VIII Semana de Ciência e Tecnologia IFMG - *Campus* Bambuí
VIII Jornada Científica

Testes estatísticos para avaliação da qualidade de geradores de números aleatórios: uma revisão de literatura

Mayron Reis Lacerda Ribeiro[1]; Samuel Pereira Dias[2]

[1] Estudante de Engenharia de Computação. Instituto Federal Minas Gerais (IFMG) *Campus* Bambuí. Rod. Bambuí/Medeiros km 5. CEP: 38900-000. Bambuí-MG. Bolsista de Iniciação Científica (PIBIC) – *Campus* Bambuí – IFMG. [2] Professor Orientador – IFMG.

RESUMO - Os números aleatórios podem ser estatisticamente definidos como independentemente diferentes entre si, pertencem a uma série numérica e não podem ser previstos por seus antecessores. Em geral, números aleatórios são usados em jogos de azar, programas de computador, sistema de segurança entre outras aplicações. Números pseudoaleatórios são aqueles que simulam o comportamento dos números aleatórios, os mesmos são gerados por métodos determinísticos e são normalmente mais utilizados por serem obtidos com maior rapidez. Os números aleatórios verdadeiros são gerados com maior qualidade criptográfica e são obtidos através de métodos não determinísticos e baseados em eventos naturais, como lançamento de moedas ou dados. Esses são aplicados em diversas áreas, principalmente aquelas que necessitam de um nível de aleatoriedade elevado para prover segurança na aplicação e não são simples de serem obtidos para fins computacionais. Geradores de números pseudoaleatórios de qualidade são bastante requisitados para fins criptográficos em que a aleatoriedade de qualidade é de suma importância. Para comprovar esta aleatoriedade dos números gerados é necessário a aplicação de uma sequência de testes estatísticos. Os testes visam provar que os números gerados possuem grau de aleatoriedade com qualidade criptográfica. O NIST, *National Institute of Standard and Technology,* recomenda uma bateria de 15 testes estatísticos. De acordo com os resultados obtidos na revisão de literatura pode-se afirmar que os testes estatísticos são fundamentais para se comprovar a aleatoriedade de qualidade desejada.

Palavras-chave: geração de números aleatórios, números pseudoaleatórios e aleatórios verdadeiros, testes estatísticos.

INTRODUÇÃO

Números aleatórios são utilizados em diversos tipos de aplicações e assumem grande importância em aplicações criptográficas. Podem ser divididos em sequências aleatórias verdadeiras ou pseudoaleatórias, conforme apresentado no Referencial Teórico deste texto. Para desenvolvimento de aplicações, utilizam-se números pseudoaleatórios, que devem apresentar uma

aleatoriedade de qualidade para fins criptográficos, a qual pode ser atestada por meio de testes estatísticos.

O objetivo deste trabalho é realizar uma breve revisão de literatura, visando elencar testes estatísticos recomendados para a avaliação da qualidade de sequências aleatórias verdadeiras ou pseudoaleatórias, para uso em aplicações criptográficas. Insere-se em um projeto de pesquisa que visa à produção de números aleatórios na plataforma Arduino para fins de criptografia, cujo estado de desenvolvimento apresenta como resultado parcial o levantamento dos testes recomendados pela literatura e pelo *National Institute of Standard and Technology* (NIST).

No Referencial Teórico são abordadas as conceituações sobre os tipos de números aleatórios e suas aplicações e, nas Conclusões, apresenta-se o que foi identificado após a revisão de literatura.

REFERENCIAL TEÓRICO

Números aleatórios são componentes de uma serie aleatória em que cada número aleatório é estatisticamente independente de outro número da sequência e, por conseguinte, imprevisível (LAWRIE; WILLIAM, 2014).

Números aleatórios verdadeiros podem ser definidos como aqueles que apresentam aleatoriedade de qualidade, por evidenciar uma sequência maior de números, eles são ideais para aplicações criptográficas e podem ser obtidos através de métodos não determinísticos. Segundo Lee (2013), a não utilização desses pode acarretar em falhas de segurança, deixando um sistema propenso a ataques criptográficos.

A melhor maneira de se obter números aleatórios verdadeiros é realizando medidas de fenômenos físicos tais como decaimento radioativo, ruído térmico em semicondutores, amostra de sons, dentre outros (GUTMANN, 1998).

Números pseudoaleatórios têm como característica a simulação dos números aleatórios. Uma sequência pseudoaleatória de números é geralmente preferível no lugar de números totalmente aleatórios em se tratando de aplicações de simulação, pois frequentemente, na depuração e teste de modelos, é desejável repetir-se o mesmo experimento de simulação, quando resultados diferentes forem requeridos (PORTNOI, 2005).

Números aleatórios e sequências aleatórias são amplamente utilizados em muitas áreas tais como, a teoria dos jogos, análise numérica e mecânica quântica, esses constituem também uma parte necessária da criptografia (SEKER, 2014). De acordo com Robson (2013), verdadeiros geradores aleatórios são uma parte crucial da vida cotidiana moderna, os sistemas que permitem fazer qualquer tipo de atividade *online* como enviar simples *e-mails* ou fazem aplicações mais complexas como consultas bancárias necessitam essencialmente de números aleatórios.

Os testes estatísticos são uma forma de comprovar matematicamente que os números gerados a partir de métodos determinísticos são realmente aleatórios (BOCHARD *et al.*, 2010). Desde 1997, o *National Institute of Standard and Technology* (NIST) recomenda uma bateria de 15 testes estatísticos que deve ser utilizada por desenvolvedores de sistemas criptográficos para verificar, em primeira instância, se o sistema em desenvolvimento se porta como um gerador de números aleatórios. Esses analisam sequências binárias arbitrariamente geradas a partir de um *hardware* ou *software* (LEE, 2013).

Os testes estatísticos podem ser descritos resumidamente no Quadro 1, de acordo com Bassham *et al.* (2010), Lee (2013) e Robson (2013). A concepção, a parametrização e a aplicação dos testes não faz parte do escopo do presente trabalho, recomendando-se a leitura destes autores para maiores informações.

Quadro 1. Testes estatísticos recomendados pelo NIST.

TESTE	DESCRIÇÃO
Frequência (*Monobit*)	Analisa a proporção de zeros e uns em toda a amostra sequencial de *bits*. Verifica se a ocorrência desses valores acontece como esperado em uma sequência aleatória de fato.
Frequência dentro de Blocos	Divide a amostra em blocos com quantidade "M" de *bits*, então analisa-se a proporção de dígitos "um" em cada bloco "M". O propósito deste teste é determinar a se a frequência de "uns" em cada bloco aproxima-se de "M".
Corridas	Este teste tem como o objetivo determinar se o número de corridas de uns e zeros de vários comprimentos é o desejado para uma sequencia aleatória. De uma forma geral este teste determina se a oscilação de *substrings* é rápida ou lenta.
Mais longa corrida de 1's em um bloco	Analisa as mais longas corridas de uns (1's) dentro de blocos de N *bits*, para isso a amostra deve ser dividida em M blocos, como o teste de frequência dentro de blocos. Neste teste os blocos são divididos de acordo com o tamanho da amostra de *bits*.
Posto para matrizes binárias	Analisa o posto de submatrizes disjuntas de uma sequência de *bits* a fim de verificar se há dependência linear entre *substrings* de comprimentos pré-fixados, *substrings* estas originadas da sequência original.
Transformação discreta de Fourier (Espectral)	O objetivo deste teste é determinar se a frequência espectral da sequência binária coincide com o que seria esperado para uma sequência verdadeiramente aleatória.

Quadro 1. Testes estatísticos recomendados pelo NIST. (cont.)

TESTE	DESCRIÇÃO
Não sobreposição de padrão	Buscar número de ocorrências de *strings* alvo pré-especificadas para detectar geradores que produzem muitas ocorrências de um dado padrão não periódico.
Sobreposição de padrão	Trabalha com as ocorrências de *strings* alvo pré-especificadas e com uma janela de m-*bits* para buscar por um padrão específico de m-*bits*, assim como o teste de não sobreposição de padrão. Se o padrão não for encontrado, a janela é transladada de um *bit* de sua posição.
Estatística universal de Maurer	Encontra o número de *bits* entre padrões, detectando se a sequência pode ou não ser significativamente comprimida sem perda de informação. Uma sequência significativamente é considerada ser não aleatória.
Complexidade linear	O objetivo deste teste e determinar se a sequência é ou não complexa o suficiente para ser considerada verdadeiramente aleatória.
Serial	Analisa a frequência de todas as possíveis sobreposições de padrões de m-*bits* na sequência toda, determinando se o número de ocorrências dos 2m padrões de sobreposição de m-*bits* é aproximadamente o mesmo do esperado para uma sequência aleatória.
Entropia aproximada	Tem foco em encontrar a frequência de todas as possíveis sobreposições de padrões de m-*bits* na sequência toda, comparando a frequência de sobreposições de dois blocos consecutivos de comprimentos m e m + 1 contra o resultado esperado para uma sequência aleatória.
Somas cumulativas	O objetivo deste teste e determinar se o máximo de somas acumuladas em uma sequência muito grande ou muito pequena; indicativo de muitos uns ou zeros no início das primeiras (últimas) etapas.
Excursões aleatórias	O teste tem como função verificar o número de ciclos dentro de uma sequência e determinar se o número de visitas a um estado demarcado excede o esperado para uma sequência verdadeiramente aleatória.
Variante de excursões aleatórias	Procura encontrar o número total de vezes que um estado particular é visitado em uma soma cumulativa de um caminho aleatório, cujo propósito é detectar desvios do número esperado de visitas para vários estados em um caminho aleatório.

A suíte de implementação dos testes pode ser encontrada em NIST (2010), assim como a documentação completa acerca destes.

CONCLUSÕES

Pode-se constatar que os números aleatórios verdadeiros são essenciais para se obter segurança através da qualidade criptográfica. Como estes são de difícil obtenção, utilizam-se sequências pseudoaleatórias, produzidas por métodos determinísticos. A geração destas sequências com qualidade criptográfica é essencial em sistemas computacionais para os mais diversos fins.

Testes estatísticos são imprescindíveis, tendo um papel fundamental para apuração da qualidade da aleatoriedade. São eles os responsáveis por validar essa característica, verificando se esses aproximam-se do comportamento de uma sequência aleatória verdadeira e o consequente uso em aplicações de segurança.

AGRADECIMENTOS

Ao IFMG *Campus* Bambuí por fomentar a pesquisa. Ao GPSisCom pela infraestrutura e contribuições de seus membros.

REFERÊNCIAS BIBLIOGRÁFICAS

BASSHAM, L. E., III.; RUKHIN, A.; SOTO, J.; NECHVATAL, J.; SMID, M.; BARKER, B.; LEIGH, S.; LEVENSON, M.; VANGEL, M.; BANKS, D.; HECKERT, A.; DRAY, J.; VO, S. **SP 800-22 Rev. 1a. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications**. Gaithersburg, MD, United States: National Institute of Standards & Technology, 2010.

BOCHARD, N.; BERNAD, F.; FISCHER, V.; VALTCHANOV, B. True-Randomness and Pseudo-Randomness in Ring Oscillator-Based True Random Number Generators. **International Journal of Reconfigurable Computing**, 13p. 2010.

GUTMANN, P. Software generation of random numbers for cryptographic purposes. In: USENIX SECURITY SYMPOSIUM, 7. 1998. San Antonio, Texas. **Anais.** San Antonio, Texas: University of Auckland, 1998.

LAWRIE, B.; WILLIAM, S. **Segurança de Computadores, Trad. 2ª Ed.** [S.l.]: Elsevier Brasil, 2014.

LEE, M. **Investigating Modern Cryptography**. 2013. 445p. Tese (Doutorado) - University of Windsor, Windsor, Ontario, Canadá.

NIST. Download Documentation and Software. 2010. Disponível em: <http://csrc.nist.gov/groups/ST/toolkit/rng/documentation_software.html>. Acesso em: 20 de out. de 2015.

PORTINOI, M. **Probabilidade, variáveis aleatórias, distribuição de probabilidades e geração aleatória.** Salvador: Universidade Salvador – UNIFACS, 2005. 30p.

SEKER, O. **A randomness test based on postulate r-2 on the number of runs**. 2014. 60p. Tese (Mestrado) - The Graduate School of Applied Mathematics of Middle East Technical University.

ROBSON, S. **A Ring Oscillator Based Truly Random Number Generator**. 2013. 97p. Dissertação (Mestrado) - University of Waterloo, Waterloo, Ontario, Canadá.