POZNAN UNIVERSITY OF TECHNOLOGY ACADEMIC JOURNALS
No 73 Electrical Engineering 2013

Rafał STĘPIEŃ*
Janusz WALCZAK*

ANALIZA WŁAŚCIWOŚCI STATYSTYCZNYCH SYGNAŁÓW PSEUDOLOSOWYCH GENERATORÓW ZBUDOWANYCH NA REJESTRACH PRZESUWNYCH

W artykule opisano wyniki testów statystycznych sekwencji wyjściowej generatora pseudolosowego zrealizowanego na rejestrze przesuwnym i liniowym sprzężeniu zwrotnym (ang. Linear Feedback Shift Register). Przedstawiono budowę generatorów LFSR oraz opisano wykorzystane testy statystyczne. Do analizy sekwencji wyjściowej generatora wykorzystano pakiet testów statystycznych DIEHARD. Omówiono sposób interpretacji danych uzyskiwanych z pakietu DIEHARD i wyniki analizy testowej sekwencji pseudolosowej.

1. WSTĘP

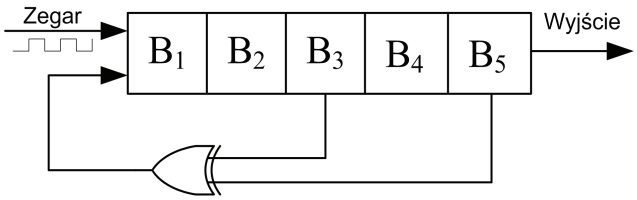
Generatory sekwencji pseudolosowych stosowane są w wielu systemach szyfrujących np. w szyfrach strumieniowych [1], w skramblingu i descramblingu systemu TETRA [2] lub w systemie A5 stosowanym w telefonii GSM [1]. Wszystkie systemy korzystające z generatorów pseudolosowych są tak bezpieczne jak bezpieczna (trudna do odtworzenia) jest sekwencja generatora. Określenie tego bezpieczeństwa nie jest łatwe [3], [4] a jedną z powszechnie stosowanych metod badania sekwencji pseudolosowych są testy statystyczne [3], [4], [5].

2. BUDOWA GENERATORÓW LFSR

Jedną z klas generatorów sygnałów pseudolosowych stanowią generatory zbudowane na rejestrach przesuwnych. Klasa ta obejmuje dwa podstawowe typy generatorów: z liniowym sprzężeniem zwrotnym LFSR (ang. Linear Feedback Shift Register) [1] oraz nieliniowym sprzężeniem zwrotnym NLFSR (ang. Non Linear Feedback Shift Register) [1]. Parametry pętli sprzężenia zwrotnego obu typów generatorów są niezmienne w czasie. Sprzężenie zwrotne opisuje się najczęściej wielomianem [1].

Na rysunku 1 pokazano przykładowy schemat generatora LFSR opisany wielomianem (1), którego potęgi zmiennej x są numerami poszczególnych bitów rejestru przesuwnego.

* Politechnika Śląska.



Rys. 1. Przykładowy generator LFSR

$$L(x) = x^5 + x^3 + 1 \tag{1}$$

Układy generatorów LFSR stosowane są w wielu aplikacjach m.in. w telekomunikacji i kryptografii [1], [6], [7]. Dodatkowe szczegóły dotyczące budowy oraz opisu generatorów LFSR można znaleźć w [8], [9].

3. TESTY STATYSTYCZNE

Sekwencje pseudolosowe stosowane w technice uzyskiwane są w wyniku działania w pełni deterministycznych algorytmów, wiec nie mogą zapewnić 100 procentowego bezpieczeństwa szyfrowania. Prawdopodobieństwo odtworzenia sekwencji pseudolosowej wykorzystanej do szyfrowania jest zależne od konstrukcji generatora. Im więcej „defektów" statystycznych i prawidłowości w sekwencji pseudolosowej generatora tym prościej złamać system kryptograficzny stosujący taki generator.

Jedną z metod oceny jakości sekwencji pseudolosowej jest ocena podobieństwa sekwencji pseudolosowej do sekwencji w pełni losowej i na tej podstawie określenie, czy stosowany do szyfrowania generator pseudolosowy zapewnia założony poziom bezpieczeństwa. Do tego celu wykorzystywane są testy statystyczne takie jak: DIEHARD, ENT, SP800-22 oraz inne, np. DIEHARDER lub wybrane testy nieparametryczne [4]. Stwierdzają one z pewnym prawdopodobieństwem losowość sekwencji pseudolosowej – nazywane jest to testowaniem hipotezy H_0 (ang. Null Hypothesis). Celem każdego z tych testów jest stwierdzenie, czy badana hipoteza H_0 jest spełniona z założonym poziomem istotności. Poziom istotności α oznacza prawdopodobieństwo odrzucenia hipotezy H_0 jeśli jest ona prawdziwa. Poziom istotności α przyjmuje się najczęściej jako wartość z zakresu 0,001 do 0,1 [3].

Błędna interpretacja wyników prowadzi do dwóch błędów w testowaniu [3]:

1. Błąd I rodzaju – uznaje się hipotezę H_0 za fałszywą gdy jest ona prawdziwa, czyli odrzuca się sekwencję losową.
2. Błąd II rodzaju – uznaje się hipotezę H_0 za prawdziwą gdy jest ona fałszywa, czyli uznaje się sekwencję z defektami statystycznymi jako sekwencję w pełni losową.

Z obu wymienionych błędów w testowaniu gorszy w swoich skutkach jest błąd II rodzaju. Popełnienie tego błędu skutkuje użyciem generatora o niewłaściwych parametrach statystycznych, co zmniejsza bezpieczeństwo systemu wykorzystującego ten generator.

3. PAKIET DIEHARD

Zestaw testów DIEHARD został opracowany przez George'a Marsagalie [10]. Składa się on z 18 testów. Dokładny opis testów wraz z źródłami w języku C można znaleźć na stronie internetowej autora [10]. Pakiet DIEHARD działa zarówno w środowisku Windows (dostępna jest gotowa aplikacja w formie pliku wykonywalnego) jak i Linux. Wynikiem działania pakietu jest liczba lub zestaw liczb – tzw. p-wartości . Liczby te nazywa się prawdopodobieństwami testowymi [3]. Oznaczają one prawdopodobieństwo popełnienia błędu I rodzaju przy testowaniu hipotezy H_0. W pakiecie DIEHARD, w zależności od rodzaju testu, zwracana jest pojedyncza p-wartość (wynik testu Kołmogorowa Smirnowa na serii p-wartości lub wynik testu χ^2 na danych uzyskanych podczas testu) lub seria p-wartości. Serię p-wartości uzyskuje się w testach, które dzielą analizowaną sekwencję na mniejsze podsekwencje i wykonują dany test statystyczny na każdej podsekwencji. Jeżeli analizowana sekwencja jest w pełni losowa to uzyskane p-wartości powinny być rozłożone równomiernie na przedziale [0;1]. W przypadku testów zwracających jedną p-wartość powinna się ona zawierać pomiędzy wartością α lub $1-\alpha$. Przy założonym poziomie istotności α sekwencja nie przechodzi testu, jeżeli [3] zachodzi nierówność

$$\text{p-wartość} < \alpha \text{ lub p-wartość} > 1-\alpha \qquad (2)$$

Zgodnie z uwagą Autora pakietu DIEHARD, nawet dobre generatory sygnałów pseudolosowych mogą dla serii p-wartości nie spełniać relacji określonej wzorem (2). Nie oznacza to, że generator ten należy odrzucić. Pakiet DIEHARD podczas testów sekwencji, która posiada defekty statystyczne, zwróci wiele p-wartości leżących bardzo blisko krańców przedziału [0;1] np. 0,002 lub 0,9971. Na tej podstawie należy odrzucić generator badanej sekwencji pseudolosowej, ponieważ rozkład tych wartości nie jest rozkładem równomiernym.

4. POMIARY STATYSTYCZNE GENERATORA LFSR

Sekwencja testowa, wykorzystana do testowania, została wygenerowana przez 32-bitowy generator LFSR. Wielomian sprzężenia zwrotnego był pierwotny [1] i określony wzorem:

$$L(x) = x^{32} + x^{31} + x^{26} + x^{18} + 1 \qquad (2)$$

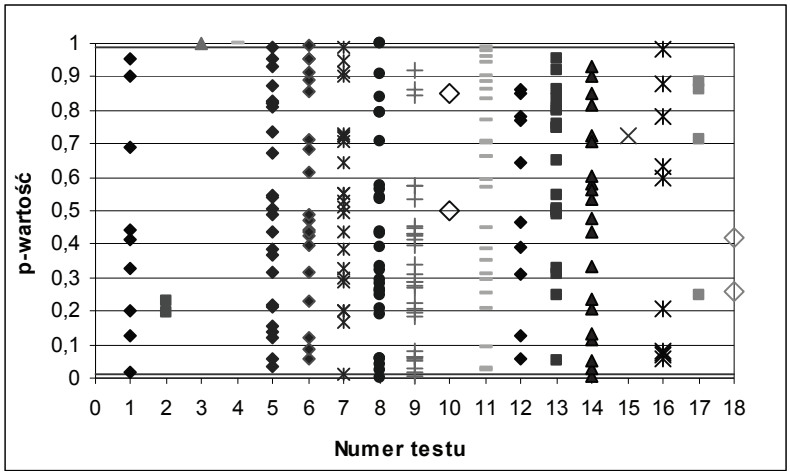
Sekwencja została wygenerowana poprzez napisany specjalnie do tego celu program w środowisku Delphi 7.0 i miała rozmiar 11MB. Ze względu na

wymagania pakietu DIEHARD sekwencja ta nie powinna być krótsza, aby nie wpływać wyniki testu. Wygenerowaną sekwencję poddano wszystkim testom zawartym w pakiecie. W tabeli 1 zebrano numery testów oraz ich nazwy, które zostały wykorzystane do opisu danych pokazanych na rysunku 2.

Tabela 1. Nazwy testów i ich numeracja na rysunku 2

Nr testu	Nazwa testu	Nr testu	Nazwa testu
1	Odstęp między datami urodzin	10	Test Policz "1" w ciągu bajtów
2	Test nakładających się 5 liczb	11	Test Policz "1" dla wybranych bajtów
3	Test rzędu macierzy 31x31	12	Test parkowania
4	Test rzędu macierzy 32x32	13	Test minimalnej odległości
5	Test rzędu macierzy 6x8	14	Test wypełnienia kulami
6	Test ciągu bitów	15	Test redukowania
7	Test OPSO	16	Test zachodzących na siebie sum
8	Test OQSO	17	Test ilości kierunków ciągów
9	Test DNA	18	Test gry w kości

Rozkład uzyskanych p-wartości wraz z zaznaczonym poziomem istotności $\alpha = 0{,}01$ przedstawiono na rysunku 2, 3 oraz 4. Dodatkowo na rysunkach 3 i 4 przedstawiono p-wartości, uzyskane dla wszystkich testów statystycznych, w zakresie 0-0,1 oraz 0,9-1, co umożliwia łatwe stwierdzenie czy dany test został spełniony czy nie.

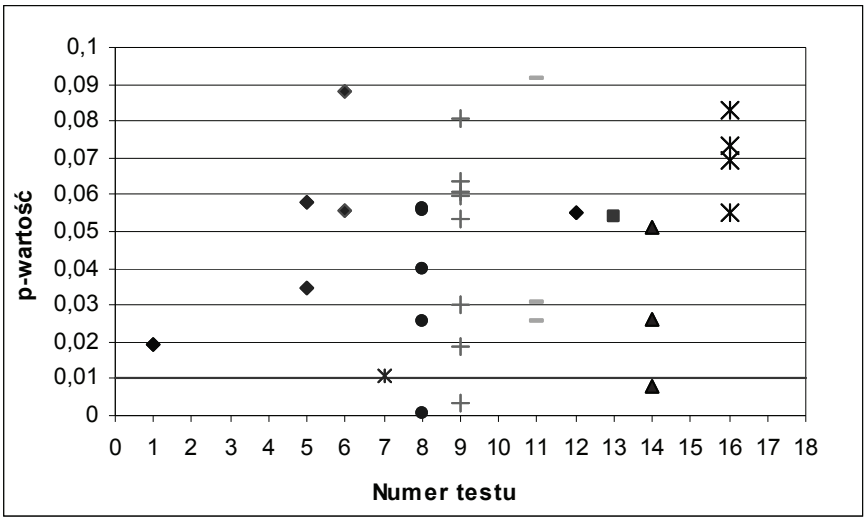


Rys. 2. Przykładowy generator LFSR

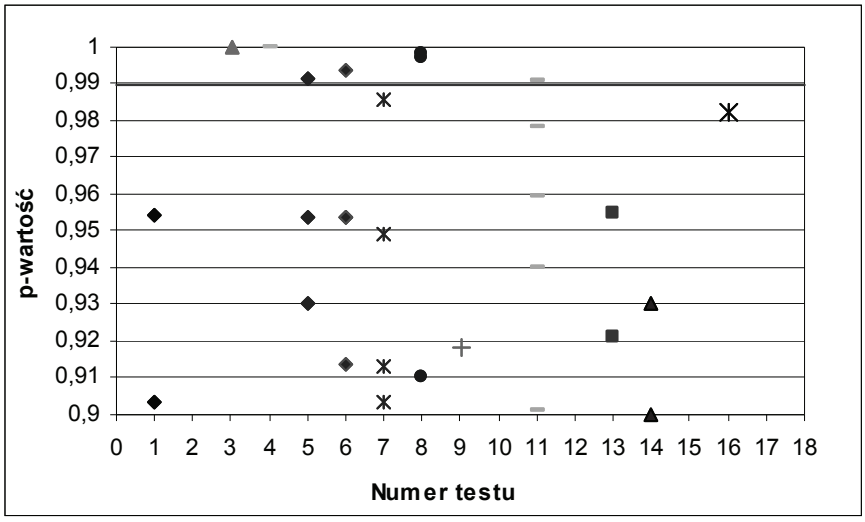
Badana sekwencja nie przechodzi następujących testów statystycznych:

- test rzędu macierzy binarnych o rozmiarze 31x31 (test numer 3),
- test rzędu macierzy binarnych o rozmiarze 6x8 (test numer 5).

Dla obu wymienionych testów obliczona przez pakiet DIEHARD p-wartość jest równa 1. Pozostałe p-wartości rozmieszczone są na zbiorze [0;1] i pozwalają ocenić, że generator spełnia dany test.



Rys. 3. Wykres p-wartości w pobliżu dolnej prostej określającej poziom istotności



Rys. 4. Wykres p-wartości w pobliżu górnej określającej poziom istotności

5. PODSUMOWANIE

W artykule opisano wyniki testów statystycznych generatora LFSR. Do badań wykorzystano testy statystyczne pakietu DIEHARD, który jest jednym z najpopularniejszych narzędzi do analizy sekwencji pseudolosowych. Do analizy wygenerowano testową sekwencję pseudolosową z wykorzystaniem programowej implementacji generatora LFSR. Uzyskane wyniki omówiono i zaprezentowano w formie wykresu.

LITERATURA

[1] Schneier B.: *Kryptografia dla praktyków*, Vol. 2, WNT, Warszawa 2002.
[2] Wesołowski K.: *Systemy radiokomunikacji ruchomej*, Wydawnictwo Komunikacji i Łączności, Warszawa 2006.
[3] Zwierko A.: *Testowanie generatorów pseudolosowych – wybrane programowe pakiety testów statystycznych*, VII Krajowa Konferencja Zastosowań Kryptografii, Warszawa, maj 2003, ss:1-20.
[4] Soto J.: *Statistical Testing of Random Number Generators*, National Institute of Standards & Technology, Proceedings of the 22nd National Information Systems Security Conference, 10/99, pp:1-12.
[5] Czernik P.: *Metodyka testowania bezpieczeństwa generatorów liczb pseudolosowych w systemach pomiarowo-sterujących*, Prace Instytutu Lotnictwa, Kwartalnik naukowy 6/2009 (201), ss: 20-34.
[6] Kotulski Z.: *Generatory liczb losowych: algorytmy, testowanie, zastosowania*, Matematyka Stosowana 2,2001, ss:1-9.
[7] Chen L, Gong G.: *Communication Systems Security*, Appendix A, 2008, pp:1-20
[8] Golomb S. W.: *Shift Register Sequences*, Laguna Hills, C A Aegean. Park Press, 1982.
[9] Walczak J., Stępień R.: *Modeling of the pseudo random signal generators using digital filters*. Proceedings of XXXIII conference IC-SPETO, 2010, pp: 85-86.
[10] Strona internetowa pakietu testów statystycznych DIEHARD: http://www.stat.fsu.edu/pub/diehard/

STATISTICAL PROPERTIES ANALISIS OF THE SHIFT REGISTERS PSEUDO RANDOM SIGNAL GENERATORS

The following article provides a description of a statistical tests results of the linear feedback shift register pseudo random signal generator (LFSR). It shows the structure of LFSR generators and it describes statistical tests which were used. Analysis of the generator output sequence was preformed in the DIEHARD statistical tests battery. The DIEHARD output data interpretation and the statistical tests of the sample pseudo random sequence were described in this article

Praca była współfinansowana ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego w ramach Projektu „SWIFT (Stypendia Wspomagające Innowacyjne Forum Technologii)” POKL.08.02.01-24-005/10.