Нова функція витрат для покращення нелінійності бієктивних S-боксів

Абстрактні S-коробки є фундаментальними в конструкції симетричних блок-шифрів. Евристичні та метаевристичні алгоритми часто застосовуються для створення S-боксів з хорошими криптографічними властивостями. Результати, досягнуті цими методами, залежать від функцій придатності, що використовуються в процесі еволюції. Цей документ пропонує нову функцію придатності, щоб розвивати S-бокси з високою нелінійністю.

Ключові слова

S-бокси - Нелінійність - Еволюційні обчислення

1. Введення

Симетричні блок-шифри знаходяться в постійному процесі проектування та оновлення завдяки існуванню численних методів криптоаналізу, які загрожують безпеці, яку пропонують ці алгоритми. Замінні коробки (S-бокси або (n, m) -функції) є однією з найбільш досліджуваних складових блокових шифрів. S-бокси - це векторні булеві функції, які вносять невизначеність в процес шифрування. Існує три добре відомі критерії для проектування таких функцій: псевдовипадкова генерація, алгебраїчні конструкції [3, 12] та еволюційні алгоритми [2, 6, 7, 9].

Публічна література містить обширний огляд еволюційних обчислювальних робіт, пов’язаних з проектуванням та генерацією S-боксів, що мають хороші криптографічні властивості. Ми представляємо коротке резюме деяких важливих результатів у цій галузі досліджень.

У 2005 році Кларк та співавт. запропонували нову функцію витрат на еволюцію S-боксів у поєднанні з модельованим відпалом для отримання S-коробок із значеннями нелінійності до 102 [6]. Пізніше, у 2010 році, Tesaˇr виконує широке налаштування параметрів функції витрат Кларка, що в поєднанні зі спеціальним генетичним алгоритмом, який називається автором Genetic and Tree, дозволяє отримати 8×8 S-боксів з нелінійністю до 104 [14]. У 2013 р. Казимиров та співавт. представляють модифікований метод градієнтного спуску для отримання S-коробок з нелінійністю 104 та високим алгебраїчним опором [11]. Іванов та ін. проводять експеримент із модифікованим імунним алгоритмом із використанням трьох різних функцій витрат, включаючи налаштовану версію функції витрат Кларка, запропоновану Тесар(Tesaˇr), і функцію над диференційним спектром S-боксів для досягнення нелінійності 104 та диференційної однорідності 6 [9]. Ті самі автори використовують генетичні алгоритми, що працюють у зворотному режимі, для еволюції високих нелінійних бієктивних S-боксів розмірами від 8×8 до 16×16 [10]. У 2016 році Пічек(Picek) та співавт. представляють нову функцію витрат для еволюції високих нелінійних бієктивних S-боксів [13]. Mart´ınez-D´ıaz працює з локальними алгоритмами пошуку, щоб розвивати S-бокси з поліпшеним опором бічного каналу, що мають хороші значення нелінійності [7]. Болуфе та Тамайо(Boluf´e and Tamayo) використовують гібридні евристичні методи та машинне навчання для еволюції S-коробок, враховуючи властивості нелінійності та порядку прозорості [2]. Ахмад(Ahmad) та ін. запропонувати використовувати оптимізацію рою частинок для отримання 8×8 S-боксів з високими значеннями нелінійності [1].

Деякі з вищезазначених статей орієнтуються на оптимізацію більш ніж однієї властивості S-боксів [2, 7, 9, 10], тоді як інші повністю присвячені вивченню нелінійності [1, 6, 13, 14]. Як і ці статті, ця робота присвячена вдосконаленню значення нелінійності S-боксів шляхом визначення нової функції придатності для розвитку високих нелінійних бієктивних векторних булевих функцій.

Нова функція витягує інформацію з усього спектра Уолша-Адамара без використання будь-яких зовнішніх параметрів, таких як функція витрат Кларка [6] (реальні значення параметрів X і R) та функція витрат Піцека(Picek)[13] (параметр N). В цілому, ми аналізуємо тісний взаємозв'язок між значеннями нашої нової функції та властивістю нелінійності, показуючи, що єдине використання нашої функції, незалежно від значення нелінійності для допомоги в еволюційному процесі, призводить до генерації високих нелінійних бієктивних S-боксів.

2. Попередні етапи та методи

Одна  - це відображення з простору  у , де   . Коли  дорівнює 1,  називається булевою функцією, а якщо, то  називається векторною булевою функцією або S-боксом. Для цілей цієї статті ми використовуємо представлення S-боксу як списку значень (таблиця пошуку) в діапазоні від 0 до .

Перетворення Уолша-Адамара будь-якої  визначається як [4]:

 (1)

де . Лінійність , що позначається як , визначається максимальним абсолютним значенням перетворення Уолша-Адамара

 (2)

Нелінійність однієї дорівнює мінімальній нелінійності всіх нетривіальних лінійних комбінацій від компонентних функцій  [4]

 (3)

Підставляючи (2) у (3), можна виразити нелінійність  з точки зору його лінійності як

 (4)

Будь-яка  збалансована, кожне значення  відображається однаковою кількістю  разів. Коли , зазвичай буває  бієктивним (збалансованим), тобто кожне значення на виході з'являється рівно один раз. Збалансовані  - це перестановки у [4, 12]. Максимально досяжна нелінійність для збалансованих обмежена наступним виразом:

 (5)

У разі рівності такі функції називаються функціями майже зігнутих (Almost Bent - AB). Зверніть увагу, що функції AB існують лише тоді, коли  непарна [4]. Коли  парне, найкраща відома нелінійність досягається оберненою функцією і дорівнює [3]

 (6)

З рівнянь (4), (5) та (6) можна вивести мінімальні межі лінійності для збалансованих 

 (7)

 (8)

2.1 Функції, пов’язані з нелінійністю

Як показано в рівнянні (3), нелінійність S-боксу залежить від максимального абсолютного значення спектру Уолша-Адамара. Більшість статей, які використовують лише цінність властивості як функцію придатності в процесі розвитку, не можуть досягти нелінійності більше 100 у просторі 8×8 S-боксу. Це може статися, оскільки нелінійність містить лише інформацію про надзвичайну цінність спектру Уолша-Адамара, без вилучення будь-яких даних про інші значення спектру. Якщо ретельно проаналізувати рівняння (3), прямо зауважити, що зменшення значень усього спектру призводить до збільшення остаточної нелінійності S-боксів.

Беручи до уваги вищезазначений факт, Кларк та співавт. вводить функцію витрат, яка враховує всі значення спектра Уолша-Адамара для булевих функцій [5], розширених пізніше до S-боксів [6]. Ми позначаємо цю функцію як WHS і вона визначається як [6]

 (9)

де X і R є дійсними параметрами, такими як  і  [6]. Максимальна нелінійність, отримана Кларком та співавт. для 8×8 S-боксів з використанням WHS та модельованого відпалу було 102 [6]. Пізніше, в 2010 році, Тесар зробив велике налаштування параметрів на WHS, отримавши, що найкращий набір параметрів - R = 7; X = 21, після встановлення евристичного ранжування понад тисячі пар значень для X і R [14]. Найкращою нелінійністю, досягнутою налаштованою версією WHS, представленою Тесар, та алгоритмом, названим тим самим автором, Genetic and Tree, було 104 у випадку 8×8 S-боксу [14]. Основний недолік WHS, за думкою Picek та співавт. [13], полягає у виборі параметрів R і X, оскільки для отримання результатів із покращеною нелінійністю потрібно пройти фазу налаштування параметрів у кожному просторі S-бокс.

Щоб уникнути попередньої фази налаштування параметрів, Picek та співавт. запропонували представлення спектру Уолша-Адамара як гістограми частот для всіх абсолютних значень спектра [13]. Гістограма представлена ​​вектором, що має в i-й складовій кількість коефіцієнтів асоційоване з абсолютним значенням  із спектру Уолша-Адамара S-боксу. Нехай  - остання складова вектора з ненульовим значенням. Максимальне абсолютне значення спектра Уолша-Адамара дорівнює ; тому лінійність і нелінійність S-боксу задається цим значенням.

Тоді функція, що викликається у цьому документі PCF, визначається як [13]

 (10)

де  - k-та складова нульового індексованого вектора , яка представляє згадану гістограму частот, а . Помноживши на доданок , автори роблять вигляд, що дають певний рейтинг впливу кожного коефіцієнта на кінцеве значення функції. Наприклад, максимальне абсолютне значення спектра Уолша-Адамара помножується на , що означає, що воно є найбільш впливовим на кінцевий результат ПХФ. Загалом, Пічек та співавт. припускають, що значення параметра дорівнює 10, дане тим фактом, що якщо один S-бокс не має десяти рівнів коефіцієнтів (), то враховуються всі коефіцієнти спектра, коли розраховується PCF.

Пічек та співавт. провели експеримент з генетичним алгоритмом (GA) [8], алгоритм генетики та дерева (GaT) з [14] та їх власний алгоритм локального пошуку, що пояснюється далі в цій роботі. Переваги PCF щодо WHS чітко оцінюються для 8×8 S-боксів, особливо при переході між нелінійністю 102 і 104. У той час як налаштована версія WHS повинна переглянути більше трьох мільйонів рішень перед тим, як отримати нелінійність 104, за допомогою функції PCF автори [13] змогли створити S-бокси з нелінійністю 104, переглядаючи в середньому менше 200000 рішень.

2.2 Експериментальна установка

Ми використовуємо два різні еволюційні алгоритми для експериментів, проведених у Розділі 3. Перший з них - це локальний алгоритм пошуку (LSA), представлений у [13], а другий - простий алгоритм сходження на схил (HC), обидва пояснено в розділах 2.2. 2 та 2.2.3 відповідно.

2.2.1 Загальні параметри

Ми виконуємо 30 незалежних прогонів на кожному згаданому вище алгоритмі. Критерієм припинення одного алгоритму є досягнення фіксованої кількості оцінок рішення, або якщо найкращий знайдений S-бокс має нелінійність, яка дорівнює значенню, вказаному на вході алгоритму, зосередивши наше дослідження у випадку 8×8 S-блоків. На основі результатів, представлених у [6, 13, 14] для 8×8 S-блоків, ми вирішили, що критерієм зупинки для кожного алгоритму є оцінка максимум у 9000000 рішень або досягнення нелінійності 104.

2.2.2 Алгоритм локального пошуку

Локальний алгоритм пошуку з [13] отримує як вхідний сигнал випадковий S-бокс із простору. На кожній ітерації алгоритм виробляє  нових рішень, створених з заданими операторами мутації. Оператор мутації випадковим чином вибирає k різних позицій у поданні рішення, а потім випадковим чином переставляє елемент у вибрані позиції. Щоб досягти найкращих результатів, Picek та співавт. забезпечили оператори мутації .

Як вхідні дані до алгоритму LS надається набір операцій мутації. Кожен оператор мутації визначається двома параметрами *k* і *l*, де *k* - кількість позицій, елементи яких мають бути переставлені, а l визначає, скільки разів застосовуватиметься цей оператор мутації до поточного рішення у кожній ітерації [13] (таблиця 6, розділ 5, сторінка 11). З *M* знайдених рішень вибирається найкраще і встановлюється як поточне рішення.

2.2.3 Алгоритм сходження на пагорб

Ідея алгоритму високого сходження досить проста. Алгоритм бере як вхід випадковий S-бокс з простору. Потім, хоча кількість рішень для оцінки не вичерпується, а нелінійність найкращого рішення нижче бажаного, алгоритм повторює той самий цикл. По-перше, обираються дві випадкові позиції , а значення *x* та *y* поміняні місцями. Нарешті, якщо новий S-бокс , згенерований за допомогою мутації підкачки, кращий за поточне найкраще рішення алгоритму *S*, відповідно до умов відповідності задачі, тоді *S* замінюється на .

3. Результати та обговорення

Як ми вже згадували в розділі 2.1, незручність використання WHS полягає у відборі параметрів *X* і *R*, оскільки вони повністю незалежні від спектру Уолша-Адамара. З іншого боку, функцію PCF є більш точною для вимірювання поведінки коефіцієнтів у спектрі. Однак функція PCF також залежить від зовнішнього параметра *N*, який, хоча і обґрунтовано, не має відношення до коефіцієнтів у спектрі Уолша-Адамара. Зауважимо, що автори [13] аналізують ефективність PCF при  і . Якщо , результати, отримані за допомогою PCF, є тими, що згадані в розділі 2.1, але коли , незважаючи на авторський звіт S-боксів з нелінійністю 104, жоден з алгоритмів не міг багаторазово виробляти S-бокси з такою нелінійністю. Після швидкого огляду можна зрозуміти, що втручання будь-якого зовнішнього параметра, не пов'язаного зі спектром Уолша-Адамара, призводить до різних результатів для різної конфігурації параметрів функцій витрат.

3.1 Вилучення інформації із спектру Уолша-Адамара

З рівнянь (7) та (8) відомо про мінімально досяжне значення максимального абсолютного значення спектру Уолша-Адамара n×n бієктивних S-боксів. Позначимо множину , коли *n* - непарне і , коли *n* - парне. Набір *C* містить усі можливі абсолютні значення, менші або рівні найкращій лінійності, що досягається для бієктивних S-боксів. Наша нова функція заснована на принципі, що максимальне абсолютне значення спектру Уолша-Адамара не може бути нижчим за (7) або (8), а коли це значення мінімальне, тоді всі абсолютні значення коефіцієнтів спектра міститься в множині *C*. Таким чином, ми визначаємо нашу функцію, яка називається WCF, як

 (11)

Враховуючи будь-яке абсолютне значення спектра Уолша-Адамара в *С*, значення WCF не змінюється, оскільки добуток, що включає абсолютне значення , дорівнює нулю. Отже, остаточне значення WCF виводиться з тих коефіцієнтів, абсолютне значення яких перевищує рівняння (7) та (8). Таким чином, мінімальне значення WCF досягається, коли максимальне абсолютне значення спектра Уолша-Адамара задовольняє рівність із рівнянь (7) та (8), де WCF дорівнює нулю.

Таблиця 1. Середні оцінки рішень для досягнення бажаної нелінійності.



3.1.1 Оцінка WCF як функції витрат

Як ми пояснюємо в розділі 2.2.1, кожен алгоритм зупиняється після перегляду дев'яти мільйонів рішень або якщо найкраща досягнута нелінійність дорівнює 104. Для виконання алгоритмів, представлених у розділі 2.2, ми використовуємо кортеж  для порівняння, якщо S-бокс *A* кращий, ніж S-бокс *B*, наступним чином:

1. 

2. , якщо умова 1 не виконується.

У таблиці 1 наведені середні оцінки рішень із використанням функцій витрат WHS, PCF та WCF. Результати, наведені в таблиці 1 для функцій WHS та PCF, є результатами, представленими в [13] для найкращої конфігурації параметрів обох функцій витрат.

Щоб встановити деяке порівняння продуктивності трьох функцій вартості, ми використовуємо конфігурацію локального алгоритму пошуку для WHS,  та , локальний алгоритм пошуку, а також генетичний алгоритм, використовуваний Picek та співавторами, не може багаторазово виробляти S-бокси з нелінійністю 104. Навпаки, всі алгоритми, представлені в [13], змогли створити S-бокси, що мають нелінійність 104, якщо замість них використовується функція витрат PCF. У приватному випадку алгоритму локального пошуку нелінійність 104 досягається в середньому за менш ніж 200000 оцінок рішення. Нарешті, ми розглядаємо результати, досягнуті за допомогою функції, представленої в цій роботі. При використанні WCF алгоритм локального пошуку зменшує на **11, 32.5** та **49.1** відсотка середню кількість оцінок рішення для отримання нелінійності 100, 102 та 104 відповідно порівняно з результатами, отриманими за допомогою PCF.

  
Рисунок 1. Найкраща кількість оцінок рішень для отримання нелінійності 104 з використанням алгоритму сходження на пагорб та кортежу  як функції витрат.

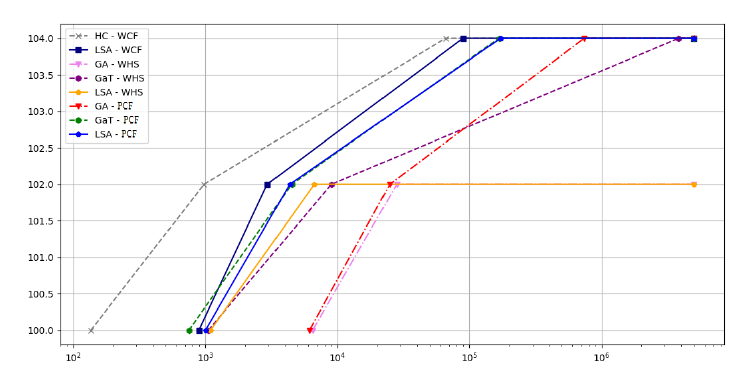
Згадуючи інформацію з таблиці 1, можна помітити, що алгоритм сходження на пагорб виробляє високоякісні рішення, аналізуючи в середньому значно меншу кількість S-боксів. Ця ефективність є результатом зосередження зусиль алгоритму над одним рішенням. На рисунку 1 показано найкращу кількість оцінок рішень для отримання нелінійності 104 з використанням алгоритму сходження на пагорб та WCF як функції витрат. Потім, на основі результатів, представлених у таблиці 1, та інформації з рисунків 1 і 2, ми гарантуємо, що поєднання алгоритму підйому на пагорб і нашої функції призводить до дуже ефективної процедури для отримання високих нелінійних S-боксів.

3.1.2 Оцінка WCF як функції придатності

Експерименти, проведені в попередньому розділі, були розроблені для використання WCF як функції витрат для поліпшення нелінійності S-боксів результатів еволюційних алгоритмів. У поточному розділі ми відновлюємо результати, досягнуті після того, як функції WHS та WCF були прийняті як функції придатності для керування еволюційним процесом, незалежно від значення нелінійності аналізованих S-боксів. Ми не включаємо функцію PCF в наші експерименти, оскільки інтуїтивно зрозуміло, що PCF є функцією, що не залежить від нелінійності.

Враховуючи переваги алгоритму сходження на пагорб, щоб знайти хороші рішення швидше, ніж будь-який інший алгоритм, згаданий у цій роботі, ми пропонуємо наступний експеримент:

1. Виберіть функцію придатності  із набору функцій    
2. Випадково згенеруйте 8×8 S-бокс  
3. Встановіть максимальну кількість оцінок рішень 2,5 мільйона  
4. Встановіть бажану нелінійність 104  
5. Виконайте алгоритм сходження на пагорб із початкового S-боксу з кроку 2 та критеріями заміни рішення між двома S-боксами , якщо , поки не будуть виконані умови 3 або 4.

  
Рисунок 2. Збіжність алгоритмів, представлених у таблиці 1 із відповідною функцією витрат.

Виявляється, для WHS найкраща нелінійність, досягнута алгоритмом сходження на пагорб, нижча ніж 100, що є нелінійністю S-боксу, яка в більшості випадків дається як вхідний алгоритм. Загалом, ми розглядаємо, що кінцеві S-бокси мають значення нелінійності нижче або дорівнюють 90 у більшості експериментів, менше ніж 80 для деяких отриманих S-боксів. Це показник того, що, як і PCF, функція WHS залежить від значення нелінійності аналізованого S-боксу, і повинна використовуватися як функція витрат поряд з нелінійністю S-боксу.

Ми спостерігаємо різну поведінку під час аналізу результатів, отриманих за допомогою функції, представленої в цій роботі. На ранніх етапах алгоритму сходження на пагорб виробляються S-бокси зі значеннями нелінійності, рівними випадково сформованим рішенням. Однак такий процес не споживає значної кількості оцінок рішення до того, як алгоритм почне створювати високі нелінійні S-бокси. Отримавши результати експериментальної фази, ми гарантуємо, що S-бокси, що мають нелінійність 102, неодноразово виробляються алгоритмом після 10000 оцінок рішення. Навіть коли алгоритм сходження на пагорб досягає 2,5 мільйонів оцінок рішення, перш ніж досягти нелінійності 104 у більш ніж половині експериментів, є деякі випадки, коли були виявлені S-бокси, що мають це значення нелінійності. Таблиця 2 відображає результати після 30 незалежних експериментів із використанням алгоритму сходження на пагорб, використовуючи WHS та WCF як функції придатності.

Таблиця 2. Нелінійність, досягнута в кожному експерименті із загального числа 30 алгоритмів сходження на пагорб з використанням функції придатності WHS та WCF.

C:\Users\yuliy\Downloads\Cтеганография\Перевод\2.png

Експерименти, проведені в розділах 3.1.1 та 3.1.2, показують, що наша нова функція може бути використана як функція витрат, так і як функція придатності для еволюції S-боксів з високими значеннями нелінійності, пропонуючи хороші показники для отримання бажаних результатів.

4. Висновки

У цій роботі представлена ​​нова функція оптимізації нелінійності. Показано, що нова функція зменшує кількість оцінок рішення для досягнення хороших значень нелінійності за допомогою еволюційних алгоритмів та існуючого зв'язку між значенням функції та нелінійністю S-блоків. Подальша робота буде спрямована на вивчення зменшення обчислювальних витрат цієї нової функції та можливості досягнення вищих значень нелінійності, починаючи із випадково сформованих S-боксів.

Подяка

Автори хочуть подякувати Карлосу Мігелю Леґону Перезу та Рейньє Антоніо де ла Крус Джименес за корисні коментарі та огляди для покращення якості цієї статті.

Список літератури

[1] Musheer Ahmad, Ishfaq Ahmad Khaja, Abdullah Baz, Hosam Alhakami, and Wajdi Alhakami. Particle swarm optimization based highly nonlinear substitution-boxes generation for security applications. IEEE Access, 8:116132–116147, 2020.

[2] Antonio Boluf´e-R¨ohler and Dania Tamayo-Vera. Machine learning based metaheuristic hybrids for s-box optimization. Journal of Ambient Intelligence and Humanized Computing, pages 1–14, 2020.

[3] Lilya Budaghyan, Claude Carlet, and Alexander Pott. New classes of almost bent and almost perfect nonlinear polynomials. IEEE Transactions on Information Theory, 52(3):1141–1152, 2006.

[4] Claude Carlet, Yves Crama, and Peter L Hammer. Vectorial boolean functions for cryptography., 2010.

[5] John A Clark, Jeremy L Jacob, and Susan Stepney. Searching for cost functions. In Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No. 04TH8753), volume 2, pages 1517–1524. IEEE, 2004.

[6] John A Clark, Jeremy L Jacob, and Susan Stepney. The design of s-boxes by simulated annealing. New Generation Computing, 23(3):219–231, 2005.

[7] Ismel Mart´ınez D´ıaz. B´usqueda local de s-cajas con alta varianza del coeficiente de confusi´on. Tesis de maestr´ıa, Universidad de la Habana, 2019.

[8] Agoston E Eiben, James E Smith, et al. Introduction to evolutionary computing. Springer, 2003.

[9] Georgi Ivanov, Nikolay Nikolov, and Svetla Nikova. Cryptographically strong s-boxes generated by modified immune algorithm. In International Conference on

Cryptography and Information Security in the Balkans, pages 31–42. Springer, 2015.

[10] Georgi Ivanov, Nikolay Nikolov, and Svetla Nikova. Reversed genetic algorithms for generation of bijective sboxes with good cryptographic properties. Cryptography and Communications, 8(2):247–276, 2016.

[11] Oleksandr Kazymyrov, Valentyna Kazymyrova, and Roman Oliynykov. A method for generation of highnonlinear s-boxes based on gradient descent. IACR Cryptology ePrint Archive, 2013:578, 2013.

[12] Kaisa Nyberg. On the construction of highly nonlinear permutations. In Workshop on the Theory and Application of of Cryptographic Techniques, pages 92–98. Springer, 1992.

[13] Stjepan Picek, Marko Cupic, and Leon Rotim. A new cost function for evolution of s-boxes. Evolutionary computation, 24(4):695–718, 2016.

[14] Petr Tesaˇr. A new method for generating high nonlinearity s-boxes. Radioengineering, 19(1):23–26, 2010.