

prof. dr. hab. inż. Józef Pieprzyk

Instytut Podstaw Informatyki

Polska Akademia Nauk

ul Jana Kazimierza 5

01-248 Warszawa

&

Senior Principal Research Scientist

Data61, CSIRO

Sydney, NSW, Australia

email: josef.pieprzyk@csiro.au

Recenzja Rozprawy Doktorskiej

16 czerwca 2025

Tytuł: **Pakiet testów losowości do badania prymitywów kryptograficznych**

Autor: *mgr inż. Krzysztof Mańk*

Wprowadzenie

Rozprawa doktorska podejmuje istotny i aktualny problem oceny jakości losowości w kontekście zastosowań kryptograficznych. Już we wstępie autor podkreśla kluczową rolę losowości jako podstawy bezpieczeństwa współczesnych systemów kryptograficznych — od generatorów liczb losowych, poprzez protokoły szyfrowania, aż po podpisy cyfrowe i funkcje skrótu. Błędna ocena jakości losowości może prowadzić do poważnych luk bezpieczeństwa, dlatego rozwój efektywnych, niezależnych testów statystycznych ma fundamentalne znaczenie praktyczne i teoretyczne.

Rozdział I stanowi wprowadzenie do zagadnień związanych z losowością. Autor przedstawia formalne definicje losowości, pseudolosowości i kryptograficznej pseudolosowości, a także omawia klasy generatorów liczb losowych: deterministycznych, kryptograficznych oraz generatorów z funkcjami skrótu. Wyjaśniono też relację między właściwościami statystycznymi a bezpieczeństwem.

Rozdział II zawiera szczegółowy przegląd klasycznych testów losowości – od prostych testów pojedynczych bitów, przez testy równomierności rozkładów, analizę permutacji, analizę entropii, po testy oparte na transformacjach Fouriera i testy odległości. Przedstawiono kryteria poprawności testów oraz omówiono wpływ parametrów testowania.

W Rozdziale III autor analizuje istniejące pakiety testowe, takie jak DIEHARD, NIST STS, Dieharder, ENT i TestU01. Oceniono ich strukturę, zakres testów i ograniczenia, wskazując m.in. na nadmiarowość, niską skalowalność, ograniczenia względem długości wejściowych danych czy brak pełnej automatyzacji analizy wyników.

Rozdział IV przedstawia główny wkład autora – propozycję nowego, trójpoziomowego pakietu testów. Wyróżnia on trzy etapy analizy: testowanie pojedynczych próbek, analiza rozkładu wyników dla wielu próbek oraz testy metatestów. Pakiet ten pozwala na automatyczne przetwarzanie wyników i weryfikację statystyczną ich istotności.

W Rozdziale V przeprowadzono analizę skuteczności testów względem długości danych i liczby próbek. Dzięki eksperymentom zaproponowano eliminację testów o niskiej sile detekcji błędów losowości, co pozwala znacząco zoptymalizować czas testowania bez utraty jakości.

Rozdział VI zawiera analizę korelacji pomiędzy testami – liniową, nieliniową oraz wielowymiarową. Wskazano

grupy testów silnie skorelowanych, których wspólne stosowanie nie wnosi istotnie nowej informacji. Autor proponuje zredukowany zestaw testów minimalnie skorelowanych, które zapewniają szerokie pokrycie właściwości statystycznych.

Całość uzupełniają dwa załączniki zawierające gotowe, zoptymalizowane zestawy testów dla danych o długości 1 kiB i 1 MiB, gotowe do zastosowania w praktyce. Praca stanowi istotny wkład w rozwój narzędzi do testowania losowości w zastosowaniach kryptograficznych i może służyć jako podstawa do dalszych prac nad automatyzacją oceny bezpieczeństwa prymitywów kryptograficznych.

Oryginalny Wkład

Rozprawa doktorska zawiera wiele interesujących i oryginalnych przyczynków w dziedzinie oceny losowości. Autor, opierając się na własnych badaniach teoretycznych i eksperymentalnych, zaproponował szereg nowych testów, metod analizy oraz usprawnień istniejących rozwiązań. Część z tych wyników została już opublikowana, inne są dopiero przedmiotem procedur recenzyjnych lub mają potencjał publikacyjny. Poniżej przedstawiono podzbiór wkładów autorskich, które – w mojej ocenie – zasługują na szczególne wyróżnienie. Lista bazuje na zestawieniu zawartym przez autora w Podsumowaniu (strony 193–194).

- TRÓJPOZIOMOWA PROCEDURA TESTOWA I ANALIZA KORELACJI TESTÓW – Autor zaproponował nowatorską metodologię trójpoziomowego testowania generatorów losowych, składającą się z: testowania pojedynczych ciągów (poziom pierwszy), agregacji wyników przy użyciu testów drugiego poziomu, oraz oceny stabilności i spójności zestawów testowych (poziom trzeci). Dodatkowo przeprowadzono analizę korelacji między testami, umożliwiającą ich grupowanie i eliminację nadmiarowych procedur. Ta część pracy zasługuje na publikację w renomowanym czasopiśmie międzynarodowym.
- PRZYSPIESZENIE REALIZACJI TESTÓW NAJMNIEJSZEJ ODLEGŁOŚCI, M NAJBLIŻSZYCH PAR I BICKELA-BREIMANA – W pracy opracowano zoptymalizowane wersje trzech klasycznych testów geometrycznych, znacząco skracając ich czas działania poprzez eliminację nadmiarowych operacji oraz zastosowanie efektywnych struktur danych do wyszukiwania najbliższych sąsiadów. Ta część została opublikowana jako K. Mańk, *Speeding Up Minimum Distance Randomness Tests*, *Journal of Telecommunications and Information Technology*, 2022.
- DOKŁADNE WARTOŚCI PARAMETRÓW DLA TESTU MAURERA – Oryginalny wkład polegał na empirycznym wyznaczeniu dokładnych wartości oczekiwanych i wariancji statystyki testowej Maurera, w miejsce wcześniejszych przybliżeń teoretycznych. Przyczynek ten został opublikowany jako K. Mańk, *Wyznaczenie dokładnych wartości parametrów dla testu Maurera*, *Biuletyn WAT*, Warszawa, 2008.
- METODA WYZNACZANIA ROZKŁADÓW W TESTACH DRZEWA BINARNEGO – Autor opracował i przetestował nową metodę wyznaczania rozkładów statystyk dla testów opartych na strukturze drzewa binarnego, które wcześniej nie były formalnie analizowane. Przyczynek ten zaowocował publikacją K. Mańk, *Upgrading Frequency Test for Overlapping Vectors and Fill Tree Tests*, in revision, *IJET*, 2024.
- TEST CZĘSTOŚCI DLA NAKŁADAJĄCYCH SIĘ WEKTORÓW – Rozwinięto klasyczny test częstości o przypadek nakładających się fragmentów danych, wskazując, że nieuwzględnienie zależności pomiędzy fragmentami prowadzi do błędnych ocen. Opublikowano to jako K. Mańk, *Efficient Evaluation of Frequency Test for Overlapping Vectors Statistic*, *Cyberprzestępczość i ochrona informacji – Bezpieczeństwo w internecie*, tom II, 2013.
- PARAMETRY ROZKŁADU DLA TESTU POCHODNEJ CIĄGU LICZB – Wyznaczono empiryczny rozkład statystyki testowej dla testu pochodnej ciągów liczbowych, co pozwoliło na dokładniejsze wyznaczanie progów istotności. Szczegóły przedstawiono w opracowaniu wewnętrznym K. Mańk, *Wyznaczenie rozkładu średniej wartości odstępu w teście pochodnej ciągu liczb*, projekt ROTOR, 2017.
- TESTY DLA 2 I 3 LICZB – NOWE PROPOZYCJE – Autor zaproponował nowe testy oparte na statystykach zbudowanych odpowiednio z dwóch i trzech kolejnych elementów ciągu. Oprócz ich formalnych definicji, wyprowadzono również rozkłady teoretyczne dla idealnie losowych danych wejściowych. Testy te

stanowią potencjalną bazę do rozbudowy pakietów testowych.

Ogólne Uwagi i Komentarze

Rozdziały są merytorycznie poprawne, jednak miejscami zbyt obszerne i wewnętrznie niejednorodne (np. rozdziały II i III zawierają zarówno treści teoretyczne, jak i oceny praktyczne, co zaciera granicę między przeglądem a krytyką). Prezentację pracy można znacząco poprawić. Poniższe zalecenia proszę traktować jako sugestie — ich wdrożenie nie jest obowiązkowe, zwłaszcza jeśli oznaczałoby gruntowne przepisanie dużych fragmentów pracy. Oto kilka propozycji i komentarzy:

- Praca zawiera opis dużej liczby testów, które zostały uporządkowane zgodnie z typem danych wejściowych: bloków bitów, ciągów nienakładających się wektorów, ciągów nakładających się wektorów oraz ciągów bitów. Oczywiście nie jest to jedyny możliwy porządek. Dobrym rozwiązaniem byłoby wprowadzenie na początku Rozdziału II jasnej taksonomii testów — ze względu na typ danych wejściowych, rodzaj badanej własności statystycznej, zastosowaną metodę analizy itp. Taksonomia ta powinna zostać krótko omówiona i konsekwentnie wykorzystana w dalszej części pracy.
- Aby zwiększyć przejrzystość prezentacji, warto uporządkować testy w rozdziałach zgodnie z przyjętą hierarchią klasyfikacyjną (np. typ danych → badana własność → metoda analizy). Zaleca się również wprowadzenie tabeli podsumowującej, przypisującej każdy test do konkretnej kategorii według zdefiniowanej taksonomii. W tabeli można dodatkowo zaznaczyć, które testy są oryginalne, autorskie lub zmodyfikowane przez autora. Takie tabele mogą podsumowywać wybrane grupy testów lub stanowić ogólne zestawienie umieszczone na końcu pracy bądź w załączniku.
- Wszystkie testy przedstawione w pracy powinny być opisywane według spójnego schematu. Przykładowy szablon opisu może obejmować następujące elementy: cel testu, typ danych wejściowych (to pole można pominąć, jeśli testy zostały wcześniej odpowiednio sklasyfikowane), opis procedury testowej, statystyka testowa, metoda wyznaczania wartości statystycznej, kryterium decyzji, złożoność obliczeniowa, zastosowania testu, uwagi autora i modyfikacje. Oczywiście jest to jedynie sugestia, ale ujednolicona prezentacja znacznie poprawi czytelność pracy.
- Rozdział III powinien nosić tytuł *"Przegląd i Krytyka Wybranych Pakietów Testowych"* i zawierać wyważoną dyskusję na temat zalet i wad poszczególnych pakietów. Zasadne wydaje się zamieszczenie tabeli porównawczej prezentującej m.in. liczbę testów zawartych w danym pakiecie, typ obsługiwanych danych wejściowych, zakres długości przetwarzanych ciągów, możliwość wykorzystania metatestów, łatwość automatyzacji itp. W tabeli można również wskazać mocne i słabe strony poszczególnych pakietów w kontekście zamierzonych zastosowań.
- Autor stosuje podwójne odnośniki do literatury: raz w formie przypisu w stopce, gdzie dana pozycja literatury jest wyszczególniona, oraz drugi raz w bibliografii końcowej. Takie podejście nie jest zgodne z powszechnie przyjętymi standardami akademickimi. Trzeba jednak przyznać, że rozwiązanie to ma pewne zalety — ułatwia czytelnikowi szybki dostęp do źródła bez konieczności przeszukiwania całej bibliografii.

Podsumowanie

Rozprawa doktorska mgr inż. Krzysztofa Mańka stanowi starannie opracowaną i merytorycznie dojrzałą pracę, która świadczy o bardzo dobrej znajomości zarówno teoretycznych podstaw testów statystycznych stosowanych w ocenie losowości, jak i praktycznych wymagań, jakie stawiane są współczesnym pakietom testowym. Autor wykazał się dużym zaangażowaniem i systematycznością – imponuje zwłaszcza zakres implementacji testów oraz wnikliwa ich weryfikacja i ocena, często wykraczająca poza standardowe podejścia. Należy również z uznaniem odnotować fakt, że część wyników została już opublikowana w recenzowanych czasopismach krajowych i zagranicznych. Zachęcam doktoranta do dalszej aktywności publikacyjnej, w szczególności do przedstawienia wyników analizy pakietów testowych w renomowanych czasopismach międzynarodowych z zakresu kryptografii i inżynierii oprogramowania.

Biorąc pod uwagę jakość naukową pracy, oryginalność wyników oraz wkład własny doktoranta, wnoszę do Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Wojskowej Akademii Technicznej im. Jarosława Dąbrowskiego w Warszawie o przyjęcie rozprawy doktorskiej mgr inż. Krzysztofa Mańka oraz dopuszczenie go do dalszych etapów przewodu doktorskiego.



Józef Pieprzyk
IPI PAN oraz Data 61 CSIRO

Dokument podpisany cyfrowo przez *Certum Asseco Data Systems*

Lista Błędów Językowych i Stylistycznych

Poniżej przedstawiono listę błędów językowych, stylistycznych oraz literówek znalezionych w pracy. Każdy wpis zawiera numer strony, typ błędu, fragment zawierający problem oraz sugestię poprawy.

Strona	Linia	Rodzaj błędu	Fragment	Sugestia poprawy
4	14	Oznaczenie	$a..b$	$[a, b]$ albo $\{a, \dots, b\}$
4	-4	Oznaczenie	$\Gamma(z)$	Brakujący znak różniczki dt w całce
18	-3	Składniowy	tak jak to było w definicji przekształcają	zgodnie z definicją przekształcają
19	3	Składniowy	ponieważ niesie informację o algorytmie jako takim	niepotrzebny pleonazm „jako takim”
23	-15	Ortograficzny	Shamira, czy RSA, oceniamy jako skrajnie niepraktyczne...	Usuń przecinek po <i>Shamira</i>
23	-13	Ortograficzny	togo problemu	Popraw na: <i>tego problemu</i>
23	-10	Składniowy	którym naszym zdaniem nie poświęcono...	Dodaj przecinki: <i>którym, naszym zdaniem, nie...</i>
27	16	Brak odnośnika	D_n^+ i D_n^-	podaj ich definicje
47	28	Styl	udało się nam wyznaczyć wiele parametrów...	Lepszy styl: <i>wyznaczono wiele parametrów...</i>
57	-1	Brak definicji	parametr p	definicja p ?
93	7	Gramatyczny	Wartości oczekiwana całości jest również bardzo bliska 0.	Popraw na: <i>Wartość oczekiwana całości...</i>
93	9	Składniowy	przewagę będzie miał w nim rozkład...	Lepszy szyk: <i>Rozkład [...] będzie miał przewagę</i>
116	6	Składniowy	zaś suma elementów ...	suma elementów ...
119	19	Składniowy	Zawiera ona ten sam błąd jak dla poprzedniego testu.	Popraw na: <i>Zawiera ten sam błąd, co poprzedni test.</i>
127	10	Styl	Definicja idealnej pary, zaproponowana (...) posiada pewną słabość	Zastąp: <i>posiada</i> $\rightarrow$ <i>ma</i>