POZNAN UNIVERSITY OF TECHNOLOGY ACADEMIC JOURNALS
No 78 Electrical Engineering 2014

Rafał STĘPIEŃ*

WYKORZYSTANIE PAKIETU TESTÓW STATYSTYCZNYCH NIST STS 2.1.1 DO TESTOWANIA SEKWENCJI GENERATORÓW DLFSR

W artykule opisano wyniki testów statystycznych sekwencji wyjściowych generatora pseudolosowego zrealizowanego na rejestrze przesuwnym i dynamicznym liniowym sprzężeniu zwrotnym (ang. Dynamic Linear Feedback Shift Register - DLFSR). Do analizy sekwencji wykorzystano pakiet testów statystycznych NIST STS 2.1.1. Ten pakiet testów statystycznych posłużył do przebadania dwóch sekwencji wyjściowych dwóch różnych generatorów DLFSR. Uzyskane wyniki przedstawiono w formie wykresu oraz omówiono.

SŁOWA KLUCZOWE: sygnały pseudolosowe, generatory DLFSR, testy statystyczne, STS-2.1.1

1. WSTĘP

Generatory ciągów pseudolosowych znajdują zastosowanie w wielu dziedzinach techniki [1]. Są niezbędnym elementem algorytmów statystycznych, znajdują zastosowania w kryptografii oraz w telekomunikacji [1]. Jedną z klas generatorów sygnałów pseudolosowych są generatory DLFSR [2]. Generator DLFSR ma zmienną w czasie strukturę pętli sprzężenia zwrotnego (w przeciwieństwie do generatorów NLFSR oraz LFSR, które mają statyczną w czasie pętlę sprzężenia zwrotnego). Ta cecha generatora DLFSR umożliwia znaczne wydłużenie okresu generowanej sekwencji pseudolosowej oraz poprawienie jej parametrów statystycznych. Dodatkowe szczegóły dotyczące budowy oraz opisu generatorów DLFSR można znaleźć w [2, 3].

2. PAKIET TESTÓW STATYSTYCZNYCH NIST STS 2.1.1

Sekwencje generowane przez generatory z rejestrami przesuwnymi nie są sekwencjami w pełni losowymi. Mają one skończoną długość, a przez co są okresowo powtarzalne. Także każdy kolejny bit wyjściowy sekwencji generatora zbudowanego na rejestrze przesuwnym jest uzyskiwany w pełni deterministyczny sposób, określony algorytmem generacji sekwencji pseudolosowej [1, 4, 5, 6].

* Politechnika Śląska.

Sekwencja pseudolosowa może przyjmować bardziej lub mniej dokładnie statystyczne cechy sygnału losowego. W celu pomiaru podobieństwa sekwencji pseudolosowej do sekwencji losowej stosuje się analizę statystyczną za pomocą różnych narzędzi, takich jak np. testy statystyczne lub testy indywidualne, które stwierdzają podobieństwo statystyczne sekwencji pseudolosowej do sekwencji losowej [7, 8, 9, 10, 11].

Jednym z pakietów testów statystycznych, służącym do analizy sekwencji pseudolosowych jest pakiet STS-2.1.1 autorstwa National Institute of Standard and Technology (NIST). Pakiet ten jest jednym z najważniejszych narzędzi do określania bezpieczeństwa informatycznego dotyczącego generacji ciągów losowych i pseudolosowych.

Do podstawowych zalet pakietu NIST zalicza się:

- dostępne kody źródłowe oraz gotowa aplikacja dla środowiska Windows,
- duża ilość testów statystycznych,
- dostępna dokładna dokumentacja oraz sposób interpretacji wyników.

Do podstawowych wad pakietu NIST zalicza się:

- utrudnioną analizę porównawczą dwóch zestawów wyników,
- sposób zapisu danych wyjściowy pakietu (wynik każdego testu do osobnego pliku),
- bardzo długi czas testowania (w zależności od długości sekwencji).

Pakiet NIST składa się z 15 testów statystycznych. Dostępna jest dokładna dokumentacja [11] dotycząca działania pakietu oraz opisu matematycznego zastosowanych testów statystycznych. Dostępne są również kody źródłowe pakietu w języku C oraz gotowa aplikacja przeznaczona dla użytkowników systemu Windows [12].

Pakiet NIST umożliwia testowanie sekwencji zapisanej w pliku dyskowym oraz jednego z dziewięciu zaimplementowanych generatorów. W tym artykule wykorzystano wyłącznie możliwość analizy sekwencji zapisanych w plikach dyskowych.

Autorzy pakietu NIST zalecają, aby sekwencja wejściowa miała długość mieszczącą się pomiędzy 10^3 a 10^7 bitów. Podczas przeprowadzania testów zaobserwowano, że w przypadku pliku wejściowego o rozmiarze 110 MB (32 bitowa sekwencja o długości 28672000 słów) pakiet NIST nie jest w stanie wykonać testu analizy widmowej DFT. W takich przypadkach nie zmniejszano pliku z sekwencją wyjściową tylko ustalano przy uruchomieniu pakietu z linii poleceń rozmiar sekwencji na 10^7 bitów. Pozostałe testy były wykonywane prawidłowo.

Zestaw testów statystycznych pakietu NIST oraz przypisane im numery testów, występujące na rysunkach 1 – 4, przedstawiono w tabeli 1.

Pakiet NIST wymaga znacznie dłuższego czasu przeznaczonego na analizę takiego samego pliku co np. pakiet DIEHARD [3]. Analiza pliku 110 MB na procesorze Intel Core i5 zajmuje około 5 minut.

Tabela 1. Testy pakietu NIST oraz odpowiadające im numery

Numer testu	Nazwa testu
1	Test entropii
2	Blokowy test częstości
3	Test skumulowanych sum
4	Test analizy widmowej DFT
5	Test częstości
6	Test złożoności liniowej
7	Test najdłuższej sekwencji
8	Test niezachodzących na siebie wzorców
9	Test zachodzących na siebie wzorców
10	Test losowych wycieczek
11	Alternatywny test losowych wycieczek
12	Test rzędu macierzy
13	Test sekwencji
14	Test częstości wielobitowych wzorców
15	Test kompresji

Pakiet testów NIST zwraca wyniki każdego z testów statystycznych jako tzw. *p*-wartości (nazywane również prawdopodobieństwem testowym) [7, 11]. Dany test statystyczny uznaje się za zdany, jeżeli *p*-wartości danego testu statystycznego są większe niż poziom istotności testu. W przypadku pakietu NIST w wersji STS-2.1.1 poziom istotności został ustalony na wartość domyślną $\alpha = 0{,}01$.

Wyniki wszystkich testów statystycznych zapisywane są do plików tekstowych w katalogu pakietu NIST. W celu ich dalszej analizy napisano oprogramowanie, które pobiera wszystkie *p*-wartości zapisane w plikach tekstowych i zwraca je w formie dogodnej do importu w arkuszu kalkulacyjnym.

3. WYNIKI BADAŃ GENERATORA DLFSR

Przebadano dwa 32-bitowe generatory DLFSR opisane wielomianami pierwotnymi oraz funkcjami przełączającymi zamieszczonymi w tabeli 2. Długość badanych sekwencji wynosiła 110 MB. Generatory były zaimplementowane programowo w języku Borland Delphi 7.0.

Literą *c* oznaczono numer cyklu zegarowego. Jeżeli spełniony jest jeden z warunków określony wartością funkcji przełączającej to funkcja sprzężenia zwrotnego generatora pracuje z zestawem odczepów określonym wielomianem ($L_2(x)$). W innym przypadku funkcja sprzężenia zwrotnego wykorzystuje odczepy opisane wielomianem ($L_1(x)$). Stałe występujące w obu funkcjach przełączających

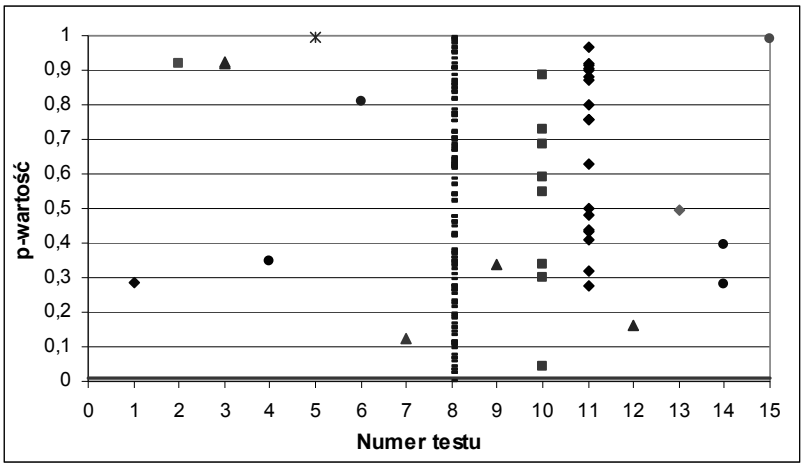
zostały dobrane doświadczalnie [2, 3]. Na rysunku 1 przedstawiono wyniki testów statystycznych sekwencji pseudolosowej generatora DLFSR1.

Tabela 2. Parametry generatorów DLFSR

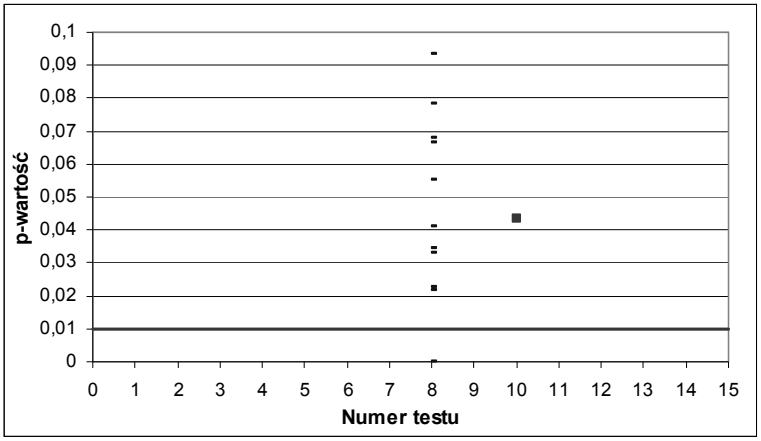
	Generator DLFSR1	**Generator DLFSR2**
Funkcja przełączająca	(c mod 15)=0 lub (c mod 63)=0	(c mod 15)=0 lub (c mod 64)=0
Wielomiany sprzężenia zwrotnego	$L_1(x) = x^{32} + x^{31} + x^{26} + x^{18} + 1$ $L_2(x) = x^{32} + x^{19} + x^{18} + x^{13} + 1$	

Sekwencja generatora DLFSR1 spełnia prawie wszystkie testy pakietu NIST, co potwierdza wysokie parametry statystyczne wygenerowanej sekwencji testowej. Jedna wartość testu numer 8 (test niezachodzących na siebie wzorców) leży poniżej linii określającej poziom istotności testu równy $\alpha = 0{,}01$, por. rysunek 2. Sekwencja nie zdaje tego testu dla jednego z wzorców generowanych przez pakiet testów statystycznych.

W celu dokładniejszej analizy należy przeprowadzić testy kilku innych sekwencji wygenerowanych przez ten sam generator. Każda kolejna sekwencja powinna być wygenerowana od innego warunku początkowego rejestru przesuwnego generatora. Można również wykorzystać inne testy statystyczne, np. DIEHARD lub ENT [3], [7] w celu weryfikacji parametrów statystycznych generatora.

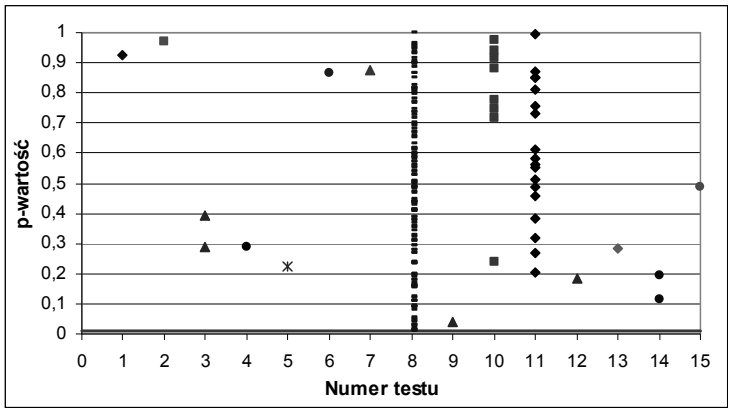


Rys. 1. Wykres *p*-wartości uzyskanych z analizy sekwencji generatora DLFSR1

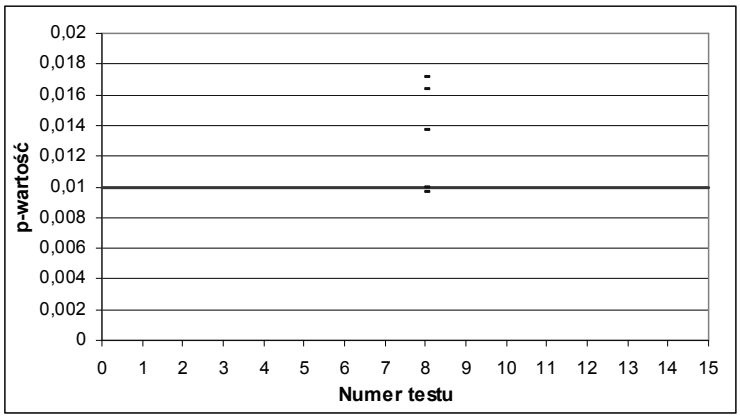


Rys. 2. Dolny fragment wykresu *p*-wartości uzyskanych z analizy sekwencji generatora DLFSR1

Na rysunku 3 przedstawiono wyniki testów statystycznych sekwencji generowanej przez generator DLFSR2. Sekwencja generatora DLFSR2 zdaje wszystkie testy statystyczne pakietu NIST. Tak jak w poprzednim przypadku wątpliwości może budzić wynik testu numer 8 (test niezachodzących na siebie wzorców). Jedna z wartości uzyskiwana w wyniku testowania leży w pobliżu prostej określającej poziom istotności $\alpha = 0{,}01$. Na rysunku 4 przedstawiono dolny fragment wyników pakietu NIST. Z rysunku 4 wynika, że najniższa wartość 8 testu leży poniżej linii poziomu istotności testu, jej dokładna wartość to 0,009623. Wynik tego testu należy uznać za lepszy niż w przypadku sekwencji generatora DLFSR1, a o akceptacji konstrukcji generatora DLFSR, ze zmierzonymi parametrami statystycznymi, powinien zadecydować konstruktor.



Rys. 3. Wykres *p*-wartości uzyskanych z analizy sekwencji generatora DLFSR2



Rys. 4. Dolny fragment wykresu *p*-wartości uzyskanych z analizy sekwencji generatora DLFSR2

Wyniki analizy sekwencji pseudolosowych generatorów DLFSR wskazują, że spełniają one prawie wszystkie testy statystyczne pakietu NIST STS 2.1.1. W celu pełnej analizy oraz dopuszczenia generatora DLFSR do zastosowań kryptograficznych konieczna jest dokładniejsza analiza statystyczna w innych testach statystycznych np. DIEHARD [7] oraz porównanie uzyskanych wyników z innymi typami generatorów ciągów pseudolosowych np. NLFSR lub LFSR [3].

4. PODSUMOWANIE

W artykule opisano wykorzystanie pakietu testów statystycznych STS 2.1.1 do testowania sekwencji pseudolosowych generowanych przez generatory z dynamicznym liniowym sprzężeniem zwrotnym – DLFSR. Krótko omówiono budowę generatorów DLFSR oraz przeprowadzono dokładniejsze omówienie wykorzystanego w artykule pakietu testów statystycznych. Przebadano sekwencje dwóch 32-bitowych programowych generatorów DLFSR. Uzyskane wyniki omówiono i przedstawiono w formie wykresów. Z uzyskanych wyników wyciągnięto wnioski.

LITERATURA

[1] Schneier B.: Kryptografia dla praktyków, Vol. 2, WNT, Warszawa 2002.
[2] Stępień R., Walczak J.: Application of the DLFSR generators in spread spectrum communication, 19th International Conference "MIXDES Design of Integrated Circuits and Systems" ,MIXDES-2012, Warszawa, maj 2012, pp:555-558.
[3] Stępień R., Walczak J.: Comparative Analysis of Pseudo Random Signals of the LFSR and DLFSR Generators, proceedings of 20th International Conference "MIXED Design of Integrated Circuits and Systems", MIXDES-2013, Gdynia, czerwiec 2013, pp: 598-602.

[4] Patidar V, Sud K,K.: A Novel Pseudo Random Bit Generator Based on Chaotic Standard Map and its Testing, Electronic Journal of Theoretical Physics, No.20, 2009, pp: 327–344.
[5] Park S.K., Miller K.W. : Random number generators: good ones are hard to find, Communications of the ACM, Volume 31, Issue 10, October 1988, pp:1192 – 1201.
[6] Haag M.: Introduction to random signal and processes, Connections Project, 2005.
[7] Zwierko A.: Testowanie generatorów pseudolosowych – wybrane programowe pakiety testów statystycznych, VII Krajowa Konferencja Zastosowań Kryptografii, Warszawa, maj 2003, ss:1-20.
[8] 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.
[9] Rashidah K., Maarof M.A.: Randomness Analysis of Pseudorandom Bit Sequences, International Conference on Computer Engineering and Applications, IPCSIT vol.2 IACSIT Press, Singapore, 2011, pp:390-394.
[10] Kotulski Z.: Generatory liczb losowych: algorytmy, testowanie, zastosowania, Matematyka Stosowana 2, 2001, ss:1-9.
[11] Rukhin A i inni, A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, National Institute of Standards and Technology, rev 1a, april 2010.
[12] Strona internetowa pakietu STS 2.1.1 http://csrc.nist.gov/groups/ST/toolkit/rng/documentation_software.html

APPLICATION OF THE STATISTICAL TEST SUITE STS 2.1.1 TO TESTING OF THE DLFSR SEQUENCES

The following article provides a description of a statistical tests results of the dynamic linear feedback shift register generator - DLFSR. The generators' sequences were analyzed in the NIST statistical test suite STS 2.1.1. This test suite was used to analyze two pseudo random sequences generated by the two different 32 bits DLFSR generators. Obtained results were discussed and shown in a form of diagrams.