Касаткин А. М., Касаткина Л. М., Талаев С. А. Автоматы и разумное поведение. –Киев: Наук думка, 1973. – 370 с.

УДК 681.3.06:519.248.681

ТЕСТ НА ПРОФІЛЬ ЛІНІЙНОЇ СКЛАДНОСТІ

Людмила Завадська, Максим Семибаламут
ФТІ НТУУ «КПІ»

Анотація: **Наведено новий тест для оцінки якості випадкових послідовностей, який базується на профілі лінійної складності. Тест оперує кількістю стрибків лінійної складності і працює у певних випадках значно ефективніше за відповідний тест з набору NIST.**
Summary: **A new randomness test, based on linear complexity profile is proposed. It takes into account the number of linear complexity jumps and shows in some cases much better results than the NIST linear complexity test does.**
Ключові слова: **Криптографія, випадкові послідовності, тести оцінки якості, лінійна складність, профіль лінійної складності.**

I Вступ

Задача оцінки якості випадкових та псевдовипадкових послідовностей посідає важливе місце у сучасній теоретичній та прикладній криптографії. Висока якість вихідної послідовності є необхідною умовою стійкості потокових систем шифрування; якісні випадкові послідовності потрібні для генерування ключів і різноманітних випадкових параметрів симетричних та асиметричних криптосистем тощо.

Під якісною бінарною послідовністю розуміють послідовність бітів, близьку за своїми характеристиками до чисто випадкової послідовності – послідовності незалежних бітів з рівними ймовірностями 0 та 1. Існують набори тестів для оцінки якості послідовностей, найвідомішим серед яких є набір статистичних тестів NIST (National Institute of Standards and Technology) [1]. Проте задача побудови нових, більш ефективних критеріїв для виявлення відхилень від випадковості залишається актуальною.

Важливою характеристикою бінарної випадкової послідовності є лінійна складність – довжина найкоротшого регістра зсуву з лінійним зворотним зв'язком, що може згенерувати дану послідовність. Однак високої лінійної складності ще недостатньо для того, аби послідовність була якісною з точки зору випадковості. Вона, наприклад, може бути звичайною рекурентою, до якої в кінці додано коротку послідовність випадкових бітів, і попри все це мати високу лінійну складність. Зокрема набір тестів NIST містить і тест на лінійну складність, який не враховує усіх нюансів її поведінки і тому пропускає певні типи неякісних послідовностей. Вхідна послідовність розбивається на відносно короткі відрізки, для кожного з яких за допомогою алгоритму Берлекемпа-Мессі [2, 3] підраховується значення лінійної складності без урахування специфіки її зміни на цьому відрізку. Отримані значення перевіряються за χ^2-критерієм на відповідність певному розподілу. У даній статті пропонується побудова та обґрунтування нового тесту, який враховує кількість стрибків лінійної складності на відрізках, і у певних випадках працює краще за тест на лінійну складність з набору NIST.

Під час роботи над матеріалом авторами було виявлено в Інтернеті анонс роботи [4]. На жаль, в нас немає доступу до тексту статті, тому нам невідомо, яку статистику використовували її автори. Можна лише констатувати, що тест LP було розроблено незалежно від [4].

II LP-тест

Профіль лінійної складності випадкової послідовності являє собою графік зміни лінійної складності L_n залежно від довжини послідовності n. Для чисто випадкової послідовності профіль стрибкоподібно зростає, коливаючись певним чином навколо прямої

$$L_n = n/2, \tag{1}$$

причому великі стрибки L_n малоймовірні внаслідок обмеженості дисперсії лінійної складності [5]. На рисунку 1, запозиченому з [5], зображено типовий вид профілю лінійної складності якісної випадкової, а також лінійної рекурентної послідовностей. Якщо в послідовності присутні достатньо довгі відрізки лінійної рекуренти, які чергуються з відрізками якісної послідовності, або лінійна рекурента зрідка спотворюється, це має призводити до великих стрибків профілю лінійної складності. В свою чергу, наявність великих стрибків лінійної складності спричиняє зменшення їх кількості. Тож авторами пропонується як основну

характеристику випадковості послідовності розглядати M_n – кількість стрибків профілю лінійної складності на відрізку довжини *n*. Незначна модифікація алгоритму Берлекемпа-Мессі дозволяє поряд з лінійною складністю обчислювати й кількість її стрибків для послідовності довільної довжини *n*.

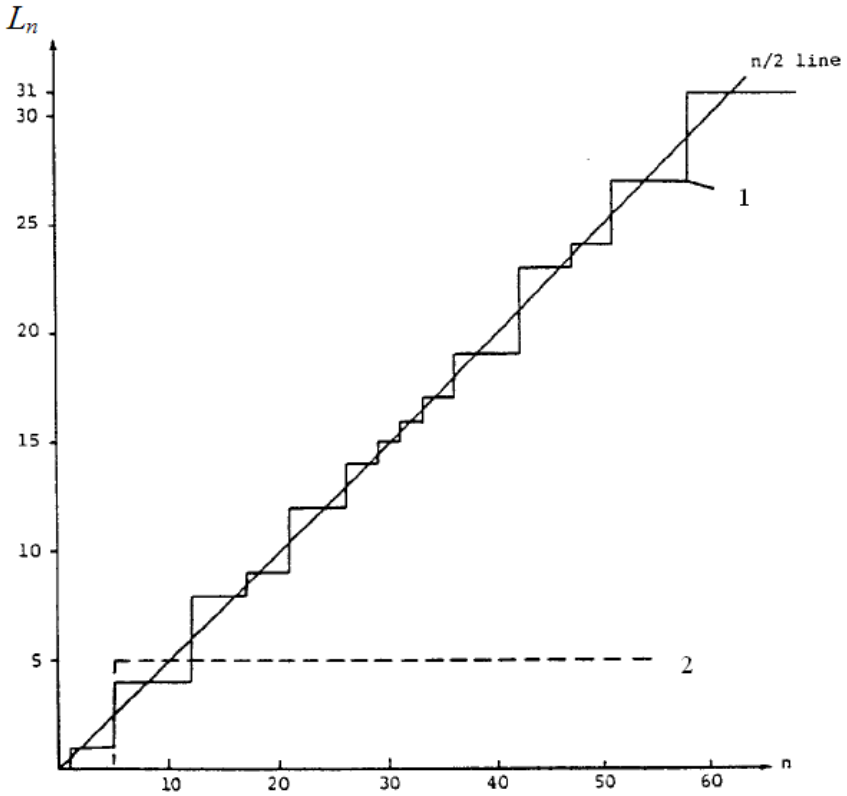

Рисунок 1 – Профілі лінійної складності якісної послідовності, отриманої шляхом підкидання симетричної монети (1), та лінійної рекурентної послідовності (2)

Для побудови критерію перевірки гіпотези H_0 про чисту випадковість тестованої послідовності необхідно мати принаймні асимптотичний при $n \to \infty$ розподіл обраної статистики за умови вірності H_0. За значенням характеристики M_n легко знайти величину N_n – кількість моментів t на відрізку послідовності довжини n, в яких виконується рівність $L_t = t/2$, тобто кількість моментів перетину (без стрибка) профілем лінійної складності прямої (1). На рисунку 2 t_i, t_{i+1} – два послідовні моменти з вказаною властивістю. Розподіл же N_n, як показано нижче, знаходиться доволі просто.

Нагадаємо властивості профілю лінійної складності [2], [3], [5]:

1) лінійна складність – неспадна функція часу; $L_0 = 0$;

2) стрибок профілю лінійної складності можливий лише при $L_n \le n/2$;

3) якщо $L_{n+1} > L_n$, то $L_{n+1} = n + 1 - L_n$, тобто стрибки лінійної складності симетричні відносно прямої (1).

Враховуючи наведені властивості, неважко бачити, що за гіпотези H_0 моменти t_i є моментами регенерації випадкового процесу $L_n - n/2$. Дійсно, внаслідок симетричності профілю лінійної складності відносно прямої (1) горизонтальні відрізки профілю претинають цю пряму у парні моменти часу. Таким чином, між кожними двома послідовними стрибками лінійної складності неодмінно існує момент t_i. Після настання події $L_t = t/2$ наступний стрибок складності може виникнути у будь-який момент з імовірністю ½

незалежно від передісторії. Величина стрибка залежить лише від часу, що пройшов від попереднього моменту t_i, а після стрибка значення лінійної складності не може змінитися аж до наступного моменту t_{i+1}. Таким чином, поведінка процесу $L_n - n/2$ після моменту t_i не залежить від його поведінки до цього моменту, а також від значення t_i.

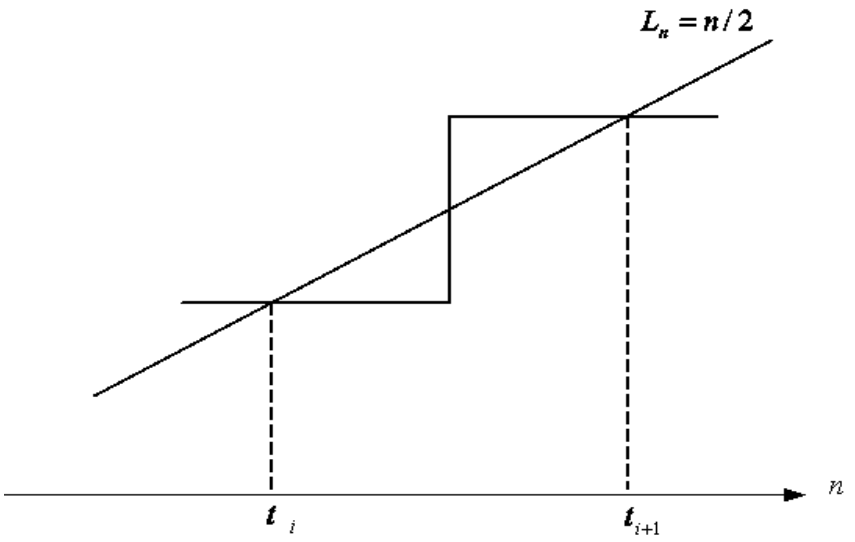

Рисунок 2 – Моменти регенерації випадкового процесу L_n - $n/2$

Величини M_n та N_n пов'язані між собою очевидним співвідношенням:

$$N_n = M_n - 1, \text{ якщо } L_n > n/2, \text{ інакше } N_n = M_n. \quad (2)$$

Для знаходження розподілу кількості N_n моментів регенерації на проміжку довжини n скористаємося результатами, викладеними в [6], з яких випливає, що N_n має асимптотично нормальний розподіл з середнім n/m та дисперсією nD/m^3, де m - математичне сподівання, а D –дисперсія часу між двома послідовними моментами регенерації. Відомо [5], що для чисто випадкової послідовності $m = 4$; нескладні підрахунки, виконані авторами, дають $D = 8$. Таким чином, випадкова величина

$$S_n = (N_n - n/4)/\sqrt{n/8} \quad (3)$$

має стандартний нормальний асимптотичний розподіл при виконанні гіпотези H_0 і $n \to \infty$.

Випадкова величина (3) покладена в основу тесту, названого авторами LP-тестом (Linear Profile-тест), який працює наступним чином. Тестована послідовність довжини n розбивається на k блоків довжини $r = [n/k]$ (останній блок відкидається, якщо він коротший за r). Як і у тесті NIST, пропонується обирати n порядка 10^6, $500 \le r \le 5000$ (проте це питання вимагає додаткового дослідження). Для кожного блоку за допомогою алгоритму Берлекемпа-Мессі підраховується кількість стрибків профілю лінійної складності $M_r(i)$ і за формулами (2), (3) – відповідні значення $S_r(i),\ i = 1,...,k.$ Обчислюється значення статистики

$$S = \sum_{i=1}^{k} S_r(i)/\sqrt{k},$$

яка також, як неважко бачити, має стандартний нормальний розподіл, якщо таким чином розподілені величини $S_r(i)$. За обраним рівнем значущості α визначається критична множина $X_\alpha = (-\infty, -t_{1-\alpha/2}) \cup (t_{1-\alpha/2}, +\infty)$, де $t_{1-\alpha/2}$ – квантиль стандартного нормального розподілу порядка $1-\alpha/2$. При $S \notin X_\alpha$ послідовність проходить тест, інакше тестована послідовність вважається невипадковою.

Швидкість роботи побудованого тесту приблизно така сама, як і швидкість тесту на лінійну складність з пакету NIST, так як в основі обох тестів лежить алгоритм Берлекемпа-Мессі; проте реалізація LP-тесту дещо простіша, адже в ньому статистика має стандартний розподіл на відміну від специфічного розподілу статистики у тесті NIST.

III Експериментальні результати

Було проведено порівняльний експериментальний аналіз тесту LP та тесту на лінійну складність з набору NIST. Якісні послідовності формувалися за допомогою BBS-генератора та модифікації вбудованого у C++ генератора випадкових чисел, що задається наступним чином:

double rnd = (static_cast<double>(rand())+1)/(RAND_MAX+1).

Вихідні послідовності останнього проходять усі тести NIST.

Серед 200 якісних послідовностей довжиною 1000000 бітів кількість тих, що не пройшли LP-тест, так само як і кількість послідовностей, відбракованих тестом на лінійну складність з набору NIST, виявилась у межах, що узгоджуються з обраним рівнем значущості $\alpha = 0{,}01$. У той же час всі 200 лінійних рекурентних послідовностей такої самої довжини з різними періодами були відкинуті обома тестами.

В результаті подальших досліджень було виявлено, що тест LP набагато ефективніший принаймні на наступних типах неякісних вхідних послідовностей (тут шум означає інвертування кожного біту з імовірністю *p*):

a. Лінійні рекурентні послідовності з шумом.
b. Послідовності, сформовані шляхом регулярного або випадкового чергування відрізків різних лінійних рекурентних послідовностей.
c. Послідовності, сформовані шляхом регулярного або випадкового чергування відрізків лінійних рекурентних послідовностей та відрізків, утворених за допомогою гарного генератора псевдовипадкових чисел.
d. Послідовності, сформовані як зазначено у попередньому пункті, з шумом.
e. Послідовності, сформовані з лінійних рекурентних послідовностей шляхом випадкового видалення бітів.

В наступних таблицях розміщені деякі результати порівняльного аналізу; «+» означає, що неякісна послідовність пройшла тест, «–» - бракування послідовності.

Таблиця 1 – Результати тестування лінійних рекурентних послідовностей з шумом, *p = 0.01*

Довжина блоку	Тест LP				Тест NIST			
	Номер послідовності							
	1	2	3	4	1	2	3	4
500	-	-	-	-	-	-	-	-
1000	-	-	-	-	+	-	+	+
2000	-	-	-	-	+	+	+	+
3000	-	-	-	-	+	+	+	+
4000	-	-	-	-	+	+	+	+
5000	-	-	-	-	+	+	+	+

Таблиця 2 – Результати тестування лінійних рекурентних послідовностей з шумом, *p = 0.02*

Довжина блоку	Тест LP				Тест NIST			
	Номер послідовності							
	1	2	3	4	1	2	3	4
500	-	-	-	-	+	+	+	+
1000	-	-	-	-	+	+	+	+
2000	-	+	-	+	+	+	+	+
3000	-	+	+	+	+	+	+	+
4000	+	+	+	+	+	+	+	+
5000	+	+	+	+	+	+	+	+

Таблиця 3 – Результати тестування послідовностей, утворених шляхом випадкового чергування лінійних рекурент та якісного генератора псевдовипадкових чисел, з шумом, $p = 0.01$

Довжина блоку	Тест LP			Тест NIST		
	Номер послідовності					
	1	2	3	1	2	3
500	-	-	-	-	-	+
1000	-	-	-	+	+	+
2000	-	-	-	+	+	+
3000	-	-	-	+	+	+
4000	-	-	-	+	+	+
5000	-	-	-	+	+	+

Таблиця 4 – Результати тестування послідовностей, утворених шляхом випадкового чергування лінійних рекурент та гарного генератора псевдовипадкових чисел, з шумом, $p = 0.02$

Довжина блоку	Тест LP			Тест NIST		
	Номер послідовності					
	1	2	3	1	2	3
500	-	-	-	+	+	+
1000	-	-	-	+	+	+
2000	+	+	-	+	+	+
3000	+	+	+	+	+	+
4000	+	+	+	+	+	+
5000	+	+	+	+	+	+

IV Висновки

Запропоновано новий тест оцінки випадковості, що базується на властивостях профілю лінійної складності випадкової послідовності. Статистика тесту LP заснована на кількості стрибків лінійної складності на відрізках вхідної послідовності певної довжини. На деяких типах неякісних послідовностей новий тест показує кращі результати, ніж тест на лінійну складність з набору NIST; крім того, тест LP має більш просту реалізацію через нормально розподілену статистику, на відміну від специфічного розподілу у тесті NIST. Швидкість роботи обох тестів приблизно однакова, адже обидва використовують для розрахунків алгоритм Берлекемпа-Мессі.

Література: ***1.*** *NIST Special Publications 800-22, A Statistical Test Suite for the Validation of Random Number Generators and Pseudo Random Number Generators for Cryptographic Applications. – 2000.* ***2.*** *J. L. Massey. Shift-register synthesis and BCH decoding // IEEE Trans. Information Theory , 1969. – IT-15 (1). – pp. 122–127.* ***3.*** *A. J. Menezes, P. C. van Oorschot, S. A. Vanstone. Handbook of Applied Cryptography. – CRC Press, 1996. – 816 p.* ***4.*** *K. Hamano, F. Sato, H. Yamamoto. A new randomness test based on linear complexity profile // IEICE Trans. on Fundamentals of Electronics, Communications and Computer Science, E92.A(2009). – No1. – pp.166-172.* ***5.*** *R. A. Rueppel. Analysis and Design of Stream Ciphers. – N.Y.: Springer-Verlag, 1986. – 236 p.* ***6.*** *В. Феллер. Введение в теорию вероятностей и ее приложения. В 2-х томах. Т1. – М.: МИР, 1967 – 498 с.*

УДК 681.3.06

УНИВЕРСАЛЬНОЕ ХЕШИРОВАНИЕ ПО РАЦИОНАЛЬНЫМ ФУНКЦИЯМ АЛГЕБРАИЧЕСКИХ КРИВЫХ В КУБИЧЕСКОМ ПОЛЕ

Геннадий Халимов
Харьковский национальный университет радиоэлектроники

Аннотация: **Представлено универсальное хеширование по рациональным функциям алгебраических кривых с большим числом точек в кубическом поле.**
Summary: **Present the universal hashing of rational functions of algebraic curves with many points in a cubic field.**