УДК 004.2

МАРКОВСКИЙ А.П.,
АБУ-УСБАХ А.Н.,
РОМАНЕЦ Н.Н.

МЕТОД ДООПРЕДЕЛЕНИЯ ЧАСТИЧНО-ЗАДАННЫХ БУЛЕВЫХ ФУНКЦИЙ ДЛЯ ОБЕСПЕЧЕНИЯ ИХ ЛАВИННЫХ СВОЙСТВ

В статье предложен метод построения булевых балансных функций, которые соответствуют критерию строго лавинного эффекта. Особенность решаемой задачи состоит в том, что формируемая функция частично задана. Суть проектирования состоит в доопределении частично заданной функции таким образом, чтобы она удовлетворяла критерию лавинного эффекта. Подробно описана формализованная процедура конструирования балансных функций, обладающих строгим лавинным эффектом. Приведен пример синтеза функции.

In this paper the method for designing of Boolean balanced function that satisfies the Strict Avalanche Criterion (SAC) are presented. A peculiaruty of solving task is that designed function is partially given. The goal of designing consist of finish building of partially difined function such way that it will satisfies the Strict Avalanche Criterion (SAC). The formalized procedure for construction partial derivative balanced SAC-functions is described in detail. Examples of function design are given.

Введение

Интенсивное расширение информационной интеграции на основе компьютерных сетей является одним из основных фактором повышения эффективности управления во всех областях человеческой деятельности.

Развитие информационной интеграции в значительной мере зависит от эффективности реализации функций защиты данных и контроля прав доступа к информационным ресурсам. Это определяет необходимость постоянного развития и совершенствования систем защиты информации.

К настоящему времени в основе большинства таких систем лежат криптографические механизмы, базирующиеся на аналитически неразрешимых математических задачах теории чисел, эллиптических кривых и булевых функций [1]. Использование последних играет особенно важную роль в современных системах защиты информации, поскольку вычисление булевых преобразований выполняется на 3-4 порядка быстрее по сравнению с сложными мультипликативными операциями модулярной арифметики, выполняемыми над числами, длина которых на порядок превышает разрядность процессоров.

Наиболее важным, с точки зрения практического использования, является свойство строгого лавинного эффекта (Strict Avalanche Criterion), которое характеризуется максимальным значением дифференциальной энтропии изменения значения функции при инвертировании любой из переменных, на которых определена эта функция.

Булевы функции, обладающие свойством строгого лавинного эффекта играют ключевую роль при создании широкого класса алгоритмов защиты информации, поскольку это свойство обеспечивает устойчивость к нарушению защиты дифференциальным и линейным криптоанализом [2] .

Важной проблемой практического использования булевых функций, обладающих свойством лавинного эффекта является получение таких функций.

Таким образом, задача получения булевых функций, обладающих важными для криптографических средств защиты информации свойствами максимума полной и дифференциальной энтропии, является важной и актуальной.

Обзорный анализ методов получения SAC-функций

Булева функция $f(x_1,...,x_n)$ от n переменных определена на 2^n возможных наборах значений, принадлежащих множеству Z, и принимает значения на множестве $\{0,1\}$. Функция соответствует критерию максимума полной энтропии, или, иными словами, явля-

ется балансной, если она с одинаковой вероятностью принимает значения нуля и единицы:

$$\sum_{x_1,...,x_n \in Z} f(x_1,...,x_n) = 2^{n-1} \quad (1)$$

Булева функция $f(x_1,...,x_n)$ соответствует критерию максимума условной энтропии, или критерию строго лавинного эффекта (SAC), если изменение значения какой либо одной из n переменных приводит к изменению значения функции с вероятностью 0.5:

$\forall x_j, j = 1,...,n$:

$$\sum_{x_1,...,x_n \in Z} f(x_1,...,x_j,...,x_n) \oplus f(x_1,...,\overline{x}_j,...,x_n) = 2^{n-} \quad (2)$$

Наиболее важными критериями качества методик получения балансных булевых SAC-функций с точки зрения их практического применения следует считать:

- затраты вычислительных ресурсов (машинного времени и памяти) на реализацию процесса синтеза;
- количество функций от n переменных, которые могут быть синтезированными (до сегодняшнего дня проблема определения общего количества балансных SAC-функций для n переменных есть открытой [1]);

Принимая во внимание практическую важность проблемы автоматизированного синтеза балансных SAC–функций для современных средств защиты информации, за последние 15 лет предложен ряд подходов по решению этой проблемы.

Все известные методы получения SAC–функций ориентированы на их синтез “с нуля”, то есть исходят из того, что при построении булевой функции не накладывается дополнительных условий на их значения.

Однако, в последние годы началось активное использование булевых функций для решения задач защиты информации, традиционно решаемых с использованием технологий теории чисел. В частности, появились работы, в которых булевы преобразования используются для реализации криптографической концепции “нулевых знаний”, аутентификации абонентов, несимметричной криптографии [3]. Очевидным преимуществом использования булевых функций для перечисленных задач является существенный (на 2-3 порядка) выигрыш в скорости реализации функций защиты информации.

Для большинства из приведенных задач процесс построения булевых функций включает два этапа:

- определение значений функции на части наборов значений переменных исходя из условий, накладываемых криптографическим преобразованием;
- доопределение функции на оставшихся наборах переменных так, чтобы она удовлетворяла определенным криптографическим критериям, основными из которых является максимум полной и дифференциальной энтропии.

При реализации второго этапа возникает задача доопределения частично-заданных булевых функций таким образом, чтобы они были балансными и удовлетворяли критерию лавинного эффекта (SAC).

Высоконелинейные балансные SAC-функции могут быть получены путем деконкатенации bent-функций [2], тем не менее получение самых bent-функций от большого количества переменных представляет собой довольно сложную проблему, решение которой требует затрат значительных вычислительных ресурсов, как в плане затрат машинного времени, так и в плане нужных объемов памяти.

Значительно меньших ресурсов требует реализация рекурсивного получения булевых балансных SAC-функций от n переменных с использованием порождающих четырех SAC-функций (не балансных) от $n-1$ переменных [1]. Однако этот метод также не пригоден для доопределения частично-определенных SAC-функций, поскольку на порождающие функции также накладывается условие частичной определенности.

Наиболее совершенный на сегодняшний день метод получения балансовых SAC-функций предложен К.Куросава и Т.Сатохом [2].

Анализ показывает, что метод разрешает получать лишь незначительную часть балансных SAC-функций от их общего количества.

Метод не пригоден для доопределения частично-заданных функций, поскольку при этом накладываются нелинейные условия на компоненты вектора Q, которые могут быть выполнены только в процессе перебора.

Таким образом, приведенный выше краткий обзор известных на сегодняшний день подходов к проблеме синтеза балансных SAC-функций показывает, что они не позволяют решать новую задачу – построение SAC-функций при наличии ограничений в виде предварительного задания значений функций на части наборов переменных. Для решения этой задачи необходимо разработать новый метод.

Целью работы является создание метода доопределения значений частично-заданных булевых функций так, чтобы они удовлетворяли условию строгого лавинного эффекта.

Способ доопределения таблицы истинности частично-заданной функции для обеспечения ее лавинных свойств

Исходным для решения поставленной задачи является булева функция $f(x_1,...,x_n)$ от n переменных, определенная на некотором количестве M наборов.

Вполне очевидно, что в зависимости от значения M, задача доопределения функции так, чтобы она удовлетворяла SAC, может быть решена с некоторой вероятностью. Причем, чем ближе значение M к максимальному числу наборов – 2^n, на которых определена функция, тем меньше упомянутая вероятность.

Исходя из этого, основная цель разрабатываемого подхода – получение функций, в наибольшей степени удовлетворяющих критерию SAC, а также исследование зависимости лавинного эффекта от наперед заданного M.

В основу разрабатываемой методики положен тот факт, что функция $f(x_1,...,x_n)$ удовлетворяет критерию SAC, если она и ее частичные производные $df/dx_1,...,df/dx_n$ являются балансными, т.е. количество нулей и единиц в таблице истинности функции с ее производными одинаковое.

Исходной для задачи синтеза является заполненная на M наборах таблица истинности функции.

Суть предлагаемой методики состоит в целенаправленном заполнении таблицы истинности функции $f(x_1,...,x_n)$ так, чтобы в максимальной степени сбалансировать (уравнять) число единичных и нулевых значений как самой функции, так и каждой из ее частичных производных $df/dx_1,...,df/dx_n$.

Предлагаемый способ доопределения функции с целью обеспечения ее лавинных свойств сводится к следующей последовательности действий:

1) По известным значениям функции $f(x_1,...,x_n)$ вычисляются значения частичных производных $df/dx_1,...,df/dx_n$. Для определения производной $df/dx_j=\varphi(x_1,..,x_{i-1},x_{i+1},..x_n)$ по переменной x_j, $j\in\{1,..,n\}$ анализируются все 2^{n-1} наборов переменных $x_1,..,x_{i-1},x_{i+1},..,x_n$. Каждому из этих наборов соответствует два набора всех n переменных: $x_1,..,x_{i-1},0,x_{i+1},..,x_n$ и $x_1,..,x_{i-1},1,x_{i+1},..,x_n$. Если, при расчете производной, функция $f(x_1,...,x_n)$ определенна на обоих наборах и имеет одинаковые значения, то значение производной принимает нулевое значение, если функция имеет разные значения, то значение производной принимает единичное значение. В случае, когда функция определена только на одном из двух наборов, соответствующее значение производной учитывается как частично определенное значение. Для обозначения частично определенного значения производной вводится символ ‘*’. Если функция не определена на обоих наборах, то и значение производной – не определенное и обозначается как ’-‘.

2) Для оценки возможных вариантов значений функции $f(x_1,...,x_n)$ каждое не определенное значение функции $f(x_1,...,x_n)$ заменяется нулевым и единичным значениями, и вычисляется оценка влияния такой замены на свойства сбалансированости как самой функции, так и ее производных.

3) Функция доопределяется значением на наборе, интегральная оценка которых имеет максимальное значение.

4) Производится коррекция значений функций и ее производных с учетом вновь определенного значения функции на выбранном наборе. Если значения функции определены не на всех наборах – возврат на повторное выполнение пп.2., иначе – конец.

Изложенное иллюстрируется следующим примером. Пусть задана частично-определенная булева функция от четырех переменных: $f(x_1,x_2,x_3,x_4)$ (далее f), которая определена на M=8 наборах. Таблица истинности

частично-заданной функции представлена в виде последовательности ее значений на наборах переменных от 0000 до 1111: *f*={0, -, 0, 1, -, -, -, 0, 0, 0, -, -, 1, -, 1, -}, где символом **'-'** обозначены не определенные значения функции.

Следуя методике, поэтапно доопределяется частично-заданная булева функция так, чтобы она удовлетворяла критерию SAC.

Этап 1: Выполняется вычисление производных по заданным значениям функции *f*.

Табл. 1. Таблица истинности частично-определенной функции f и ее производных

$x_1\,x_2\,x_3\,x_4$	*f*	df/dx_1	df/dx_2	df/dx_3	df/dx_4
0000	0	0	*	0	*
0001	-	*	-	*	*
0010	0	*	*	0	1
0011	1	*	1	*	1
0100	-	*	*	-	-
0101	-	-	-	*	-
0110	-	*	*	-	*
0111	0	*	1	*	*
1000	0	0	1	*	0
1001	0	*	*	*	0
1010	-	*	*	*	-
1011	-	*	-	*	-
1100	1	*	*	0	*
1101	-	-	-	-	*
1110	1	*	*	0	-
1111	-	*	1	-	-

Как указывалось выше, определенные значения производных обозначаются '0', если соответствующие значения функции одинаковы, или '1', если значений функции разные. Частично определенные значения производных обозначаются символом '*', неопределенные – '-'. Значения производных для функции *f* приведены в таблице 1.

Этап 2: итерационно, вместо каждого не определенного значения функции подставляются нуль и единица (пробные значения) и вычисляются производные, значения для которых изменились в результате подстановки (таблица 2).

Затем оцениваются результаты замены. Для этого подсчитывается количество единиц и нулей функции и ее производных. В функции *f* определено три единицы и пять нулей, то есть на текущем шаге итерации при выборе следующего значения функции необходимо отдать преимущество единице.

Табл. 2. Таблица пробных значений функции f и ее производных

$x_1\,x_2\,x_3\,x_4$	*f*	df/dx_1	df/dx_2	df/dx_3	df/dx_4
0001	1	1	*	0	1
	0	0	*	1	0
0100	1	0	1	*	*
	0	1	0	*	*
0101	1	*	*	1	*
	0	*	*	0	*
0110	1	0	1	*	1
	0	1	0	*	0
1010	1	1	0	1	*
	0	0	1	0	*
1011	1	0	*	1	*
	0	1	*	0	*
1101	1	*	1	*	0
	0	*	0	*	1
1111	1	1	*	*	0
	0	0	*	*	1

Для производной df/dx_1 количество нулей равно двум, а количество единиц равно нулю, значит, приемлемым есть единичное значение. По аналогии определяется, что для производной df/dx_2 требуемое значение нуль, для df/dx_3 – единица, а для df/dx_4 – допустимо любое значение, так как количество единиц и нулей равно. В результате изложенного, формируется таблица 3 оценок пробных значений функции и производных.

В этой таблице приемлемые значения обозначаются символом '+', неприемлемые – '-', допустимые – '×'.

Этап 3: Анализируя значения таблицы 3, определяется набор, на котором наибольшее количество требуемых значений. Количество '+' в первой строчке равно двум, во второй – единице и так далее. Наибольшее количество приемлемых значений в девятой строчке для варианта f(1010)=1, то есть выбранное значение функции *f* записывается в таблицу истинности и считается определенным вариантом.

Этап 4: выполняется коррекция значений функций и ее производных с учетом опреде-

ленного значения функции на выбранном наборе.

Табл. 3. Таблица оценок пробных значений функции f и ее производных

$x_1\,x_2\,x_3\,x_4$	f	df/dx_1	df/dx_2	df/dx_3	df/dx_4
0001	+	+	×	-	×
	-	-	×	+	×
0100	+	-	-	×	×
	-	+	+	×	×
0101	+	×	×	+	×
	-	×	×	-	×
0110	+	-	-	×	×
	-	+	+	×	×
1010	+	+	+	+	×
	-	-	-	-	×
1011	+	-	×	+	×
	-	+	×	-	×
1101	+	×	-	×	×
	-	×	+	×	×
1111	+	+	×	×	×
	-	-	×	×	×

Перечисленные этапы повторяются до тех пор, пока в функции f не будет определенна на всех 2^n наборах. Результатом выполнения примера есть следующая булева функция f ={0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 1}, которая удовлетворяет критерию SAC.

Анализ влияния меры определенности функции на лавинные свойства

Важным аспектом практического использования разработанной методики доопределения частично-заданной булевой функции для обеспечения максимума полной и дифференциальной энтропии играет установление зависимости между мерой априорной определенности функции и возможностью решения указанной выше задачи.

В качестве меры определенности значений функции выступает ξ – отношение числа M наборов, на которых функция задана к общему их числу – 2^n, выраженное в процентах: $\xi = M \cdot 2^{-n} \cdot 100\%$.

Ясно, что чем больше значение ξ, тем меньше вероятность P_{SAC} того, что возможно доопределение функции такое, что она будет балансной и удовлетворять критерию *SAC*. Очевидным, также является и то, что указанная вероятность зависит и от числа n переменных, на которых определена функция.

К настоящему времени не существует аналитического выражения для исчисления количества булевых SAC-функций от n переменных [1].

В качестве грубой оценки их числа используют мультипликативную модель, которая предполагает, что свойства лавинности по каждой из переменных не зависимы. Исходя из упомянутой модели, вероятность P_n того, что функция $f(x_1,x_2,\ldots,x_n)$ удовлетворяет SAC по всем n переменным определяется как произведение вероятностей того, что функция удовлетворяет SAC по каждой из переменных:

$$P_n=P_{x1}\cdot P_{x2}\cdot\ldots\cdot P_{xn} = {P_{x1}}^n \qquad (4)$$

Число K_1 функций, удовлетворяющих SAC по одной переменной, например x_i, $i\in\{1,..,n\}$ достаточно просто вычислить, используя известные формулы комбинаторики. Действительно, можно рассмотреть 2^{n-1} возможных значений n-1 переменной, исключая x_i. Для того, чтобы функция была балансной и удовлетворяла SAC по переменной x_i, необходимо, чтобы на половине (2^{n-2}) этих пар наборов функция при изменении x_i меняла свое значение на противоположное, на одной четвертой (2^{n-3}) пар наборов принимала строго единичное значение вне зависимости от x_i и на оставшихся (2^{n-3}) наборах принимала нулевое значений. Выбор пар наборов, на которых функция меняет свое значение может быть произведен $C_{2^{n-1}}^{2^{n-2}}$ способами, среди оставшихся 2^{n-2} пар наборов выбор 2^{n-3} на которых функция принимает единичное значение, можно выполнить $C_{2^{n-2}}^{2^{n-3}}$ способами. Кроме того, на каждом из пары 2^{n-2} наборов возможно 2 варианта значения функции. Таким образом, общее количество K_1 балансных SAC-функций по одной переменной определяется формулой:

$$K_1 = C_{2^{n-1}}^{2^{n-2}} \cdot C_{2^{n-2}}^{2^{n-3}} \cdot 2^{n-2} \approx \frac{2^{3\cdot n+2.5}}{\pi} \qquad (5)$$

Соответственно, вероятность P_{x1} того, что балансная функция удовлетворяет SAC по одной фиксированной переменной равно:

$$P_{X1} = \frac{K_1}{2^{2^n}} \approx \frac{1}{\pi} \cdot 2^{-2^n+3\cdot n-2.5} \qquad (6)$$

Из формул (4) и (6) следует, что с ростом числа n переменных вероятность P_n того, что функция $f(x_1,x_2,\ldots,x_n)$ удовлетворяет SAC по всем n переменным уменьшается.

Зависимость вероятности P_{SAC} того, что в результате предложенной процедуры доопределения частично-заданная на $M=\xi\cdot 2^n$ наборах функция станет балансной и будет удовлетворять *SAC*, от значения ξ носит достаточно сложный характер. Это обусловлено тем, что на успешность доопределения функции значительное влияние оказывает не только число M наборов, на которых функция задана, но и от выбора указанных наборов. Ввиду указанных обстоятельств зависимость P_{SAC} от ξ исследовалась экспериментально. Результаты этих исследований приведены в таблице 4.

При нарушении защиты данных, построенной на основе булевых SAC-функций, используются методы линейного и дифференциального криптоанализа, базирующиеся на статистической обработке. Это делает целесообразным учет при анализе результативности разработанной процедуры доопределения функций не только строго SAC-функций, но и функций, близких к ним, и статистически не отличимых от них.

Для оценки степени близости булевой функции $\varphi(x_1,\ldots,x_n)$ к SAC-функции используется мера δ, численно равная отношению минимального числу R наборов, изменение значений функции $\varphi(x_1,\ldots,x_n)$ на которых делает ее удовлетворяющей критерию SAC к общему числу возможных наборов 2^n:

$$\delta = R\cdot 2^{-n}\cdot 100\%. \qquad (7)$$

Результаты исследований – зависимости выраженной в процентах статистической успешности (то есть получения функции, удовлетворяющей SAC с погрешностью δ) предложенной методики доопределения булевых функций от меры ξ предварительной ее заданности, числа n переменных и погрешности δ приведены в таблице 4.

Таблица 4. Зависимость успешности (в процентах) разработанной методики от влияющих на нее факторов.

ξ	δ	Результативность в процентах		
		***n*=4**	***n*=6**	***n*=8**
30	0	85.5	49,8	30.4
	1	85.9	55.5	54.9
	2	86.4	62.2	88.3
	3	87.3	69.2	100
	4	87.6	75.7	100
	5	88.1	82.4	100
40	0	83.7	49.1	25.6
	1	84.2	55.7	53.7
	2	84.6	62.4	87.3
	3	85.2	69.3	100
	4	85.8	75.6	100
	5	86.3	82.2	100
50	0	72.2	47,4	23.4
	1	72.9	54.2	52.6
	2	73.7	61.1	86.2
	3	74.8	67.8	100
	4	75.7	74.6	100
	5	76.7	81.4	100
60	0	47.5	43.1	22.4
	1	49.1	50.3	54.7
	2	50.6	57.4	90.1
	3	52.2	64.6	100
	4	53.8	71.7	100
	5	55.3	78.9	100
70	0	34.3	25,4	20.3
	1	35.1	30.5	53.5
	2	36.8	36.9	86.6
	3	38.6	43.4	100
	4	40.4	49.9	100
	5	42.2	56.3	100
80	0	16.3	7,5	14.4
	1	18.1	12.2	38.0
	2	20.2	17.2	61.7
	3	22.1	22.1	85.6
	4	23.9	27.3	90.1
	5	25.7	32.2	100

Выводы

В результате проведенных исследований разработана методика итеративного доопределения частично-заданных булевых функций, обеспечивающая их балансность и лавинные свойства, важные с точки зрения использования таких функций в системах защиты информации.

Предложенная методика доведена до уровня готовых к практическому использованию программных продуктов. Экспериментальными и теоретическими исследованиями доказано, что предложенная методика обеспечивает эффективное решение новой задачи автоматизированного проектирования компонент систем защиты информации, поз-

воляет на 1-2 порядка уменьшить время синтеза частично-заданных лавинных функций по сравнению с перебором.

Разработка может быть использована для создания перспективных средств защиты информации повышенной производительности.

Список литературы

1. Марковский А.П., Эль-Хами И., Рябуха Л.Р. “Метод одержання булевих балансних функцій, які задовольняють критерію чіткого лавинного ефекту“, Наукові Вісті НТУУ “КПІ“, 2001, № 2, с.31-40.
2. Kurosawa K., Satoh T. Design of SAC/PC(l) of Order k Boolean Functions and Three Other Cryptographic Criteria. // Advanced in Cryptology – Eurocrypto’97 Proceeding, Lecture Notes in Computer Science 1233 – 1997-P.433-449.
3. Самофалов К.Г., Марковский А.П. Комбинаторный подход к получению булевых функций, обладающих строгим лавинным эффектом // Электронное моделирование.- 2004,- Том. 26, - № 3, - с.27-40

Поступила в редакцию 8.12.2009