауссовских шумов. – Харьков, 1988. – 48с. / Препринт №363. Институт радиофизики и электроники АН УССР /.

Черкасский инженерно-технологичный институт *Поступила в редколлегию 09.03.2000*

УДК 621.391

А. С. ГАВРИШ

АЛГОРИТМЫ ОЦЕНКИ ПАРАМЕТРОВ ПОЛИГАРМОНИЧЕСКОГО СИГНАЛА НА ФОНЕ НЕГАУССОВСКИХ ПОМЕХ

Результаты, полученные в работе [1] для оценки параметров гармонического сигнала на фоне негауссовских помех, позволяют говорить об их ценности для различных областей науки и техники. Очевидно, что при рассмотрении более сложной модели полезного сигнала можно значительно расширить круг решаемых с ее помощью задач. Следовательно, если в качестве сложного сигнала рассмотреть совокупность гармонических функций, то логично предположить, что оценки параметров сложного сигнала будут обладать схожими точностными характеристиками оценок параметров простого сигнала [1].

В данной работе рассматривается задача нахождения коэффициентов усеченного тригонометрического ряда Фурье (полигармонического сигнала), принимаемого на фоне аддитивной негауссовской помехи. Такая модель полезного сигнала нашла широкое применение в системах высокоскоростной передачи информации [2], а также с ее помощью можно аппроксимировать произвольный сигнал [3]. Подчеркнем, что в базис по гармоническим функциям раскладывается не случайный процесс, а только детерминированный сигнал. С физической точки зрения данную модель можно интерпретировать как наложение спектральных линий полезного сигнала на гладкий спектр широкополосной негауссовской помехи [4].

Известно, что любой периодический сигнал может быть представлен в виде тригонометрического ряда Фурье:

$$S(t) = c_0 + \sum_{q=1}^{\infty} (a_q \sin \omega_q t + b_q \cos \omega_q t), \qquad (1)$$

при этом частоты всех составляющих кратны основной частоте f_0, а именно:

$$\omega_q = 2\pi q f_0, \quad f_0 = \frac{1}{T}, \quad q = 1,2,\ldots$$

где T - период сигнала. В выражении (1) коэффициент c_0 называется постоянной составляющей, а остальные слагаемые называются гармониками.

Коэффициенты c_0, a_q и b_q вычисляются по формулам:

$$c_0 = \frac{1}{T}\int_{-T/2}^{T/2} S(t)dt, \quad a_q = \frac{2}{T}\int_{-T/2}^{T/2} S(t)\sin \omega_q t dt, \quad b_q = \frac{2}{T}\int_{-T/2}^{T/2} S(t)\cos \omega_q t dt. \qquad (2)$$

Естественно, на практике не возможно вычислить бесконечное количество коэффициентов, поэтому вместо ряда рассматривается полином. Как и в предыдущей работе [1], для упрощения задачи, рассмотрим сигнал с дискретным временем:

$$S_\nu = c_0 + \sum_{q=1}^{r} (a_q \sin \omega_q \delta\nu + b_q \cos \omega_q \delta\nu), \quad \nu = \overline{1,n}. \qquad (3)$$

где δ - период дискретизации сигнала; ν - отсчёты (моменты времени наблюдения).

В реальных каналах связи принимаемый сигнал чаще всего наблюдается в аддитивной смеси с помехой. Математической моделью такого взаимодействия может служить выражение:

$$x_\nu = S_\nu + n_\nu, \quad \nu = \overline{1,n}, \qquad (4)$$

где в качестве полезного сигнала S_ν будем рассматривать сигнал вида (3), содержащий конечную последовательность гармоник, и находить оценку векторного параметра $\vec{\vartheta} = \{c_0, a_1, \ldots a_r, b_1, \ldots b_r\}$.

В выражении (4) рассматриваемая аддитивная помеха n_ν является негауссовской случайной величиной, которая описывается конечной последовательностью кумулянтов или кумулянтных коэффициентов [5]: $Mn_\nu = 0$, $Mn_\nu^2 = \chi_2$, $Mn_\nu^3 = \chi_3$ и т.д.

На выборочные значения сигнала накладываются ограничения, которые, с одной стороны, легко реализуются в практических приложениях, а с другой стороны приводят к упрощению алгоритма нахождения оценок параметров. При этом точностные характеристики получаемых оценок параметров сохраняют свои замечательные свойства.

Будем считать, что шаг дискретизации δ выбран так, что для любой нечетной степени k справедливы соотношения:

$$\sum_{\nu=1}^{n}\sin_\nu^k = 0, \quad \sum_{\nu=1}^{n}\cos_\nu^k = 0, \tag{5}$$

а для четных степеней:

$$\sum_{\nu=1}^{n}\sin_\nu^2 = \sum_{\nu=1}^{n}\cos_\nu^2 = \frac{n}{2}, \quad \sum_{\nu=1}^{n}\sin_\nu^4 = \sum_{\nu=1}^{n}\cos_\nu^4 = \frac{3}{8}n. \tag{6}$$

Также будем считать, что шаг дискретизации выбран так, что кроме равенств (5), (6) для любых $p,k = 2,3,\ldots$ выполняются еще и следующие равенства:

$$\sum_{\nu=1}^{n}\sin_\nu p\cos_\nu k = 0, \quad \sum_{\nu=1}^{n}\sin_\nu p\sin_\nu k = 0, \quad p \neq k, \quad \sum_{\nu=1}^{n}\cos_\nu p\cos_\nu k = 0, \quad p \neq k. \tag{7}$$

Условия (5) - (7) обеспечивают ортогональность составляющих полигармонического сигнала.

В (5) - (7) в целях сокращения записи опущены аргументы у синуса и косинуса и сделаны следующие обозначения:

$$\sin_\nu = \sin\omega\delta\nu, \qquad \cos_\nu = \cos\omega\delta\nu,$$

$$\sin_\nu p = \sin p\omega\delta\nu, \qquad \cos_\nu k = \cos k\omega\delta\nu, \qquad p,k = 2,3,\ldots$$

Эти обозначения будем использовать и в дальнейшем.

Далее, по заданным значениям выборки $\vec{X} = \{x_1, x_2, \ldots x_n\}$, используя метод максимизации полинома [5], будут синтезированы алгоритмы совместной оценки параметров полигармонического сигнала $\vec{\vartheta} = \{c_0, a_1, \ldots a_r, b_1, \ldots b_r\}$ при воздействии негауссовской помехи, статистические характеристики которой известны точно.

Синтезируем линейный алгоритм измерения параметров полигармонического сигнала при воздействии помех. Согласно метода максимизации полинома оценка искомого векторного параметра при $s = 1$ находится из совместного решения $(2r+1)$ уравнений вида:

$$\sum_{\nu=1}^{n} k_{1\nu}^{(m)}(\vec{\vartheta})[x_\nu - S_\nu(\vec{\vartheta})]\big|_{\vec{\vartheta}=\hat{\vec{\vartheta}}} = 0, \quad m = \overline{1,p}, \tag{8}$$

где $p = 2r+1$ - размерность оцениваемого векторного параметра.

В каждом m-ом уравнении (8) коэффициент $k_{1\nu}^{(m)}(\vec{\vartheta})$ находится из решения линейного алгебраического уравнения [5] и соответственно равен:

$$k_{1\nu}^{(m)}(\vec{\vartheta}) = \frac{\overline{\frac{\partial}{\partial\vartheta_m}S_\nu}}{\chi_2} = A_\nu\chi_2^{-1}, \tag{9}$$

где

$$A_\nu = \begin{cases} m, & при \quad m = 1, \\ \sin_\nu(m-1), & при \quad m = \overline{2,(r+1)}, \\ \cos_\nu(m-r-1), & при \quad m = \overline{(r+2),p}. \end{cases} \tag{10}$$

Далее, весовые коэффициенты $k_{1\nu}^{(m)}(\vec{\vartheta})$ вида (9) подставляются в систему уравнений максимизации полинома (8). Легко показать, что ввиду ограничений (5)-(7) каждое получаемое уравнение зависит только от одного оцениваемого параметра, а конечные выражения для оценок параметров полигармонического сигнала имеют вид:

$$\hat{c}_0 = \frac{1}{n}\sum_{\nu=1}^{n} x_\nu, \quad \hat{a}_q = \frac{2}{n}\sum_{\nu=1}^{n} x_\nu \sin_\nu q, \quad \hat{b}_q = \frac{2}{n}\sum_{\nu=1}^{n} x_\nu \cos_\nu q, \quad q = \overline{1,r}. \tag{11}$$

При сравнении выражений (11) и (2) легко усматривается их тождественность. Таким образом, оценки параметров полигармонического сигнала на фоне помех, найденные методом максимизации полинома при $s = 1$, совпадают с коэффициентами детерминированного тригонометрического полинома.

Рассмотрим случай, когда оценка векторного параметра $\vec{\vartheta} = \{c_0, a_1, \ldots a_r, b_1, \ldots b_r\}$ находится из решения системы $(2r+1)$ уравнений максимизации полинома при $s = 2$:

$$\sum_{\nu=1}^{n} k_{1\nu}^{(m)}(\vec{\vartheta})[x_\nu - S_\nu(\vec{\vartheta})] + \sum_{\nu=1}^{n} k_{2\nu}^{(m)}(\vec{\vartheta})[x_\nu^2 - S_\nu^2(\vec{\vartheta}) - \chi_2]\big|_{\vec{\vartheta}=\hat{\vec{\vartheta}}} = 0, \quad m = \overline{1,p}, \tag{12}$$

где каждая пара коэффициентов $k_{1\nu}^{(m)}(\vec{\vartheta})$ и $k_{2\nu}^{(m)}(\vec{\vartheta})$ соответствующего m-го уравнения находится из решения системы двух линейных алгебраических уравнений вида:

$$\begin{cases} k_{1\nu}^{(m)}(\vec{\vartheta})F_{(1,1)\nu}(\vec{\vartheta}) + k_{2\nu}^{(m)}(\vec{\vartheta})F_{(1,2)\nu}(\vec{\vartheta}) = A_\nu, \\ k_{1\nu}^{(m)}(\vec{\vartheta})F_{(1,2)\nu}(\vec{\vartheta}) + k_{2\nu}^{(m)}(\vec{\vartheta})F_{(2,2)\nu}(\vec{\vartheta}) = 2A_\nu S_\nu, \end{cases} \tag{13}$$

Из анализа выражений (10) и (13) видно, что левая часть системы уравнений (13) будет сохранять свой вид вне зависимости от того для какого из $(2r+1)$ уравнений находятся коэффициенты $k_{1\nu}^{(m)}(\vec{\vartheta})$ и $k_{2\nu}^{(m)}(\vec{\vartheta})$, в то время как выбор функций в правой части искомой системы уравнений четко зависит от номера уравнения максимизации полинома, для которого находятся соответствующие коэффициенты.

Используя правило Крамера легко показать, что решением системы уравнений (13) будут коэффициенты:

$$\mathrm{k}_{1\nu}^{(m)}(\vec{\vartheta}) = \frac{\mathrm{A}_\nu}{\Delta_2}\chi_2^2(\gamma_4 + 2\mathrm{S}_\nu\gamma_3\chi_2^{-0,5} + 2), \quad \mathrm{k}_{2\nu}^{(m)}(\vec{\vartheta}) = -\frac{\mathrm{A}_\nu}{\Delta_2}\chi_2^{1,5}\gamma_3, \quad \mathrm{m} = \overline{1,\mathrm{p}}, \tag{14}$$

где главный определитель системы уравнений (13) имеет вид:

$$\Delta_2 = \chi_2^3(\gamma_4 + 2 - \gamma_3^2).$$

При подстановке коэффициентов (14) в систему уравнений (12), и учитывая соотношения (5)-(7), после несложных алгебраических преобразований легко получить конечные выражения уравнений максимизации полинома для нахождения совместной оценки параметров полигармонического сигнала, принимаемого на фоне негауссовских помех. Для наглядности и удобства записи, полученные уравнения для оценки векторного параметра $\vec{\vartheta} = \{c_0, a_1, \ldots a_r, b_1, \ldots b_r\}$ целесообразно разбить на группы. Классификация уравнений осуществляется по виду подставляемых в них коэффициентов (14) и, в конечном счете, зависит от функции A_ν вида (10). Из этого следует, что искомые уравнения можно представить в виде трех выражений.

В случае если максимизация полинома осуществляется по компоненте c_0, то первое уравнение системы имеет вид:

$$c_0^2 + 0{,}5\sum_{q=1}^{r}(a_q^2 + b_q^2) + c_0\left(\alpha - \frac{2}{n}\sum_{\nu=1}^{n} x_\nu\right) - \frac{2}{n}\sum_{\nu=1}^{n} x_\nu \sum_{q=1}^{r}(a_q \sin_\nu q + b_q \cos_\nu q) + \\ + \frac{1}{n}\sum_{\nu=1}^{n} x_\nu^2 - \alpha\frac{1}{n}\sum_{\nu=1}^{n} x_\nu - \chi_2\big|_{\vec{\vartheta}=\hat{\vec{\vartheta}}} = 0, \tag{15}$$

где $\alpha = \chi_2^{0,5}\gamma_3^{-1}(\gamma_4 + 2)$.

Вторую группу составляют r уравнений максимизации полинома, в каждом из которых отыскивается экстремум полинома для соответствующей компоненты a_m, $m = \overline{1,r}$:

$$a_m c_0 + 0{,}5\sum_{q=1}^{m-1} a_q b_{m-q} + 0{,}5\sum_{q=1}^{r-m}(a_{q+m}b_q - a_q b_{q+m}) + 0{,}5\alpha a_m - \frac{2}{n}\sum_{\nu=1}^{n} x_\nu \sin_\nu m \times$$

$$\times[c_0 + \sum_{q=1}^{r}(a_q \sin_\nu q + b_q \cos_\nu q)] + \frac{1}{n}\sum_{\nu=1}^{n} x_\nu^2 \sin_\nu m - \alpha\frac{1}{n}\sum_{\nu=1}^{n} x_\nu \sin_\nu m\big|_{\vartheta=\hat{\vartheta}} = 0, \quad q = 1,2,... \tag{16}$$

Оставшиеся r уравнений будут относиться к третьей группе, среди которых каждое m-ое можно представить в виде выражения:

$$b_m c_0 + 0{,}5\sum_{q=1}^{[M]}(b_q b_{m-q} - a_q a_{m-q}) + 0{,}5(M - [M])(b_{m/2}^2 - a_{m/2}^2) +$$

$$+ 0{,}5\sum_{q=1}^{r-m}(a_q a_{q+m} + b_q b_{q+m}) + 0{,}5\alpha b_m - \frac{2}{n}\sum_{\nu=1}^{n} x_\nu \cos_\nu m \times \tag{17}$$

$$\times[c_0 + \sum_{q=1}^{r}(a_q \sin_\nu q + b_q \cos_\nu q)] + \frac{1}{n}\sum_{\nu=1}^{n} x_\nu^2 \cos_\nu m - \alpha\frac{1}{n}\sum_{\nu=1}^{n} x_\nu \cos_\nu m\big|_{\vartheta=\hat{\vartheta}} = 0,$$

где $M = \dfrac{m-1}{2}$, $m = \overline{1,r}$, $q = 1,2,...$

В выражении (17) используется обозначение [M], указывающее на то, что рассматривается антье (целая часть) от числа M.

Легко заметить, каждое из уравнений вида (15)-(17) зависит от всех компонент исследуемого векторного параметра и является нелинейным относительно оцениваемых параметров. Таким образом, искомые оценки $\{\hat{c}_0, \hat{a}_1, ... \hat{a}_r, \hat{b}_1, ... \hat{b}_r\}$ находятся из совместного решения $(2r+1)$ уравнений (15)-(17). Ввиду нелинейности уравнений не представляется возможным записать выражения оценок исследуемых параметров в явном виде, поэтому для решения искомой системы уравнений необходимо использовать численные методы.

Таким образом, оценки, найденные методом максимизации полинома при $s = 2$ являются новыми и существенно отличаются от оценок (2).

Список литературы: **1.** *Кунченко Юрій, Гавриш Олександр.* Оцінка параметрів гармонічного коливання при негауссівських завадах. // Праці 3-ої міжнародної конференції «Оброблення сигналів і зображень та розпізнавання образів». Київ. 1996. С. 57-60. **2.** *Зяблов В.В., Коробков Д.Л., Портной С.Л.* Высокоскоростная передача сообщений в реальных каналах. М.: Радио и связь, 1991. 228 с. **3.** *Бабак В.П., Хандецький В.С., Шрюфер Е.* Обробка сигналів: Підручник. - К.: Либідь, 1996. 392 с. **4.** *Zhou G., Giannakis G.B.* Polyspectral Analysis of Mixed Processes and Coupled Harmonics. // IEEE Trans. Inform. Theory. 1996. Vol. 42, no.3. P. 943-958. **5.** *Кунченко Ю.П., Лега Ю.Г.* Оценка параметров случайных величин методом максимизации полинома. К.: Наукова думка, 1992. 180 с.

Черкасский инженерно-технологический институт *Поступила в редколлегию 16.12.1999*

УДК 621.391

В.А. ПИСЬМЕНЕЦКИЙ, канд. техн. наук, А.В. БОРОДИН, П.И. ПЛАТОНОВ

АНАЛИЗ ИНФОРМАТИВНЫХ ПРИЗНАКОВ ДЛЯ ИДЕНТИФИКАЦИИ ИМПУЛЬСНЫХ СИГНАЛОВ ПО СПЕКТРАЛЬНЫМ ПРИЗНАКАМ

Важным этапом алгоритмов распознавания является выделение минимума информативных параметров и последующее сжатие объема обрабатываемой информации. Как известно, переход в спектральную область, основанный на интегрирующем преобразовании исследуемого процесса, обеспечивает сжатие объема информации и, соответственно, времени обработки. При этом сжатие объема информации возрастает с уменьшением эффективной ширины спектра радиосигналов, что открывает возможности обработки в квазиреальном времени.

Набор информативных параметров определяется многими факторами, среди которых наиболее существенны метод и погрешность преобразования. Отметим, что задача выделения минимального количества информативных параметров при сохранении максимума количества информации об анализируемом сигнале для априорно неизвестных моноимпульсных сигналов на фоне непериодических помех в настоящее время не нашла окончательного однозначного решения. Вместе с тем, большинство разработанных алгоритмов анализа сигналов основано на применении численных методов, которые реализуются цифровой фильтрацией на аппаратном или программном уровнях. В данной работе распознаваемые сигналы представляют собой радиоимпульсы с различной амплитудной огибающей, в частности, будем рассматривать наиболее часто применяемые импульсные радиосигналы с огибающей типа функции Хэмминга и колоколообразной огибающей $e^{-x^2/2a^2}$. Сумма их мгновенных спектров S(ω) представляет монотонную зависимость с несколькими максимумами.

Традиционно идентификация по спектральным признакам строится на базе отсчетов огибающей S(ω) мгновенных спектров. Однако, простое выделение отсчетов спектра как информативных параметров не всегда сопровождается их устойчивостью при случайных изменениях амплитудных значений сигналов.

Анализ показывает, что более надежными и более устойчивыми признаками являются:

- частоты максимумов спектра f_m;
- ширина спектра на различных относительных уровнях Δf_n, где n - уровень отсчета (0,7; 0,5);
- коэффициент прямоугольности K_n огибающей спектра $K_{n_{NM}} = \Delta f_N / \Delta f_M$, где N, M - уровни отсчета в пределах $N \in (0{,}5..0{,}8)$, $M \in (0{,}001..0{,}01)$.

Перечисленные признаки можно представить как множества: $x_{1_m} \in (f_m); x_{2_m} \in (\Delta f_n)$; $x_{3_m} \in (K_{n_{NM}})$.

Таким образом общее количество признаков может быть достаточно большим, так как

$x_{1_m} \in (x_1; x_2 ... x_l)$, где l - количество максимумов спектра;

$x_{2_m} \in (x_1; x_2 ... x_j)$, где j - количество уровней отсчета ширины спектра;

$x_{2_m} \in (x_1; x_2 ... x_i)$, где i - количество уровней, по которым вычисляется коэффициент прямоугольности $K_{n_{NM}}$.

Спектр одного сигнала имеет один максимум и гладкую огибающую и, следовательно, первое множество признаков будет иметь только одно значение, то есть $x_{1m} = x_1$.

Из второго и третьего множеств признаков первоначально будем использовать по одному признаку и, соответственно, задача распознавания заключается в выделении сигналов одного класса Ω_1 или Ω_2, ($\Omega_1(x_1, x_2, x_3)$ - класс радиосигналов с огибающей типа функции Хэмминга, а $\Omega_2(x_1, x_2, x_3)$ - класс радиосигналов с колоколообразной огибающей), когда они одновременно поступают на устройство распознавания.

Как известно, если вероятность принятия правильного решения окажется ниже заданной, количество признаков X нужно увеличить.

Поскольку анализируемые радиосигналы представляют короткие радиоимпульсы, для формирования их мгновенных спектров применяется дисперсионный Фурье-процессор. Его отклик при достаточно большом числе отсчетов мгновенного спектра представляет преобразование Фурье входного воздействия. Соответственно на выходе дисперсионного Фурье процессора получим

$$|S(\omega)| = \sqrt{\left|S_1^2(\omega,\omega_{0_1},T_1)\right| + \left|S_2^2(\omega,\omega_{0_2},T_2,\tau)\right| + 2\cdot S_2(\omega,\omega_{0_1},T_1)\cdot S_2(\omega,\omega_{0_2},T_2,\tau)\cdot\cos(\varphi)}, \quad (1)$$

где $S_1\left(\omega,\omega_{0_1},T_1\right) = 0{,}54A_1\,\mathrm{sinc}\left(\frac{\omega-\omega_{0_1}}{2}T_1\right) + 0{,}23A_1\,\mathrm{sinc}\left(\frac{\omega-\omega_{0_1}}{2}T_1+\pi\right) +$

$$+ 0{,}23A_1\,\mathrm{sinc}\left(\frac{\omega-\omega_{0_1}}{2}T_1-\pi\right), \quad (2)$$

– мгновенный спектр радиоимпульса с огибающей типа функции Хэмминга,

$$S_2\left(\omega,\omega_{0_2},T_2\right) = \sqrt{2\pi}A_2 a\cdot e^{-\frac{(\omega-\omega_{0_2})^2}{2}\cdot a^2}, \quad (3)$$

где $a = T_2\Big/\left(2\sqrt{2\ln\frac{1}{g}}\right)$; $g = 0.01, 0.005, 0.001$. – мгновенный спектр колоколообразного радиоимпульса, $\varphi = \omega\tau$ – фазовый сдвиг с учетом временного запаздывания второго сигнала относительно первого.

Приведенное соотношение (1) учитывает как различие радиоимпульсов по длительности, так и временное запаздывание относительно друг друга.

Спектральные характеристики рассматриваемых сигналов классов Ω_1 и Ω_2 без временного запаздывания приведены на рис.1 (кривая 1), где S_m – нормированный суммарный спектр. Влияние временного запаздывания на спектр суммарного сигнала для случаев относительного запаздывания $\tau = 2T_1$ иллюстрируется кривой 2 на рис.1.

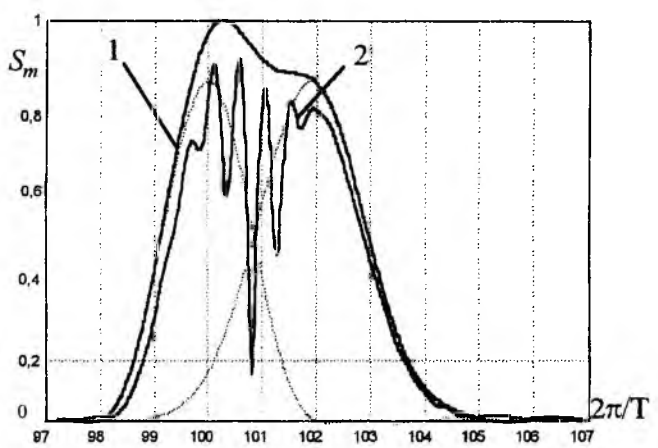


Рис. 1

Параметры реальных сигналов описываются стохастическими зависимостями. Для исследования вероятностных характеристик предложенных информативных признаков будем предполагать, что распределение плотности вероятностей $P_2(x_1)$ центральной частоты ω_{0_2} радиоимпульса класса Ω_2 описывается нормальным законом: $\omega_{0_2} = \omega_{0_1} + \Delta + \sigma\cdot Rnd$, где $\Delta = \left(\omega_{0_2} - \omega_{0_1}\right)$ - смещение начального положения максимума ω_{0_2} спектра радиосигнала класса Ω_2 относительно максимума спектра радиосигнала класса Ω_1; *Rnd* - случайное число с нормальным распределением; σ - дисперсия.

С учетом этого определялись плотности вероятностей трех информативных признаков для радиоимпульса класса Ω_1.

Для общего случая принимаем равные вероятности появления обоих классов сигналов: $P(\Omega_1) = P(\Omega_2) = 0{,}5$.

Рассмотрим первый признак x_1 (центральную частоту), начальное значение которого для класса Ω_1 равно $\omega_{0_1} = 100 \cdot 2\pi/T$ и, соответственно для класса Ω_2 — $\omega_{0_2} = 101 \cdot 2\pi/T$.

Исследования плотности вероятности признака $P_1(x_1)$ выполнялись при следующих условиях:

- имеет место влияние фазы каждого сигнала на суммарный сигнал (рис.1);
- расстояние между максимумами спектров изменяется в пределах $(\{0{,}5...1{,}5\} \cdot 2\pi/T)$;
- имеется различие амплитуд сигналов ($A_1/A_2 = \{0{,}001 ... 1; 1 ... 1000\}$);
- имеется различие длительностей сигналов ($T_1 = \{0{,}5...1{,}5\} \cdot T_2$);
- изменение дисперсии x_1 одного из сигналов равно ($\{0{,}2...1\} \cdot 2\pi/T$).

На рис.2, (кривая 1) изображен график плотности вероятности центральной частоты $P_1(x_1)$ без временного запаздывания ($\tau = 0$) при одинаковых амплитудах основного сигнала и сигнала-помехи. При вводе временного запаздывания сигнала-помехи относительно основного сигнала ($\tau = 2T_1$) форма графика плотности вероятности изменяется (рис. 2, кривая 2) из-за появления осцилляций, характерных для спектра сигнала, представляющего собой сумму сигналов с временным запаздыванием (рис.1, кривая 2).

При увеличении дисперсии x_1 сигнала-помехи до величины $2\pi/T$ на графике $P_1(x_1)$ (рис.2, кривая 3) наблюдается появление всплесков. Это объясняется совмещением спектров сигнала-помехи и основного сигнала или значительным их удалением, что вносит некоторые погрешности при измерении центральной частоты суммарного сигнала.

При уменьшении расстояния между максимумами спектров основного сигнала и сигнала-помехи плотность вероятности напоминает нормальный закон распределения, так как в этом случае сигнал-помеха, фактически перекрывает основной сигнал. В случае же, когда сигнал-помеха удален от основного сигнала, график плотности вероятности имеет один всплеск

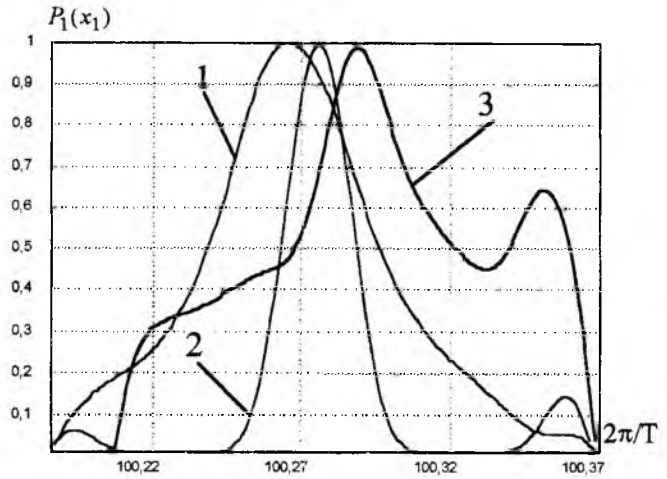


Рис. 2

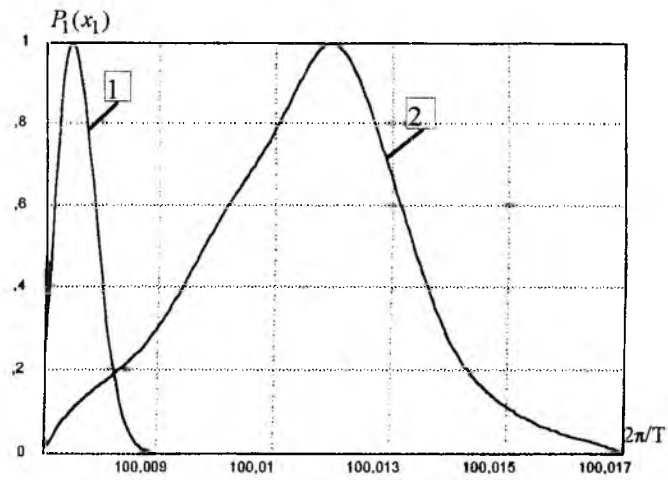


Рис. 3

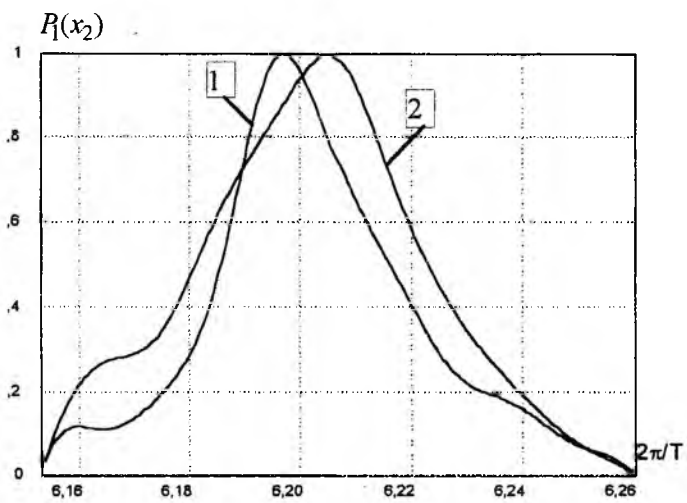


Рис. 4

с очень узкой формой из-за того, что сигнал-помеха фактически не влияет на основной сигнал. Если амплитуда основного сигнала существенно превышает амплитуду сигнала-помехи ($D = A_1/A_2 = 50$), то максимум $P_1(x_1)$ смещается в сторону максимума спектра основного сигнала, что иллюстрируется на рис.3 (кривая 1).

Когда же амплитуда сигнала-помехи превышает амплитуду основного сигнала ($D = A_1/A_2 = 1/50$), $P_1(x_1)$ близка к нормальному закону распределения и ее максимум смещается в сторону максимума спектра сигнала-помехи (рис.3, кривая 2).

При уменьшении, как и при увеличении длительности основного сигнала, плотность вероятности его центральной частоты фактически не изменяется; при уменьшении длительности сигнала-помехи максимум плотности вероятности центральной частоты суммарного сигнала смещается в сторону максимума спектра основного сигнала, так как уменьшается влияние сигнала-помехи на основной сигнал; в случае же увеличения длительности сигнала-помехи на графике плотности вероятности появляются всплески.

При рассмотрении второго признака x_2 (ширины спектра) учитывались начальные условия для признака x_1. На рис. 4, кривая 1 изображен график $P_1(x_2)$ без учета временного запаздывания. При вводе временного запаздывания форма огибающей фактически не изменяется, однако максимальное значение смещается в сторону увеличения (рис. 4, кривая 2). Та же картина наблюдается при увеличении расстояния между максимумами спектров основного сигнала и сигнала-помехи и увеличении дисперсии центральной частоты сигнала-помехи, а также при изменении отношения амплитуд сигналов.

Анализ третьего признака x_1 (коэффициента прямоугольности) показывает, что он устойчив как к изменению длительности сигналов (рис. 5), так и других параметров сигналов, рассмотренных при анализе первых двух признаков, кроме того он инвариантен по отношению к изменению полосы пропускания и центральной частоты.

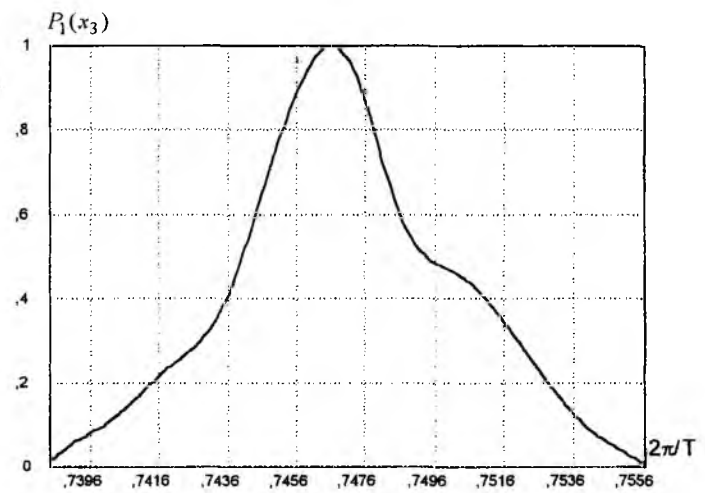


Рис. 5

Таким образом при решении задачи распознавания импульсных радиосигналов можно использовать такие признаки как: центральная частота, полоса пропускания и коэффициент прямоугольности спектральной характеристики сигнала.

Список литературы: **1.** *Измерение* несущей частоты наносекундных импульсов спектральным методом / *Письменецкий В.А., Бородин А.В., Платонов П.И.*: Харьк. техн. ун-т радиоэлектроники. - Харьков, 1996. - 10 с.: ил.- Библиогр.: 2 назв. - Рус. Деп. в ГНТБ Украины. **2.** *Горелик А.Л. Скрипкин В.А.* Методы распознавания. Москва. Высш. шк. 1989. 232с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.02.2000

УДК 621.327:519.216

С. М. ПЕРВУНІНСЬКИЙ, канд. техн. наук, Р. М. ДІДКОВСЬКИЙ

ВИЗНАЧЕННЯ ДВОВИМІРНИХ КУМУЛЯНТНИХ ФУНКЦІЙ СИГНАЛІВ З ФАЗОВОЮ МОДУЛЯЦІЄЮ ТА ЇХ СПЕКТРІВ

В задачах, пов'язаних з аналізом нелінійних перетворень модульованих високочастотних сигналів, часто виявляється недостатньою інформація про сигнали, що міститься в їх кореляційних функціях та енергетичних спектрах. Більш повна інформація про сигнал у вигляді моментних і кумулянтних функцій вищого порядку необхідна також в задачах аналізу і синтезу нелінійних систем обробки сигналів, що приймаються в шумах негауссового типу. Задача визначення моментних функцій вищих порядків для фазомодульованих дискретних сигналів розглядалась в [1]. В даній роботі досліджується неперервний випадок. Методика, яка застосовується для вирішення поставленої задачі, дозволяє отримати вирази, використання яких є більш простим при розв'язанні практичних задач.

Сигнал з фазовою модуляцією (ФМ) опишемо виразом

$$s(t) = A_0 \cos\left(\omega_0 t + \varphi_0 + \Delta\varphi\lambda(t)\right), \tag{1}$$

де A_0, ω_0, φ_0 – відповідно амплітуда, частота і початкова фаза сигналу; $\Delta\varphi$ – сталий коефіцієнт; $\lambda(t)$ – модулююча функція.

Визначимо моментні функції $p+q$-го порядку виразом

$$\begin{aligned} m_{p,q} &= m_{p,q}\left(t_1^{[p]}, t_2^{[q]}\right) = m_{p,q}\left(\underbrace{t_1,\ldots,t_1}_{p};\underbrace{t_2,\ldots,t_2}_{q}\right) = M\left(s^p(t_1)\cdot s^q(t_2)\right) = \\ &= A^{p+q} M\left(\cos^p\left(\omega_1+\varphi+\lambda_1\right)\cdot\cos^q\left(\omega_2+\varphi+\lambda_2\right)\right), \end{aligned} \tag{2}$$

де M – символ математичного сподівання, $A = A_0$, $\varphi = \varphi_0$; $\omega_i = \omega_0 t_i$, $\lambda_i = \Delta\varphi\lambda(t_i)$, $i = 1,2$.

Розглянемо випадки: 1) p; q – парні, $p \le q$; 2) p; q – непарні, $p \le q$; 3) p – парне, q – непарне.

1) Випадок парних p і q. Нехай $p = 2k$, $q = 2n$, $k, n \in Z_+$, $k \le n$. З (2) маємо:

$$m_{2k,2n} = A^{2(k+n)} M\left(\cos^{2k} g_1 \cdot \cos^{2n} g_2\right), \tag{3}$$

де $g_i = \omega_i + \varphi + \lambda_i$, $i = 1,2$. З (3), враховуючи, що [2]

$$\cos^{2n} x = \frac{1}{2^{2n}}\left(2\sum_{i=0}^{n-1} C_{2n}^i \cos 2(n-i)x + C_{2n}^n\right), \tag{4}$$

отримаємо

$$\begin{aligned} m_{2k,2n} = \left(\frac{A}{2}\right)^{2(k+n)} M\Bigg(\left(2\sum_{i=0}^{k-1} C_{2k}^i \cos 2(k-i)g_1 + C_{2k}^k\right)\times \\ \times\left(2\sum_{i=0}^{n-1} C_{2n}^i \cos 2(n-i)g_2 + C_{2n}^n\right)\Bigg). \end{aligned}$$

Після розкриття внутрішніх дужок будемо мати наступне:

$$\begin{aligned} m_{2k,2n} = \left(\frac{A}{2}\right)^{2(k+n)} M\Bigg(C_{2k}^k C_{2n}^n + 2C_{2k}^k \sum_{i=0}^{n-1} C_{2n}^i \cos 2(n-i)g_2 + \\ + 2C_{2n}^n \sum_{i=0}^{k-1} C_{2k}^i \cos 2(k-i)g_1 + \\ + 4\sum_{i=0}^{k-1} C_{2k}^i \cos 2(k-i)g_1 \sum_{i=0}^{n-1} C_{2n}^i \cos 2(n-i)g_2\Bigg). \end{aligned}$$

При обчисленні математичного сподівання по φ другий і третій доданки дають нуль, тому виконавши множення рядів у четвертому доданку, можемо записати, що

$$m_{2k,2n} = \frac{A}{2}^{2(k+n)} \left(C_{2k}^{k} C_{2n}^{n} + \right.$$

$$\left. + M\ 4 \sum_{i=0}^{k-1} \sum_{j=0}^{n-1} C_{2k}^{i} \cos 2(k - i) g_1 \quad C_{2n}^{j} \cos 2(n - j) g_2 \right) .$$

Шляхом перетворення добутку косинусів на суму отримаємо:

$$m_{2k,2n} = \left(\frac{A}{2}\right)^{2(k+n)} \left(C_{2k}^{k} C_{2n}^{n} + 4M \left(\sum_{i=0}^{k-1} \sum_{j=0}^{n-1} C_{2k}^{i} C_{2n}^{j} \times \right. \right.$$

$$\times \frac{1}{2} \left(\cos 2\left((n - j - k + i)\varphi + (n - j) f_2 - (k - i) f\right) + \right.$$

$$\left. \left. \left. + \cos 2\left((n - j + k - i)\varphi + (n - j) f_2 + (k - i) f_1\right)\right)\right)\right),$$

де $f_i = \omega_i + \lambda_i,\ i = 1,2$.

З урахуванням того, що тільки ті доданки залишаються відмінними від нуля при інтегруванні по φ, в яких коефіцієнт біля φ дорівнює нулю, можемо записати:

$$m_{2k,2n} = \left(\frac{A}{2}\right)^{2(k+n)} \left(C_{2k}^{k} C_{2n}^{n} + \right.$$

$$\left. + 2M \left(\sum_{i=0}^{k-1} C_{2k}^{i} C_{2n}^{n-k+i} \cos 2(k - i)(f_2 - f_1) \right) \right).$$

Перегрупувавши доданки і використовуючи формулу косинуса суми, отримаємо:

$$m_{2k,2n} = \left(\frac{A}{2}\right)^{2(k+n)} \left(C_{2k}^{k} C_{2n}^{n} + 2M \left(\sum_{i=0}^{k-1} C_{2k}^{i} C_{2n}^{n-k+i} \times \right. \right.$$

$$\left. \left. \times \left(\cos 2(k - i)\omega_\tau \cos 2(k - i)\lambda_\tau - \sin 2(k - i)\omega_\tau \sin 2(k - i)\lambda_\tau \right) \right) \right) =$$

$$= \left(\frac{A}{2}\right)^{2(k+n)} \left(C_{2k}^{k} C_{2n}^{n} + 2 \left(\sum_{i=0}^{k-1} C_{2k}^{i} C_{2n}^{n-k+i} \times \right. \right.$$

$$\left. \left. \left. \times \left(\cos 2(k - i)\omega_\tau M\left(\cos 2(k - i)\lambda_\tau\right) - \sin 2(k - i)\omega_\tau M\left(\sin 2(k - i)\lambda_\tau\right) \right) \right) \right),$$

де $\omega_\tau = \omega_2 - \omega_1,\ \lambda_\tau = \lambda_2 - \lambda_1$.

Розглянемо функцію $\psi_c(l) = M\left(\cos l\lambda_\tau\right)$. Враховуючи, що $\cos x = \sum_{i=0}^{\infty} (-1)^i \frac{x^{2i}}{(2i)!}$, маємо:

$\psi_c(l) = \sum_{j=0}^{\infty} (-1)^j \frac{l^{2j} M(\lambda_\tau)^{2j}}{(2j)!}$. Розклавши $\lambda_\tau^{2j} = (\lambda_2 - \lambda_1)^{2j}$ за формулою бінома Ньютона та внісши знак математичного сподівання в середину дужок матимемо:

$$\psi_c(l)=\sum_{j=0}^{\infty}\left((-1)^j\frac{(l\Delta\varphi)^{2j}}{(2j)!}\sum_{\nu=0}^{2j}(-1)^{\nu}C_{2j}^{\nu}m_{\nu,2j-\nu}^{(\lambda)}\left(t_1^{[\nu]},t_2^{[2j-\nu]}\right)\right).$$

Аналогічно отримаємо:

$$\psi_s(l)=M(\sin l\lambda_\tau)=$$

$$=\sum_{j=1}^{\infty}\left((-1)^{j+1}\frac{(l\Delta\varphi)^{2j-1}}{(2j-1)!}\sum_{\nu=0}^{2j-1}(-1)^{\nu}C_{2j}^{\nu}m_{\nu,2j-\nu}^{(\lambda)}\left(t_1^{[\nu]},t_2^{[2j-\nu]}\right)\right).$$

Введемо позначення $\Psi(l)=\cos l\omega_\tau\psi_c(l)-\sin l\omega_\tau\psi_s(l)$.

Отже, $m_{2k,2n}=\left(\frac{A}{2}\right)^{2(k+n)}\left(C_{2k}^{k}C_{2n}^{n}+2\left(\sum_{i=0}^{k-1}C_{2k}^{i}C_{2n}^{n-k+i}\Psi(2(k-i))\right)\right)$.

2) Випадок непарних p і q. Нехай $p=2k-1$, $q=2n-1$, $k,n\in N$, $k\le n$. З (2) з урахуванням ого, що [2]

$$\cos^{2n-1}x=\frac{1}{2^{2n-1}}\sum_{i=0}^{n-1}C_{2n}^{i}\cos 2(n-i)x, \tag{5}$$

маємо аналогічно:

$$m_{2k-1,2n-1}=2\left(\frac{A}{2}\right)^{2(k+n-1)}\sum_{i=0}^{k-1}C_{2k-1}^{i}C_{2n-1}^{n-k+i}\Psi(2k-2i-1).$$

3) Випадок непарного p і парного q. Нехай $p=2k-1$, $q=2n$, $n\in Z_+$, $k\in N$. Тоді з (2), раховуючи (4) і (5), отримаємо:

$$m_{2k-1,2n}=2\left(\frac{A}{2}\right)^{2(k+n-1)}M\left(\sum_{i=0}^{k-1}\sum_{j=0}^{n-1}C_{2k-1}^{i}C_{2n}^{j}\times\right.$$

$$\times\left(\cos\left((2n-2j-2k+2i+1)\varphi+(2n-2j)f_2-(2k-2i-1)f_1\right)+\right.$$

$$\left.\left.+\cos\left((2n-2j+2k-2i-1)\varphi+(2n-2j)f_2+(2k-2i-1)f_1\right)\right)\right).$$

Коефіцієнти біля φ в даному виразі не дорівнюють нулю при будь-яких цілих значеннях k, n, i, j. Тому при інтегруванні по φ в результаті отримуємо нуль.

Використовуючи відомі [3] формули зв'язку кумулянтних і моментних функцій, можна отримати начення також кумулянтних функцій даного сигналу. Наведемо значення перших кумулянтних ункцій до 8-го порядку включно (враховуючи, що перестановка індексів не змінює значення умулянта):

$$\kappa_{0,1}=0;\ \kappa_{0,2}=\frac{A^2}{2};\ \kappa_{0,3}=0;\ \kappa_{0,4}=-\frac{3A^4}{8};\ \kappa_{0,5}=0;\ \kappa_{0,6}=\frac{5A^6}{4};$$

$$\kappa_{0,7}=0;\ \kappa_{0,8}=-\frac{1155A^8}{128};\ \ldots$$

$$\kappa_{1,1}=\frac{A^2}{2}\Psi(1);\ \kappa_{1,2}=0;\ \kappa_{1,3}=-\frac{3A^4}{8}\Psi(1);\ \kappa_{1,4}=0;\ \kappa_{1,5}=\frac{5A^6}{4}\Psi(1);$$

$$\kappa_{1,6}=0;\ \kappa_{1,7}=-\frac{1155A^8}{128}\Psi(1);\ \ldots \tag{6}$$

$$\kappa_{2,2}=\frac{A^4}{8}\left(\Psi(2)-4\Psi^2(1)\right);\ \kappa_{2,3}=0;\ \kappa_{2,4}=\frac{A^6}{4}\left(-\Psi(2)+6\Psi^2(1)\right);$$

$$\kappa_{2,5}=0;\ \kappa_{2,6}=\frac{165A^8}{128}\left(\Psi(2)-8\Psi^2(1)\right);\ \ldots$$

$$\kappa_{3,3}=\frac{A^6}{32}\left(\Psi(3)+9\Psi(1)-18\Psi(1)\Psi(2)+48\Psi^3(1)\right);$$

$$\kappa_{3,4}=0;\ \kappa_{3,5}=\frac{15A^8}{128}\left(-\Psi(3)-8\Psi(1)+28\Psi(1)\Psi(2)-96\Psi^3(1)\right);\ \ldots$$

$$\kappa_{4,4}=\frac{A^8}{128}\left(\Psi(4)+64\Psi(2)-576\Psi^2(1)-32\Psi(1)\Psi(3)+576\Psi(2)\Psi^2(1)-\right.$$
$$\left.-1152\Psi^4(1)-36\Psi^2(2)\right);\ \ldots\ .$$

Зауважимо, що при $t_2=t_1$ буде: $\Psi(l)=1,\ l\in Z$.

З (6) досить легко можна визначити кумулянтні коефіцієнти $\gamma_{i,j}=\frac{\kappa_{i,j}}{\kappa_{0,2}^{(i+j)/2}}$, $i,j=0,1,2,\ \ldots$. А саме:

$$\gamma_{0,1}=0;\ \gamma_{0,2}=1;\ \gamma_{0,3}=0;\ \gamma_{0,4}=-\frac{3}{2};\ \gamma_{0,5}=0;\ \gamma_{0,6}=10;\ \gamma_{0,7}=0;\ \gamma_{0,8}=-\frac{1155}{8};\ \ldots$$

і так далі.

З наведених виразів очевидним є негауссовий характер розподілу імовірностей випадкових процесів, що відповідають високочастотним фазомодульованим сигналам. Крім того, звідси випливає стаціонарність у широкому розумінні даних процесів не залежно від процесу $\lambda(t)$. У випадку стаціонарності у вузькому розумінні модулюючої функції $\lambda(t)$ такої властивості набуває і сигнал $s(t)$.

Якщо $\lambda(t)$ – процес з функцією щільності розподілу ймовірностей симетричною відносно нуля (отже, і нульовим математичним сподіванням), то $\psi_S(l)=0$, так як в розкладі $\psi_S(l)$ містяться лише моменти непарного порядку. Тоді функція $\Psi(l)$ набуває вигляду $\Psi(l)=\cos\omega_\tau\Psi_C(l)$.

Припустимо, що $\lambda(t)$ — стаціонарний гауссовий випадковий процес, покладемо для нього $\kappa_{0,2}^{(\lambda)}=\kappa_{2,0}^{(\lambda)}=D_\lambda;\ \kappa_{1,1}^{(\lambda)}=D_\lambda r(\tau)$ (всі інші дорівнюють нулю). Відомо [4], що за цих умов:

$$\psi_c(l)=M\left(\cos l\Delta\varphi\left(\lambda(t_2)-\lambda(t_1)\right)\right)=\operatorname{Re}M\left(\exp\left(jl\Delta\varphi\left(\lambda(t_1+\tau)-\lambda(t_1)\right)\right)\right)=$$
$$=\exp\left(-l^2D\left(1-r(\tau)\right)\right),\ D=\Delta\varphi^2D_\lambda.$$

Тоді $\Psi(l)=\cos l\omega_0\tau\cdot\exp\left(-l^2D\left(1-r(\tau)\right)\right)$.

Розглянемо функцію $\mu(\xi,\zeta)=\cos\xi\omega_0\tau\cdot\exp\left(-\zeta D\left(1-r(\tau)\right)\right)$.

За умови гауссовості процесу $\lambda(t)$ будуть справедливі наступні співвідношення:

1. $\Psi(\xi)=\mu\left(\xi,\xi^2\right)$;

2. $\Psi(\xi)\Psi(\zeta)=\frac{1}{2}\left(\mu\left(\xi-\zeta,\xi^2+\zeta^2\right)+\mu\left(\xi+\zeta,\xi^2+\zeta^2\right)\right)$;

3. $\Psi(\xi)\Psi(\zeta)\Psi(\varepsilon)=\frac{1}{4}\left(\mu\left(\xi+\zeta-\varepsilon,\xi^2+\zeta^2+\varepsilon^2\right)+\right.$
$+\mu\left(-\xi+\zeta+\varepsilon,\xi^2+\zeta^2+\varepsilon^2\right)+$
$+\mu\left(\xi-\zeta+\varepsilon,\xi^2+\zeta^2+\varepsilon^2\right)+$
$\left.+\mu\left(\xi+\zeta+\varepsilon,\xi^2+\zeta^2+\varepsilon^2\right)\right)$;

$$4.\ \Psi^4(\xi)=\frac{1}{8}\left(3\mu\left(0,4\xi^2\right)+4\mu\left(2\xi,4\xi^2\right)+\mu\left(4\xi,4\xi^2\right)\right).$$

Застосовуючи вказані властивості в формулах (6), отримаємо значення кумулянтних функцій для випадку гауссового характеру модулюючої функції. Запишемо, наприклад, діагональні елементи матриці кумулянтних функцій $\kappa_{p,p}$.

$$\kappa^r_{1,1}(\tau)=\frac{A^2}{2}\mu(1,1);\ \kappa_{2,2}=\frac{A^4}{8}\left(\mu(2,4)-2\mu(2,2)-2\mu(0,2)\right);$$

$$\kappa^r_{3,3}(\tau)=\frac{A^6}{32}\left(\mu(3,9)-9\mu(3,5)+12\mu(3,3)-9\mu(1,5)+36\mu(1,3)+9\mu(1,1)\right);$$

$$\kappa^r_{4,4}(\tau)=\frac{A^8}{128}\left(\mu(4,16)-16\mu(4,10)-18\mu(4,8)+144\mu(4,6)-144\mu(4,4)-\right.$$
$$-16\mu(2,10)+288\mu(2,6)-512\mu(2,4)-288\mu(2,2)-$$
$$\left.-18\mu(0,8)+144\mu(0,6)-432\mu(0,4)-288\mu(0,2)\right);\ \ldots\ .$$

Побудуємо спектри отриманих функцій за певного вигляду кореляційної функції $r(\tau)=r_0(\tau)$ (цей факт будемо позначати верхнім індексом r_0 біля імен функцій). Якщо відомий Фур'є-образ $w^{r_0}(\xi,\zeta)$ функції $\mu^{r_0}(\xi,\zeta)$, то спектри $W^{r_0}_{p,q}(\omega)$ кумулянтних функцій $\kappa^{r_0}_{p,q}(\tau)$ отримуються формальною заміною κ і τ відповідно на W і ω в лівій частині рівності і μ на w – в правій.

Для знаходження спектра функції $\mu(\xi,\zeta)$ розкладемо її в ряд, використовуючи $e^x=\sum_{j=0}^{\infty}\frac{x^j}{j!}$.

Отримаємо

$$\mu(\xi,\zeta)=e^{-\zeta D}\cos\xi\omega_0\tau\sum_{j=0}^{\infty}\frac{\left(\zeta Dr(\tau)\right)^j}{j!}.$$

Нехай тепер $r(\tau)=r_1(\tau)=e^{-a\tau^2}$. Для знаходження $w^{r_1}(\xi,\zeta)$ використаємо те, що Фур'є-образом функції

$$\rho(\xi,j,\tau)=\cos\xi\omega_0\tau\left(e^{-a\tau^2}\right)^j$$

є функція

$$\mathrm{P}(\xi,j,\omega)=\frac{\pi}{\sqrt{4\pi aj}}\left(\exp\left(-\frac{(\omega-\xi\omega_0)^2}{4aj}\right)+\exp\left(-\frac{(\omega+\xi\omega_0)^2}{4aj}\right)\right).$$

Звідки отримаємо наступне:

$$w^{r_1}(\xi,\zeta)=e^{-\zeta D}\left(\pi\left(\delta(\omega-\xi\omega_0)+\delta(\omega+\xi\omega_0)\right)+\sum_{j=1}^{\infty}\frac{(\zeta D)^j}{j!}\mathrm{P}(\xi,j,\omega)\right).$$

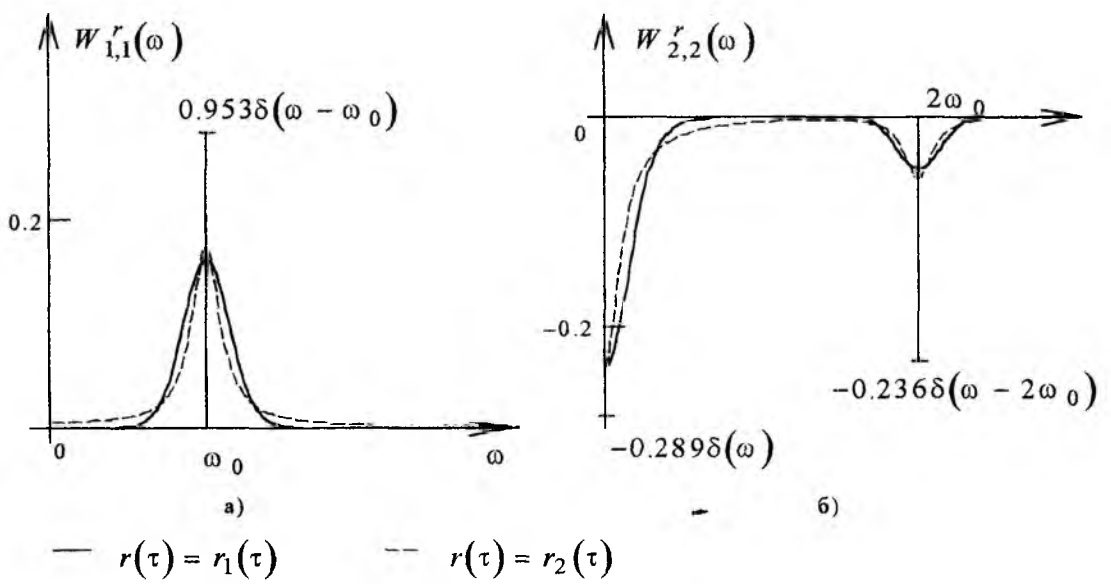
Розглянемо інший варіант кореляційної функції інформаційного повідомлення $r(\tau)=r_2(\tau)=e^{-a|\tau|}$. Помноживши відомий з [5] для цього випадку спектр кореляційної функції $\kappa^{r_2}_{1,1}(\tau)$ на $A^2/2$ та замінивши ω_0 на $\xi\omega_0$ і D на ζD, отримаємо

$$w^{r_2}(\xi,\zeta)=\frac{e^{-\zeta D}}{a}\left(\pi\left(\delta\left(\frac{\omega-\xi\omega_0}{a}\right)+\delta\left(\frac{\omega+\xi\omega_0}{a}\right)\right)+\sum_{j=1}^{\infty}\frac{(\zeta D)^j}{(j-1)!}\times\right.$$
$$\left.\times\left(\frac{1}{j^2+a^{-2}(\omega-\xi\omega_0)^2}+\frac{1}{j^2+a^{-2}(\omega+\xi\omega_0)^2}\right)\right).$$

Застосовуючи чисельні методи можна побудувати графіки спектрів кумулянтних функцій. На рисунку зображено спектри $W^r_{1,1}$ і $W^r_{2,2}$ при $r(\tau)$ рівному $r_1(\tau)$ і $r_2(\tau)$. Враховуючи симетрію графіків відносно осі ординат, подано лише їх частини, що лежать у правій півплощині.

Аналіз отриманих результатів показує, що моментні і кумулянтні функції непарних порядків високочастотних сигналів з фазовою модуляцією дорівнюють нулю, отже мають також і нульові спектри. Кумулянти одновимірних перерізів залежать лише від амплітуди несучого коливання і не залежать від модулюючої функції, мають спектр у вигляді дельта-функції на нульовій частоті (для кумулянтів парного порядку).

Обмежуючись стаціонарними гауссовими процесами з нульовим математичним сподіванням в якості інформаційного повідомлення, можна вказати на такі особливості спектрів. Знак спектра кумулянтної функції парного порядку додатній, якщо $(p+q)/2$ число непарне і навпаки, де (p,q) – порядок кумулянта. Пелюстки спектра зосереджені на частотах 0, $2\omega_0$, $4\omega_0$, ... ,$2k\omega_0$, якщо порядок кумулянта $(2k,2n)$, $k \le n$, і на частотах ω_0, $3\omega_0$, $5\omega_0$, ... $(2k-1)\omega_0$, якщо порядок кумулянта $(2k-1,2n-1)$, $k \le n$. Кожна пелюстка симетрична відносно вузлової частоти. Спектр носить дискретно-неперервний характер. Спектри кумулянтних функцій $\kappa^r_{1,1}$ (а) і $\kappa^r_{2,2}$ (б):



Отримані аналітичні вирази, що виражають двовимірні моментні і кумулянтні функції сигналу з фазовою модуляцією через моментні функції модулюючого сигналу. Наведено кілька прикладів побудови спектрів кумулянтних функцій, за умови, що інформаційне повідомлення є центрований гауссовий випадковий процес з відомою кореляційною функцією. Ця інформація може бути використана як для побудови оптимальної структури систем нелінійної обробки сигналів, так і при імітаційному моделюванні таких систем.

Список літератури: 1. *Федоров В.Б.* Мгономерные семиинварианты дискретно модулированных сигналов // Радиотехника и. электроника 1998. т. 43. №6. с. 696-702. **2.** *Градштейн И.С., Рыжик И.М.* Таблицы интегралов, сумм, рядов и произведений. - М.: Наука, 1971. - 1108 с. **3.** *Ширяев А.Н.* Вероятность. - М.: Наука, 1989. - 640 с. **4.** *Тихонов В.И.* Нелинейные преобразования случайных процессов. - М.: Радио и связь, 1986. - 296 с. **5.** *Мидлтон Д.* Введение в статистическую теорию связи. В 2-х т.: Пер. с англ. / Под ред. *Б.Р. Левина*. - М.: Сов. радио, Т.1, 1961. - 782 с.; Т.2, 1962. – 832 с.

Черкаський інженерно-технологічний інститут

Надійшла до редколегії 13.12.1999

УДК 621.391.14

В.В. ДАНИЛОВ, канд. техн. наук, С.В. ИВАНОВ.

УЧЁТ ВЛИЯНИЯ СВЯЗУЮЩИХ СЛОЕВ НА ЭФФЕКТИВНОСТЬ ПРЕОБРАЗОВАНИЯ АКУСТООПТИЧЕСКОГО УСТРОЙСТВА

Производительность оптических систем обработки информации, в основном, определяется быстродействием элементов ввода и вывода результатов обработки. В качестве элементов ввода информации могут использоваться устройства на различных принципах функционирования, однако, для обработки радио- и оптических сигналов, преимущественное распространение получили [1] акустооптические устройства (АОУ). Базовым элементом любого АОУ является акустооптическая ячейка (АОЯ). Устройства, реализованные на основе АОЯ, различаются определенными модуляционной характеристикой и оптикой, формирующей лазерный пучок. В системах спектрального и корреляционного анализа радиосигналов такие АОЯ получили название акустооптических модуляторов, в системах оптической памяти - акустооптические затворы и дефлекторы, и т.д. Физические принципы функционирования АОЯ приведены в [2]. Быстродействие акустооптических устройств в зависимости от функционального назначения системы определяется по разному, однако, в общем случае оно является функцией частотной зависимости эффективности преобразования АОЯ энергии и частоты управляющего радиосигнала, в энергию и пространственное положение лазерного пучка. Основные элементы АОЯ показаны на рис. 1. Здесь ПП- пьезоэлектрическая пластина, выполняющая функцию электроакустического преобразователя (ЭАП) управляющего радиосигнала; СЗП- светозвукопровод, в определенной области которого, происходит акустооптическое взаимодействие лазерного пучка и упругих (акустических) волн, генерируемых ПП. ЭАП с СЗП связан термокомпрессионной сваркой, посредством связующих слоев, на основе различных материалов.

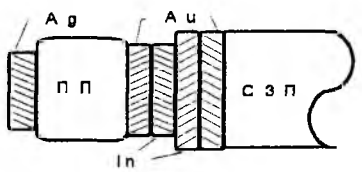


Рис.1

Различие акустических импедансов слоев влияет на акустическое согласование ПП и СЗП, и, приводит к сокращению полосы рабочих частот конкретного АОУ и снижению эффективности его преобразования. Процесс определения частотной зависимости эффективности преобразования АОЯ, в силу ограниченности обьема статьи, разделим на две стадии. В первой, рассмотрим задачу о передаче энергии упругой волны тонкими (четвертьволновыми) слоями связущих материалов в светозвукопровод, в полосе рабочих частот АОУ. Во второй - более общий случай, - частотную зависимость эффективности преобразования энергии радиосигнала в энергию лазерного пучка.

В данной работе предлагается методика инженерного расчета частотной зависимости акустического импеданса элементов многослойной структуры, обобщенная на случай затухания упругой волны в отдельных ее элементах, а также коэффициента передачи акустической энергии от ЭАП к СЗП.

Рассмотрим структуру (рис.2), связанную со светозвукопроводом и состоящую из N механически связанных (абсолютно) слоев, каждый из которых имеет: толщину $d_i = X_{i+1} - X_i, (i = 1,\ldots,N; X_{N+1} = 0)$, $z_i = \rho_i V_i$-волновое акустическое сопротивление i-го связущего слоя, ρ_i, V_i- плотность материала и скорость упругой волны в материале i-го свзущего слоя, α_i- коэффициент затухания в i-м связущем слое, z_{pp}, z_{szp}- волновые акустические сопротивления ПП и СЗП соответственно.

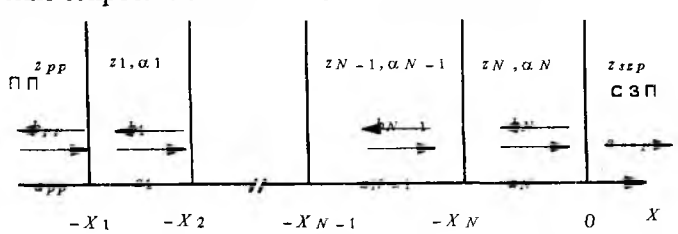


Рис. 2

Предположим, что акустическая волна распространяется вдоль оси OX нормально к границам раздела всех плоскопараллельных слоев, трансформация вида волн из продольных в поперечные и наоборот отсутствует, влияние взаимной диффузии смежных материалов на резкость границы слоев пренебрежимо мало. На каждой границе раздела X_i будут возникать отраженная и прошедшая волны, потенциал которых равен b_{i-1}, a_i и удовлетворяет граничным условиям равенства колебательных скоростей и давлений [3]:

$$v_{i-1}^{(i)} + v_{i-1}^{(r)} = v_i^{(d)}, p_{i-1}^{(i)} + p_{i-1}^{(r)} = p_i^{(d)}, \tag{1}$$

где $v_{i-1}^{(i)}, p_{i-1}^{(i)}$ - скорость и давление в падающей волне, $v_{i-1}^{(r)}, p_{i-1}^{(r)}$ - скорость и давление в отраженной волне (в $i-1$ -м слое), $v_i^{(d)}, p_i^{(d)}$ - скорость и давление в прошедшей волне в i -м слое. Связь между колебательной скоростью v, давлением p и потенциалом φ дается соотношениями

$$v = -\frac{\partial \varphi}{\partial X}, \quad p = \rho \frac{\partial \varphi}{\partial t}, \tag{2}$$

где ρ - плотность материала. Решение задачи о распространении упругих волн в слоистой структуре ,сводится к решению краевой задачи вида

$$\begin{cases} \dfrac{d^2\varphi_i}{dX^2} + K_i^2 \varphi_i = 0, \rho_{i-1} \dfrac{\partial a_{i-1}}{\partial t} - \rho_{i-1} \dfrac{\partial b_{i-1}}{\partial t} = \rho_i \dfrac{\partial a_i}{\partial t} \\ \dfrac{\partial a_{i-1}}{\partial X} - \dfrac{\partial b_{i-1}}{\partial X} = \dfrac{\partial a_i}{\partial X}; \varphi_0(-X_1) = \{a_{pp}, b_{pp}\}, \varphi_{N+1}(0) = \{a_{szp}, 0\} \end{cases}, \tag{3}$$

где $\varphi_i = \{a_i, b_i\}; i = 1,..,N$; K_i - волновое число: $K_i = k_i - j\alpha_i$, $k_i = \omega / V_i$, ω - циклическая частота колебаний упругой волны (УВ), α_i - коэффициент затухания УВ в i -м слое, $j = \sqrt{-1}$. Решением краевой задачи (3) является совокупность прямых и обратных волн, амплитуды которых: $\{a_i, b_i\}$. Искомый коэффициент передачи мощности структуры равен $D = a_{szp}^2 / a_{pp}^2$. Амплитуда упругих колебаний в СЗП равная a_{szp} , согласно [3], находится из системы $2(N+1)$ линейных алгебраических уравнений. Более общим, с точки зрения физики явления, является определение коэффициента передачи D, путем рекуррентного вычисления входного акустического импеданса структуры, обобщенного на случай затухания УВ. Введем амплитуду смещения в волне ξ, связанную с колебательной скоростью v и силой F, соотношениями $v = \partial \xi / \partial t, F = \rho S V^2 (\partial \xi / \partial X)$. Величина ξ удовлетворяет волновому уравнению (3) . Решения для ξ, в i-ом слое, будем искать в виде суммы прямой и обратной волн:

$$\xi_i = \left(\overset{\Gamma}{\xi}_i e^{-jk_i X} + \overset{S}{\xi}_i e^{jk_i X} \right) e^{j\omega t} e^{-\alpha_i X}. \tag{4}$$

Выражения для колебательных скоростей и сил, нормированных на волновое сопротивление i -го слоя, в прямой и обратной волнах i -го слоя, найдем из соотношений

$$\begin{cases} v_i = \left. \dfrac{\partial \xi_i}{\partial t} \right|_{X=X_i}, v_{i+1} = \left. \dfrac{\partial \xi_i}{\partial t} \right|_{X=X_{i+1}} \\ f_i = -SV_i \left. \dfrac{\partial \xi_i}{\partial X} \right|_{X=X_i}, f_{i+1} = -SV_i \left. \dfrac{\partial \xi_i}{\partial X} \right|_{X=X_{i+1}} \end{cases}, \tag{5}$$

где S - площадь сечения слоев. Введем, для i -го слоя, локальную систему координат $OX_i: [X_i, X_{i+1}] \to [-d_i, 0]$. Подставляя решение (4) в соотношения (5), с учетом OX_i и $\delta_i = k_i d_i$, получаем следующую систему уравнений:

$$\begin{cases} v_i = j\omega\left(\overset{r}{\xi}_i e^{j\delta_i} + \overset{s}{\xi}_i e^{-j\delta_i}\right)e^{j\omega t}e^{\alpha_i d_i}, v_{i+1} = -j\omega\left(\overset{r}{\xi}_i + \overset{s}{\xi}_i\right)e^{j\omega t} \\ f_i = SV_i\left[\left(jk_i + \alpha_i\right)\overset{r}{\xi}_i e^{j\delta_i} - \left(jk_i - \alpha_i\right)\overset{s}{\xi}_i e^{-j\delta_i}\right]e^{j\omega t}e^{\alpha_i d_i} \\ f_{i+1} = SV_i\left[\left(jk_i + \alpha_i\right)\overset{r}{\xi}_i - \left(jk_i - \alpha_i\right)\overset{s}{\xi}_i\right]e^{j\omega t} \end{cases} \quad (6)$$

Входной акустический импеданс i - го слоя $Z_i = f_i / v_i$ равен:

$$Z_i = z_i \frac{Z_{i+1}\left(\cos\delta_i - B_i \sin\delta_i\right) + jz_i\left(1 + B_i^2\right)\sin\delta_i}{z_i\left(\cos\delta_i - B_i \sin\delta_i\right) + jZ_{i+1}\sin\delta_i}, \quad (7)$$

где $z_i = k_i V_i S / \omega$ - волновое акустическое сопротивление i-го слоя, $Z_{i+1} = f_{i+1} / v_{i+1}$ - акустический импеданс $i+1$-го слоя, $B_i = \alpha_i / k_i$. Представляя $Z_i = X_i + jY_i$ и $Z_{i+1} = X_{i+1} + jY_{i+1}$, найдем вещественную и мнимую части импеданса:

$$X_i = z_i^2 X_{i+1}\left[1 + 2B_i\left(B_i \sin\delta_i - \cos\delta_i\right)\right] / A, \quad (8)$$

$$Y_i = z_i\left\{\left[\left(1 + B_i^2\right)z_i - Y_{i+1}B_i\right]\sin\delta_i + Y_{i+1}\cos\delta_i\right\} / A \times$$
$$\times \frac{\left[z_i \cos\delta_i - \left(B_i z_i + Y_{i+1}\right)\sin\delta_i\right] - z_i X_{i+1}^2\left(\cos\delta_i - B_i \sin\delta_i\right)\sin\delta_i}{A}, \quad (9)$$

где $A = \left[z_i \cos\delta_i - \left(B_i z_i + Y_{i+1}\right)\sin\delta_i\right]^2 + X_{i+1}^2 \sin^2\delta_i$. Последовательно применяя рекуррентные соотношения (8,9) можно найти акустический импеданс структуры в сечении $X = -X_1$ (см. рис.2).

В соответствии с выражениями (10,11), были реализованы на ЭВМ алгоритм и программа вычисления входного импеданса многослойной структуры, акустически нагруженной на СЗП. В качестве материалов связующих слоев был выбран ряд металлов [4,5]: платина (Pt), хром (Cr), золото (Au), алюминий (Al), индий (In), медь (Cu). Значения скоростей УВ (поперечных и продольных) , а также коэффициентов затухания, для указанных материалов связующих слоев, взяты из работ [4,5]. Учитывая линейность частотной зависимости коэффициента затухания УВ в металлах [6], параметр B_i равен $B_i = \alpha_i V_i \Omega / 8\pi$, $\Omega = f / f_0$- относительная частота, f - текущая частота рабочего диапазона, $f_0 = \sqrt{f_н f_в}$ - среднегеометрическая частота рабочего диапазона, $f_н, f_в$- нижняя и верхняя частоты рабочего диапазона. Толщина каждого слоя d_i равна четверти длины УВ на частоте f_0. Фаза УВ определяется из соотношения $\delta_i = \pi\Omega / 2$. На рис. 3-8 представлены расчетные значения входного импеданса связующих слоев Z (нормир. на волновое сопротивление СЗП) для следующих структур(продольные волны):

$ZnO\text{-}Al\text{-}Al_2O_3$ (рис.3); $LiNbO_3\text{-}Cr\text{-}Cu\text{-}In\text{-}Cu\text{-}Cr\text{-}TeO_2$ (рис.6);
$ZnO\text{-}Au\text{-}Al_2O_3$ (рис.4); $LiNbO_3\text{-}Cr\text{-}Al\text{-}In\text{-}Al\text{-}Cr\text{-}TeO_2$ (рис.7); (10)
$ZnO\text{-}Pt\text{-}Al_2O_3$ (рис.5); $LiNbO_3\text{-}Cr\text{-}Au\text{-}In\text{-}Au\text{-}Cr\text{-}TeO_2$ (рис.8).

Кривые 1 и 2 (см. рис. 3-8) каждого графика, показывают вещественные и мнимые части импедансов соответственно.

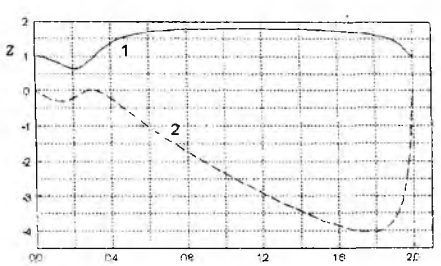


Рис. 3

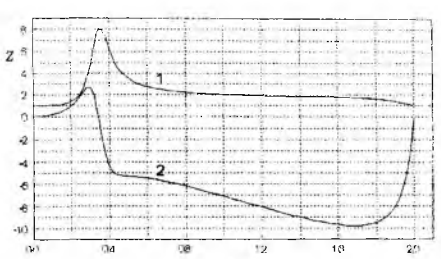


Рис. 4

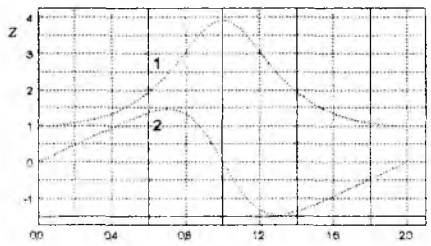


Рис. 5

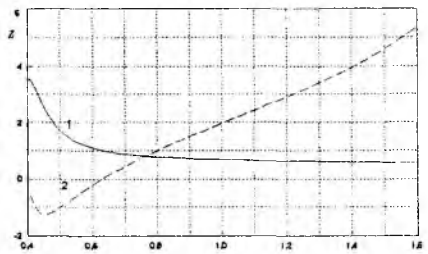


Рис. 6

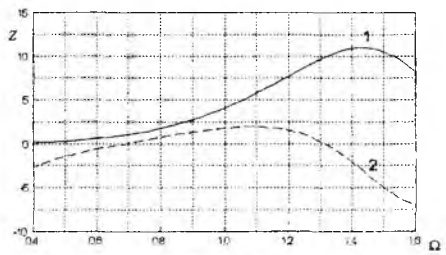


Рис. 7

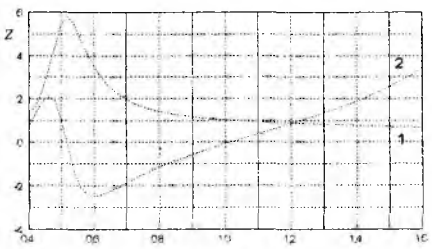


Рис. 8

При прохождении акустической волны через границу раздела двух сред с акустическими импедансами Z_1 и Z_2 коэффициент передачи (интенсивности) равен: $d_I = 4Z_1Z_2 / (Z_1 + Z_2)^2$. Коэффициент передачи акустической мощности пьезопреобразо-вателем, через согласующие слои с импедансом $Z = X + jY$, равен:

$$D = 4z_{pp}X / \left[(z_{pp} + X)^2 + Y^2 \right], \qquad (11)$$

где z_{pp} - волновое акустическое сопротивление ПП.

Частотные зависимости коэффициентов передачи мощности структур представлены:

- на рис. 9, кривые 1,2,3 описывают $D = D(\Omega)$ соответственно для структур типа ZnO-Al-Al_2O_3, ZnO-Au-Al_2O_3, ZnO-Pt-Al_2O_3;
- на рис. 10-12 для структур: $LiNbO_3$-Cr-Cu-In-Cu-Cr-TeO_2, $LiNbO_3$-Cr-Al-In-Al-Cr-TeO_2 и $LiNbO_3$-Cr-Au-In-Au-Cr-TeO_2, соответственно.

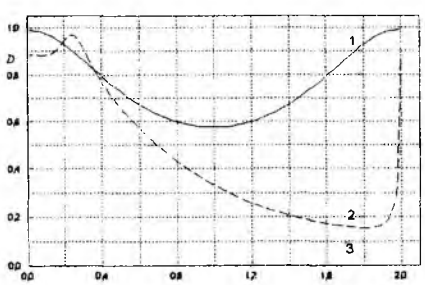


Рис. 9

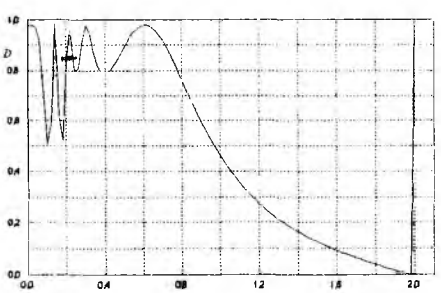


Рис. 10

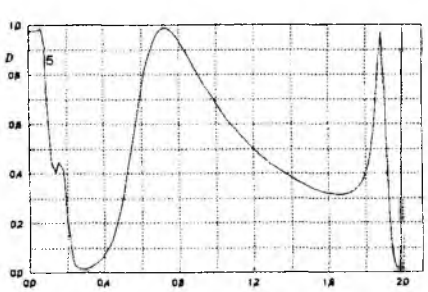


Рис. 11

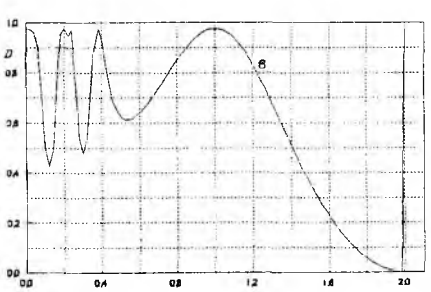


Рис. 12

Входной электрический импеданс механически нагруженной ПП посредством связующих слоев на СЗП, в окрестности полуволнового резонанса аппроксимируют [2] импедансом электрической эквивалентной схемы, представленной на рис.13:

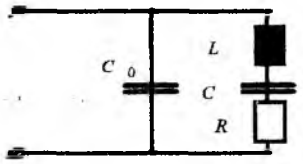


Рис. 13

Где C_0 - статическая емкость зажатого пьезопреобразователя; R - сопротивление описывающее активные потери в ПП, при этом электрическая добротность ПП равна: $Q_e = 1/\omega_0 RC$, где ω_0 - циклическая среднегеометрическая частота частотной характеристики согласующей цепи (в данной работе не рассматривалась). Акустическая добротность системы: "ПП-согласующие слои" равна:

$$Q_a = \pi z_{pp} / (2\mathrm{Re}Z), \qquad (12)$$

где $\mathrm{Re}Z$ - реальная часть входного акустического импеданса согласующих слоев. Поскольку $\mathrm{Re}Z$ - частотно- зависимая функция то $Q_a = Q_a(\Omega)$.

Таким образом, общий вид зависимостей рис. 9-12, обусловлен как количеством слоев, так и различием их акустических волновых сопротивлений. Число слоев определяет количество экстремумов, а различие волновых акустических сопротивлений влияет на отношение экстремальных значений. Для получения плоской частотной характеристики передачи акустической энергии, необходимо, импеданс СЗП трансформировать к значению импеданса ПП, что можно реализовать подбором соответствующих согласующих слоев. Следует отметить, что задача максимальной передачи акустической мощности, в заданной полосе частот, является задачей синтеза многослойных структур и выходит за пределы данной работы.

Приведенные рекуррентные соотношения для вычисления трансформированного акустического импеданса согласующих слоев, получены с учетом их механической нагрузки на СЗП. Реализованы алгоритм и программа расчета коэффициента передачи акустической мощности системы: ПП-согласующие слои- СЗП. Внесена поправка в методику [2], связанная с влиянием частотно-зависимой нагрузки комплексного характера на акустическую добротность ПП.

Список литературы: **1.** *Оптическая* вычислительная техника. ТИИЭР, 1984, т.72, стр.8-256. **2.** *О.Б. Гусев, С.В. Кулаков, Б.П. Разживин, Д.В. Тигин.* Оптическая обработка радиосигналов в реальном времени.- М.: Радио и связь, 1989. **3.** *В.А. Шутилов.* Основы физики ультразвука.- Л.: Изд. Ленинградского университета, 1980. **4.** *J.D. Larson, D.K. Winslow.* Ultrasonically welded piezoelectric transdusers. IEEE trans. on sonics and ultrasonics. 1971, vol. SU-18, №3, pp. 142-151. **5.** *А.И. Морозов, В.В. Проклов, Б.А. Станковский.* Пьезоэлектрические преобразователи для радиоэлектронных устройств.- М.: Радио и связь, 1981. **6.** *Дж. Такер, В. Рэмптон.* Гиперзвук в физике твердого тела.- М.: Мир, 1975.

Донецкий государственный университет
НИИ комплексной автоматизации

Поступила в редколлегию 25.11.99

УДК 621.472

А.А. КОНОВАЛЬЦЕВ, канд. техн. наук, Ю.А. ЛУЧАНИНОВ, М.А. ОМАРОВ, канд. техн. наук,
В.М. ШОКАЛО, д-р. техн. наук,

ПРИМЕНЕНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ БЕСПРОВОДНЫХ СИСТЕМ ПЕРЕДАЧИ ЭНЕРГИИ СВЧ-ЛУЧОМ

Практическое воплощение идеи беспроводной передачи энергии СВЧ-лучом стало возможным в 60-е годы в связи с развитием мощных источников генерирования СВЧ-энергии, появлением основополагающих работ Грубо по максимизации КПД передачи между двумя апертурами [1] и в связи с изобретением Брауном ректенны [2], преобразующей энергию электромагнитных волн в энергию постоянного тока. Результатом этих работ явилось создание нового типа энергетических систем - систем передачи энергии с помощью СВЧ-луча (СПЭСЛ).

Возможность создания эффективных СПЭСЛ и целесообразность их практического использования подтверждена в настоящее время результатами проведения целого ряда демонстрационных экспериментов. Четыре из них отмечены в работе [3] как важнейшие достижения в области беспроводной передачи энергии.

Прежде всего - это первая демонстрация полета летательного аппарата, запитываемого СВЧ-лучом [2]. В этом опыте, который был выполнен под руководством Брауна, использовалась ректенна с кремниевыми выпрямительными диодами, имевшими низкий КПД. Этого недостатка удалось избежать при второй демонстрации полета вертолета, на котором была установлена ректенна с диодами Шоттки [4]. В этом эксперименте впервые использовались электронные системы наведения СВЧ-луча на апертуру ректенны, которые должны быть неотъемлемой частью СПЭСЛ.

Третья важная демонстрация - это лабораторный опыт по беспроводной передаче энергии, в результате которого была доказана возможность достижения КПД СПЭСЛ 54% с учетом КПД преобразования энергии первичного источника постоянного тока в энергию СВЧ [4]. Достигнутый результат является пока наилучшим среди известных экспериментальных данных.

Наиболее крупномасштабный полигонный эксперимент по передаче 30 кВт СВЧ-мощности на расстояние в одну морскую милю (1,6 км) был проведен в 1975 г. [4]. Ректенна содержала примерно 5 тысяч приемно-выпрямительных элементов, каждый из которых преобразовывал 6Вт СВЧ мощности. Измеренный КПД ректенны равнялся 82% на частоте 2388 МГц.

Проведенные демонстрационные эксперименты позволили осознать реальность осуществления беспроводной передачи энергии с помощью СВЧ-луча и приступить к разработке большого числа различных проектов СПЭСЛ, в том числе используемых для подпитки высотных платформ [4-7], космических аппаратов [8-12], реактивных двигательных установок межорбитальных буксиров [13-15] с поверхности Земли, с низкой или геостационарной орбиты. Обсуждалась также возможность энергоснабжения наземных потребителей от солнечных космических электростанций (СКЭС) [16-19], от удаленных наземных источников [20,21] и др.

Особое место среди перечисленных проектов занимают работы по специальной научно-исследовательской программе Министерства энергетики США (ДОЕ) и НАСА в 1977-1980 гг. [18]. Результаты этих работ позволили сформировать основные представления о физических и технических особенностях солнечных космических электростанций. Краткие сведения о результатах работ по программе ДОЕ/НАСА и характеристики базового варианта СКЭС приведены в табл. 1 и на рис. 1.

В программе ДОЕ/НАСА рассматривалась возможность создания в начале XXI века системы из 60 энергоспутников: по две СКЭС в год с единичной мощностью 5 ГВт каждая. Считалось, что СКЭС может стать одним из крупномасштабных источников энергии в XXI веке. Однако, так как строительство их требует существенных затрат, то правительством США было решено не форсировать практическую реализацию СКЭС, а продолжить исследование в областях, наиболее существенных для СПЭСЛ.

Дальнейшее развитие СПЭСЛ (80-90-е гг.) шло в двух направлениях. Первое направление - это создание СПЭСЛ небольшой мощности, применяемых в системах идентификации транспорта, грузов и товаров в складских помещениях, а также используемых в качестве дистанционно управляемых источников энергоснабжения аппаратуры конфиденциальной связи [22].

Второе направление - разработка крупномасштабных СПЭСЛ. Характерной особенностью последних работ в этом направлении является переход от исследований в дециметровом диапазоне

волн к исследованиям в сантиметровом и миллиметровом диапазонах [3]. Повышение рабочей частоты приводит к существенному сокращению габаритов СПЭСЛ. Например [23], при проектировании системы энергоснабжения высотной платформы (высота 20,0 км) с диаметром ректенны 11,5 м передающая антенна при рабочей частоте СПЭСЛ 2,45 ГГц должна иметь диаметр 193м, а на частоте 35 ГГц эффективная передача энергии СВЧ обеспечивается при диаметрах апертур передающей антенны и ректенны всего 13,5 м.

Таблица 1

Параметр	Численное значение
Общее число СКЭС	60
Выходная мощность единичной СКЭС	5 ГВт
Площадь панелей солнечных батарей	5 × 10 км^2
Масса энергоспутника (включая 25%-ную надбавку на возможные ошибки):	
СКЭС с Si –фотобатареями	50 000 тонн
СКЭС с GaAs –фотобатареями	34 000 тонн
Диаметр передающей антенны	1 км
Размеры приемной антенны на широте 350	10 × 13 км
Рабочая длина волны	12,245 см
Максимальная плотность мощности на приемной антенне	0,23 кВт/м^2
Максимальная плотность мощности на передающей антенне	23 кВт/м^2
Форма распределения поля на передающей антенне:	
фазовое распределение	квадратичное, с фокусировкой на центр приемной антенны
амплитудное распределение	гауссовское, со спадением к краям апертуры на 10 дБ
Поляризация поля	линейная
КПД передачи энергии	95%
Общий КПД тракта передачи энергии	63%
Общая сумма капитальных затрат на 20-летний период, включая запуск первой СКЭС	102,4 млрд. $[1)]
Стоимость создания каждой следующей СКЭС	11,3 млрд. $
Ресурс работы СКЭС	30 лет
Время полной окупаемости СКЭС	1 – 6 лет

[1)] или 25 млрд. $ без стоимости разработки транспортных средств

Определенным итогом исследований по СПЭСЛ в последнее десятилетие являются работы [3,24], в которых определены пути применения и стратегия развития СПЭСЛ. По современным воззрениям крупные СПЭСЛ - это системы, с помощью которых может быть осуществлен переход от двумерных энергетических сетей на Земле к трехмерным энергетическим комплексам будущего [3]. Один из возможных вариантов трехмерного энергетического комплекса показан на рис 2. Он состоит из солнечных космических электростанций на геостационарной (1) и низкой (5) орбитах, транспортирующих энергию на космические (2,3,4) и наземные (9) потребители, и наземной СПЭСЛ (8), служащей для энергоснабжения с Земли космических кораблей с электрореактивными двигателями (ЭРД) и беспилотных летательных аппаратов (БЛА) (6,7).

Основные этапы развития СПЭСЛ, обсуждаемые в [24], приведены в таблице 2. Предполагается, что наиболее быстро будут завершены разработки БЛА (высотных стационарных платформ, самолетов, дирижаблей и др.). Спектр применений БЛА очень широкий. Разведка, экологические наблюдения, ретрансляция информации и т.д. По данным специалистов, система связи на основе БЛА будет в 2-3 раза дешевле спутниковой. Исследование БЛА охватывает широкий круг вопросов, связанных с формированием облика этих аппаратов и определением рациональных областей их применения. Часть этих исследований уже проведена в России [25]. В результате выяснено, что наиболее целесообразным типом БЛА для передачи энергии на борт СВЧ-лучом является самолет с большим удлинением крыла или дирижабль.

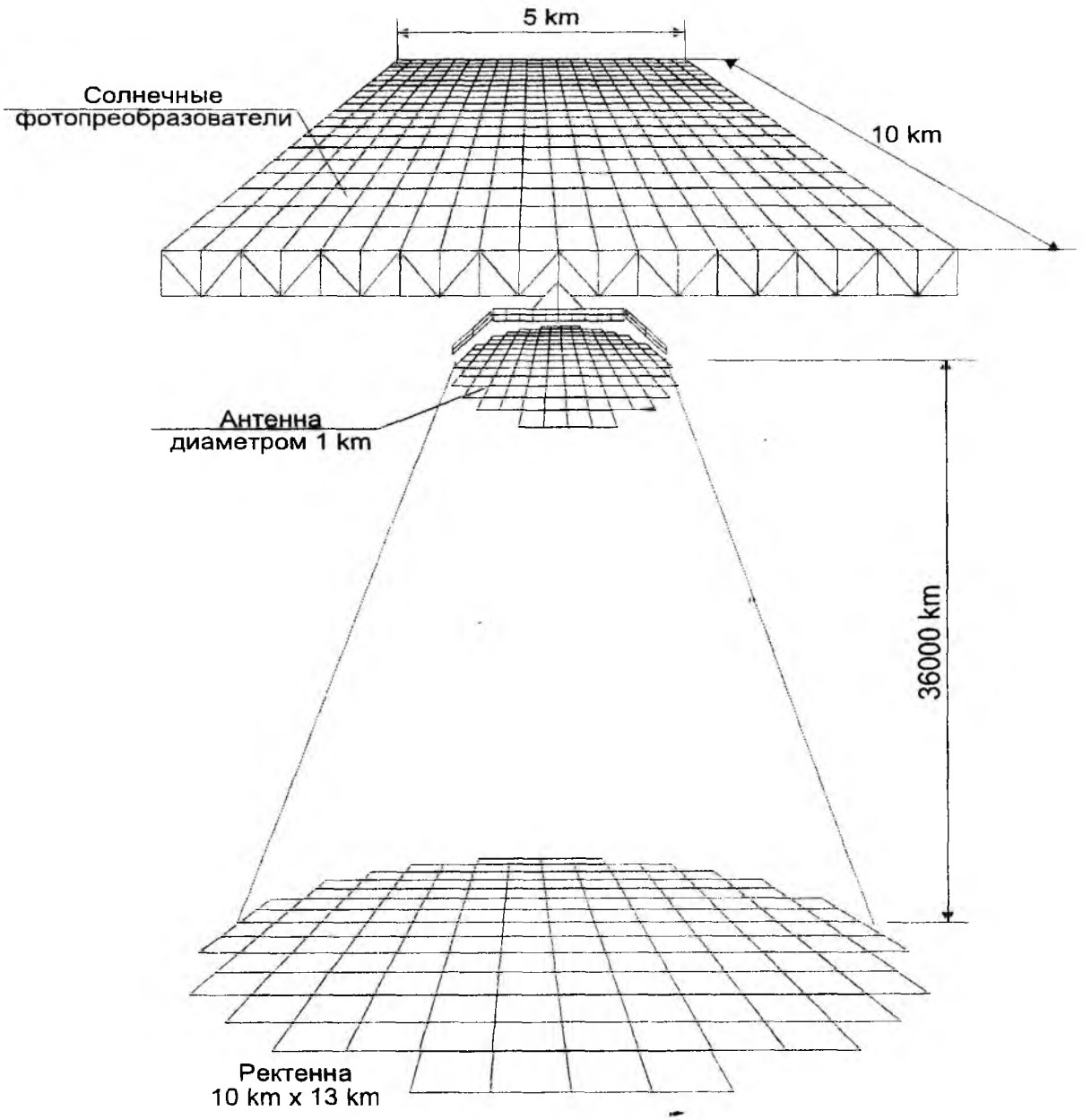


Рис. 1

Над созданием БЛА (программа SHARP) успешно работают канадские специалисты. Предполагаемые характеристики создаваемой ими системы следующие [5]. БЛА должен летать по замкнутой траектории на высоте около 20 км и снабжаться энергией с Земли СВЧ-пучком на частоте 2,45 ГГц. В качестве передающей антенной решетки будут использоваться 260 параболоидов с комбинированным управлением лучом, диаметр передающей решетки - 70 м, излучаемая мощность - 500кВт. Диаметр пучка на высоте 20 км составит 30 м, т.е. примерно будет равен размеру самолета. Величина плотности мощности на приемной антенне - 500 Вт/м, на ее выходе предполагается снимать около 300 кВт мощности постоянного тока, необходимых для питания электромотора.

Возможность практической реализации проекта SHARP подтверждена запуском масштабной модели БЛА [26,27] (рис. 3). БЛА запитывался волнами СВЧ круговой поляризации, поэтому ректенна состояла из двух ортогональных решеток (рис. 4), функционирующих самостоятельно при облучении линейно поляризованным полем.

В последнее время широкие исследования по созданию БЛА типа высотных платформ проводятся в Японии [28]. Вариант японской многофункциональной высотной платформы изображен на рис. 5.

Таблица 2

995	2000	2005	2010	2015	2030 Годы
С поверхности Земли	ГСО/НО (низкая орбита)	Луна и окололунное пространство	Планеты	Земля	Солнечные системы
Демонстрация передачи энергии СВЧ лучом Самолет с подпиткой СВЧ-энергией Другие	Энергия для космических шаттлов Снабжение космических станций и одноорбитных платформ Снабжение космических транспортов на НО	Снабжение энергией космических транспортных систем и лунных баз Снабжение энергией ретрансляционных спутников на геостационарной орбите (ГСО)	Снабжение энергией транспортных систем и баз на: - Фобосе; - Деймосе; - Марсе; - астероидах	Передача энергии с ГСО для использования на Земле Использование лунных материалов	Использование внеземной энергии и материалов для пользы человечества Передача энергии с Луны

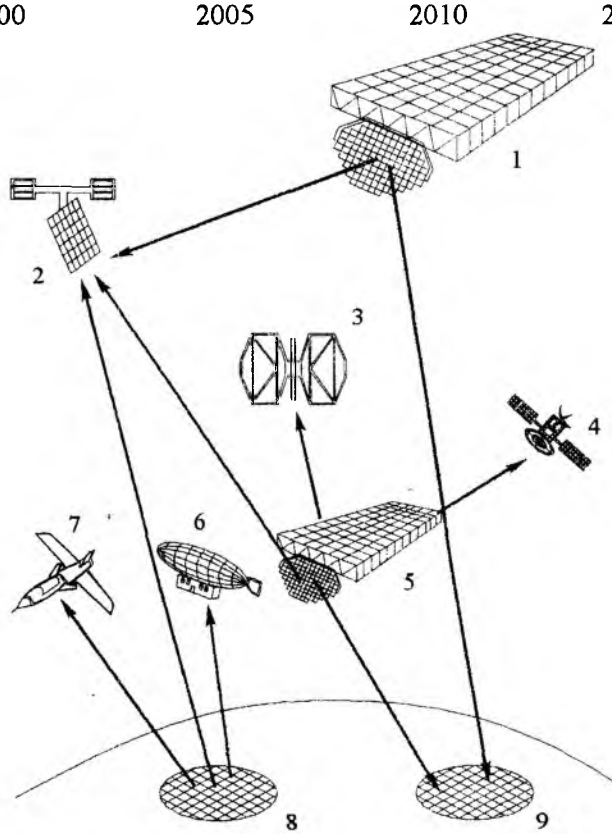


Рис. 2

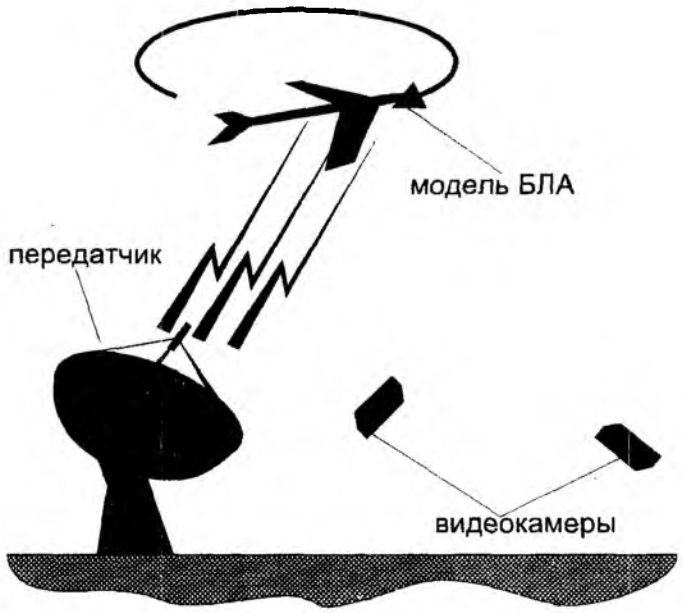
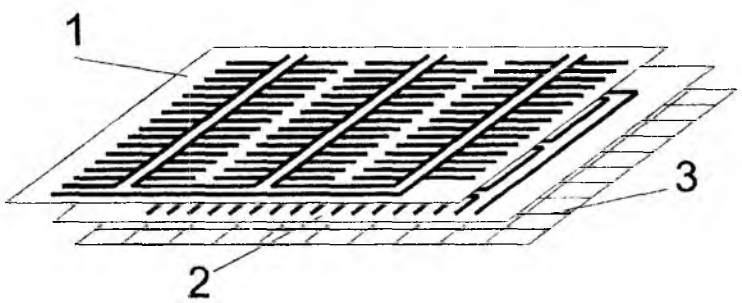


Рис. 3



1, 2 –платы ортогональных решеток ПВЭ, 3 – экран

Рис. 4

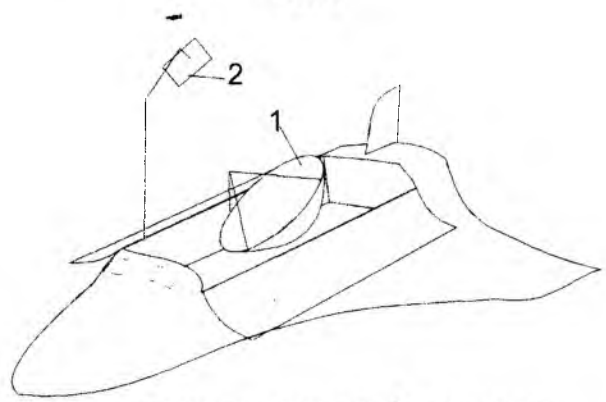
К числу ближайших задач развития СПЭСЛ относится также и передача энергии на Земле из богатых энергией районов в труднодоступную местность (например, поселения на Аляске [24]), или передача энергии от возобновляемых источников (солнечных, ветровых, станций прилива) к наземным потребителям.

Более дальняя перспектива передачи энергии СВЧ-лучом - это создание космической энергетики для обеспечения жизнедеятельности человека в космосе и осуществления электродвижения космических аппаратов. В силу ряда уникальных свойств, СПЭСЛ являются наиболее приспособленными для космоса энергетическими системами. Среди них возможность быстро изменять направление передачи и отсутствие потерь энергии при ее передаче через вакуум космоса [4].

Перечисленные свойства позволяют рассмотреть проекты снабжения энергией ряда космических объектов с одной орбитальной станции и поставить вопрос о создании сверхдальних СПЭСЛ с рабочей частотой 100 ГГц или 300 ГГц, функционирующих на трассе Луна - Космос [23]. Развитие космических СПЭСЛ требует проведения целого ряда дорогостоящих экспериментов в космосе. Первый планируемый из них - это демонстрация в ближайшем будущем передачи энергии с помощью СПЭСЛ, установленной в грузовом отсеке транспортного корабля (рис. 6) [29]. Расстояние передачи 10-20 м на частоте 2,45 ГГц при мощности передатчика 250 Вт и КПД всей системы 60%.

Браун считает [3], что передача энергии СВЧ-лучом - технологический прорыв в организации электродвижения космических аппаратов. Прежде всего, это обусловлено параметрами ректенн, которые как источник мощности постоянного тока имеют весьма низкое отношение выходной мощности к массе - 1 кВт/кг. По оценкам специалистов, применение СПЭСЛ для питания ЭРД позволит создать транспортную систему, которая будет иметь на порядок меньше отношение массы корабля к массе полезного груза, чем космический аппарат с химическим двигателем.

Рис. 5



1 – передающая антенна; 2 – ректенна

Рис. 6

Применение крупномасштабных СПЭСЛ сопряжено с необходимостью решения целого ряда экономических, экологических, социальных проблем и проблем электромагнитной совместимости. Наиболее детально эти проблемы рассмотрены применительно к СКЭС в работе [30], как части программы CDEP Департамента энергетики США и NASA. СКЭС сравнивались с альтернативными энергетическими системами, включая угольные, ядерные и наземные фотоэлектрические, по показателям цены единицы вырабатываемой энергии, влияния на здоровье и безопасность людей, окружающую среду и электромагнитные системы. Оценки показали следующее:

- стоимость энергии, вырабатываемой СКЭС, лежит в пределах колебаний цен на энергию альтернативных энергетических технологий;
- все рассматриваемые технологии имеют заметное, хотя и разное, негативное влияние на здоровье людей и состояние окружающей среды;
- суммарное количество земли, требуемое для полного сырьевого цикла, приблизительно одинаково для всех энергетических технологий, однако для размещения ректенн СКЭС и наземных фотоэлектрических систем может понадобиться большая площадь; выходом из этого положения является сооружение ректенн на прибрежных плавучих структурах.

За последние 3-4 года исследования по проблеме беспроводной передачи энергии существенно расширились. Результаты их рассмотрены на Второй международной конференции по беспроводной передаче энергии, состоявшейся в 1995 г. в г.Кобе. Здесь был доложен ряд концептуальных докладов, в

которых рассматривались сценарии построения энергетических станций в Космосе и на Луне. Некоторые проекты уже разработаны в деталях. Следует отметить, что в Японии начаты работы по созданию солнечного энергетического спутника SPS-2000 [31], который предназначен для энергоснабжения некоторых районов Танзании и Папуа Новой Гвинеи. Расчетный КПД проектируемой СКЭС равен 0,5 % и вычислялся он как отношение мощности, дошедшей до наземного потребителя, к мощности солнечной энергии, поглощаемой космическими фотобатареями с КПД преобразования 10%.В 10% оценен собственно КПД СПЭСЛ (потери перехвата энергии СВЧ-луча, потери в СВЧ-генераторе и в ректенне). При расчетах также предполагалось, что доставка энергии по проводам от ректенны до реальных потребителей будет осуществляться с КПД 50%.

Приведенный краткий обзор состояния разработок и перспектив создания СПЭСЛ указывает на актуальность рассматриваемой проблемы и ее все усиливающееся влияние на процесс развития мировой энергетики.

Список литературы: 1. *Goubau G., Schwering F.* On the guided propagation of electromagnetic wave beams // IRE Trans. Antennas Propagation. 1961. v. AP-9. p.248-256. **2.** *Brown W.C.* Experimental involving a microwave beam to power and position a helicopter // IEEE Trans., v. AES-5, 1969, No.9, p.692. **3.** *Brown W.C., Eves E.E.* Microwave power transmission and its application to space // IEEE Trans. 1992. v. MTT-40. No.8. p. 1239-1250. **4.** *Brown W.C.* The history of power transmission by radio waves // IEEE Trans. 1984. v. MTT-32. No.9. p. 1230-1242. **5.** *Fisher A.* Secret of perpetual flight? Beam-power plane // Popular Science. 1988. v.232. No.1. p. 62-65. **6.** *Morris C.E.* Microwave powered, unmanned, high-altitude airplanes // Journal of aircraft. 1984. v.21. No.12. p.966-970. **7.** *Brown W.C.* Microwave powered, long duration, high-altitude platform // Internal. microwave symp. N.Y.: IEEE. 1986. p.507-510. **8.** *Glaser P.E.* Microwave power transmission for use in space // Microwave Journal. 1986. No.12. p.44-58. **9.** *Arndt G.D., Kerwin E.M.* Application of earth-orbit power transmission // Space power. 1986. No.12. p.44-58. **10.** *Landis G.A.* A new space station power system // Acta astronautica. 1988. v.17. No.9. p.975-977. **11.** *Chang K., McCleary J.C., Pollock M.A.* Feasibility study of 35 GHz microwave transmission in space // Solar power. 1989. v.8. No.3. p. 365-370. **12.** *Hoffert M.I., Miller G., Kadiramangalam M., Ziegler W.* Earth-to-satellite microwave power transmission // Journal of propulsion and power. 1989. v.5. No.6. p.750-758. **13.** *Minovith M.A.* Solar powered, self-refueling, microwave propelled interorbital transportation system // AIAA Paper. 1983. No.1446. **14.** *Brown W.C.* Earth to space DC power transmission system utilizing a microwave beam as a source of energy for electric propelled interorbital vehicles // AIAA Paper. 1985. No.2045. **15.** *Brown W.C.* All electronic propulsion - key to future spaceship design // AIAA Paper. 1988. No. 3170 **16.** *Glaser P.E.* Power from the Sun: its future // Science. 1968. v.162. P.857-861. **17.** *Ванке В.А., Лопухин В.М., Саввин В.Л.* Проблемы солнечных космических электростанций // Успехи физических наук. 1977. Т.123. Вып.4. С.633-656. **18.** *Satellite power system concept development and evaluation program* // Reference system report. Wash.: DOE/ER. 1978. (DOE/ER - 0023). **19.** *Книжник Р.С., Кочубей А.Н.* Передача энергии пучком СВЧ радиоволн и солнечные космические электростанции // Зарубежная радиоэлектроника. 1983. № 7. С.75-84. **20.** *Rogers T.F.* Reflector satellites for solar power // IEEE Spectrum. 1981. V.18. No.7. p.38-43. **21.** *Angelini A.M.* On the possibility of intercontinental power transmission via satellite // Space power. 1988. V.7. No.2. P.175-186. **22.** *Пат.3745569 США, МКИ G01s 9/56*, Remotely powered transponder // Works G.A., Murray J.C., Ostroff E.D., Freedman N. (USA), July 10, 1973. **23.** *Koert P., Cha J.T.* Millimeter wave technology for space power beaming // IEEE Trans. on Microwave Theory and Techniques. 1992. V.40. No.6. P. 1251-1258. **24.** *Glaser P.E.* An overview of the solar power satellite option // IEEE Trans. on Microwave Theory and Techniques. 1992. V.40. No.6. P. 1230-1238. **25.** *Разработка, изготовление и испытания ЛА с передачей энергии СВЧ лучом. Этап 1. Исследование возможности создания ЛА различных типов и назначения с передачей энергии на борт с помощью СВЧ луча*; Отчет о НИР / Московский авиационный институт; Руководитель А.К. Чурусов. 60500. Москва. 1993. 127с. **26.** *Sohlesak J.J., Alden A., Ohno T.* SHARP (Stationary high altitude platform): rectenna and low altitude tests // Globecom 85: IEEE Glob. Telecommun. conf. New Orleans. 1985. V.2. P. 960-964. **27.** *Jull G.W., Lillemark A., Turner R.M.* SHARP (Stationary high altitude platform): telecommunication missions and systems // Globecom 85: IEEE Glob. Telecommun. conf. New Orleans. 1985. V.2. P. 955-959. **28.** *Ito T., Fujino Y., Fujita M.* Fundamental experiment of a rectenna array for microwave power reception // IEICE Trans. Commun. 1993. V. E76-B. No.12. P.1508-1513. **29.** *Chang K., Patton A.D., Kennedy M.O. and others* Demonstration of microwave power transmission in space // Int. Symp. on SPS, Paris, 1991. P.343-347. **30.** *M.R.Riches* A comparative assessment of the reference satellite power system with selected current, Near-term and Advanced Energy Technologies, Department of Energy, Conf. Report, 800491, P.66-67, 1980. **31.** *SPS 2000 News Letter*, Institute of Space and Astronautical Science, No.13, October 1995.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 04.02.2000

УДК 537.874.4

В.А. ПЕТРОВ, канд. физ.-мат. наук, В.М. КАРТАШОВ, канд. техн. наук

АНАЛИЗ СТРУКТУРНО-ФИЗИЧЕСКОЙ МОДЕЛИ РАССЕЯНИЯ ВОЛН В ТУРБУЛЕНТНОЙ АТМОСФЕРЕ

В системах дистанционного зондирования атмосферы (звуковыми или радиоволнами) рассеяние волн в турбулентной среде происходит за короткое время, в течение которого неоднородность среды можно считать стационарной. Усредненные по множеству реализаций характеристики рассеивающего объема среды и рассеянного сигнала достаточно подробно изучены [1,2]. Анализ эквивалентной структуры одной реализации случайной неоднородности позволяет получить дополнительные сведения о структуре среды и динамике процессов в рассеивающем объеме.

Пусть стационарная неоднородная среда характеризуется параметром $\varepsilon_1(\vec{r}) = \varepsilon_0 + \varepsilon_2(\vec{r})$, где $\vec{r}$ - радиус-вектор точки с координатами x, y, z, $|\vec{r}| = \left(x^2 + y^2 + z^2\right)^{\frac{1}{2}}$. Область среды, в которой происходит рассеяние, ограничена объемом $V(\vec{r})$, причем $V(\vec{r}) = 1$ внутри объема и $V(\vec{r}) = 0$ за его пределами. Обозначим $V(\vec{r}) \cdot \varepsilon_2(\vec{r}) = \varepsilon(\vec{r})$ и будем считать, что рассеяние обусловлено флуктуациями $\varepsilon(\vec{r})$, причем $|\varepsilon(\vec{r})| << \varepsilon_0$, где ε_0 - среднее по ансамблю реализаций. Полагаем, что для любой реализации справедливо соотношение

$$\int_{-\infty}^{\infty} \iiint |\varepsilon(x,y,z)|^2 \, dxdydz = E < \infty. \qquad (1)$$

При условии (1) существует трехмерное преобразование Фурье $G(\vec{k}) = G(k_x, k_y, k_z)$ функции $\varepsilon(x,y,z)$. Здесь $\vec{k}$ - волновой вектор, $|k| = (k_x^2 + k_y^2 + k_z^2)^{\frac{1}{2}}$.

Как показано в работе [3], учитывая селективность взаимодействия монохроматической волны со средой, можно условно выделить в спектре $G(\vec{k})$ область W составляющих, участвующих в рассеянии. Если вектор рассеяния $\vec{B}$ направлен вдоль оси k_x, то область W ограничена значениями пространственных частот в интервале $(-\infty < k_y < \infty, \quad -\infty < k_z < \infty, \quad b - \beta \le k_x \le b + \beta)$, где $b = |\vec{B}|$, $\beta = \Delta k_x / 2$, $\vec{B} = \vec{a}_s - \vec{a}_0$, $\vec{a}_s$ и $\vec{a}_0$ - волновые векторы рассеянной и основной (падающей) волны, Δk_x - полоса селектируемых частот и $\Delta k_x << b$. Полагая равными нулю все составляющие $G(\vec{k})$, не участвующие в рассеянии, обратным преобразованием Фурье получим эквивалентную структуру рассеивающего объема среды:

$$\varepsilon_s(\vec{r}) = \int_{-\infty}^{\infty} \int G(b, k_y, k_z) e^{2\pi j(k_y y + k_z z)} dk_y dk_z \left[\int_{b-\beta}^{b+\beta} e^{2\pi j k_x x} dk_x + \int_{-b-\beta}^{-b+\beta} e^{2\pi j k_x x} dk_x \right]. \qquad (2)$$

Интегрирование в (2) выполняется по двум областям, содержащим комплексно-сопряженные значения $G(\vec{k})$ и $G^*(\vec{k})$ в предположении, что в пределах полосы Δk_x значения $G(\vec{k})$ не зависят от k_x.

Из выражения (2) следует $\varepsilon_s(\vec{r}) = f(x) \cdot f(y,z)$, где

$$f(x) = \frac{2}{\pi x} \sin(\pi \Delta k_x x) \cos(2\pi b x) = F(x) \cos(2\pi b x), \qquad f(y,z) = \varepsilon_g(b,y,z) = |\varepsilon_g(b,y,z)| \exp\{j\psi_g(y,z)\}.$$

Функция $f(x)$ описывает линейную решетку, т.е. квазипериодическую структуру с периодом $1/b$ и огибающей $F(x)$, скорость изменения которой связана с Δk_x. В свою очередь, Δk_x по порядку величины соизмерима с $1/L$, где L - размер области $V(\vec{r})$ в направлении вектора $\vec{B}$.

Комплексная функция $f(y,z)$ определяет амплитуду $|\varepsilon_g|$ и начальную фазу ψ_g колебаний в каждой элементарной решетке, имеющей координаты y и z при заданном $b = |\vec{B}|$.

Таким образом, эквивалентная структура рассеивающего объема турбулентной среды представляет собой совокупность линейных решеток со случайными амплитудами и начальными фазами, зависящими только от структуры реальной среды. Поскольку $\Delta k_x << b$, каждую из линейных решеток можно рассматривать как выборку узкополосного случайного процесса $\varepsilon_{si}(x)$.

Радиус поперечной корреляции решеток ρ_k в плоскости y,z можно найти по эффективной ширине спектра в плоскости k_y, k_z, причем порядок ρ_k можно оценить исходя из общих свойств трехмерной спектральной плотности.

Обозначим $\Phi(\vec{k}) = < G(\vec{k}) \cdot G^*(\vec{k}) >$, где знак $< \cdot >$ означает статистическое усреднение. По теореме Парсеваля

$$\int_{-\infty}^{\infty}\iint \Phi(\vec{k})dk_x dk_y dk_z = \int_{-\infty}^{\infty}\iint < \varepsilon^2(x,y,z) > dxdydz = E < \infty . \qquad (3)$$

Будем считать, что поле $\varepsilon(x,y,z)$ статистически изотропно. Тогда совокупность значений $\Phi(\vec{k})$ образует в пространстве волновых векторов $\vec{k}$ сферическое поле, т.е. $\Phi(\vec{k})$ остается постоянной на сфере радиуса $|\vec{k}|$.

В сферических координатах

$$2\int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin\theta d\theta \int_0^{\infty} \Phi(k,\theta,\varphi)k^2 dk = E . \qquad (4)$$

Для сходимости несобственного интеграла в (4) необходимо, чтобы спектральная плотность $\Phi(k,\theta,\varphi)$ при $k \to \infty$ убывала быстрее, чем k^{-3}.

Применительно к турбулентной атмосфере $\Phi(k)$ при $k > 0$ аппроксимируют функцией $\Phi(k) = Ak^{-n}$, где A – постоянная, $n = 11/3 \approx 3{,}66$.

Эффективную ширину спектра Δ в сечении (b, k_y, k_z) найдем из уравнения

$$\pi\Delta^2 \cdot \Phi(b,0) = \int_0^{2\pi} d\varphi \int_0^{\infty} \Phi(b,\chi,\varphi)\chi d\chi , \qquad (5)$$

где $\pi\Delta^2$ - площадь основания цилиндра высотой $\Phi(b,0)$; χ, φ - полярные координаты в плоскости сечения; $\chi = (k_y^2 + k_z^2)^{\frac{1}{2}}$, $\Phi(b,0)$ - спектральная плотность на оси k_x при $k_x = b$.

Используя аппроксимацию $\Phi(k) = Ak^{-n}, (n > 3)$ запишем уравнение (5):

$$\pi\Delta^2 \cdot Ab^{-n} = \int_0^{2\pi} d\varphi \int_0^{\infty} A(b^2 + \chi^2)^{-\frac{n}{2}} \chi d\chi , \qquad (6)$$

откуда

$$\Delta = b\left(\frac{2}{n-2}\right)^{\frac{1}{2}} . \qquad (7)$$

Интервал поперечной корреляции линейных решеток

$$/0\,2\rho_k \approx \frac{1}{\Delta} = \left(\frac{n-2}{2}\right)^{\frac{1}{2}} \frac{1}{b} = \left(\frac{n-2}{2}\right)^{\frac{1}{2}} \frac{\lambda_0}{2}, \tag{8}$$

где λ_0 - длина акустической или электромагнитной волны.

Таким образом, ρ_k зависит от скорости убывания спектральной плотности и при $n \approx 4$ радиус корреляции $\rho_k \approx \lambda_0/4$.

Если характерный размер области $V(\vec{r})$ в плоскости (y,z) равен L_1, то можно оценить максимальное число N некоррелированных решеток в объеме $V(\vec{r})$: $N \approx (L_1/2\rho_k)^2$.

Очевидно, в конкретной реализации турбулентной среды линейные решетки могут образовывать упорядоченные структуры большего поперечного размера, чем $2\rho_k$, но вероятность таких образований быстро падает с ростом поперечного размера.

Таким образом, при обратном рассеянии $(\varphi \approx \pi)$ рассеивающий объем можно рассматривать как совокупность «блестящих точек» со случайными параметрами. Модель радиолокационной цели в виде конечного числа блестящих точек подробно изучена [4]. Некоторые результаты исследования этой модели можно, таким образом, распространить и на "гладкие" неоднородные среды, если известны статистические характеристики линейных решеток. Нужно, однако, иметь в виду и существенные различия моделей. Во-первых, линейные решетки нельзя считать изотропными излучателями. Во-вторых, при углах рассеяния $\varphi < \pi$ линейные решетки "видны" под углами $(\pi - \varphi)/2$, и "блестящие точки" становятся "блестящими линиями", параллельными вектору рассеяния. Наконец, величину парциального (рассеянного одной решеткой) сигнала, как показано ниже, нельзя однозначно связать со структурными параметрами линейной решетки.

Введем эффективную длину l решетки

$$l = \frac{1}{F_m(x)} \int_{x_1}^{x_2} F(x)dx = \frac{1}{F_m(x)} \int_{x_1}^{x_2} \frac{2}{\pi x} \sin(\pi x \Delta k_x)dx, \tag{9}$$

где $F(x)$- огибающая решетки, $F_m(x)$ значение $F(x)$ в максимуме; x_1 и x_2 - координаты ближайших минимумов $F(x)$.

Из (9) следует,

$$l = \frac{2Si(\pi)}{\pi \Delta k_x} \approx 1{,}178 \frac{1}{\Delta k_x}, \tag{10}$$

где $Si(\cdot)$ - интегральный синус.

Если амплитуда U_i парциального рассеянного сигнала пропорциональна величине колебаний параметра среды ε_{gi}, то можно записать

$$U_i = \mu U_0 \varepsilon_{gi} l_i = \mu U_0 \varepsilon_g(y_i, z_i) l_i, \tag{11}$$

где $\mu << 1$, $\mu\varepsilon_{gi}$ - коэффициент отражения единицы эффективной длины i-той решетки; U_0 - амплитуда падающей волны; ε_{gi} - случайная величина, средний квадрат которой пропорционален Δk_x. Действительно,

$$< \varepsilon_g^2 >= \int_{-\infty}^{\infty} \int_{b-\beta}^{b+\beta} \Phi(k_x, k_y, k_z) dk_x dk_y dk_z. \tag{12}$$

Поскольку $\beta = \Delta k_x / 2 << b$, можно в первом приближении считать, что в пределах полосы частот Δk_x спектральная плотность не зависит от k_x.

Тогда из (12) получается

$$< \varepsilon_g^2 >= \int_{-\infty}^{\infty} \Phi(b, k_y, k_z) dk_y dk_z \int_{b-\beta}^{b+\beta} dk_x = E_b \Delta k_x, \tag{13}$$

где E_b - суммарная «энергия» пространственных гармоник в сечении $k_x = b$ при $\Delta k_x = 1$.

Выражение (11) можно записать с учетом (10) и (13):

$$U_i = C_i U_0 (\Delta k_x)^{-\frac{1}{2}}, \quad C_i << 1, \tag{14}$$

где C_i - коэффициент пропорциональности. Из (14) видим, что в среднем амплитуда парциального сигнала U_i тем больше, чем уже полоса Δk_x. В этом проявляется селективность взаимодействия падающей волны с турбулентной средой. Основной вклад в величину рассеянного сигнала вносят решетки с максимальной эффективной длиной l.

Однако, для любого фиксированного значения U_i, как видно из (11), имеет место неопределенность, которую можно выразить соотношением

$$\varepsilon_{gi} l_i = const \ , \tag{15}$$

причем ε_{gi} - случайная величина, распределение вероятностей которой, с учетом (9), соответствует распределению максимумов огибающей узкополосного случайного процесса [5]. Поэтому всегда существует конечная вероятность присутствия в рассеивающем объеме среды эквивалентных структур в виде относительно коротких решеток с большими значениями $|\varepsilon_{gi}|$, а границы области селектируемых пространственных гармоник Δk_x оказываются нечеткими. Последнее существенно при моделировании рассеивающей среды и анализе динамики неоднородности.

Таким образом, анализ рассеяния монохроматических волн одной реализацией неоднородной среды в ограниченном объеме $V(x, y, z)$ естественным образом приводит к эквивалентной структуре среды в виде совокупности линейных решеток. Структура эквивалентных пространственных образований в плоскости перпендикулярной вектору рассеяния при $\varphi \approx \pi$ определяется только характеристиками среды. При углах рассеяния $\varphi \approx \pi$ совокупность линейных решеток проявляет себя аналогично набору "блестящих точек" с соответствующими параметрами.

Соотношение неопределенности (15) дает основание считать, что в эквивалентной структуре реальной неоднородности присутствуют линейные решетки с разной эффективной длиной.

Возможность раздельного наблюдения локальных областей интенсивного рассеяния определяется только угловой разрешающей способностью приемного устройства.

Список литературы: **1.** *Рытов С.М., Кравцов Ю.А., Татарский В.И.* Введение в статистическую радиофизику. Ч.2. М.: Наука, 1978. 436 с. **2.** *Татарский В.И.* Распространение волн в турбулентной атмосфере. М.: Наука, 1967. 548 с. **3.** *Петров В.А., Цветкова В.С.* Физические модели обратного рассеяния волн в турбулентной атмосфере // Радиотехника. 1991. Вып.97. С. 37-44. **4.** *Островитянов Р.В., Басалов Ф.А.* Статистическая теория радиолокации протяженных цепей. М.: Радио и связь, 1982. 232 с. **5.** *Тихонов В.И.* Выбросы случайных процессов. М.: Наука, 1970. 392 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000

УДК 621.372.8

О.В. КУЛАКОВ, канд. техн. наук, Н.И. ПЯТАК, канд. физ.-мат. наук., с.н.с.

СОБСТВЕННЫЕ КОЛЕБАНИЯ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ ПОПЕРЕЧНО НАМАГНИЧЕННОГО ФЕРРИТОВОГО РЕЗОНАТОРА В ИЗЛОМЕ ПРЯМОУГОЛЬНОГО ВОЛНОВОДА

Волноводно-ферритовые (и как частный случай – волноводно-диэлектрические) резонаторы применяются в волноводной технике в качестве селективных, частотно перестраиваемых систем [1-4]. Использование условия биортогональности [5] позволило существенно уточнить методику расчета таких систем, упростило процесс построения новых избирательных устройств. Учитывая, что волноводный тракт обычно представляет собой довольно протяженную структуру с многочисленными изгибами, представляется целесообразным построение резонансных систем именно в них.

Рассмотрим изгиб под углом 90° (излом) запредельного прямоугольного волновода в Н-плоскости. В общем случае ширина волновода после излома может изменяться, но волновод должен при этом оставаться запредельным. Предположим, что область излома полностью заполнена поперечно (т.е. вдоль оси z) намагниченным ферритом (продольное сечение структуры в плоскости $x0y$ приведено на рис.1). Предполагая идеальную проводимость стенок волновода, ограничимся анализом H_{m0} волн ($\frac{\partial}{\partial z} = 0$, поле описывается тремя компонентами $E_z(x,y)$, $H_x(x,y)$, $H_y(x,y)$). Как известно [6], намагниченный феррит характеризуется скалярной величиной диэлектрической проницаемости ε и тензорной величиной магнитной проницаемости $\tilde{\mu} = \begin{pmatrix} \mu & i\cdot\mu_a & 0 \\ -i\cdot\mu_a & \mu & 0 \\ 0 & 0 & \mu_{II} \end{pmatrix}$.

Решение задачи проведем широко применяемым методом частичных областей [7]. Особенностью применения этого метода в данной работе является выделение области связи по аналогии с тем, как это сделано в [8]. Разобьем продольное сечение волновода на три частичные области: область I (область связи) - $0 \le x \le a$, $0 \le y \le b$; область II - $0 \le x \le a, b \prec y \prec \infty$; область III - $-a \prec x \prec \infty, 0 \le y \le b$. Единственная электрическая компонента поля в каждой из частичных областей должна удовлетворять волновому уравнению

$$\frac{\partial^2 E_z(x,y)}{\partial x^2} + \frac{\partial^2 E_z(x,y)}{\partial y^2} + (\frac{2\cdot\pi}{\lambda})^2 \cdot \varepsilon \cdot \mu_\perp \cdot E_z(x,y) = 0,$$

где $\mu_\perp = \mu - \frac{\mu_a^2}{\mu}$ - эффективная магнитная проницаемость поперечно намагниченного феррита; λ - длина волны в свободном пространстве, граничным условиям ($E_z(x,y) = 0$ на металлических поверхностях) и условию излучения (области II и III запредельны). Представим ее в виде:

$$E_z^I(x,y) = \sum_{n=1,2,3}^{\infty} \frac{\Psi_{nb}(y)}{\alpha_n} \cdot [A_n^+ \cdot e^{i\cdot\Gamma_{nb}\cdot x} + A_n^- \cdot e^{-i\cdot\Gamma_{nb}\cdot x}] +$$

$$+ \sum_{m=1,2,3}^{\infty} \frac{\Psi_{ma}(x)}{\beta_m} \cdot [B_m^+ \cdot e^{i\cdot\Gamma_{ma}\cdot y} + B_m^- \cdot e^{-i\cdot\Gamma_{ma}\cdot y}],$$

$$E_z^{II}(x,y) = \sum_{m=1,2,3}^{\infty} \Psi_{ma}(x) \cdot C_m \cdot e^{-\gamma_{ma}\cdot(y-b)},$$

$$E_z^{III}(x,y) = \sum_{n=1,2,3}^{\infty} \Psi_{nb}(y) \cdot D_n \cdot e^{-\gamma_{nb}\cdot(x-a)}.$$

где $A_n^{\pm}, B_m^{\pm}, C_m, D_n$ - амплитудные коэффициенты разложения поля;

$\Psi_{ma}(x) = \sin(\frac{m \cdot \pi}{a} \cdot x)$, $\Psi_{nb}(y) = \sin(\frac{n \cdot \pi}{b} \cdot y)$ - собственные функции электрического поля волноводов, образующих излом;

$\alpha_n = \sin(\Gamma_{nb} \cdot a)$, $\beta_m = \sin(\Gamma_{ma} \cdot b)$ - нормировочные множители;

$\Gamma_{ma} = \sqrt{(\frac{2 \cdot \pi}{\lambda})^2 \cdot \varepsilon \cdot \mu_\perp - (\frac{m \cdot \pi}{a})^2}$, $\Gamma_{nb} = \sqrt{(\frac{2 \cdot \pi}{\lambda})^2 \cdot \varepsilon \cdot \mu_\perp - (\frac{n \cdot \pi}{b})^2}$ - постоянные распространения (действительные величины для распространяющихся волн и мнимые величины для затухающих волн);

$\gamma_{ma} = \sqrt{(\frac{m \cdot \pi}{a})^2 - (\frac{2 \cdot \pi}{\lambda})^2}$, $\gamma_{nb} = \sqrt{(\frac{n \cdot \pi}{b})^2 - (\frac{2 \cdot \pi}{\lambda})^2}$ - постоянные затухания (всегда действительные величины).

Отличные от нуля компоненты магнитного поля выражаются через компоненту электрического поля $E_z(x, y)$ следующим образом:

$$H_x(x, y) = \frac{1}{(2 \cdot \pi/\lambda) \cdot \mu_\perp} \cdot (-\frac{\mu_a}{\mu} \cdot \frac{\partial E_z(x, y)}{\partial x} + i \cdot \frac{\partial E_z(x, y)}{\partial y}),$$

$$H_y(x, y) = \frac{1}{(2 \cdot \pi/\lambda) \cdot \mu_\perp} \cdot (-\frac{\mu_a}{\mu} \cdot \frac{\partial E_z(x, y)}{\partial y} - i \cdot \frac{\partial E_z(x, y)}{\partial x}).$$

После применения условия непрерывности тангенциальных компонент электрического и магнитного полей на границах раздела частичных областей для получения системы функциональных уравнений используется условие биортогональности [5]:

$$\int_S \{[\vec{E}_n, \vec{H}_m^+] + [\vec{E}_m^+, \vec{H}_n]\} d\vec{s} = 0.$$

Процедуру применения данного условия поясним для границы раздела I и II областей. Условия непрерывности при $y = b$ имеют вид $E_z^I(x, y) = E_z^{II}(x, y)$, $H_x^I(x, y) = H_x^{II}(x, y)$. Домножаем первое уравнение на $\Phi_{ka}^{\pm}(x)$ ($\Phi_{ka}^{\pm}(x) = \pm \frac{i \cdot \Gamma_{ka}}{\mu_\perp} \cdot \Psi_{ka}(x) + i \cdot \frac{\mu_a}{\mu \cdot \mu_\perp} \cdot \frac{\partial \Psi_{ka}(x)}{\partial x}$ - собственные функции магнитного поля для прямых и обратных волн в волноводе шириной a), второе – на $\Psi_{ka}(x)$, складываем полученные уравнения и интегрируем по x в пределах от 0 до a. При этом учитываем интегралы, которые являются условиями биортогональности для рассматриваемой структуры:

$$\int_0^a (\Psi_{ma}(x) \cdot \Phi_{ka}^{\pm}(x) + \Psi_{ka}(x) \cdot \Phi_{ma}^{\pm}(x)) \cdot dx = \pm \frac{i \cdot \Gamma_{ka} \cdot a}{\mu_\perp} \cdot \delta_{mk},$$

$$\int_0^a (\Psi_{ma}(x) \cdot \Phi_{ka}^{-}(x) + \Psi_{ka}(x) \cdot \Phi_{ma}^{+}(x)) \cdot dx = 0,$$

где $\delta_{mk} = \begin{cases} 1, m = k \\ 0, m \neq k \end{cases}$.

Проводя математические преобразования получаем следующую сдвоенную бесконечную систему линейных алгебраических уравнений (СЛАУ) относительно неизвестных амплитуд A_n^+ и B_m^+:

$$\begin{cases} \dfrac{2\cdot\pi^2\cdot k\cdot(-1)^k}{a^2\cdot b}\cdot\sum\limits_{n=1,2,3}^{\infty}\dfrac{n\cdot(-1)^n\cdot A_n^+}{(\dfrac{k\cdot\pi}{a})^2-\Gamma_{nb}^2}= \\ =\sum\limits_{m-1,2,3}^{\infty}\{[\Gamma_{ka}\cdot ctg(\Gamma_{ka}\cdot b)+\gamma_{ka}\cdot\mu_\perp]\cdot\delta_{mk}-i\cdot\dfrac{\mu_a}{\mu}\cdot\dfrac{k\cdot\pi}{a}\cdot T_{mk}\}\cdot B_m^+ \\ \sum\limits_{n-1,2,3}^{\infty}\{[\Gamma_{kb}\cdot ctg(\Gamma_{kb}\cdot a)+\gamma_{kb}\cdot\mu_\perp]\cdot\delta_{nk}+i\cdot\dfrac{\mu_a}{\mu}\cdot\dfrac{k\cdot\pi}{b}\cdot T_{nk}\}\cdot A_m^+= \end{cases}$$

$$=\frac{2\cdot\pi^2\cdot k\cdot(-1)^k}{b^2\cdot a}\cdot\sum_{m=1,2,3}^{\infty}\frac{m\cdot(-1)^m\cdot B_m^+}{(\dfrac{k\cdot\pi}{b})^2-\Gamma_{ma}^2}$$

где $T_{mk}=\begin{cases}0, при (m+k)-четное \\ \dfrac{4\cdot m}{\pi\cdot(m^2-k^2)}-при(m+k)-нечетное\end{cases}$

Приравнивая нулю определитель СЛАУ, получаем уравнение для нахождения собственных длин волн низшего магнитного вида колебаний анализируемой структуры. Правильность полученного выражения проверяется предельным переходом $\mu_\perp \to 1$, $\mu_a \to 0$ (т.е. к случаю диэлектрического заполнения излома волновода). Получаемая при этом СЛАУ совпадает со СЛАУ из [9], описывающей собственные колебания вида H_{220} диэлектрического резонатора в волноводном разветвлении.

Заметим, что в задачах о волноводных структурах с намагниченным ферритом нельзя ограничиваться выделением четных и нечетных видов колебаний, как это обычно делается в задачах о структурах с диэлектриками (например, [9]). Наличие недиагональных элементов тензора магнитной проницаемости $\vec{\mu}$ приводит к связи четных и нечетных видов колебаний. Этот факт несколько усложняет электродинамический анализ.

Численное исследование скорости сходимости алгоритма при увеличении порядка определителя (числа учитываемых волн) показало, что она остается высокой практически во все диапазоне изменения параметра $\mu_\perp$. Причина этого, как будет показано ниже, заключается в неизменности характера колебаний в данном диапазоне. Одноволновое приближение (N=1) соответствует решению определителя второго порядка, двухволновое – четвертого, трехволновое – шестого и так далее.

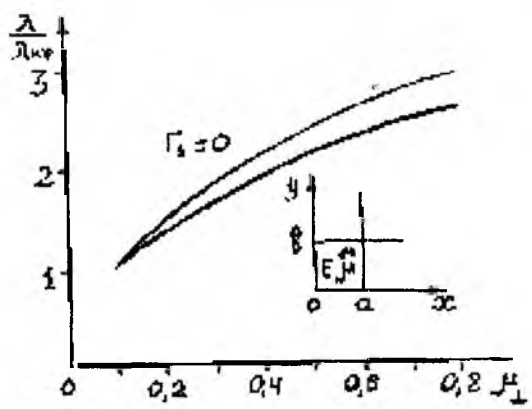


Рис. 1

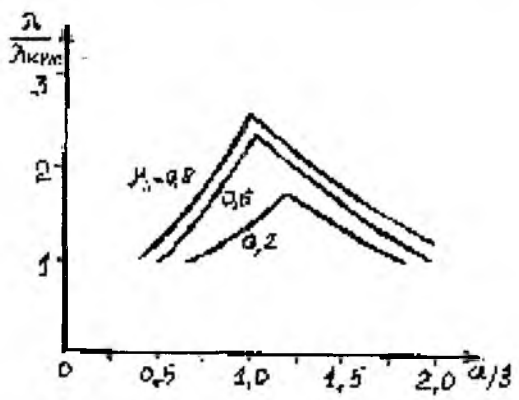


Рис. 2

На рис. 1 приведены рассчитанные для феррита марки 3СЧ6 зависимости нормированных собственных длин волн $\frac{\lambda}{\lambda_{кр}}$ *(где* $\lambda_{кр}=2\cdot a$ *-критическая длина волны незаполненного волновода)*

низшего магнитного вида колебаний излома от степени намагничения феррита (которую характеризует параметр $\mu_\perp$) для пяти учитываемых волн при $\frac{a}{b}=1$. Здесь же нанесена кривая $\Gamma_1(\mu_\perp)=\sqrt{\varepsilon\cdot\mu_\perp}=0$, иллюстрирующая момент перехода основной волны волноводного типа $(\Gamma_1^2 \succ 0)$ в поверхностную волну $(\Gamma_1^2 \prec 0)$ [10]. В [9] для колебания H_{220} поля диэлектрического резонатора в крестообразном волноводном разветвлении была выделена точка $\varepsilon \approx 1{,}11$ перехода затухающих колебаний в области с диэлектриком в распространяющиеся. Для феррита марки 3СЧ6 эта точка соответствует $\mu_\perp \approx \frac{1{,}11}{11}=0{,}1$. Вычисления подтвердили эту точку перехода. При уменьшении $\mu_\perp$ от значения 0,1 условия существования низшего магнитного вида колебаний рассматриваемой структуры перестают выполняться. Таким образом, практически во всем возможном диапазоне изменения величины внешнего магнитного поля низший магнитный вид колебаний излома находится в диапазоне распространяющегося основного типа волны.

На рис.2 построены зависимости $\frac{\lambda}{\lambda_{кр\,\max}}$ ($\lambda_{кр\,\max}$ - максимальная критическая длина волны незаполненных волноводов) от отношения $\frac{a}{b}$ (степень несимметрии) для различных значений параметра $\mu_\perp$. Увеличение величины внешнего магнитного поля приводит к нарушению условий существования собственных колебаний при увеличении степени несимметрии излома.

Полученные в работе результаты являются теоретической основой для построения малогабаритных электрически перестраиваемых частотно-избирательных систем в изломах волноводных трактов прямоугольного поперечного сечения, а также неразрушающего измерения (контроля) параметров ферритовых образцов на СВЧ.

Список литературы: **1.** *Коробкин В.А., Пятак Н.И., Двадненко В.Я. и др.* Перестраиваемые ферритовые фильтры на основе волноводно-диэлектрических резонаторов // Известия вузов. Радиоэлектроника. 1983. Т. 26. № 1.- С. 25-31. **2.** *Шестопалов В.П., Кириленко А.А., Масалов С.А., Сиренко Ю.К.* Резонансное рассеяние волн. В 2-х томах. Киев: Наукова думка, 1986. **3.** *Капилевич Б.Ю., Трубехин Е.Р.* Волноводно-диэлектрические фильтрующие структуры: Справочник. Москва: Радио и связь, 1990. 272 с. **4.** *Кулаков О.В., Пятак Н.И.* Электромагнитные колебания поперечно намагниченного ферритового параллелепипеда в прямоугольном волноводе // Известия вузов. Радиоэлектроника. 1995. Т. 38. № 2. С. 51-56. **5.** *Walker L.R.* Orthogonality relation for gyrotropic wave guides // Journal of Applied Physics. 1957. Vol. 28. No. 3. P. 377. **6.** *Гуревич А.Г.* Магнитный резонанс в ферритах и антиферромагнетиках. Москва: Наука, 1973. 591 с. **7.** *Миттра Р. Ли С.* Аналитические методы теории волноводов. Москва: Мир, 1974. **8.** *Пятак Н.И., Кулаков О.В.* Собственные колебания поля намагниченного ферритового резонатора в крестообразном волноводном разветвлении // Известия вузов. Радиоэлектроника. 1994. Т. 37. № 10. С. 58-65. **9.** *Коробкин В.А., Осинцев В.В.* Собственные электромагнитные колебания поля диэлектрического резонатора в волноводном разветвлении // Радиотехника и электроника. 1985. Т. 30. Вып. 3. С.417-421. **10.** *Микаэлян А.Л.* Теория и применение ферритов на сверхвысоких частотах.- Москва: Госэнергоиздат, 1963. 664 с.

Харьковский Национальный университет им. В.Н.Каразина

Поступила в редколлегию 07.02.2000

УДК 621.376.54

Н.Р. ПОПОВ, канд. техн. наук, М.В. ГУНБИН, канд. техн. наук, А.И. ГАПОН, канд. техн. наук, П.А. КАЧАНОВ, канд. техн. наук

АНАЛИЗ ДВУХКОНТУРНОЙ СИСТЕМЫ ТЕРМОСТАБИЛИЗАЦИИ УСТРОЙСТВ РАДИОЭЛЕКТРОННОЙ АППАРАТУРЫ

Эксплуатация систем термостабилизации показала, что применение кварцевого резонатора с большим температурным коэффициентом частоты (ТКЧ) в качестве датчика температуры обеспечивает стабильность температуры в термостате с точностью $10^{-3^{\circ}}C$ в течение длительного времени при изменении температуры окружающей среды в широких пределах. Выбором определенного угла среза кварца достигается линейная зависимость частоты датчика f_{∂} от температуры в широком диапазоне изменения частоты. В существующих одноконтурных системах термостабилизации, использующих пьезокварцевые датчики температуры, не реализуются возможности пьезодатчиков по точности стабилизации температуры. Точность системы термостабилизации можно значительно повысить, если для этой цели применить двухконтурную систему стабилизации. В этом случае объект управления системы термостабилизации можно представить в виде, изображенном на рис. 1,а, а его одномерная эквивалентная модель на рис. 1,б, где резисторы отражают тепловые сопротивления участков, а конденсаторы – теплоемкости этих участков. Здесь $R_к, R_т, R_н, R_с, R_в; C_к, C_т, C_н, C_с, C_в$ - соответственно тепловые сопротивления и теплоемкости кожуха, термоизолятора, нагревателей, стенок камер и воздуха. Для одного из конкретных объектов управления их значения приведены в таблице.

i	1	2	3	4	5	6	7	8	9	10
R_i, Ом	14,4	6,5	$11,86 \cdot 10^{-4}$	6,157	5,788	$10,27 \cdot 10^{-4}$	10,684	9,369	$2,2 \cdot 10^{-4}$	40
C_i, UF	0,47	0,145	63,86	59,68	22,33	73,7	68,74	54,62	138,4	-

Структурная схема двухконтурной системы термостабилизации изображена на рис. 2 и состоит из каналов грубого и точного терморегулирования. Пьезокварцевый датчик температуры I возбуждается в схеме автогенератора 2, с выхода которого напряжение через формирователь длительности и амплитуды импульсов поступает на схему сравнения частот 4. На другой вход схемы сравнения частот поступает через формирователь импульсов напряжение опорного генератора 3, стабилизированного кварцевым резонатором. Схема сравнения частот вырабатывает импульсы длительностью $\tau_и = const$ с периодом повторения

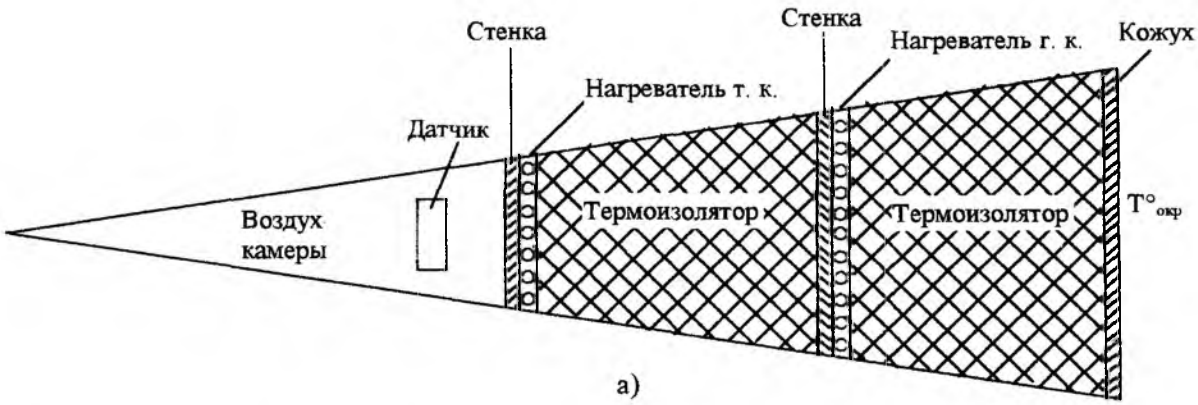
$$T_i = F^{-1}, \tag{1}$$

где $F = f_{оп} - f_{\partial}$ - разность частот опорного генератора и датчика соответственно. При изменении в процессе терморегулирования разности частот F меняется скважность импульсов, а, следовательно, средняя мощность нагревателя канала грубого терморегулирования. Импульсные напряжения генераторов датчика и опорного подаются также на схему определения знака разности их частот. Таким образом, в этом контуре управления осуществляется частотно-импульсная модуляция [1].

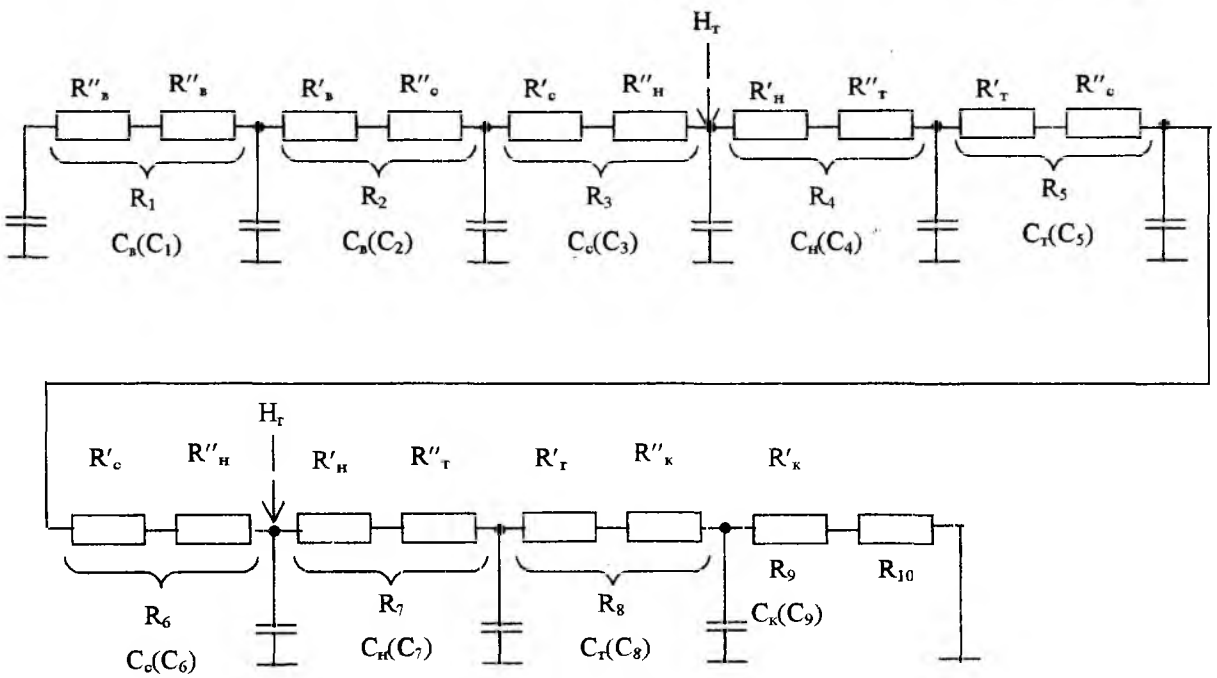
Если в процессе нагрева термостата отношение

$$F \cdot f_{оп}^{-1} \le 2 \cdot 10^{-6}, \tag{2}$$

то через цепь синхронизации происходит автозахват частоты генератора датчика частотой опорного генератора, и далее в зоне захвата частоты регулирование температуры осуществляется точным контуром по величине рассогласования фаз с последующим формированием широтно-импульсного сигнала, поступающим на нагреватель точного канала. Такой метод измерения малой разности частоты датчика и опорного генератора с преобразованием в фазовый сдвиг φ называют фазогенераторным. Он позволяет значительно повысить точность измерения разности частот. Эти преобразователи не являются интегрирующими, поэтому утверждение об астатическом терморегулировании по частоте является необоснованным [2].


Стенка
Стенка
Нагреватель г. к.
Кожух
Нагреватель т. к.
Датчик
Воздух камеры
Термоизолятор
Термоизолятор
T°окр

а)


Hт
R″в R″в R′в R″с R′с R″н R′н R″т R′т R″с
R1 R2 R3 R4 R5
Св(C1) Св(C2) Сс(C3) Сн(C4) Ст(C5)
Hг
R′с R″н R′н R″т R′г R″к R′к
R6 R7 R8 R9 R10
Сс(C6) Сн(C7) Ст(C8) Ск(C9)

б)

Рис. 1

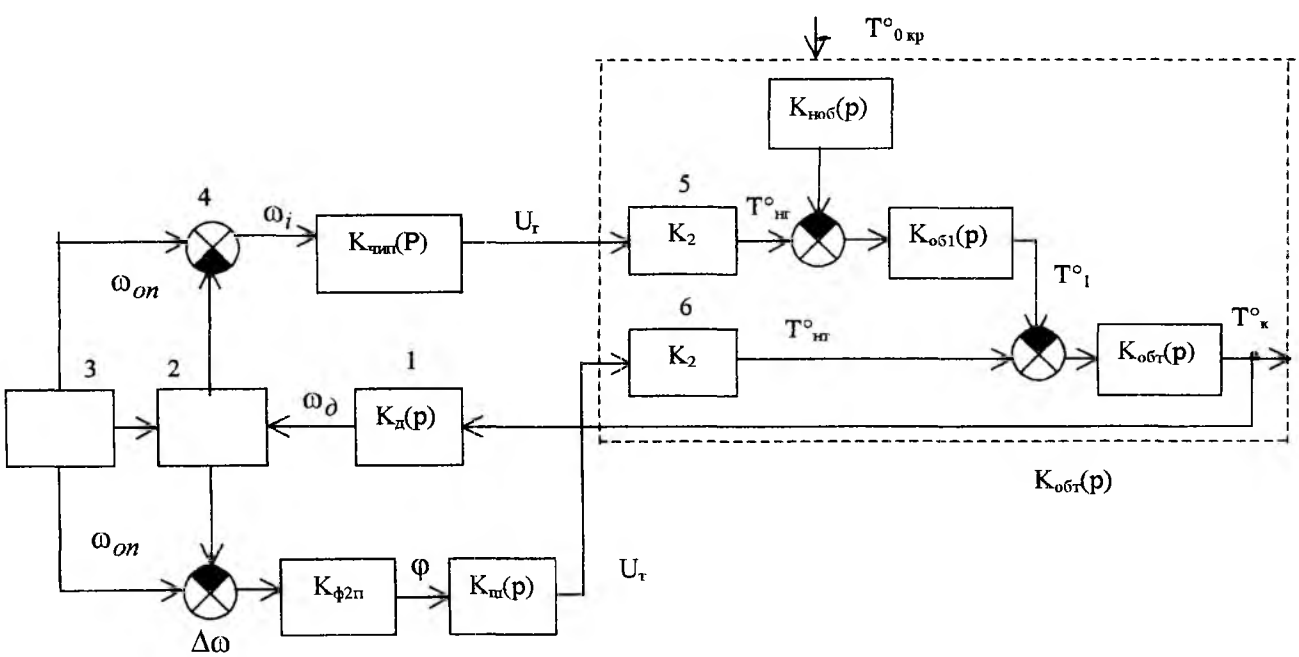

T°0 кр
Kнаб(p)
4
ωi
Kчип(P)
Uт
5
K2
T°нг
Kоб1(p)
T°1
6
K2
T°нг
Kобт(p)
T°к
3
2
1
ωд
Kд(p)
Kобт(p)
ωоп
ωоп
Kф2п
φ
Kп(p)
Uт
Δω

Рис. 2

Рассмотрим статические и динамические свойства двухконтурной системы термостабилизации.

Динамические характеристики ОУ в соответствии с рис. 1,б представлены тремя участками, передаточные функции которых имеют вид:

участок кожух – нагреватель грубого канала

$$K_{ноб}(p) = 1/[T_7T_8T_9p^3 + (T_8T_9 + T_7T_9 + T_7T_8 + T_9R_8C_7 + T_7R_9C_8)p^2 + (T_9 + T_8 + T_7 + R_9C_8 + R_9C_7 + R_8C_7)p + 1] = \frac{1}{(231909p^3 + 387838p^2 + 1925p + 1)}, \quad (3)$$

где $T_7 = R_7C_7; T_8 = R_8C_8; T_9 = R_9C_9;$

участок нагреватель грубого канала – нагреватель точного канала

$$K_{об1}(p) = 1/[T_4T_5T_6p^3 + (T_5T_6 + T_4T_6 + T_4T_5 + T_6R_5C_4 + T_4R_6C_5)p^2 + (T_6 + T_5 + T_4 + R_6C_5 + R_5C_4)p + 1] = \frac{1}{(360{,}86p^3 + 47495{,}6p^2 + 842{,}94p + 1)}, \quad (4)$$

где $T_6 = R_6C_6; T_5 = R_5C_5; T_4 = R_4C_4;$

участок нагреватель точного канала – датчик

$$K_{ОБТ}(p) = 1/[T_1T_2T_3p^3 + (T_2T_3 + T_1T_3 + T_2T_3 + T_3R_2C_1 + T_1R_3C_2)p^2 + (T_1 + T_2 + T_3 + R_3C_2 + R_3C_1 + R_2C_1)p + 1] = \frac{1}{(0{,}45p^3 + 6p^2 + 422p + 1)}, \quad (5)$$

где $T_3 = R_3C_3; T_2 = R_2C_2; T_1 = R_1C_1.$

Передаточная функция пьезокварцевого датчика температуры определена по экспериментальной кривой разгона

$$K_д(p) = K_4(1 + 16p)^{-1}, \quad (6)$$

где $K_4 = \frac{f_д}{T^\circ_д}$ - статический коэффициент преобразования датчика (ТКЧ).

Передаточная функция формирующего звена схемы сравнения частот (ЧИП) в грубом канале [I]

$$K_{ЧИП}(p) = K_4E\frac{1 - \ell^{\tau_u p}}{p}\ell^{-pT_i}, \quad (7)$$

где τ_u, T_i, E - соответственно длительность импульса, период следования и амплитуда выходных импульсов схемы сравнения частот.

При $f_{on} = 5{,}1 \cdot 10^6$ Гц минимальная разностная частота (перед моментом автозахвата) может быть в пределах $F = 5{,}1 \div 10$ Гц или $\omega_i = 31 \div 63c^{-1}$. Таким образом, минимальная частота дискретизация в ЧИП грубого канала равна

$$\omega_i = 31 \div 63c^{-1}. \quad (8)$$

Фазогенераторный преобразователь (измеритель малой девиации частоты), как показано в [2], при такой величине f_{on} может быть представлен безинерционным звеном с коэффициентом преобразования

$$K_{ФГП} = \frac{-2Q(m+1)}{\omega_{до}}, \quad (9)$$

где Q - эквивалентная добротность резонансной схемы синхронизации;

m - коэффициент связи генераторов;

$\omega_{до}$ - частота собственных колебаний синхронизированного генератора датчика.

Передаточная функция формирующего звена широтно-импульсного преобразователя, построенного на основе фазогенераторного преобразователя [2]

$$K_{Ш}(p) = \frac{E}{2\pi} \cdot \frac{1-\ell^{-\tau_\varphi p}}{\tau_\varphi p}, \tag{10}$$

где $\tau_\varphi = \frac{\varphi}{2\pi} T$ - длительность выходных импульсов широтно-импульсного модулятора;

T - период дискретизации, равный периоду промежуточной частоты f_n

$$T = \frac{1}{f_n} = \frac{1}{9 \cdot 10^3} = 1{,}1 \cdot 10^{-4} c, \tag{11}$$

при этом

$$\omega_n = 2\pi f_n = 56{,}52 \cdot 10^3 c^{-1}. \tag{12}$$

Поскольку объект управления имеет постоянные времени, значительно превышающие частоты дискретизации в ЧИП и ШИП, то исследование системы, в отличие от [3] , произведем непрерывными методами.

Для анализа устойчивости системы термостабилизации определим устойчивость грубого и точного каналов, для чего определим передаточные функции разомкнутых и замкнутых контуров:

грубого канала

$$K_г(p) = \frac{K_г}{(360{,}86p^3 + 47495{,}6p^2 + 842{,}94p + 1)(1+16p)(0{,}45p^3 + 6p^2 + 422p + 1)} = \frac{B_г(p)}{A_г(p)}; \tag{13}$$

$$K_{ог}(p) = \frac{B_г(p)}{A_г(p) + B_г(p)} = \frac{B_г(p)}{C_г(p)}; \tag{14}$$

$$C_г(p) = A_г(p) + B_г(p), \tag{15}$$

где $K_г = K_д K_{ЧИП} K_2 K_{обг}; K_{обг} = K_{об1} K_{ОБТ} = 1;$

точного канала

$$K_T(p) = \frac{K_T}{(1+16p)(0{,}45p^3 + 6p^2 + 422p + 1)} = \frac{B_T(p)}{A_T(p)}; \tag{16}$$

$$K_{OT}(p) = \frac{B_T(p)}{A_T(p) + B_T(p)} = \frac{B_T(p)}{C_T(p)}; \tag{17}$$

$$C_T(p) = A_T(p) + B_T(p), \tag{18}$$

где $K_T = K_д K_{ФГП} K_2 K_{ОБТ}; K_{ОБТ} = 1.$

Легко показать, что при данных параметрах статические критические коэффициенты усиления разомкнутых контуров из условия нахождения этих контуров на границе устойчивости равны

$$K_{гкр} = 18; K_{Tкр} = 2897.$$

Отсюда видно, что, с учетом запаса устойчивости по модулю, соответствующие коэффициенты могут быть выбраны по величине

$$K_г = 10; K_T = 1000.$$

Для анализа точности системы термостабилизации по возмущающему воздействию – температуре окружающей среды $T^\circ_{окр}(t)$ - определим изображение ошибки точного канала

$$\Delta T^\circ_T(p) = -\frac{\Delta T^\circ_г(p)}{1 + K_T(p)} = \frac{T^\circ_{окр}(p) K_{ноб}(p) K_{об1}(p) K_{ОБТ}(p)}{[1 + K_г(p)][1 + K_T(p)]}, \tag{19}$$

где $\Delta T^\circ_г(p) = -\frac{T^\circ_{окр}(p) K_{ноб}(p) K_{об1}(p) K_{ОБТ}(p)}{1 + K_г(p)} = T_1(p)$ - изображение ошибки грубого

канала от возмущающего воздействия. $T^\circ_1(t)$ является возмущающем воздействием для точного канала.

Изображение по Лапласу возмущающего воздействия в статическом режиме имеет вид

$$T^\circ_{окр}(p) = \frac{T^\circ_{окр}}{p}. \tag{20}$$

Абсолютная величина статической ошибки системы определяется с использованием теоремы о конечном значении функции

$$\Delta T^\circ_{Tc} = \lim_{p\to 0} p \cdot \Delta T^\circ_T(p) = -\frac{T^\circ_{окр}}{(1+K_г)(1+K_T)}. \tag{21}$$

Из полученного выражения статической ошибки следуют требования к выбору необходимых коэффициентов усиления разомкнутых контуров для обеспечения ее допустимого значения в зависимости от величины воздействия $T^\circ_{окр}$.

Исходя из вышесказанного, следует, что относительная ошибка системы

$$\varepsilon = \frac{\Delta T^\circ_{Tc}}{T^\circ_{окр}} = \frac{1}{(1+K_г)(1+K_T)} \tag{22}$$

может иметь порядок не хуже 10^{-4}.

При изменении температуры по линейному закону $T^\circ_{окр}(t) = \Omega t$ со скоростью нарастания Ω для определения установившегося значения динамической (кинетической) ошибки удобнее воспользоваться выражениями, полученными с помощью коэффициентов ошибки, полученных при разложении передаточной функции ошибки $K_{\varepsilon T}(p)$ в ряд Маклорена

$$K_{\varepsilon T}(p) = \frac{K_{ОБТ}(p)}{1+K_T(p)} = S_{oT} + \frac{S_{1T}}{1!}p + \frac{S_{2T}}{2!}p^2 + \ldots, \tag{23}$$

где $S_{oT} = \frac{1}{(1+K_T)}; \frac{S_{1T}}{1!} = \frac{16K_T - 422}{(1+K_T)^2}$.

Поскольку ошибка грубого канала $\Delta T^\circ_г(t)$ является воздействием для точного канала, то выражение кинетической ошибки точного канала имеет вид

$$\Delta T^\circ_T(t) = S_{oT}\Delta T^\circ_г(t) + \frac{S_{1T}}{1!}(\Delta T^\circ_г(t))' = \frac{1}{(1+K_г)(1+K_T)} \cdot \left[\Omega t + \Omega\left(\frac{-1909K_г - 1265}{1+K_г} + \frac{16K_T - 422}{1+K_T}\right)\right], \tag{24}$$

где

$$\Delta T^\circ_г(t) = S_{oг}T^\circ_{окр}(t) + \frac{S_{1г}}{1!}(T^\circ_{окр}(t))' = \frac{1}{1+K_г}\Omega t + \frac{-1909K_г - 1265}{(1+K_г)^2}\Omega. \tag{25}$$

Из выражения кинетической ошибки можно получить допустимую скорость изменения температуры окружающей среды.

Список литературы: **1.** *Перепелкин С.Р., Попов Н.Р., Гунбин М.В., Кадулин В.И.* Математическая модель частотно-импульсного преобразователя с пьезокварцевым резонатором // Системы сбора и обработки измерительной информации. 1985. Вып. 6. С. 96-100. **2.** *Перепелкин С.Р., Попов Н.Р., Гунбин М.В., Кадулин В.И.* Математическая модель широтно-импульсного модулятора на основе фазо-генераторного преобразователя // Вестник Харьк. политех. ин-та. Автоматика и приборостроение. 1985. № 221. Вып. 11. С. 16-19. **3.** *Воронов В.Г., Перепелкин С.Р., Попов Н.Р.* Анализ устойчивости системы термостабилизации с пьезокварцевым датчиком // Системы сбора и обработки измерительной информации. 1985. Вып. 6. С. 87-96.

Харьковский государственный политехнический университет

Поступила в редколлегию 14.02.2000

УДК 621.383.52

В.А. АНТОНОВА, канд. техн. наук, В.Н. БОРЩЕВ, д-р техн. наук,
А.М. ЛИСТРАТЕНКО, Н.И. СЛИПЧЕНКО, канд. техн. наук

КОНСТРУКТИВНО-ТЕХНОЛОГИЧЕСКИЕ ОГРАНИЧЕНИЯ КРЕМНИЕВЫХ МИКРОПОЛОСКОВЫХ PIN ПРИЕМНИКОВ ИЗЛУЧЕНИЯ

В начале 80-х годов получил развитие новый класс полупроводниковых детекторов – микрополосковые детекторы. Их появление было стимулировано потребностями физики высоких энергий и элементарных частиц. Уникальные возможности сверхточных измерений координат с помощью таких детекторов обусловили их быстрое развитие и появление большого числа перспективных разработок как в области самих координатных детекторов, так и в области специализированной микроэлектроники к ним [1,2].

Практически все используемые в настоящее время в физике элементарных частиц полупроводниковые детекторы выполнены на основе кремния, в том числе и микрополосковые детекторы. Микрополосковый приемник излучений представляет собой набор *p-n* переходов в виде узких параллельных полос, которые формируются в исходных кремниевых высокоомных и особо чистых пластинах (обычно *n*-типа с удельным сопротивлением 4 – 8 кОм*см). Шаг полос определяется задачами эксперимента и обычно составляет от 25 до 100 мкм.

Пространственное разрешение кремниевого микростripового приемника зависит от того, в каком качестве он используется. Если он используется только как датчик соударений, то пространственное разрешение связано с шагом полос P_s соотношением:

$$\sigma = P_s / \sqrt{12} . \qquad (1)$$

При аналоговом считывании разрешение зависит от нескольких параметров. Наиболее важными являются электронный шум, шаг полос, шаг считывания и напряжение смещения, определяющее горизонтальную диффузию зарядов. Горизонтальная диффузия распределяет заряд, создаваемый пролетающей частицей, по нескольким стрипам.

Используя алгоритмы нелинейной зарядовой интерполяции, можно получить оптимальное разрешение, которое в конечном счете ограничено только шумом и физическими процессами создания заряда [3]. Пространственное разрешение в этом случае описывается выражением:

$$\sigma = \beta\left(N/S\right)^{*} P_s , \qquad (2)$$

где S/N – отношение сигнал – шум; P_s – шаг полос; β – эмпирический коэффициент пропорциональности, значение которого лежит в пределах 4 ÷ 10.

Как видно из выражения (2), существенное влияние на пространственное разрешение микрополоскового приемника оказывают его шумовые характеристики. Можно выделить наиболее важные параметры приемника, которые влияют на шумовую характеристику считывающей электроники:

– емкость считывающей полосы по отношению к земле или виртуальной земле соседних микрополос;

– ток утечки (дробовой шум);

– номинал резистора смещения.

Емкость является определяющим фактором для последовательного шума, а ток утечки и номинал резистора смещения вносят основной вклад в параллельный шум [4]. Из сказанного выше можно определить требования к конструктивно-технологическим решениям кремниевых микрополосковых приемников:

– конструкция приемника должна обеспечивать минимально возможный шум считывающей электроники, для чего необходимы малая межполосковая емкость, низкие токи утечки и максимально возможное значение сопротивления смещения. Однако сопротивление смещения не должно превышать величин, которые приводят к изменению падения напряжения между соседними полосами, чтобы избежать ошибок в определении координат из-за горизонтального градиента электрического поля;

– для получения высокого разрешения также и для низкоимпульсных частиц многократное рассеяние должно быть минимальным. Это требует уменьшения толщины приемника до значений, ограничиваемых требованиями к механической прочности как при изготовлении кристалла, так и при его монтаже.

Последние технические достижения в производстве кремниевых микрополосковых детекторов связаны с разработкой схем согласования и схем смещения [5]. Разработаны и освоены промышленностью бескорпусные микросхемы, содержащие многоканальные малошумящие зарядово-чувствительные

усилители и коммутаторы последовательного вывода аналоговой информации на 64-128 каналов. При этом традиционно соединения между электронными компонентами, несущими платами и самим кремниевым детектором осуществляются с использованием проводов. Эта технология хорошо известна и относительно дешева. Однако она предъявляет жесткие требования к конструкции детекторных модулей, что делает невозможным их использование в некоторых системах, применяемых в новых экспериментах в физике частиц высоких энергий [6]. Проволочная коммутация позволяет соединять только близко расположенные части, механически зафиксированные друг относительно друга и находящиеся на одной и той же стороне. В системах слежения, где используется открытой вся плоскость детекторов, трудно расположить электронику предварительной обработки в одной плоскости. В этом случае используются промежуточные несущие платы, соединенные между собой кабелями. Однако проволочные соединения остаются хрупкими, а соединительные платы являются неактивной и нежелательной добавкой материала в системе.

В настоящее время проводятся работы по созданию детекторных модулей, в которых соединения между компонентами осуществляется с помощью пленочных шлейфовых микрокабелей [7]. Легкие гибкие алюминиевые микрокабели на полиимидном носителе, которые можно присоединять непосредственно к компонентам, позволяют сделать соединения трехмерными. Более того, т. к. устраняется необходимость в множестве дополнительных соединений при сборке детекторных модулей, ожидается значительное улучшение не только массо-габаритных характеристик, но и значительное повышение надежности детекторных модулей.

Целью данной работы является расчет и разработка топологии микрополоскового **PIN** приемника излучения, предназначенного для экспериментальных исследований и изучения влияния конструктивных особенностей и электрических характеристик пленочных микрокабелей на чувствительность, быстродействие и разрешающую способность кремниевых микрополосковых детекторных модулей нового поколения [6,7]. Кроме того, разрабатываемая топология должна обеспечить возможность отработки основных технологических процессов сварки, приклейки и последовательности операций сборки микрокабелей и других компонентов детекторного модуля для эксперимента ALICE [8].

В ходе выполнения работ по подготовке эксперимента ALICE разработана специальная аналого-цифровая микросхема A128C, рассчитанная на первичную обработку и передачу информации в приемный тракт со 128 микрополосковыми диодами полноразмерного кремниевого **PIN** приемника. Микросхема имеет 128 выходных контактных площадок, которые расположены в четыре ряда вдоль одной стороны кристалла. Шаг размещения контактных площадок в одном ряду составляет 88 мкм [9]. Для связи микросхемы A128C с микрополосковым кремниевым **PIN** приемником разработан микрокабель на полиимидном носителе со 128 алюминиевыми проводниками толщиной 14 мкм. В зоне присоединения к контактным площадкам микрополоскового приемника шаг проводников микрокабеля составляет 95 мкм. При этом ширина проводников равна 60 мкм, а зазор между проводниками 35 мкм. Шаг согласован с шагом микрополоскового полноразмерного кремниевого **PIN** приемника, который предполагается использовать в эксперименте ALICE[7].

Требования к топологии тестового кремниевого микростripового приемника излучений должны основываться на требованиях к основным характеристикам полноразмерного приемника излучений эксперимента ALICE. Другими словами, электрофизические характеристики тестового микростripового кремниевого приемника не должны уступать характеристикам полноразмерного приемника ALICE, в том числе, по самому важному параметру – темновому току утечки при напряжении полного обеднения. Только в этом случае можно будет иметь объективную информацию о влиянии конструктивных особенностей, электрических характеристик и технологии сборки и монтажа алюминиевых микрокабелей на электрофизические характеристики микрополосковых детекторных модулей.

Таким образом, разрабатываемая топология тестового микрополоскового приемника включает:

– количество микрополосков, шт. – 128;
– шаг микрополосков, мкм – 95;
– ширина микрополосков, мкм – 60;
– зазор между микрополосками, мкм – 35;
– длина микрополоска, мм – 12;
– толщина кремниевого кристалла, мкм – 300 ± 5;
– площадь приемной области, мм^2 – 12 x 12;
– размеры кристалла приемника, мм^2 – 15 x 15.

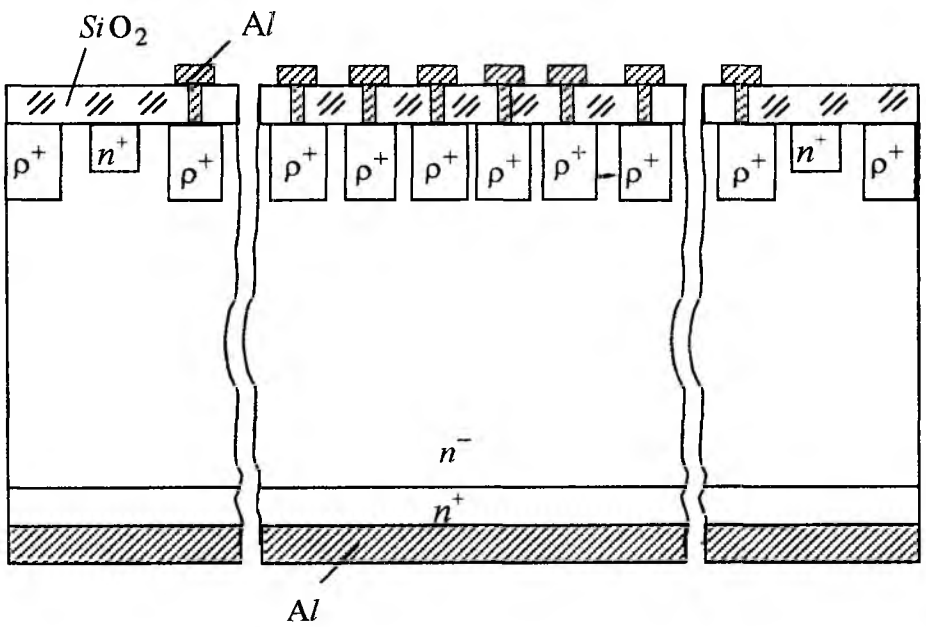
Приемная область представляет собой квадрат с размерами 12 х 12 мм. На погонной длине 12 мм укладывается 128 микрополосков с шагом 95 мкм. Такие размеры приемной области позволяют оптимизировать длину микрополосков как минимально возможную и равную 12 мм.

С использованием методики [10] проведен расчет токов утечки одного микрополоскового p-n – перехода при напряжении обратного смещения 70 В, обеспечивающем полное обеднение в кристалле кремния толщиной 300 мкм с концентрацией носителей $\sim 10^{12}$ см$^{-3}$. При выбранных геометрических размерах микрополоскового перехода 60 х 12000 мкм и напряжении обратного смещения 70 В суммарный темновой ток не превышает 0,1 нА, выходная емкость составляет ~0,15 пФ. Как следует из расчета, проведенного по методике [11], напряжение пробоя в наиболее слабых местах p - n – перехода в 2-3 раза превышает напряжение полного обеднения, таким образом топология микрополоскового приемника позволяет обеспечить его работу в условиях полного обеднения и при более высоких напряжениях.

Как известно, на поверхности высокоомной кремниевой подложки (особенно p-типа проводимости) возникают инверсионные слои, приводящие к возрастанию темновых токов утечки на несколько порядков. Для устранения этого нежелательного эффекта приемная площадка должна быть окружена охранным кольцом [12]. Электронно-дырочный переход охранного кольца должен быть изолирован от основной приемной области. Расстояние между охранным кольцом и крайними микрополосковыми p - n – переходами приемной области должно быть порядка ширины слоя объемного заряда при рабочем напряжении. Охранное кольцо должно иметь отдельный вывод. Благодаря такой конструкции инверсионный слой смыкается с охранным кольцом, в результате ток тепловой генерации в этом слое и ток утечки протекают через цепь кольца. В цепях приемной площадки, которая изолирована от охранного кольца слоем объемного заряда, протекает только ток носителей, генерированных в пределах приемной области.

Для снижения инверсионных токов нами применяется противоинверсионное стопорное кольцо. Ширина кольца и концентрация примеси в нем выбрана таким образом, чтобы носители заряда, диффундирующие через кольцо, рекомбинировали, то есть, ширина кольца должна быть больше диффузионной длины носителей заряда. В данной разработке использована также защита области реза кристалла приемника в виде диффузионного кольца вдоль периметра кристалла. Это позволило исключить возможность короткого замыкания и появления значительных обратных токов, а также снижения пробивного напряжения в случае касания выводов поврежденных защитно-пассивирующих покрытий на приемной стороне кристалла.

Топология кристалла кремниевого микрополоскового приемника излучений приведена на рисунке.



В высокоомной исходной кремниевой пластине n-типа с удельным объемным сопротивлением ρ_v = 4÷8 кОм* см с ориентацией (100) или (111) сформированы микрополосковые p^+– слои в виде решетки с шагом 95 мкм. Ширина диффузионных полос равна 60 мкм, длина 12000 мкм. Расстояние между диффузионными полосами 35 мкм. В пределах приемной области на погонной длине 12 мм расположе-

но 128 p - n – переходов, не связанных между собой электрически. Глубина p^+- слоев ~1,5÷1,8 мкм, поверхностная концентрация примеси ~ (1÷5)* 10^{19} $см^{-3}$. От крайних микрополосковых p - n – переходов и торцов p - n – переходов приемной области на расстоянии 75 мкм расположено охранное кольцо p^+– типа шириной 50 мкм с поверхностной концентрацией (1÷5)* 10^{19} $см^{-3}$.

Приемная область вместе с охранным кольцом окружена противоинверсионным стопорным кольцом шириной 800 мкм.

Стопорное кольцо, расположенное от охранного кольца на расстоянии 200 мкм, представляет собой n^+– слой с поверхностной концентрацией ~ 2*$10^{20}$$см^{-3}$. Вдоль периметра кристалла на расстоянии 200 мкм от стопорного кольца расположено защитное кольцо области реза, представляющее p^+– слой шириной 200 мкм с поверхностной концентрацией 1÷5)* $10^{19}$$см^{-3}$, которое формируется в одном технологическом цикле с микрополосковыми p-n – переходами и охранным полевым кольцом. С тыльной стороны кристалла формируется высоколегированный n^+ – слой, имеющий глубину ~ 2 мкм и поверхностную концентрацию ~ 2*$10^{20}$$см^{-3}$. К p^+– и n^+ – слоям сформированы металлические контакты на основе алюминия.

Разработанная топология кристалла кремниевого микрополоскового приемника излучений предполагает создание на каждом микрополоске связывающих емкостей. При толщине защитного слоя SiO_2 над p-n – переходом микрополоска ~ 0,1 мкм, связывающая емкость одиночного микрополоска составит 25-27 пФ.

Практическая реализация предложенных конструктивно-технологических решений позволит изготовить кремниевый микрополосковый приемник ионизирующих излучений на уровне лучших мировых образцов, а также разработать и изготовить микрополосковый детекторный модуль на основе самых последних достижений: с применением алюминиевых микрокабелей, микросхем первичной обработки типа ALICE 128C, гибридных микросборок с использованием новейших высокотеплопроводных подложек из AlN и пироуглепластиков.

Такой детекторный модуль может быть применен не только для проведения испытаний и исследования влияния конструктивных особенностей и электрических параметров пленочных микрокабелей на основные приемно-детектирующие характеристики, но и как базовый элемент для разработки и создания высокочувствительных систем с высоким разрешением в медицинских томографах, в других системах регистрации рентгеновского и различных видов ионизирующих излучений.

Список литературы: 1. *Чилингаров А.Г.* Координатные полупроводниковые детекторы в физике элементарных частиц. Новосибирск: препринт 90-113, 1990. 288с. **2.** *Лятидевский В.К.* Методы детектирования излучений. М.: Энергоатомиздат, 1987. 405 с. **3.** *Weilhammer P.* Double-sided Si strip sensors for LEP vertex // Nucl. Instr. and Meth.1994. A 342. P. 1-15. **4.** *Dabrowski W.* Charge division in silicon strip detectors with a large strip pitch // Nucl. Instr. and Meth. 1994. A 349. P. 424 - 430. **5.** *Lutz G.* Silicon radiation detectors // Nucl. Instr. and Meth. 1995. A 367. P. 21-33. **6.** *ALICE:*Technical Proposal for a Large Ion Collider Experiment at the CERN, LHCC/95-71 LHCC/P3, Geneva,1995. **7.** *ALICE:* Inner Tracking system / Technical Desing Report, CERN/LHCC 99-12, ALICE TDR 4,18 June 1999. **8.** *Haas A.P., Van den Brick A., Kuijer P., Oskamp C.J.*, NIKHEF, Utrecht, V.N. Borschov, S.K. Kiprich, V.M. Ruzhitsky SRTIIM, Kharkov. Very low mass microcables for the ALICE silicon strip detector, Proccedings of LEB 99, 1999. **9.** *Hebrard L.* Electrical characterisation of ALICE 128C: a low-power CMOS ASIC for the iadout of silicon Strip Detectors / CERN / LHCC / 98-36, Proceeding of the 4-th workshop on electronics for LHC experiments, Rome, september 21-25, 1998. **10.** *Антонова В.А., Борщев В.Н., Листратенко А.М., Слипченко Н.И.* К вопросу о новых конструктивно-технологических решениях при создании высокоэффективных фотоприемников большой площади // Радиотехника. 1999. Вып. 109. С. 114-120. **11.** *Beck S.A.,Carter A.A.* Junction depth dependence of breacdown in silicon detector diodes // Nucl. Instr. and Meth. 1996. A 373. P. 223, 226. **12.** *Тейлор П.* Расчет и проектирование тиристоров. М.: Энергоатомиздат, 1990. 208 с.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 22.11.99

СОДЕРЖАНИЕ

CONTENTS

УДК 681.3.06: 519.248.681

Улучшенный стандарт симметричного шифрования XXI века: концепция создания и свойства кандидатов / М.Ф. Бондаренко, И.Д. Горбенко, А.В. Потий, О.И. Олешко, С.А. Головашич, А.С. Бондаренко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 5-14.

Рассматривается концепция создания улучшенного стандарта шифрования. Обсуждаются требования к стандарту и накладываемые ограничения. Проводится сравнительный анализ алгоритмов-кандидатов AES с целью выбора наиболее подходящего.

Табл. 6. Ил. 0. Библиогр.: 3 назв.

УДК 681.3.06: 519.248.681

Поліпшений стандарт симетричного шифрування XXI сторіччя: концепція створення і властивості кандидатів / М.Ф. Бондаренко, І.Д. Горбенко, А.В. Потій, О.І. Олешко, С.А. Головашич, А.С. Бондаренко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 5-14.

Розглядається концепція створення поліпшеного стандарту шифрування. Обговорюються вимоги до стандарту і що накладаються обмеження. Здійснюється порівняльний аналіз алгоритмів-кандидатів AES із метою вибору найбільше підхожого.

Табл. 6. Іл. 0. Бібліогр.: 3 назв.

UDC 681.3.06: 519.248.681

The Advanced Encryption Standard of XXI centuries: the concept of creation and property of the candidates / M. Bondarenko, I. Gorbenko, A. Potiy, O Oleshko, S Golovashich, A. Bondarenko // Radiotechnics. All-Ukr. Sci. Interdep. Mag. 2000. №114. PP. 5-14.

In this paper the conception of developing Advanced Encryption Standard is highlighted. Requirements and restrictions applied to above mentioned standard are also under discussion. Comparative analysis of the AES' candidate algorithms is performed for the most acceptable one to be chosen.

6 tab. 0 fig. Ref.: 3 items.

УДК 681.3.06: 519.248.681

Сущность и результаты исследований свойств перспективных стандартов цифровой подписи X9.62-1998 и распределения ключей X9.63-199X на эллиптических кривых / М.Ф. Бондаренко, И.Д. Горбенко, Е.Г. Качко, А.В. Свинарев, Т.А. Гриненко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.15-24.

Рассматриваются основные стандарты криптографических преобразований на эллиптических кривых. Дается сравнительный анализ преобразований в полях и на эллиптических кривых. Приводятся результаты оценки стойкости преобразования на эллиптических кривых.

Табл. 10.

УДК 681.3.06: 519.248.681

Сутність і результати досліджень властивостей перспективних стандартів цифрового підпису X9.62-1998 і розподілу ключів X9.63-199X на еліптичних кривих / М.Ф. Бондаренко, І.Д. Горбенко, О.Г. Качко, А.В. Свинарьов, Т.О. Гріненко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.15-24.

Розглядаються головні стандарти криптографічних перетворень на еліптичних кривих. Дається порівняльний аналіз перетворень у полях і на еліптичних кривих. Приводяться результати оцінки стійкості перетворення на еліптичних кривих.

Табл. 10.

UDC 681.3.06: 519.248.681

Essence and results of research into the properties of the perspective elliptic curve digital signature X9.62-1998 and key distribution X9.63-199X standards / M.F. Bondarenko, I.D. Gorbenko, E.G. Kachko, A.V. Svinarev, T.A. Grinenko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. №114. P.15-24.

The principal elliptic curve cryptographic algorithm standards are considered. The comparative analysis of the numeric field and elliptic curve algorithms is given. Security of the elliptic curve algorithms is estimated.

10 tab.

УДК 681.3.06:519.248.681

Критерии и методология оценки безопасности информационных технологий/ И.Д. Горбенко, А.В. Потий, П.И. Терещенко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.25-38.

Данная статья посвящена рассмотрению стандартов информационной безопасности. В частности рассматриваются положения стандарта ISO/IEC 15408 «Критерии оценки безопасности информационных технологий» и рабочего проекта стандарта CEM-97/017 – «Общая методология оценки безопасности информационных технологий».

Ил. 3. Библиогр.: 10 назв.

УДК 681.3.06:519.248.681

Критерії та методологія оцінки безпеки информаційних технологій/ І.Д. Горбенко, О.В. Потій, П.І. Терещенко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 25-38.

Дана стаття присвячена розгляду стандартів інформаційної безпеки. Зокрема розглядаються положення стандарту ISO/IEC 15408 «Критерії оцінки безпеки інформаційних технологій» та робочого проекту стандарту CEM-97/017 – «Загальна методологія оцінки безпеки інформаційних технологій».

Іл. 3. Бібліогр.: 10 назв.

UDC 681.3.06:519.248.681

Criteria and Methodology Evaluation for Information Technology Security / I.D. Gorbenko, A.V. Poty, P.I. Tereshenko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. –№ 114. P. 25-38.

This article is devoted to examination of the Standards of Information security. Particularly, aspects of standard ISO/IEC 15408 – Evaluation criteria for IT security and project of standard CEM 97/017 – Common Evaluation Methodology for Information Technology Security are being examined.

3 fig. Ref.: 10 items.

УДК 681.3.06: 519.248.681

Обеспечение стойкости DES - подобных алгоритмов шифрования к атакам линейного криптоанализа при использовании таблиц подстановок случайного типа / В.И. Долгов, И.В. Лисицкая, С.А. Головашич, А.С. Бондаренко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 39-46.

Излагается сущность и сама методика выполнения линейного криптоанализа. На примере шифра DES изучаются подходы к оценке стойкости процедуры шифрования к известным атакам линейного криптоанализа. Определяются критерии отбора случайных таблиц подстановок, обеспечивающих устойчивость алгоритма DES к известным криптонападениям этого типа.

Табл. 0. Ил. 6. Библиогр.: 5 назв.

УДК 681.3.06: 519.248.681

Забезпечення стійкості DES – подібних алгоритмів шифрування до атак лінійного криптоаналізу при використанні таблиць підстановок випадкового типу / В. Долгов, В. Лисицька, С.А. Головашич, А.С. Бондаренко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 39-46.

Викладається сутність і сама методика здійснення пінійного криптоаналізу. На прикладі шифру DES вивчаються підходи до оцінювання стійкості процедури шифрування до відомих атак пінійного криптоаналізу. Визначаютья критерії відбору випадкових таблиць підстановок, що забезпечують стійкість алгоритму DES до відомих криптонападів цього типу.

Табл. 0. Іл. 6. Бібліогр.: 5 назв.

UDC 681.3.06: 519.248.681

The provision of DES – like encryption algorithms stability to the linear cryptoanalysis attacks in case of the random type substitution tables use / V.I. Dolgov, I.V. Lisitskaya, S.A. Golovashich, A.S. Bondarenko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. PP. 39-46.

The main idea and methods of linear cryptoanalysis performing are presented. Approaches to evaluation of the encryption procedure to the known linear cryptoanalysis attacks are analyzed in case of DES cipher. The criteria of choosing the random substitution tables are defined, the tables providing the DES algorithm stability to known cryptoattacks of the similar type.

6 fig. Ref.: 5 items.

УДК 681.3.06: 519.248.681

Уточненные критерии отбора таблиц подстановок с заданными характеристиками случайности / И.В. Лисицкая, А.С. Коряк // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 47-56.

Излагаются уточненные критерии отбора случайных подстановок и случайных таблиц подстановок для симметричных шифров, рассматриваются расчетные соотношения, лежащие в их основе. Приводятся результаты статистического анализа таблиц подстановок, полученных с помощью разработанного программного комплекса генерации долговременных ключей для шифра ГОСТ 28147-89.

Ил. 0. Табл. 8. Библиогр.: 9 назв.

УДК 681.3.06: 519.248.681

Вточнені критерії відбору таблиць підстановок із заданими характеристиками випадковості / І.В. Лисицька, О.С. Коряк // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 47-56.

Викладаються вточнені критерії відбору випадкових підстановок та випадкових таблиць підстановок до симетричних шифрів, розглядаються розрахункові співвідношення, що лежать в їх основі. Наводяться результати статистичного аналізу таблиць підстановок, які здобуті за допомогою розробленого програмного комплексу генерації довгочасних ключів до шифру ГОСТ 28147-89.

Ил. 0. Табл. 8. Бібліогр.: 9 назв.

UDC 681.3.06: 519.248.681

The refined criteria of selecting the substitution tables with specified randomness characteristic. / I.V Lisitsky, A.S. Koryak // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. №114. P. 47-56.

Refined criteria of selecting random substitutions and random substitutions tables for symmetric ciphers are presented, the calculated relations being their basis are considered. Results of the substitution tables statistical analysis obtained with the developed software complex for generating the long-term keys for All-Union state standard 28147-89 cipher.

8 tables. Refs: 9 items.

УДК 681.3.06

Ключевые группы в атаках дифференциального криптоанализа DES-подобных шифров / С.А. Головашич // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 57-62.

Рассматриваются особенности построения ключезависимых дифференциальных характеристик для DES-подобных шифров. Предлагается методика более точной, по сравнению с методикой предложенной Эли Бихамом, оценки вероятностей дифференциальных характеристик. Так, в соответствии с предлагаемой методикой, вместо фиксированного значения 2^{-47}, вероятность лучшей дифференциальной характеристики, использованной в атаке Эли Бихама, с учётом 13 циклов и

фиксированного ключа, может принимать 7 дискретных значений в диапазоне $2^{-55} - 2^{-43}$, в зависимости от ключа шифрования. При этом вероятность характеристики достигает предельных значений для ключевых групп размерностью 2^{50}.

Табл. 4. Ил. 2. Библиогр.: 3 назв.

УДК 681.3.06

Ключові групи в атаках диференціального криптоаналіза DES- подібних шифрів / С.О. Головашич // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 57-62.

Розглядаються особливості побудови ключезалежних диференціальних характеристик для DES-подібних шифрів. Пропонується методика більш точної, у порівнянні з методикою запропонованою Елі Біхамом, оцінки ймовірностей диференціальних характеристик. Так, відповідно до запропонованої методики, замість фіксованого значення 2^{-47}, ймовірність кращої диференціальної характеристики, використаної в атаці Элі Біхама, з урахуванням 13 циклів і фіксованого ключа, може приймати 7 дискретних значень у діапазоні $2^{-55} - 2^{-43}$, у залежності від ключа шифрування. При цьому ймовірність характеристики досягає граничних значень для ключових груп розмірністю 2^{50}.

Табл. 4. Іл. 2. Бібліогр.: 3 назви.

UDC 681.3.06

Key groups in differential cryptanalysis attacks of DES-like ciphers / S.A. Golovashitch // Radiotehnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P. 57-62.

In this paper we consider the features of a construction of the key-dependent differential characteristics for DES-like ciphers. We propose the evaluation technique of the differential characteristic probabilities which is more exact than a technique offered by Eli Biham. So, according to an offered technique, instead of fixed value 2^{-47}, probability of the best differential characteristic used in Eli Biham attack, for the 13 cycles and fixed key, can accept 7 discrete values in a range $2^{-55} - 2^{-43}$, depending on the encryption key. The probability of this characteristic reaches limiting values for key groups with dimensionality 2^{50}.

4 tab. 2 fig. Ref.: 3 items.

УДК 681.3.06:519.248.681

«Слабые» ключи в алгоритме шифрования ГОСТ 28147-89 / В.И.Долгов, И.В. Лисицкая, Р.В.Олейников, А.И.Шумов // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 63-68.

В статье показывается, что для исследования статистической безопасности симметричных шифров недостаточно традиционной проверки лавинного эффекта. Доказывается существование «слабых» долговременных и соответствующих им сеансовых ключей в отечественном стандарте шифрования, при использовании которых не выполняются требования статистической безопасности. Приведен пример «слабой» подстановки и соответствующие ей классы «слабых» сеансовых ключей.

Табл. 6. Ил. 1. Библиогр.: 3 назв.

УДК 681.3.06:519.248.681

«Слабкі» ключі в алгоритмі шифрування ГОСТ 28147-89 / В.І.Долгов, І.В. Лисицкая, Р.В.Олійников, О.І.Шумов // Радіотехніка 2000. Вип. 114. С. 63-68.

У статті показується, що для дослідження статистичної безпеки симетричних шифрів недостатньо традиційної перевірки лавинного ефекту. Доводиться існування «слабких» довгострокових і відповідних їм сеансових ключів у вітчизняному стандарті шифрування, при використанні яких не виконуються вимоги статистичної безпеки. Наведений приклад «слабкої» підстановки і відповідні їй класи «слабких» сеансових ключів.

Табл. 6. Іл. 1. Бібліогр.: 3 назви.

UDC 681.3.06:519.248.681

Weak keys in the GOST 28147-89 encryption algorithm / V.I. Dolgov, I.V. Lisitskaya, R.V. Oleynikov, A.I. Shumov // Radiotekhnika All-Urk. Sci. Interdep. Mag. 2000. № 114. P. 63-68.

In this paper we show that usual checking of avalanche effect is insufficient for analysis of statistical security. We prove the existence of weak S-boxes and encryption keys for it in the Ukrainian (ex-Soviet) encryption algorithm. The demands of statistical security don't discharge when cipher use weak S-boxes and keys. We show an example of the weak S-box and weak encryption keys for it.

6 tab. 1 fig. Ref.: 3 items.

УДК 681.3.06

Методы и алгоритмы ускорения вычислений в несимметричных преобразованиях на эллиптических кривых / Е.Г.Качко, А.В.Свинарев, С.А.Головашич // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 69-74.

Рассматриваются методы и алгоритмы ускоренного выполнения сложения и скалярного умножения на эллиптических кривых. Приводятся оценки вычислительной сложности выполнения указанных операций при программной реализации.

Табл. 3. Ил. 2. Библиогр.: 5 назв.

УДК 681.3.06

Методи та алгоритми прискорення обчислень в несиметричних перетвореннях на еліптичних кривих / О.Г.Качко, А.В.Свинарьов, С.О.Головашич // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 69-74.

Розглядаються методи й алгоритми прискореного виконання додавання і скалярного множення на еліптичних кривих. Приводяться оцінки обчислювальної складності виконання показаних операцій при програмній реалізації.

Табл. 3. Іл. 2. Бібліогр.: 5 назв.

UDC 681.3.06

Methods and algorithms that accelerate computations in the elliptic curve public-key transformations / O.G.Kachko, A.V.Svynaryov, C.O.Golovashych // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P. 69-74.

Methods and algorithms that accelerate elliptic curve addition and scalar multiplication are considered. The estimations of computational complexity of software implementation of stated operations are given.

3 tab. 2 fig. Ref.: 5 items.

УДК 681.3.06

Применение быстрого преобразования Фурье в криптографических преобразованиях / Д.И. Лавриненко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 75-79.

Рассматривается применение быстрого преобразования Фурье в арифметических операциях многократной точности. Анализируются условия и ограничения применимости таких преобразований. Показаны результаты теоретических расчётов вычислительной сложности алгоритма умножения с БПФ и результаты вычислительного эксперимента.

Ил. 1. Библиогр.: 5 назв.

УДК 681.3.06

Застосування швидкого перетворення Фур'є у криптографічних перетвореннях / Д.І. Лавриненко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 75-79.

Розглядається застосування швидкого перетворення Фур'є в арифметичних операціях багатократної точності. Аналізуються умови й обмеження придатності таких перетворень. Показано результати теоретичних вимірювань обчислювальної складності алгоритму множення з ШПФ і результати обчислювального експерименту.

Іл. 1. Бібліогр.: 5 назв.

UDC 681.3.06

Application of Fast Fourier Transformation in cryptographic transformations / D.I. Lavrinenko // Radiotekhnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P. 75-79.

The application of fast Fourier transformation in multiple precision arithmetic operations is considered. The conditions and restrictions of applicability of such transformations are analyzed. The results of theoretical accounts of computing complexity of algorithm of multiplication with FFT and results of computing experiment are shown.

1 fig. Ref.: 5 items.

УДК 681.3.06

Расширенное поле Галуа $GF(2^m)$. Вычислительная сложность простейших операций над расширенным полем $GF(2^m)$ / И.Д. Горбенко, С.И. Збитнев // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.80-89.

В статье проводиться сравнительный анализ основных операций над расширенным полем Галуа в нормальном и полиномиальном базисах. Приводятся алгоритмы основных операций, рассчитывается их вычислительная сложность в зависимости от числа m. Приведены расчеты вычислительной сложности удвоения точки на эллиптической кривой и даны рекомендации по применению полиномиального и нормального базисов.

Табл. 3. Ил. 2. Библиогр.: 4 назв.

УДК 681.3.06

Розширене поле Галуа $GF(2^m)$. Обчислювальна складність найпростіших операцій над розширеним полем $GF(2^m)$. / І.Д. Горбенко, С.І. Збітнєв // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 80-89.

У статті проводиться порівняльний аналіз основних операцій над розширеним полем Галуа у нормальному та поліноміальному базисах. Приводяться алгоритми основних операцій, розраховується їхня обчислювальна складність у залежності від числа m. Приведені розрахунки обчислювальної складності подвоєння точки на еліптичній кривій та дані рекомендації до застосування поліноміального та нормального базисів.

Табл. 3. Іл. 2. Бібліогр.: 4 назв.

UDC 681.3.06

Binary finite Galois field $GF(2^m)$. Computing complexity of the elementary operations over binary finite field $GF(2^m)$. / I.D. Gorbenko, S.I. Zbitnev // Radiotekhnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P. 80-89.

In this paper we show comparative analysis of main operations over binary finite Galois field in normal and polynomial bases. We show algorithms of main operations, their computing complexity is calculated depending on number m. We show the complexity of calculation of points doubling on an elliptic curve and the recommendations for use of normal and polynomial basis.

3 tab. 2 fig. Ref.: 4 items.

УДК 681.324.067

Использование одностороннего преобразования, основанного на функциях Люка в несимметричных криптосистемах / О.В. Жилин, А.И. Базилевский // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 90-96.

Рассматривается математический аппарат функций Люка и возможности его применения для построения несимметричных криптосистем. Приводятся схемы применения функций Люка для построения систем, аналогичных RSA и система класса Эль-Гамаля, а также временные характеристики разработанного программного обеспечения.

Табл. 5. Ил. 0. Библиогр.: 5 назв.

УДК 681.324.067

Застосування однонапрямкового перетворення, що базується на функціях Люка, у несиметричних криптосистемах / О.В. Жилін, А.І Базилевський // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.90-96.

Розглядається математичний аппарат функцій Люка та можливості його застосування для побудова несиметричних криптосистем. Наводяться схеми застосування функцій Люка для побудови систем, подібних RSA та системам класу Ель-Гамаля, а також швидкісні характеристики розробленого програмного забезпечення.

Табл. 5. Іл. 0. Бібліогр.: 5 назв.

UDC 681.324.067

The use of Lucas oneway functions in asymmetric cryptosystems / O.V. Zhilin, A.I. Basilevsky // Radiotekhnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P. 90-96.

The mathematical background of Lucas functions and its use in the asymmetric cryptosystems is considered. The schemes of the Lucas functions use in the similar RSA and El-Gamal-like systems are given together with the developed software time responses.

5 tab. 0 fig. Ref.: 5 items.

УДК 681.3.06

Применение технологии MMX для выполнения целочисленных арифметических операций над числами многократной точности / М.А. Кривошлык, Е.Г. Качко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.97-101.

В статье производится анализ применимости технологии MMX для выполнения целочисленных арифметических операций над числами многократной точности с целью снижения нагрузки на центральный процессор.

Библиогр.: 1 назв.

УДК 681.3.06

Застосування технології MMX для виконання цілочислових арифметичних операцій над числами багатократної точності / М.А. Кривошлик, О.Г. Качко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.97-101.

В статті виконується аналіз придатності технології MMX для виконання цілочислових арифметичних операцій над числами багатократної точності з метою зниження завантаження центрального процесора.

Бібліогр.: 1 назва.

UDC 681.3.06

MMX technology application for the integer arithmetical operations realization over the multiple-precision numbers / M.A. Krivoshlyk, E.G. Kachko // Radiotekhnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P.97-101.

The analysis of the MMX technology application for the integer arithmetical operations realization over the multiple-precision numbers to decrease the working load of the central processor has been carried out.

Ref: 1 item.

УДК 681.3

Использование DLL при разработке защищенных программ / Е. Г. Качко, М. В. Благай // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.102-106.

Выполнен анализ механизма динамического связывания библиотек в ОС Windows, и особенностей структуры исполняемых файлов, которые могут привести к ослаблению защиты программного обеспечения в целом. Анализ показал большую подверженность взлому программ, использующих динамическое связывание библиотек, связанную с особенностями структуры файлов. Предложены варианты повышения общей защищенности таких программных пакетов. Результаты могут использоваться при разработке программного обеспечения для ОС Windows, требующего повышенной защищенности.

Табл. 0. Ил. 0. Библиогр.: 0 назв.

УДК 681.3

Використання DLL при розробці захищених програм / О. Г. Качко, М. В. Благай // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.102-106.

Проведен аналіз механізму динамічного зв'язування бібліотек в ОС Windows, та особливостей структури виконуємих файлів, які можуть привести до ослаблення захищеності програмного забезпечення у цілому. Аналіз показав велику схильність до злому програм, які використовують динамічне зв'язування бібліотек, пов'язану з особливостями структури файлів. Запропоновані варіанти підвищення загальної захищеності таких програмних пакетів. Результати можуть використовуватися при розробці програмного забезпечення для ОС Windows, яке потребує підвищеної захищеності.

Табл. 0. Іл. 0. Бібліогр.: 0 назв.

UDC 681.3

Use of DLL technology in development of the protected programs / Å.G. Kachko, M.V. Blagai // Radiotekhnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P.102-106.

Analysis of the mechanism of the dynamic linkage of libraries in OS Windows is executed, and the executable file structure features, which can result in lowering of the software protection as a whole, are discussed. The analysis has shown a high susceptibility to hack-work of the software products, which use dynamic linkage of libraries, connected with features of file structure. The variants to increase the overall security of such software products are offered. Developers of the software for OS Windows can

0 tab. 0 fig. Ref.: 0 items.

УДК 681.142.35

Использование методов факторизации для оценки надежности системы шифрования RSA / Н.П. Карпинский, Я.И. Кинах // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.107-110.

В статье рассмотрена система шифрования данных RSA. Проведен криптоанализ используя метод решета числового поля. Использование этого метода позволяет однозначно определить секретный ключ системы шифрования RSA.

Ил. 1. Библиогр.: 3 назв.

УДК 681.142.35

Використання методів факторизації для оцінки надійності системи шифрування RSA / М.П. Карпінський, Я.І. Кінах // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.107-110.

В цій статті розглянута система шифрування даних RSA. Проведено криптоаналіз із використанням методу решета числового поля. Використання наведеного методу дозволяє однозначно визначити таємний ключ системи шифрування RSA.

Іл. 1. Бібліогр.: 3 назви.

UDC 681.142.35

Utilization of factorization methods for estimation the RSA coding system / M. Karpinsky, Y. Kinakh // Radiotekhnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P.107-110.

The RSA encryption algorithm has been considered. Its safety has been analyzed using the number field sieve method. The algorithm application allows to define uniquely a secret key of RSA ciphering.

УДК 681.3.06

Сервис аутентификации Kerberos / И.Д. Горбенко, Е.Г. Качко, С.А. Ковалев // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.111-115.

В статье рассматривается сервис аутентификации Kerberos, являющийся частью проекта Athena в Massachusetts Institute of Technology. Протокол получил широкое распространение и за пределами этого института. В работе описываются работа по созданию программной модели сервиса.

Ил. 1. Библиогр.: 10 назв.

УДК 681.3.06

Сервіс автентифікації Kerberos / І.Д. Горбенко, Е.Г. Качко, С.О. Ковальов // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.111-115.

У статті розглядається сервіс автентифікації Kerberos, який є частиною проекту Athena у Massachusetts Institute of Technology. Протокол набув широкого поширення і за межами цього інституту. У роботі описується зусилля з створення програмної моделі сервісу.

Іл. 1. Бібліогр.: 10 назви.

UDC 681.3.06

The Kerberos Network Authentication Service / I.D. Gorbenko, E.G. Kachko, S.A. Kovalev // Radiotekhnika. All-Urk. Sci. Interdep. Mag. 2000. № 114. P.111-115.

The Kerberos Network Authentication Service, being a part of the MIT's Athena Project is considered in this article. The protocol has received wide acceptance far beyond this Institute. The guest for the Service program model creation is described in this work.

1 fig. Ref.: 10 items.

УДК 681.3.06

Разработка криптопровайдера. Интеграция криптопровайдера в систему / Е.Г. Качко, В.А. Железняк // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.116-119.

Работа посвящена исследованию возможности совместного использования алгоритмов для национальных стандартов защиты информации и операционных систем типа WINDOWS. Разработан комплекс программ для обеспечения такой возможности

Ил. 1.Библиогр.: 1 назв.

УДК 681.3.06

Розробка криптопровайдера. Інтеграція криптопровайдера в систему / О.Г. Качко, В.А. Железняк // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С. 116-119.

Робота присвячена дослідженню можливості сумістного використання алгоритмів для національних стандартів захисту інформації та операційних систем типу WINDOWS. Розроблено комплекс програм для забезпечення такої можливості

1 мал., Бібліогр.: 1 назв.

UDC 681.3.06

Development of the Cryptoprovider. Fusion of the Cryptoprovider in System. / E.G. Kachko, V.A. Geleznjak // Radiotekhnika. All-Ukr. Sci. Inderdep. Mag. 2000. № 114. P. 116-119.

The work is devoted to research the possibility of algorithms sharing for national standards of information security and operating systems of Windows type. Complex of programs was developed to provide this possibility

1 fig, Ref.: 1 item.

УДК 681.327.8

Обзор протоколов защиты информации в открытых сетях / П.В. Колесников // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.120-123.

В статье описываются возможность создания корпоративных сетей на основании защищенных сетевых протоколов на различных уровнях модели взаимодействия открытых систем. Рассматриваются протоколы защиты информации на различных уровнях стека протоколов сети. Подробно рассматривается протокол транспортного уровня TLS. Описывается его возможности, варианты реализации и их надежность.

Табл. 2. Ил. 3. Библиогр.: 0 назв.

УДК 681.327.8

Огляд протоколів захисту інформації в відкритих мережах / П.В. Колесніков // Радіотехніка 2000. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.120-123.

У статті розглядається можливість утворення корпоративних мереж на підставі захищених протоколів мереж на різних рівнях моделі взаємодії відкритих систем. Розглядаються протоколи захисту інформації на різних рівнях стеку протоколів мережі. Докладно розглідається протокол транспортного рівня TLS. Описуються його можливості, варианти реалізації та їх безпечність.

Табл. 2. Ил. 3. Библиогр.: 0 назв.

UDC 681.327.8

Review of information protection protocols in public networks / P.V. Kolesnikov // Radiotekhnika All-Urk. Sci. Interdep. Mag. 2000. № 114. P. 120-123.

The possibility to create the corporate networks based on the protected network protocols in different layers of the opened system interaction model is described. Protocols of information protection in different layers of network protocol stack are considered. The transport layer security (TLS) protocol is covered comprehensively. Its potentialities, versions of its implementation and their reliability are described.

2 tab. 3 fig. Ref.: 0 items.

УДК 681.32

Методы сертификации микропроцессорных компонентов / В.А. Горбачев, М.А. Волк, С.Н. Саранча // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.124-128.

Рассматривается проблема поиска источников аппаратно реализуемых угроз безопасности информации. Анализируются существующие методы диагностирования микропроцессорных структур, и предлагается комбинированный метод.

Библиогр.: 3 назв.

УДК 681.32

Методи сертифікації мікропроцесорних компонентів / В.О. Горбачов, М.О. Волк, С.М. Саранча // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вып. 114. С. 124-128.

Розглядається проблема пошуку джерел загроз безпеці інформації, що здійснюються за допомогою апаратних ресурсів автоматизованої системи. Аналізуються сучасні методики діагностування мікропроцесорних структур та пропонується комбінований метод.

Бібліогр.: 3 назви.

UDC 681.32

Methods of certifying of the microprocessor's components / V.A. Gorbachev, M.A. Volk, S.N. Sarancha // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000 № 114 P. 124-128.

The problems of searching of hardware errors and information security of hardware are considered. The combine diagnostic method of microprocessor's structures is proposed.

Ref.: 3 items.

УДК 681.394.74

Оптимальный алгоритм когерентной обработки различных вариантов многочастотных групповых сигналов в современных устройствах доступа к сетям передачи данных / В.К. Стеклов, И.А. Тарасенко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып.114. С.129-133.

Рассмотрен алгоритм оптимального приёма многопозиционных сигналов многоканальных модемов. Этот алгоритм позволит использовать более высокий уровень кратности модуляции по сравнению с классическим методом приёма, что позволит значительно повысить скорость передачи данных в каналах сравнительно низкого качества.

Ил. 4. Библиогр. 4 назв.

УДК 681.394.74

Оптимальний алгоритм когерентної обробки різних варіантів багаточастотних групових сигналів в сучасних пристроях доступу до мереж передачі даних / В.К. Стеклов, І.А. Тарасенко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип 114. С.129-133.

Розглянуто алгоритм оптимального прийому багатопозиційних сигналів багатоканальних модемів. Цей алгоритм дозволяє використовувати більш високий рівень кратності модуляції у порівнянні з класичними методами прийому, що дозволяє значно підвищити швидкість паредачі даних в каналах порівняно низької якості.

Іл. 4. Бібліогр.: 4 назви.

UDC 681.394.74

Optimal algorithm of various multifrequency composite signals versions coherent procession in modern means of access to the data communication networks / V.K. Steklov, I.A. Tarasenko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.129-133.

The algorithm of the multichannel modem multiposition signals optimal reception has been considered. This algorithm will allow to use a higher level of modulation multiplicity when compared to the classical reception method, this will result in a significant increase in the data transmission rate in relatively low-grade channels.

4 fig. Ref.: 4 items.

УДК 681.394.74

Интеграция интеллектуальной и мобильных сетей при создании глобальной информационной инфраструктуры / С.Н. Скляренко, С.И. Мнищенко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.134-137.

Рассмотрены вопросы взаимодействия информационных систем в соответствии с моделью взаимодействия открытых систем, интеграция мобильных сетей и интеллектуальной сети в рамках концепции CAMEL.

Ил. 2. Библиогр. 5 назв.

УДК 681.394.74

Інтеграція інтелектуальної та мобільних мереж при створенні глобальної інформаційної інфраструктури / С.М. Скляренко, С.І. Мнищенко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.134-137.

Розглянуті питання взаємодії інформаційних систем у відповідності з моделлю взаємодії відкритих систем, інтеграція мобільних мереж та інтелектуальної мережі в рамках концепції CAMEL.

Іл. 2. Бібліогр.: 5 назв.

UDC 681.394.74

Integration of the intelligent and mobile networks when creating the global information infrastructure / S.N. Skljarenko, S.I. Mnishchenko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P. 134-137.

The problems of the information systems interaction in accordance with the open systems interaction model, integration of the mobile networks and Intelligent Net in the CAMEL concept frames are considered.

2 fig. Ref.: 5 items.

УДК 621.396

Корректирующие алгоритмы системы ФАП / А.А. Руденко, А.А. Гринь // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.138-141.

Рассмотрены корректирующие алгоритмы для повышения точности и быстродействия цифровых систем фазовой автоподстройки используемых в устройствах связи.

Ил. 2. Библиогр. 3 назв.

УДК 621.396

Корегуючі алгоритми системи ФАП / О.А. Руденко, О.О. Гринь // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.138-141.

Розглянуті корегуючі алгоритми для підвищення точності та швидкодії цифрових систем фазової автопідстройки які використовуються в пристроях зв'язку.

Іл. 2. Бібліогр.: 3 назви.

UDC 621.396

Correcting algorithms for phase lock systems / A.A. Rudenko, A.A. Grin // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.138-141.

Correcting algorithms for raising accuracy and speed of the digital phase lock systems, used in communication devices, are considered.

2 fig. Ref.: 3 items.

УДК 621.396.2

Оценка точности измерения координат объектов матричными корреляционно-экстремальными системами навигации / В.И.Антюфеев, В.Н.Быков, А.С.Вильчинский, А.М.Гричанюк, М.Г.Шокин // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114 . С. 142-147.

Апробирована методика количественной оценки потенциальной точности измерения координат объектов с помощью матричных корреляционно-экстремальных систем навигации (КЭСН). Показано, что местоположение рассматриваемых объектов навигации при поэтапном визировании может быть определено матричной КЭСН с высокой точностью.

Ил. 10. Библиогр.: 3 назв.

УДК 621.396.2

Оцінка точності вимірювання координат об'єктів матричними кореляційно-екстремальними системами навігації / В.І.Антюфеєв, В.Н.Биков, А.С.Вільчинський, А.М.Грічанюк, М.Г.Шокін // Радіотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вип. 114 . С.142-147.

Апробована методика кількісної оцінки потенційної точності вимірювання координат об'єктів за допомогою матричних кореляційно-екстремальних систем навігації (КЕСН). Показано, що місцеположення об'єктів навігації, які розглядаються при поетапному візуванні може бути визначене матричною КЕСН з високою точністю.

Іл. 10. Бібліогр.: 3 назви.

УДК 621.396.2

Objects coordinates measuring accuracy estimation by matrix map matching navigation systems / V.I. Antyufyev, V.N. Bykov, A.M. Grichanyuk, M.G. Shokin // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.142-147.

The quantitative estimation procedure of objects coordinates measuring accuracy potential estimation by means of matrix map matching navigation systems is approved. It's shown that the considered objects location can be fix by the matrix system with a high accuracy.

10 fig. Ref.: 3 items

УДК 621.396.62

Об оптимальности частотно-селективных средств авиационной радиосвязи, работающих в равномерно загруженном частотном диапазоне / Г.В. Алёшин, А.А. Трублин // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.148-150

На основе новой теории электромагнитной совместимости решена задача об оптимальном (по критерию электромагнитной совместимости) распределении усилий селективных устройств радиоэлектронных средств (работающих в равномерно загруженном диапазоне волн). Проверена оптимальность параметров избирательности современных средств авиационной радиосвязи.

Табл. 1. Ил. 3. Библиогр.: 3 назв.

УДК 621.396.62

Про оптимальність частотно-селективних засобів авіаційного радіозв'язку, що працюють в рівномірно завантаженому частотному діапазоні / Г.В. Альошин, О.А. Трублін // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.148-150.

На основі нової теорії електромагнітної сумісності вирішена задача про оптимальний (за крітерієм електромагнітної сумісності) розподіл зусиль селективних пристроїв радіоелектронних засобів (що працюють у рівномірно завантаженому діапазоні хвиль). Перевірена оптімальність параметрів вибірковості сучасних засобів авіаційного радіозв'язку.

Табл. 1. Іл. 3. Бібліогр.: 3 назви.

UDC 621.396.62

On the optimality of air radio communication frequency-selective means operating in the uniformly loaded frequency band / G.V. Aljoshin, A.A. Trublin // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.148-150.

The problem of optimum (according to electromagnetic compatibility criterion) disribution of radioelectronic means (operating in the uniformly loaded wave band) selective devices efforts has solved based on an new theory of electromagnetic compatibility. Modern air radio communication means parameters optimality has been verified.

1 tab. 1 fig. Ref.: 3 items.

УДК 621.391

Обнаружение импульсного сигнала на фоне негауссовских помех / С.С.Мартыненко // Радиотехника. Всеукр межвед. научн.-техн. сб. 2000. Вып. 114. С.151-154.

Предложены алгоритмы синтеза обнаружителей импульсного сигнала, что принимается на фоне негауссовских помех. Решающие правила построены оптимальными по критерию минимума верхних границ вероятностей ошибок при моментном и кумулянтном описании случайных величин. Проведен анализ синтезированных обнаружителей. Оценен учет негауссовости помех на качество обнаружителей.

Ил. 1. Библиогр.: 2 назв.

УДК 621. 391

Виявлення імпульсного сигналу на фоні негауссівських завад / С.С.Мартыненко // Радіотехніка. Всеукр. міжвід. наук. - техн. зб. 2000. Вип. 114. С.151-154.

Запропоновано алгоритми синтезу виявлячів імпульсного сигналу, що приймається на фоні негауссовских завад. Вирішальні правила побудовані оптимальними за критерієм мінімуму верхніх меж імовірностей помилок при моментному та кумулянтному описі випадкових величин. Проведено аналіз синтезованих виявлячів. Оцінено врахування негауссовості завад на якість вивлячів.

Іл. 1. Бібліогр.: 2 назв.

UDC 621.391

Detection of the pulse signal on the background of non-Gaussian noise / S.S. Martynenko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.151-154.

Algorithms for synthesis of the detectors of the pulse signals, being received on the background of non-Gaussian noise, are offered. The analysis of the synthesized detectors is given. The non-Gaussian noise account in the detectors quality is estimated.

1 fig. Ref.: 2 items.

УДК 621.391

Алгоритмы оценки параметров полигармонического сигнала на фоне негауссовских помех / А.С.Гавриш // Радиотехника. Всеукр. межвед. научн.-техн. сб. 2000. Вып. 114. С.155-158.

Используя метод максимизации полинома синтезирован линейный и квадратичный алгоритмы нахождения коэффициентов полигармонического сигнала при воздействии негауссовских помех. Показано, что линейные оценки параметров исследуемого сигнала, принимаемого на фоне негауссовских помех, совпадают с коэффициентами детерминированного тригонометрического полинома. При нелинейной обработке выборочных данных получаются новые оценки, являющиеся функциями от статистических характеристик негауссовской помехи.

Библиогр.: 5 назв.

УДК 621.391

Алгоритми оцінки параметрів полігармонічного сигналу на тлі негауссівськнх завад / О.С.Гавриш // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.155-158.

Використовуючи метод максимізації поліному, синтезовано лінійний і квадратичний алгоритми обчислення коефіцієнтів полігармонічного сигналу при впливі негауссівських завад. Показано, що лінійні оцінки параметрів досліджуваного сигналу, прийнятого на тлі негауссівських завад, збігаються з коефіцієнтами детермінованого тригонометричного поліному. При нелінійній обробці вибіркових даних отримуються нові оцінки, що є функціями від статистичних характеристик негауссівської завади.

Бібліогр.: 5 назв.

UDC 621.391

Algorithms for estimating parameters of the polyharmonic signal on the background of non-Gaussian noise /A.S. Gavrish //Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.155-158.

Linear and quadratic algorithms for finding coefficients of a polyharmonic signal on exposure to non-Gaussian noise were synthesized using the polynomial maximization method. It is shown that the linear estimations of the involved signal parameters received on the background of non-Gaussian noise coincide with coefficients of the determined trigonometric polynomial. In the non-linear processing of the sample data the new estimations, being functions of the non-Gaussian noise statistical characteristics, are obtained.

Ref.: 5 items.

УДК 621.391

Анализ информативных признаков для идентификации импульсных радиосигналов по спектральным параметрам / В.А. Письменецкий, А.В. Бородин, П.И. Платонов // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.159-162.

Исследуется три базовых семейства спектральных признаков импульсных радиосигналов (центральная частота, ширина спектра, коэффициент прямоугольности), сформированных с помощью дисперсионного Фурье-поцессора. Показано, что на устойчивость первого и второго признаков существенное влияние оказывает динамический диапазон амплитуд двух разрешаемых по спектру радиосигналов и взаимное временное запаздывание, а третий признак является инвариантным по отношению к их длительности.

Ил. 5. Библиогр.: 3 назв.

УДК 621.391

Аналіз інформативних ознак для ідентифікації імпульсных радіосигналів по спектральным параметрам / В.О. Письменецький, О.В. Бородин, П.І. Платонов // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.159-162.

Досліджуються три базових сімейства спектральних ознак імпульсних радіосигналів (центральна частота, ширина спектру, коефіцієнт прямокутності), які формуються за допомогою дисперсійного Фурьє процесора. Показано, що на стійкість першої та другої ознак суттєво впливає динамічний диапозон амплітуд двох розв`язаних по спектру радіосигналів та взаємне часове запізнення, а третя ознака є інваріантною по відношенню до їх протяжності.

Іл. 5. Бібліогр.: 3 назв.

UDC 621.391

Analysis of informative signs for identification of pulse radio signals from spectral parameters / V.A. Pismenetsky, A.V. Borodin, P.I. Platonov // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000.№ 114.P.159-162.

Three base sets of radio signals spectral signs (center frequency, spectrum width, rectangularity index) formed with the dispersion Fourier processor are investigated. It is shown that stability of the first and the second signs are substantially affected by the dynamic range of amplitudes of two radio signals resolved by the radio signals spectrum and mutual temporal lag, and the third sign is invariant relative to their duration.

5 fig. Ref.: 3 items.

УДК 621.327:519.216

Определение двумерных кумулянтных функций сигналов с фазовой модуляцией и их спектров / С.М. Первунинский, Р.М. Дидковский // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.163-168.

Представлены изложение метода и результаты вычисления моментных и кумулянтных функций двумерного распределения сигналов с фазовой модуляцией. Приведен пример определения спектра кумулянтных функций.

Ил. 1. Библиогр.: 5 назв.

УДК 621.327:519.216

Визначення двовимірних кумулянтних функцій сигналів з фазовою модуляцією та їх спектрів / С.М. Первунінський, Р.М. Дідковський // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.163-168.

Наведено виклад методу та результати обчислення моментних та кумулянтних функцій двомоментного розподілу сигналів з фазовою модуляцією. Наведено приклад визначення спектра кумулянтних функцій.

Іл. 1. Библиогр.: 5 назв.

UDC 621.327:519.216

Definition of two-dimensional cumulant functions of phase-modulated signals and their spectra / ÏS.M. Perwuninsky, R. M. Didkowsky // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.163-168.

Method and results of calculating the moment and cumulant functions of the phase-modulated signals two-dimensional distribution are presented. The cumulant functions spectrum definition is cited.

1 fig. Ref.: 5 items.

УДК 621.391.14

Учет влияния связующих слоев на эффективность преобразования акустооптического устройства / В.В. Данилов, С.В. Иванов // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114.С.169-173.

Предлагается методика инженерного расчета частотной зависимости акустического импеданса элементов многослойной структуры, обобщенная на случай затухания упругой волны в отдельных ее элементах, а также коэффициента передачи акустической энергии от электроакустического преобразователя к светозвукопроводу. Внесена поправка в методику расчета акустооптического модулятора, связанная с влиянием частотно-зависимой нагрузки комплексного характера на акустическую добротность пьезопреобразователя.

Ил. 13 Библиогр.:6 назв.

УДК 621.391.14

Облік впливу зв'язуючих шарів на ефективність перетворювання акустооптичного приладу / В.В. Данилов, С.В. Іванов // Радіотехніка. Всеукр. межвід. наук.-техн. зб. 2000. Вип. 114. С.169-173.

Запропоновано методику інженерного розрахунку частотної залежності акустичного імпедансу елементів богатошарої структури, узагальненная на випадок згаснення пружної хвилі у окремих її елементах, а також коефіцієнта передачі акустичної енергії від електроакустичного перетворювача до світлозвукопроводу. Внесена виправка у методику розрахунку акустооптичного модулятора, пов'язанная з впливом частотно-залежної навантаженності комплексного характеру на акустичну добротність п'єзоперетворювача.

Іл. 13 Бібліогр.:6 назв.

UDC 621.391.14

Accounting of coupling layers influence upon the acoustooptic device transformation efficiency / V.V. Danilov, S.V. Ivanov // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.169-173.

Technique for engineering calculation of the acoustic impedance frequency dependence of the multilayer structure is offered and generalized when an elastic wave attenuation occurs in its separate elements. It also allows to calculate a factor of the acoustic energy transmission from the electroacoustic converter to the waveguide. The correction, associated with the frequency dependent complex load influence on the piezotransducer quality, was introduced into the acoustooptic modulator computation method.

13 fig. Ref.: 6 items.

УДК 621.472

Применение и перспективы развития беспроводных систем передачи энергии СВЧ-лучом / А.А. Коновальцев, Ю.А. Лучанинов, М.А. Омаров, В.М. Шокало // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114.С.174-180.

Приведен краткий обзор состояния разработок и перспектив создания беспроводных систем передачи энергии СВЧ-лучом, показана актуальность рассматриваемой проблемы и ее все усиливающееся влияние на процесс развития мировой энергетики. Описаны главные направления применения в настоящем и будущем систем передачи энергии СВЧ-лучом и приведены их основные этапы развития.

Табл. 2. Рис. 6. Библиогр.: 31 назв.

УДК 621.472

Використання та перспективи розвитку бездротових систем передачі енергії НВЧ променем / А.О.Коновальцев, Ю.А. Лучанінов, М.А. Омаров, В.М. Шокало // Радіотехніка. Всеукр. межвід. наук.-техн. зб. 2000. Вип. 114. С.174-180.

Приведено короткий огляд стану розробок та перспектив безпроводових систем передачі енергії НВЧ променем, показана актуальність розглянутої проблеми та її постійно посилюючийся вплив на процес розвитку світової енергетики. Описано основні напрямки використання, у теперішній час та у майбутньому, систем передачі енергії НВЧ променем і приведені основні етапи їх розвитку.

Табл.2. Іл..6. Бібліогр.: 31 назви.

UDC 621.472

Application and prospects for development of wireless systems of power transmission systems through microwave beam / A.A.Konovaltsev, Yu.A.Luchaninov, M.A.Omarov, V.M.Shokalo // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. N 114. P.174-180.

A brief review of the state of development and prospects of creating wireless systems of power transmission through a microwave beam is given; the considered problem urgency and its increasing influence on the global power engineering development process are shown. The principle directions of present and future applications of the systems of power transmission through a microwave beam are described and the milestones of their development are presented.

2 tab. 6 fig. Ref.: 31items

УДК 537.874.4

Анализ структурно-физической модели рассеяния волн в турбулентной атмосфере / В.А. Петров, В.М. Карташов //Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.181-184.

Рассмотрена модель рассеяния акустических и электромагнитных волн естественными неоднородностями атмосферы, представленная в виде совокупности эквивалентных пространственных образований. Определены основные параметры структурно-физической модели.

Библиогр.: 5 назв.

УДК 537.874.4

Аналіз структурно-фізичної моделі розсіяння хвиль в турбулентній атмосфері / В.А. Петров, В.М. Карташов // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.181-184.

Розглянуто модель розсіяння акустичних та електромагнітних хвиль природними неоднорідностями атмосфери, яка представляє собою сукупність еквівалентних просторових утворювань. Визначені основні параметри структурно-фізичної моделі.

Бібліогр.: 5 назв.

UDC 537.874.4

Analysis of structural-physical model of wave scattering in the turbulent air // V.A. Petrov, V.M. Kartashow // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. N 114. P.181-184.

The model of acoustic and electromagnetic waves scattering with the natural atmosphere non-uniformity combined with equivalent spatial formations are considered. The main parameters of the structural-physical model are defined.

Ref.: 5 items.

УДК 621.372.8

Собственные колебания электромагнитного поля поперечно намагниченного ферритового резонатора в изломе прямоугольного волновода / О.В. Кулаков, Н.И. Пятак // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114 С.185-188.

Получено выражение для расчета собственных длин волн низшего магнитного вида колебаний Н-плоскостного излома прямоугольного запредельного волновода с поперечно намагниченным ферритовым параллелепипедом, полностью заполняющим область излома. Правильность полученного выражения подтверждается предельными переходами к задачам, решенным ранее. Проанализирована скорость сходимости алгоритма. Построены зависимости величин собственных длин волн низшего магнитного вида колебаний от геометрических и электродинамических параметров структуры.

Ил.2. Библиогр.: 10 назв.

УДК 621.372.8

Власні коливання електромагнітного поля поперек намагніченого феритового резонатора в зламі прямокутного хвилеводу / О.В. Кулаков, М.І. П'ятак // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114 С.185-188.

Одержано вираз для розрахунку власних довжин хвиль нижчого магнітного виду коливань Н-площинного зламу прямокутного позамежного хвилеводу з поперек намагніченим феритовим паралелепіпедом, що цілком заповнює область зламу. Слушність отриманого виразу підтверджується граничними переходами до задач, що вирішені раніше. Проаналізовано швидкість збіжності алгоритму. Побудовано залежності власних довжин хвиль нижчого магнітного виду коливань від геометричних і електродинамічних параметрів структури.

Іл.2. Бібліогр.: 10 назв.

UDK 621.372.8

Own oscillations of elektromagnetic field of transversally magnetized ferrite resonator in the bend of rectangular waveguide / O. Kulakov, M. Pyatak // Radiotechnika. All-Ukr. Sci. Interdep. Mag.. 2000. N.114. P.185-188.

Expression for calculation of own length wave for lowest magnetic oscillation in the H-plane bend of rectangular unlimited waveguides with transversally magnetized ferrite parallelepiped, that completely fill in region of bend, is received. Correct of this expression is proved by limited tranzmissions for problems, that already decided. The velocity of this algorithm is analysed. The dependence of value of own length wave for lowest magnetic oscillation from geometrical and electrodynamical parameters of this structure are built.

2 fig. Ref.: 10 items.

УДК 621.376.54

Анализ двухконтурной системы термостабилизации устройств радиоэлектронной аппаратуры / Н.Р. Попов, М.В. Гунбин, А.И. Гапон, П.А. Качанов // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С. 189-193.

Определены математические модели элементов и контуров системы с частотным пьезокварцевым датчиком температуры, проанализированы устойчивость грубого и точного контуров системы, выражения статической и динамической ошибок системы, отмечено отсутствие интегрирующих звеньев. Показано, что относительная статическая ошибка за счет двухконтурной структуры не хуже 10^{-4}. Из выражений кинетической ошибки получена ее зависимость от скорости изменения температуры окружающей среды.

Табл. 1. Ил. 2. Библиогр.: 3 назв.

УДК 621.376.54

Аналіз двоконтурної системи термостабілізації обладнань радіоелектронної апаратури / М.Р. Попов, М.В. Гунбін, А.І. Гапон, П.О. Качанов // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114 .С. 189-193.

Визначені математичні моделі елементів і контурів системи з частотним пьезокварцевим давачем температури, проаналізовані стійкість грубого та точного контурів системи, вирази статичної і динамічної помилок системи, відмічена відсутність інтегруючих ланок. Показано, що відносна статична помилка за рахунок двоконтурної структури не гірше 10^{-4}. З виразив кінетичної помилки долучена її залежність від швидкості зміни температури навколишнього середовища.

Табл. 1. Іл. 2. Бібліогр.: 3 назви.

UDC 621.376.54

The analysis of a double-loop system stabilization of temperature of devices of an electronic equipment / N.R. Popov, M.V. Gunbin, A.I. Gapon, P.A. Kachanov // Radioengineering. Ukrain interdepartamental reseach-technical collection. 2000. instalment. № 114. P.189-193.

The mathematical models of elements and contours of a system with frequency piezoquartz temperature transmitter are determined, the expressions of static and dynamic errors of a system, stability immunity of rough and precise contours of a system are analyzed, the absence of integrating links is marked. Is shown, that the relative static error at the expense of two contour structures not worse 10-4. From expressions of a kinetic error its velocity function from change of temperature of an environment is received.

Tab. 1. Illustrations 2. References: 3 appelations.

УДК 621.383.52

Конструктивно-технологические ограничения кремниевых микрополосковых PIN приемников излучения / В.А. Антонова, В.Н. Борщев, А.М. Листратенко, Н.И. Слипченко // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2000. Вып. 114. С.194-197.

Рассмотрены требования к конструктивно-технлогическим решениям микростриповых кремниевых приемников ионизирующих излучений. Предложена топология PIN – приемника, которая позволяет разработать детекторный модуль на основе алюминиевых микрокабелей с использованием микросхем первичной обработки нового поколения типа ALICE 128 С.

Ил. 1. Библиогр.: 12 назв.

УДК 621.383.52

Конструктивно-технологічні обмеження кремнієвих мікросмужкових PIN приймачів випроміню-вання / В.А. Антонова, В.М. Борщов, О.М. Лістратєнко, М.І. Сліпченко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. 2000. Вип. 114. С.194-197.

Розглянуто вимоги до конструктиво-технологічних рішень мікрострипових кремнієвих приймачів іонізу-ючих випромінювань. Запропонована топологія PIN-приймача, яка дозволить розробити детекторний модуль на базы алюмінієвих мікросхем первинної обробки нового покоління типу ALICE 128 С.

Іл. 1. Бібліогр.: 12 назв.

UDC 621.383.52

Constructive-process limitations of silicon microstrip PIN radiation receivers / A.A. Antonova, V.M. Borshchov, A.M. Listratenko, N.I. Slipchenko // Radiotekhnika. All-Ukr. Sci. Interdep. Mag. 2000. № 114. P.194-197.

Requirements for constructive-process solutions of microstrip silicon ionizing radiation receivers are considered. The PIN receiver topology, allowing to develop the detector module based on aluminum microcables with the recent generation primary procession microcircuits of ALICE 128 c type, is offered.

1 fig. Ref.: 12 items.