Dicle University
Journal of Engineering
https://dergipark.org.tr/tr/pub/dumf
duje.dicle.edu.tr

Araştırma Makalesi / Research Article

Kriptografide Rasgelelik Kavramı ve Gerçek Rasgele Sayı Üreteçlerinin Test Metodolojisi

Concept of Randomness in Cryptography and Testing Methodology of True Random Number Generators

Ali Murat GARİPCAN[1*], Ebubekir ERDEM[2]

[1] Fırat Üniversitesi, Bilgisayar Mühendisliği Bölümü, agaripcan6223@gmail.com, ORCID: https://orcid.org/0000-0002-9659-8785
[2] Fırat Üniversitesi, Bilgisayar Mühendisliği Bölümü, aberdem@firat.edu.tr ORCID: https://orcid.org/0000-0001-7401-4964

MAKALE BİLGİLERİ

Makale Geçmişi:

Geliş 01.11.2023
Revizyon 06.01.2024
Kabul 22.02.2024
Online 29.03.2024

Anahtar Kelimeler:

Rasgelelik, rasgele sayılar, kriptografi, rasgele sayı üreteçleri, istatistiki rasgelelik testleri.

ÖZ

Kriptografik protokollerde gizlilik, karmaşıklık ve sürdürülebilir bir güvenlik anlayışının rasgele sayılar üzerinden tesis edildiği düşünüldüğünde, kriptografi ile rasgelelik arasında kuvvetli bir bağın oluştuğu görülebilir. Bu sayıların elde edilmesine kaynaklık eden rasgeleliğin niteliği ve matematiksel yöntemlerle garanti edilebilen nicel özellikleri, kriptografik sistemlerinin performansı üzerinde belirleyici bir öneme sahiptir. Dolayısıyla rasgele sayıların elde edildiği ve Rasgele Sayı Üreteci (RSÜ) olarak da özelleştirilmiş uygun tasarım bileşenlerinin seçimi ve değerlendirmesi kriptografik güvenlik açısından önemli ve zorlu bir görevdir. Zira RSÜ'nin rasgele sayılar üzerinde yol açacağı güvenlik kusurları kriptografik sistemi bütünüyle olası saldırılara karşı savunmasız bırakacaktır. Rasgele sayı dizilerinin istatistiksel özelliklerinin tanımlanması, diğer bir deyişle bu sayı dizilerinin kriptografik amaçlar için kullanılabilirliğini doğrulamak için istatistiki testler kullanılmaktadır. Bu çalışmada kriptografik RSÜ' ler için önemli bir güvenlik kriteri olan rasgelelik kavramı ele alınmış ve bu kavramla bağlantılı istatistiki nicel gereksinimlere ve değerlendirme yöntemlerine odaklanılmıştır. Bu kapsamda hazır test paketlerinin aksine bias, korelasyon, entropi, ki-kare ve standart sapma olmak üzere beş farklı bağımsız test stratejisi kullanılmıştır. Bu testler aynı zamanda Sahada Programlanabilir Kapı Dizileri (Field Programmable Gate Array-FPGA) ortamından elde edilmiş gerçek rasgele sayı dizilerine uygulanmış ve sonuçlar kriptografik gereksinimler doğrultusunda analiz edilmiştir. Literatürde bu maksatla kullanılan çoklu test paketlerinin yanı sıra, sunulan test yöntemleri ile de rasgelelik doğrulaması için geçerli ve güvenilir sonuçların elde edilebileceğini düşünmekteyiz.

ARTICLE INFO

Article history:

Received 01.11.2023
Received in revised form 06.01.2024
Accepted 22.02.2024
Available online 29.03.2024

Keywords:

Randomness, random numbers, cryptography, random number generators, statistical randomness tests

Doi: 10.24012/dumf.1384343

* Sorumlu Yazar

ABSTRACT

Considering that the confidentiality, complexity, and sustainable security mentality in cryptographic protocols are established over random numbers, it can be seen that there is a strong connection between cryptography and randomness. The qualification of the randomness that is the source of obtaining these numbers and their measurable (quantitative) properties that mathematical methods can guarantee have decisive importance on the performance of cryptographic systems. Therefore, the selection and evaluation of suitable design components from which these numbers are obtained and customized as random number Generator (RNG) is a challenging task in terms of cryptographic security. Because the security flaws that RNG will cause on random numbers will leave the cryptographic system completely vulnerable to possible attacks. Probability theory is used to describe the statistical properties of random number sequences, in other words, to verify the usability of these sequences for cryptographic purposes. In this study, the concept of randomness, which is an important security criterion for cryptographic RNGs, is discussed and the statistical quantitative requirements and evaluation methods related to this concept are focused. In this context, unlike ready-made test packages, five different independent testing strategies were used: bias, correlation, entropy, chi-square, and standard deviation. These tests are also applied to true random number sequences obtained from the Field Programmable Gate Array (FPGA) environment and the results are analyzed consistent with cryptographic requirements. Besides to the multiple test packages used in the literature for this purpose, we think that valid and reliable results can be obtained for randomness verification with the methods presented.

Giriş

Kerckhoff aksiyomu, bir kriptografik sistemin, anahtar dışındaki tüm ayrıntıları herkes tarafından bilinse bile güvenli olacak şekilde tasarlanması gerektiğini fikrine dayanır. Bu aksiyom, daha sonraları şifreleme sistemlerin tasarım ve anlaşılabilirliği ile ilgili teorik temelleri ortaya koyan Claude Elwood Shannon tarafından "*düşman sistemi biliyor*" olarak yeniden yorumlanmıştır. Gizlilik yoluyla güvenlik temelli klasik kriptografinin kapanışını da beraberinde getiren bu yaklaşım, belirsiz bir düzenden bugün açık bir disipline doğru evirilen modern kriptografi' nin gelişim sürecine de öncülük etmiştir. Öyle ki, günümüz modern kriptografik sistemlerinin neredeyse tamamı "*gizlilik yoluyla güvenlik anlayışının*" aksine, kökleri Shannon ve Kerckhoff tarafından ortaya atılan yaygın kabul görmüş bu temel prensiplere göre tasarlanmaktadır [1]-[2].

Modern kriptografik çalışmalara kaynaklık eden bu temel prensipler özetle; bir kriptografik sistemin güvenliğinin yalnızca anahtarların seçimine bağlı olduğunu ve algoritmanın kendisi de dahil olmak üzere diğer her şeyin genel geçer bilgi olarak kabul edilmesi gerektiğini vurgular. Bu nedenle, modern kriptografik güvenlik protokollerinin neredeyse tamamı uygulanma yöntemleri herkes tarafından bilinen anahtar tabanlı deterministik birer algoritmadır. Deterministikliğin doğası gereği bir güvenlik sistemini temsil eden bu protokollerin hiç biri, çıkışında girişindeki entropi değerinden daha fazlasını üretemez. Dolayısıyla bu protokollerin neredeyse tamamında temel güvenlik varsayımı rasgele üretilmiş gizli veriler (sayılar) üzerine inşa edilir [3].

Rasgele sayılar ile bu sayıların elde edildiği mekanizmalar, kriptografik protokollerin önemli bir tasarım bileşenidir. Taraflar arasında iletişime konu gizli verilerin şifrelenmesinin yanı sıra doğrulaması amacıyla da kullanılan bu verilerin güvenilir olması ve saldırganlar tarafından kolay hesaplanabilir olmaması gerekir. Daha genel bir ifadeyle; bu protokollere yapılacak olası saldırıların etki derecesi, algoritmik karmaşıklıktan ziyade rasgele sayılar için ihtiyaç duyulan temel güvenlik gereksinimleriyle doğrudan bağlantılıdır. Saldırgan sınırsız hesaplama kaynağına sahip olsa bile bu verileri tahmin etmek için başvurabileceği en iyi yöntem, basit bir tahmin veya yazı tura atışından öteye geçmemelidir. Zira her iki durumda da rasgele sayılar için tüm kombinasyonları içeren örneklem uzayı bugün pratik olarak denenemeyecek kadar büyüktür. Literatürde kaba kuvvet saldırı olarak da bilinen bu yöntemde, 128 bitlik rasgele anahtarın tüm olasılık uzayı 2^{128} ihtimalden oluşur. Bu, günümüz bilgisayarlarının hesaplama kapasiteleri düşünüldüğünde neredeyse evrenin yaşından daha fazla bir zamana ihtiyacımız olduğunu gösterir. Güvenlik gereksinimlerinin yerine getirilmesi şartıyla örneklem uzayının geniş tutulması, rasgele sayıların kaba kuvvete dayalı tahmin olasılığını da o kadar zorlaştırır. Şifreleme ve de-şifreleme işlemlerinin dışında, kriptografik sistemlerin önemli bir çoğunluğu oturum anahtarları, imza anahtarları ve parametreleri, kimlik doğrulama protokolleri, geçici anahtarlar, sıfır bilgi ispatı, blok şifreler için başlangıç vektörleri ve yan kanal saldırılarına karşı koruma maskeleme işlemleri rasgele sayıların üretim ve kullanımına gereksinim duyan kriptografik uygulamalardan bazılarıdır [4]-[5].

Rasgele sayılar kriptografinin dışında oyun teorisi, istatistiksel analiz, kuantum mekaniği, numerik analiz, fizik ve modern bilgisayar simülasyonları gibi farklı akademik disiplinlerde kullanılmaktadır. Örneğin bazı genetik algoritmalar veya yapay sinir ağı modellerinde rasgele girdilerin oldukça işe yaradığı ve geçerli sonuçlar üretebildiği bilinmektedir. Yine Monte Carlo analizlerinin kullanıldığı entegre devrelerin üretiminde, verimlilik kriterleri ile karakteristik farklılıklara yol açan varyasyonlar ile ilgili sayısal hesaplamalar için de rasgele sayılar kullanılmaktadır. Monte Carlo yönteminde rasgele sayılar özellikle belirsizlik altındaki bir olay(lar)ın olası sonuçlarının önyargılardan bağımsız doğru tahmini açısından oldukça önemlidir. Bu kullanım alanları için rasgele sayıların tahmin edilebilirlik de dahil basit istatistiksel özellikleri yeterli olabilmektedir. Fakat söz konusu kriptografi olunca bu sayıların iyi istatistiksel özelliklerine ek olarak tahmin edilemezlik ve tekrar üretilemezlik gibi temel kriterleri karşılaması istenir. Zira kriptografik protokollerin temel güvenlik varsayımı rasgele sayıların olasılık ve istatistik teorisi ile karakterize edilen bu temel gereksinimleri ne ölçüde karşıladığına bağlıdır [5]-[6].

Kriptografik düzlemde rasgele sayılar, her bir elemanı '0' ve '1' lerden oluşan ve kabul edilebilir bir aralık tahmini için bit düzeyindeki elamanların eşit oluşma olasılığı ile (düzgün dağılımla) birbirinden bağımsız oluştuğu ve elemanları arasında gizli veya açık örüntü barındırmayan bit düzeyinde sayı dizilerini ifade eder. Bu sayılar, Gerçek Rasgele Sayı Üreteçleri (GRSÜ) ve Sözde Rasgele Sayı Üreteçleri (SRSÜ) olmak üzere iki farklı tasarım sınıfından elde edilir. SRSÜ' ler çoğu zaman tohum değerlere ihtiyaç duyan bir matematiksel fonksiyondur. Giriş ve çıkış değerleri arasında matematiksel bir ilişkinin varlığı, SRSÜ' leri tahmin edilebilir kılmaktadır. Ayrıca bu üreteçlerde tohum değerler ve algoritmanın gizli kalması sistem güvenliği açısından oldukça önemlidir. Aksi takdirde, teorik olarak sistemin kopyalanarak çıkış dizilerinin kolayca elde edilebilmesi ciddi güvenlik riskleri oluşturabilmektedir. SRSÜ' lerin aksine, bir GRSÜ radyoaktif çürüme, termal gürültü, faz seğirmesi, yarı kararlı durumlar gibi fiziksel gerçeklikten beslenerek rasgele sayılar üretebilen, çoğu zaman donanım bağımlı bir cihazdır [7]-[9].

GRSÜ' lerde temel girdi olan fiziksel kaynağın kontrol edilebilirliğini sağlayan standart bir tanımının olmaması, üretilen sayıları tahmin edilemez ve tekrar üretilemez yapar. Rasgele sayılar için tekrar üretilemezlik gibi önemli bir gereksinimi yerine getirmelerine rağmen, pratikte GRSÜ' lerin tahmin edilemezlikle bağlantılı istatistiki yeterlilikleri maalesef zayıftır. Fiziksel kaynağın entropi eksikliği, rasgelelik çıkarımı, örnekleme hızı ve gerilim dalgalanmaları gibi çevresel etkiler GRSÜ' nin hassas kriptografik uygulamalar için kullanımını sınırlandıran bu durumun temel sebebi olarak görülebilir. Saldırganların rasgele sayıların zayıf istatistiksel özelliklerinden faydalanarak yararlı çıkarımlar yapabilmesi, bu sayıların kullanıldığı kriptografik uygulamalar açısından önemli bir güvenlik zafiyetidir. Bu durum, kriptografik protokollerin algoritmik yapısı ne kadar güçlü olursa olsun, güvenirliği zayıf rasgele sayılar sisteme beklenenden daha kısa sürede başarıya ulaşması muhtemel saldırıların gerçekleşmesine yol açar [5], [10]-[11].

Rasgele sayı üreteçlerini ve dolayısıyla rasgele sayıların kriptografik amaçlar için kullanılabilirliğinin test edilmesi, 1960' lı yıllardan günümüze kadar pek çok araştırmacının ilgisini çekmiştir. Bu noktada sık kullanılan yöntemler üreteç çıkışlarının, bilimsel ölçütler çerçevesinde, istatistiksel analizine dayanmaktadır. Kerckhoff aksiyomu düşünüldüğünde, rasgeleliğin kriptografide güvenlik bağlamında özel önem atfedilmiş güçlü bir silah olduğu görülebilir. Dolayısıyla kullanılan sayı dizilerinin rasgelelik açısından istatistiki gereksinimleri doğru analiz edilmeli ve değerlendirilmelidir. Bu aşamada, olasılık teorisi ve istatistik söz konusu yeterliliklerin matematiksel olarak doğrulanabilmesi için gerekli bilimsel araç ve yöntemler sunmaktadır. Her ne kadar test edilen sayı dizilerinin gerçek rasgele olup olmadığı veya elde edilme biçimleri hakkında geçerli kanıt sunmasalar da, bu yöntemlerle üreteçlerin zayıf ve üstün yönleri de saptanabilir [12].

Bu çalışmada kriptografik RSÜ' lerin temel değerlendirme kriterlerinden biri olan rasgelelik kavramı detaylıca ele alınmış ve rasgelelik varsayımının istatistiki açıdan doğrulanmasıyla ilgili bağımsız test yöntemlerine ve açıklayıcı bilgilere yer verilmiştir. Kriptografik rasgeleliğin bir dizi olasılık terimi ile karakterize edilmesi nedeniyle çalışma içerisinde hipotez tabanlı test yöntemlerinin olası hataları, farklı kritik önem sevilerine göre teorik dağılımları ve birbiriyle olan uyum ve ilişkileri de ele alınmıştır. Bu doğrultuda sunulan test yöntemlerinin deneysel doğrulaması için [13]' te önerilen ve tasarım detaylarına Bölüm 3' te yer verilen GRSÜ mimarisi kullanılmıştır. Bu mimariden elde edilen çıkış dizilerinin istatistiki yeterliliklerini doğrulamak için bias, korelasyon, entropi, düzgün dağılım, ki-kare, Dieharder ve standart sapma analizleri yapılmıştır. İstatistiki testlerle gerçek rasgelelik ile iyi sözde rasgelelik ayrımı yapılamaz. Bu nedenle çalışma içerisinde test altındaki sayı dizilerinin güvenilir bir kaynaktan elde edildiği ile geçerli kanıtlar sunabilmek için standart sapma analizine dayalı alternatif bir test yöntemi de sunulmuştur.

Kriptografik RSÜ' lerin istatistiksel yeterliliklerinin bilimsel açıdan doğru yöntem ve araçlarla test edilmesine dönük benzer bir çalışmanın literatürdeki eksikliği çalışmanın temel motivasyon kaynağını oluşturmaktadır. Bu doğrultuda çalışma içerisinde rasgeleliğin kriptografik açıdan önemini vurgulamak ve eksik ve hatalı değerlendirmeler neticesinde sistemde ciddi güvenlik zafiyetine yol açabilecek kriptografik RSÜ' lerin bilimsel açıdan doğru yöntem ve araçlarla test etmenin önemini vurgulanmıştır. Yazarlar olarak çalışmanın bu yönüyle kriptografik düzlemde yapılacak akademik çalışmalara fikir ve uygulama bazında kaynaklık edeceğine inanmaktayız.

Çalışmanın geri kalan kısmı şu şekilde organize edilmiştir: Bölüm 2' de gerçek rasgelelik ile bağlantılı temel istatistiki gereksinimler sunulmuştur. Bölüm 3' te analiz işlemleri için kullanılan rasgele sayıların elde edildiği GRSÜ mimarisi ile ilgili teorik ve teknik bilgilere yer verilmiştir. Bölüm 4' te bias, entropi, otokorelasyon, lineer karmaşıklık kavramları ve bu kavramların doğrulaması için kullanılan bağımsız test tekniklerine yer verilmiştir. Aynı bölüm içerisinde ek olarak test tekniklerinin uygulandığı rasgele sayılar için elde edilen sonuçlar tartışılmış ve bu sonuçlar, Dieharder istatistiki rasgelelik testleriyle uyumuna yer verilmiştir. Son bölümünde çalışma, elde edilen sonuçları itibariyle özetlenmiş ve gelecek çalışmalara da yer verilerek sonlandırılmıştır.

Kriptografide Gerçek Rasgelelik Kavramı ve Temel İstatistiki Gereksinimler

Kriptografik amaçlar için kullanılacak GRSÜ' lerde için temel güvenlik varsayımı, üretilen sayıların fiziksel bir gerçekliğe bağlı olmasının yanı sıra bu sayıların istatistiki özellikleriyle de güçlü bir şekilde ilişkilidir. Fakat literatürde rasgele sayılar bu için karakteristik gereksinimleri karşılayabilen kriptografik açıdan güçlü bir RSÜ' yü elde etmenin basit bir metodolojisi yoktur. Özellikle fiziksel rasgeleliğin bir sonucu olarak ortaya çıkan ve son işlem teknikleriyle giderilmeye çalışılan istatistiki zayıflık problemi, GRSÜ' lerin önemli bir eksikliğidir. Aksi takdirde zayıf istatistiksel karakteristiğe bağlı rasgele sayılar üzerindeki herhangi bir öngörülebilirlik, tüm sistemde bir zayıflığa yol açabilmektedir. Dolayısıyla kriptografik amaçlar için kullanılacak RSÜ' lerin tasarım ve analizinin doğru yapılabilmesi, gerçek rasgelelik kavramının felsefik tanımının doğru anlaşılmasıyla birlikte, bu tanımla uyumlu kesin ve nicel yaklaşımların kullanılmasıyla mümkündür.

Düzensizlik (karmaşıklık), öngörülemezlik (tahmin edilemezlik) gibi biçimsel tanımlarla ifade edilmeye çalışılan gerçek rasgelelik kavramı, disiplinlere göre farklı anlamlar içerebilmektedir. Sıradan bir insan için düzenlilik kavramı görsel olarak kolayca ifade edilebilen bir olgu iken, rasgelelik görsel olarak kolayca ifade edilebilen düzenlilik kavramının tam tersi yani, düzensizlik olarak ifade edilebilir. Örneğin bu bakış açısıyla ormandaki ağaçların yerleşimi, depremlerin oluşumu ve atmosferik olayların değişimi gibi evrendeki pek çok olayın düzensiz olarak yani gerçek rasgele meydana geldiği söylenebilir. Tahmin edilemezlik açısından ise rasgelelik kavramı, geçmişi bilinen bazı olayların gelecek sonuçlarının tahmin edilebilir olduğu deterministikliğin tam tersi bir kavram olarak düşünülebilir. Dolayısıyla rasgele olduğu varsayılan olayların gelecek veya önceki durumları arasında tahmin edilebilirliğe yol açan bir bellek etkisi olmamalı ve bu olaylar birbirinden tamamen bağımsız olarak meydana gelmelidir [12], [14].

Gerçek rasgelelik kavramının var olup olmadığına dair felsefik soruları da barındıran farklı tanımları da mevcuttur. Fakat kriptografide esas alınan tanımı, daha ziyade güvenlikle bağlantılı olan öngörülemezlik kavramıyla örtüşmektedir. Bu bağlamda; gerçek rasgelelik, bazen bir değerin olası çıktılarının yeniden üretilemeyeceği fiziksel işlemleri kapsadığı sonucunu doğurur. Sıradan bir bakış açısıyla, havaya atılan hilesiz bir madeni paranın beklenen olası sonuçlarının gerçek rasgele olduğu söylenebilir. Fakat paranın hileli olması durumunda olasılık dağılımına ait sonuçlar, bu sonuçların tahminini mümkün kılan sistemik bir hata payı ile meydana gelir. Olasılıksal dağılımın muhtemel çıkışlardan biri etrafında yoğunlaşmasına yol açan bu hata payı, havaya atılan paranın gelecekteki sonuçlarının tahmin edilebilirliğini, rasgelelik dizilimini açısından bozabilir. Özetle, hilesiz bir madeni bir paranın havaya atılması gibi düşük seviyeli rasgele olaylardan elde edilebilecek 128 bitlik bir sayı dizisi her ne kadar gerçek rasgele olsa da kriptografik kullanım açısından yetersiz olabilir. Ya da tamamen adil gerçek rasgele bir rulet çarkını çevirdiğinizde, gerçekleşme olasılığı düşük olayların bilinirlik yönüyle daha fazla bilgi içerdiği düşünüldüğünde, kaybedeceğinizi tahmin etmek pek te zor olmasa gerek. Oysaki sayı dizisinin rasgelelik dağılımı

kriptografik güvenlik açısından yeterince tahmin edilemez olmayı gerektirir [15]-[16].

Rasgelelik ile tahmin edilemezlik arasında matematiksel tanımı Denklem 1' de verilen ve entropi olarak adlandırılan matematiksel ayırt edici bir kavram daha bulunur. Bir rasgele değişkenin belirsizlik ölçüsü olarak da bilinen entropi, dizi elemanlarının oluşma olasılığının ideal değere yani $1/2$' ye yakın dağılımları için tepe değerine ulaşılır, yani 1' e eşit olur. $X_n = x_1x_2x_3 \ldots . x_n \in \{0,1\}^n$ şeklindeki n bitlik rasgele sayı dizisinde entropinin yüksek olması, bu sayı dizisindeki öncül veya ardıl herhangi bir bitin $1/2$' den daha yüksek olasılıkla tahmin edilemeyeceğini garanti eder. Bu ideal olasılığın sayı dizisindeki tüm bitler için geçerli olması veya olmaması durumuna bağlı olarak n bitten oluşan gerçek bir rasgele sayı dizisinin toplam entropisi, $0 - n$ bit arasında bir değere sahip olabilir [1], [7].

Tahmin edilemezliği garanti edebilmek için rasgele sayı dizilerinde entropi olabildiğince dizi boyutuna yakın bir değer olması istenir. Örneğin 100 bitlik bir rasgele sayı dizisi için 20 bitlik düşük entropi değeri, aynı zamanda bu sayı dizisinin $1/2^{20}$' lik bir olasılıkla doğru tahmin edilebileceğini gösterir. Diğer bir deyişle bu sayı dizisinin muhtemel bütün olasılıkları (olasılık kümesi) 2^{20} adet deneme-yanılma ile bulunabilir. Oysaki rasgele sayıların kullanıldığı günümüz modern kriptografik protokollerinde gizlilik, tüm olasılıklar kümesinin pratikte denenemeyecek kadar büyük (en az 2^{100} bit ve üzeri) olmasını gerektirir. Dolayısıyla tahmin edilemezlik açısından GRSÜ' lerde bit başına entropinin yüksek olması ve çıkış olarak üretilen sayı dizilerinin toplam entropisinin olabildiğince dizi boyutuna eşit olması istenir [4], [10].

$$H(x) = -\sum_{i=0}^{n-1} p_i \,.\log_2 p_i \qquad (1)$$

Kriptografik RSÜ' ler için temel değerlendirme kriteri olan güvenlik ile bağlantılı tasarım gereksinimleri Tablo 1' de verilmiştir. Tablo 1' deki karakteristik gereksinimler, RSÜ' lerin kullanıldıkları kriptografik uygulamaların önem seviyesine göre değişiklik gösterebilmektedir. Örneğin R1 ve R2 gereksinimleri fiziksel GSRÜ' ler için özelleştirilmiş tasarım gereksinimleri olmakla birlikte, hassas kriptografik uygulamalar için mutlaka yerine getirilmelidir. Bu gereksinimler kapsamında rasgele sayı dizilerinin saldırganlar için yararlı çıkarımlar sağlayabilecek hatalı olasılıksal dağılımlar (bias) ve güçlü istatistiki bağımlılıklar (korelasyon) içermemesi istenir. Kriptografik güvenlikle bağlantılı Tablo 1' deki gereksinimler, her ne kadar iyi tanımlanmış ve anlaşılmış olsa da pratikte bu gereksinimleri başarıyla yerine getirebilen kusursuz bir RSÜ metodolojisi yoktur [1], [5], [9].

Herhangi bir üretecin veya bu üreteçten elde edilmiş sayı dizilerinin rasgelelik açısından değerlendirilmesi, sezgisel bir tanımdan ziyade olasılık teorisi ve istatistiki araç ve yöntemleri uygulamaktan geçer. Çünkü pratikte bir üretecin gözlemlenemeyen özelliğini temsil eden gerçek rasgeleliğin, kriptografide mantıksal ve fiziksel kanıtı yoktur. Dolayısıyla sayı dizilerinin gerçek rasgele mi yoksa bir düzen içerisinde üretilip üretilmediğine karar vermek için kullanılan testler, olasılık ve istatistik teorisinden yararlanır.

Tablo 1. Kriptografik rasgele sayı üreteçleri için temel karakteristik gereksinimler

Gereksinim Adı	Gereksinimin Açıklaması
R1 Gereksinimi:	Kriptografik uygulamalarda kullanılacak rasgele sayılar iyi istatistiksel özelliklere sahip olmalıdır.
R2 Gereksinimi:	Rasgele sayıların alt dizilerinin bilmesi halinde, saldırganın öncül ve ardıl rasgele sayıları hesaplanmasına veya yüksek doğrulukla tahmin etmesine izin verilmemelidir.
R3 Gereksinimi:	Bir RSÜ' nin bilinen mevcut iç durum değerinden yola çıkarak veya iç durum bilgisine ihtiyaç duymadan önceki üretilen rasgele sayıları, yüksek doğrulukla tahmin edebilmek veya hesaplayabilmek mümkün olmamalıdır.
R4 Gereksinimi:	Bir RSÜ' nin bilinen mevcut iç durum değerinden yola çıkarak veya iç durum bilgisine ihtiyaç duymadan gelecek rasgele sayıları, yüksek doğrulukla tahmin edebilmek veya hesaplayabilmek mümkün olmamalıdır.

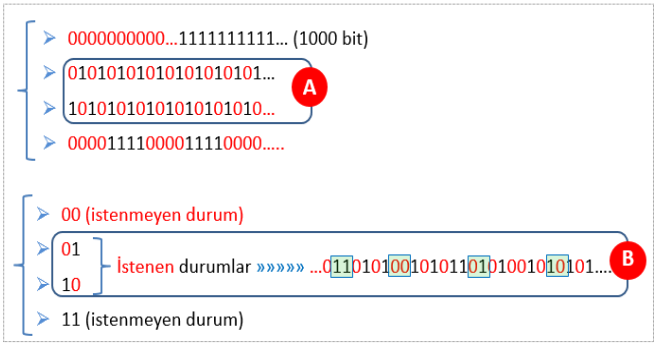
R1 gereksinimi kapsamında GRSÜ' nin ideal gerçek bir $X_n \in \{0,1\}^n$ rasgele sayı dizisinin istatistiki açıdan Denklem 2 ve 3' teki olasılıksal dağılımlar ile üretmesi beklenir [10], [17].

$$\varepsilon = |Prob\ (X_n = 0) - 0.5| = |Prob(X_n = 1) - 0.5| \qquad (2)$$

$$Prob(Transitions) = \frac{1}{2} \qquad (3)$$

İyi bir rasgele sayı dizisinde, bit düzeyendeki elamanların oluşma olasılığının birbirine eşit veya çok yakın olması istenir. Bit düzeyindeki bu sayı dizilerinin oluşma olasılığı için ideal istatistiki değer $1/2$' dir. Diğer bir deyişle rasgele sayı dizileri mümkünse bias olarak da isimlendirilen eşit oluşma olasılığının dışında bir sapma payıyla oluşmamalıdır. Bit düzeyindeki bir rasgele sayı dizisinin olası sonuçları arasında oluşan bias, Denklem 2' deki bağıntı ile ifade edilebilir. Dizi elemanlarının $1/2$' lik ideal oluşma olasılığına sahip bir istatistiki dağılımla oluştuğu durumlarda ε =0 olur ve sayı dizilerinin biassız oluştuğu kabul edilir.

Fakat çoğu durumda rasgele sayılar için Denklem 1' deki ideal istatistiki tanımın tek başına yerine gelmesi kriptografik açıdan yeterli olmayabilir. Örneğin Şekil 1 (A)' da rasgele sayılar için istatistiki açıdan her ne kadar ideal dağılım ölçüsü sağlanmış olsa da benzer tekrarlı örüntülerden oluştukları için lineer karmaşıklıkları oldukça düşüktür. Oysaki tahmin edilemezlik açısından rasgelelik, önceki tüm durumları bilinen bir X_n, $n = 0,1,2,3 \ldots$ şeklindeki sayı dizisinde X_{n+1}. elemanı tahmin etmek için ihtiyaç duyulan hesaplama karmaşıklığına dayanır. Bu durum, istatistiki gereksinimlerle karakterize edilen gerçek rasgeleliğin önemli bir tamamlayıcı unsurudur [12]-[18].



Şekil 1. Lineer karmaşıklık açısından rasgele sayılar için olumlu ve olumsuz dağılım örneği

Şekil 1 (B)' de tasvir edilen Denklem 2' deki bir diğer özellik için, ortalama her ikinci rasgele sayının bit değerinde, bir önceki üretilen rasgele sayının bit değerinden farklı bir geçişin $(0 \rightarrow 1, 1 \rightarrow 0)$ olması beklenir. Sayı dizisi içerisinde bu rasgele geçişler için iki bitlik "01" ve "10" kombinasyonları kabul edilebilir durumlardır. Dizi elemanlarının eşit oluşma olasılığının yanı sıra ortalama her iki bitte bir kendini tekrar etmesi istenen bit seviyedeki bu geçişlerin toplamının sayı dizisinin toplam uzunluğuna oranı da $1/2$ olmalıdır. Fakat kabul edilebilir kombinasyonların dizi içerisinde tamamen gerçek rasgele bir formda dağılması istenir. İstenen durumların gerçek rasgele bir formda dağıldığı n bitlik yeterince uzun bir rasgele sayı dizisinde, "11" ve "00" gibi bias oluşumuna neden olan durumlar da dahil "00", "01", "10" ve "11" gibi 4-bitlik tüm kombinasyonların elde edilebileceği unutulmamalıdır [19]-[20].

Denklem 2 ve 3' teki ideal istatistiki tanımlar doğrultusunda, Şekil 2' deki örnekte ilk senaryo için herhangi bir bitini bildiğimiz bir gerçek rasgele sayı dizisinin sonraki iki bitini tahmin etmeye çalışalım ve bunu n defa tekrar ettiğimizi varsayalım. İdeal bir rasgele sayı dizisinin istatistiki açıdan düzgün bir dağılıma sahip olması, iki bitlik $(00, 01, 10, ve\ 11)$ olası kombinasyonların dizi içerisinde eşit oluşma olasılığıyla meydana gelmesiyle mümkündür. Tüm ikili kombinasyonların frekansının eşit olduğu bu durum, aynı zamanda rasgele sayılar için tahmin edilebilirliğin istatistiki açıdan en zor olduğu durumdur. Şekil 2 (B) ve (C)' deki gibi istatistiki dengesizlik, rasgele sayıların ayrıcalıklı saldırganlar tarafından yüksek doğrulukla tahmini mümkün kılacak yararlı bilgiler sunabilmektedir. Örneğin Şekil 2 (B)' de tahmin edilecek sonraki ikili kombinasyonun "00" olma olasılığı %50 iken, 2 (C)' de ise %70 gibi istenmeyen yüksek bir orandır. Fakat bu olasılığın tüm ikili kombinasyonlar için Şekil 2 (A)' daki ideal bir istatistiki dağılımda sadece %25 olduğu unutulmamalıdır. Benzer durum Şekil 3' teki bir bitlik tahmin için de geçerlidir. Benzeri şekilde dengesiz bir istatistiki dağılımla oluşmuş Şekil 3 (C)' de bir rasgele sayı dizisinde, sonraki bit değerinin '1' olma olasılığı %80' dir. Oysaki düzgün istatistiki dağılımla oluşmuş gerçek rasgele sayılar, bu sayılar üzerinde bir saldırganın basit bir yazı tura atma olasılığından veya kör bir tahminden daha fazlasına müsaade etmemelidir.

Denklem 2 ve 3' te tanımlı iyi istatistiksel özelliklerinin yanı sıra rasgele sayılar için bir diğer önemli karakteristik gereksinim ise bağımsızlık varsayımıdır. Bağımsız olduğu varsayılan rasgele sayı dizileri, bu sayıların yüksek doğrulukla tahminini mümkün kılacak güçlü istatistiki bağımlılıklar yani korelasyon içermemelidir. Nitekim pratikte fiziksel bir GRSÜ için pek te mümkün olmayan bağımsızlık varsayımı aynı zamanda biasın temel sebeplerinden biri olarak görülebilir. Örneğin oluşma olasılıkları $E(X)$ ve $E(Y)$ olan X ve Y gibi n bitlik iki rasgele sayı dizisi için oluşma olasılığı, μ ve korelasyon katsayısı, ρ olsun. Bu dizilerin XOR' lanmısya elde edilen yen dizinin $(X \oplus Y)$ oluşma olasılığı,

$$E(X \oplus Y) = \frac{1}{2} - 2\left(\mu - \frac{1}{2}\right)^2 - 2\rho\mu(1-\mu) \quad (4)$$

Denklem 4' teki gibi hesaplanır. Eğer $E(X)$ ve $E(Y)$ için μ değeri $1/2$' ye yaklaştıkça Denklem 4 aşağıdaki gibi yeniden yazılabilir.

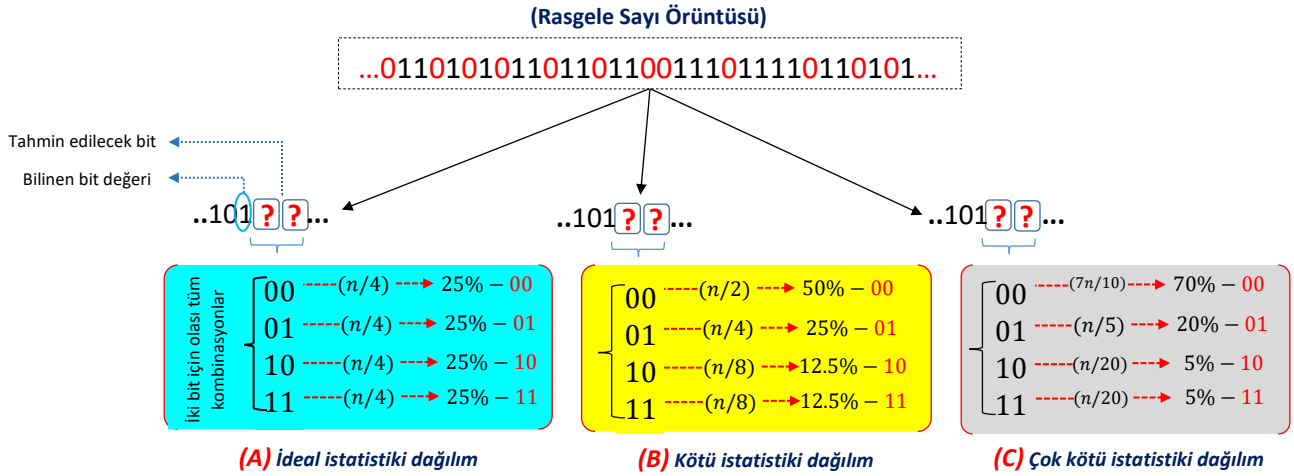
$$E(X \oplus Y) \approx \frac{1}{2}(1-\rho) \quad (5)$$

Denklem 5' ten iki dizi arasındaki korelasyonun aynı zamanda biasın oluşumuna da yol açtığı görülebilir. X ve Y dizilerinin korelasyonuz oluştuğu durumlarda $\rho = 0$' dır ve $E(X \oplus Y) \approx \frac{1}{2}$ ' ye yakınsar. Denklem 4 ve 5' teki gözlemler korelasyon katsayısının bias üzerindeki etkisini açıkça göstermektedir. Bu durumda kriptografik rasgelelik açısında bias, korelasyon ve tahmin edilemezlik kavramlarının birbiriyle doğrudan bağlantılı olduğu açıkça söylenebilir.

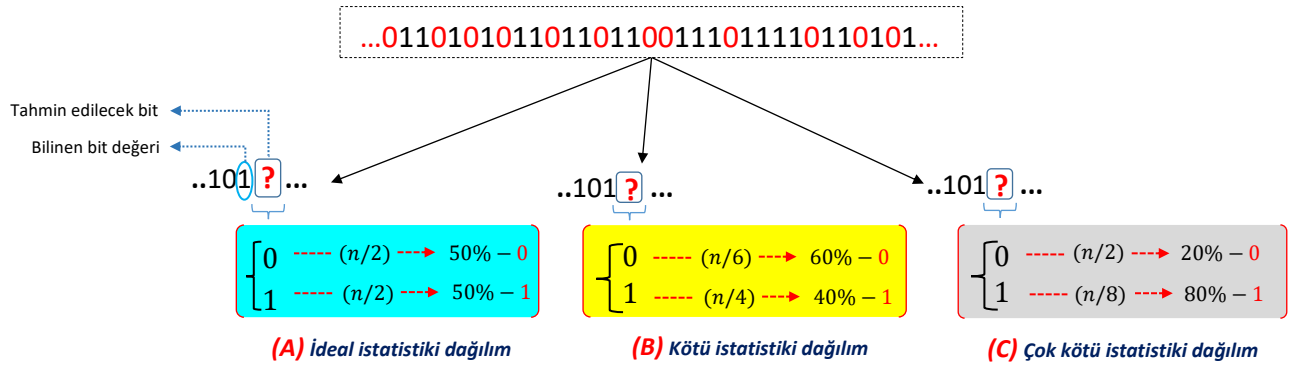
Gerçek Rasgele Sayı Üreteci Mimarisi

Çalışma içerisinde test amaçlı kullanılan rasgele sayı dizileri, blok mimarisi ve davranışsal modellemesi sırasıyla Şekil 5 ve 6' da verilen serbest salınımlı halka osilatörlerin (HO) kullanıldığı GRSÜ mimarisinden gerçek zamanlı olarak elde edilmiştir. İyi istatistiksel özelliklerinin yanı sıra düşük frekanslarda çalışabilme, yüksek çıkış bit hızı, sayısal mantık cihazlarına kolay entegre edilebilme ve tasarım esnekliği gibi öne çıkan özellikleri, bu mimarinin tercihinde belirleyici sebeplerdir.

GRSÜ' nin gürültü kaynağı, kendi içerisinde her biri 13 adet gecikme elemanından (invertör) oluşan paralel bağlı 114 adet HO' dan oluşur. Şematik yapısı Şekil 4 (a)' da verilen HO' lar giriş ve çıkış arasında bir geri besleme yolu ile ardışıl

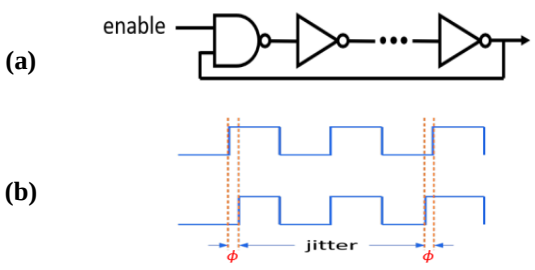


Şekil 2. İki bitlik tahmin için rasgelelik ve tahmin edilemezlik arasındaki istatistiki ilişki



Şekil 3. Bir bitlik tahmin için rasgelelik ve tahmin edilemezlik arasındaki istatistiki ilişki

bağlı tek sayıda eviriciden oluşur. Mantık devrelerinin üretim ve çalışma şartlarına bağlı farklılıklar nedeniyle HO çevrimindeki her bir mantık kapısı, kararsız yayılım ve yönlendirme gecikmesine sahiptir. Bu durum, Şekil 4 (b)' deki gibi üretilen saat sinyallerinin periyodik düzensizliğine yol açarak HO' ların frekanslarının farklı olmasına neden olur. Her çevrimde saat sinyallerinin yükselen ve düşen kenar geçişlerinde meydana gelen zamansal gecikmenin yönü ve miktarı pratikte öngörülemezdir. Zamana bağlı artış gösterme eğilimindeki bu kararsız durum, yerel (termal gürültü ve titreme gürültüsü vb.) ve global (cihazın çalışma şartlarına bağlı güç kaynağı ve sıcaklık değişimleri vb.) değişkenlerin etkisiyle gerçek rasgele meydana gelir [13], [21].



Şekil 4. (a), halka osilatörün kombinasyonel yapısı, (b) jitter oluşumu

GRSÜ mimarisinin donanımsal gerçekleştirimi için Altera Cyclone IV FPGA geliştirme kartı kullanılmıştır. GRSÜ' nin her bir bileşenin davranışsal modellemesi ve sentezleme işlemleri için Quartus uygulama geliştirme platformu kullanılmıştır. Bu platform üzerinde devre elemanlarının mantıksal tasarım ve simülasyon işlemleri için Hardware Description Language (VHDL) donanım tanımlama dilinden yararlanılmıştır. Bu teknik bilgiler ışığında kullanılan GRSÜ mimarisinin Quartus ortamında davranışsal modellemesi Şekil 6' daki gibidir.

Şekil 6'daki mimarinin çalışma prensibi özetle şu şekildedir: Sistemde osilatörlerin yüksek salınımlı çıkışları, çok girişli ve tek çıkışlı bir XOR devresi ($xorcircuit114$) yardımıyla birleştirilmiştir. Ardından birleştirilerek non-periyodik bir işarete indirgenen yüksek frekanslı bu çıkışlar, iki ayrı D-türü flip-flop ($DFF1$ ve $DFF2$) yardımıyla örneklenerek saf gerçek rasgele sayılar elde edilir. Aynı non-periyodik girdi için örnekleme işlemi 50 MHz' lik referans saat sinyalinin düşen ve yükselen kenar geçişlerinde ayrı ayrı yapılır. İki ayrı fazda üretilen ve karakteristik olarak birbirinden farklı bit düzeyindeki bu rasgele sayılar, iki farklı birleştirme devresine ($combiner\_circuit1$ ve $combiner\_circuit2$)

atanarak 8-bitlik vektörel işaretlere dönüştürülerek, XOR devresine (xor_8bit) giriş olarak uygulanır.
Son aşamada ise; saf rasgele sayıların oluşturduğu bu 8 bitlik vektörler, s-box tabanlı düşük alan-enerji gereksinimli, sıkıştırmasız son işlem tekniğine tabi tutulur. Bu aşamada 8 bitlik saf rasgele sayılar, [22]' de önerilen ve içeriği Tablo 2 ve 3' te verilen kaos tabanlı sabit s-box' ların oluşturduğu statik hafıza bloklarında ($sbox1$ ve $sbox2$) yer değiştirme işlemine tabi tutulur. Yer değiştirme işleminde, 8 bitlik giriş vektörlerinin ilk dört biti satır, son dört biti ise sütun değeri olarak alınır. Çıkış vektörleri ($GRSU\_out1$ ve $GRSU\_out2$) ise, bu değerlerin Tablo 2 ve 3 üzerindeki kesişimlerine denk gelen 0-255 arasındaki tam sayı değerleriyle yer değiştirilerek elde edilir. S-box' ların içeriklerinin birbirinden farklı olmasından dolayı, aynı 8 bitlik giriş vektörlerine karşılık elde edilen 8 bitlik çıkış vektörleri de birbirinden farklıdır. GRSÜ mimarisi her çevrimde, gürültü kaynağından elde edilen 8 bitlik saf rasgele sayı dizisi için s-box tabloları üzerinden her biri 8 bitlik toplamda 16 bitlik gerçek rasgele çıkış dizisi üretir. Sistemde çıkışlar rasgelelik analizi için Şekil 5' teki gibi hafıza mimarisine kaydedilmiştir. Hafıza içerikleri, cihaz çalışır durumda iken Joint Test Action Group (JTAG) arabirimi aracılığıyla gerçek zamanlı olarak elde edilmiştir.

Tablo 2. Henon harita tabanlı kaotik s-box

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
0	99	161	159	152	130	108	234	90	252	240	194	40	85	204	57	81
1	149	206	214	88	15	62	55	105	116	61	83	225	74	135	118	218
2	249	134	126	1	2	227	44	72	229	52	199	29	226	172	69	238
3	205	7	45	32	187	10	53	76	21	26	175	107	146	171	98	169
4	200	35	39	67	110	3	113	170	125	5	165	112	155	198	163	236
5	254	97	91	123	168	96	222	241	124	27	68	212	251	141	129	102
6	223	71	215	59	239	34	211	43	109	122	4	213	48	144	228	158
7	217	232	156	242	188	87	147	28	127	114	42	101	84	136	209	64
8	31	253	100	18	184	93	231	12	120	51	220	192	244	245	202	132
9	63	150	250	9	142	54	193	145	60	185	49	210	50	65	111	30
A	237	151	181	47	115	143	160	246	70	94	186	148	180	189	58	247
B	106	24	208	174	157	137	82	14	219	154	128	25	22	75	41	36
C	139	8	235	164	140	248	37	138	182	191	121	255	216	177	11	79
D	167	23	73	162	104	86	166	178	33	133	78	56	131	190	183	46
E	66	77	179	221	119	176	0	224	203	196	230	103	19	201	233	92
F	173	16	195	197	20	207	6	80	95	89	117	13	153	243	17	38

Tablo 3. Chen sistem tabanlı kaotik s-box

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
0	11	12	55	156	97	136	92	130	183	159	89	158	184	13	23	57
1	99	245	93	242	160	116	249	142	146	141	28	226	244	78	69	112
2	85	178	207	231	110	135	7	58	202	239	100	45	129	220	113	238
3	108	102	210	51	193	48	230	194	21	95	248	111	246	192	243	204
4	39	247	236	132	35	218	61	88	222	38	47	134	227	235	166	201
5	252	6	180	87	138	16	144	104	105	131	26	203	59	91	64	206
6	139	127	198	62	18	50	70	80	73	175	53	71	76	161	221	254
7	40	211	181	219	234	37	170	119	43	128	255	151	241	189	10	123
8	195	27	223	205	217	197	30	200	188	49	101	216	75	162	25	63
9	121	60	84	34	164	149	187	171	126	176	31	191	2	165	212	143
A	67	125	19	224	81	4	208	174	52	118	44	41	66	148	14	250
B	225	150	145	185	9	137	232	77	117	168	182	251	167	36	0	72
C	240	214	74	96	20	190	154	213	215	106	209	196	153	90	65	82
D	253	177	115	169	22	233	120	157	68	42	1	133	3	163	33	56
E	179	83	54	199	79	109	186	122	15	5	114	147	46	228	173	94
F	32	103	229	107	237	155	98	124	172	140	24	17	86	29	8	152

Rasgeleliğin İstatistiksel Doğrulama Metodolojisi Ve Kullanılan Test Teknikleri

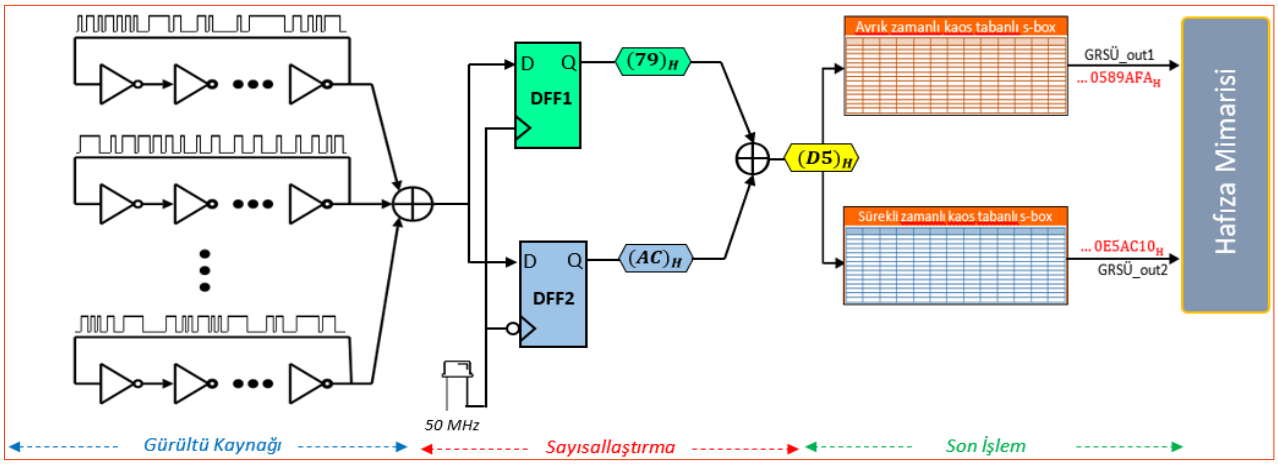
Olasılık terimleriyle karakterize edilebilen ve istatistiksel bir özellik olan rasgelelik, herhangi bir sayı dizisinin özelliklerinin gerçek bir rasgele sayı dizisinin beklenen özellikleriyle karşılaştırabilmesine imkân sağlar. Literatürde herhangi bir sayı dizisinin rasgele olup olmadığına karar verebilmek için yaygın olarak hipotez tabanlı test teknikleri kullanılmaktadır. Belirli bir anlamlılık seviyesinin (önem derecesinin) dikkate alındığı bu testlerin tamamı, rasgelelik varsayımını doğrulayan belirli bir geçersiz hipotezin (H_0) doğruluğunu test etmek için formülüze edilmiştir. Alternatif hipotez (H_a) ise dizinin rasgele olmadığını ileri sürer. Her bir test kriterinin başarısız olduğu diğer bir deyişle alternatif hipotezin kabul edildiği durumlarda H_0 hipotezi reddedilmiş olur. Mevcut her test için bir rasgelelik istatistiği seçilir ve bu istatistik geçersiz hipotezin reddine karar vermek için kullanılır.

Kriptografi de dahil olmak üzere rasgele sayıların üretim ve kullanımına ihtiyaç duyulan farklı uygulamaların güvenliği, kullanılan rasgele sayıların kalitesiyle doğrudan bağlantılıdır. Rasgele sayıların istatistiksel yeterliliklerinin doğru saptanabilmesi için bu sayı dizilerinin yeterince uzun olması gerekir. Gerçek rasgele olup olmadıklarının yanı sıra bu sayı dizilerinin tahmin edilemezlik ve bağımsızlık varsayımları istatistiki araç veya yöntemlerle mutlaka doğrulanmalıdır. Bu maksatla önerilmiş hipotez tabanlı farklı testler bulunmaktadır. Test yöntemleri, H_0 hipotezini doğrulamak için uygulandıkları sayı dizileri üzerinde, bu dizilerin rasgele olmadığına işaret edecek örüntüleri tararlar [13].

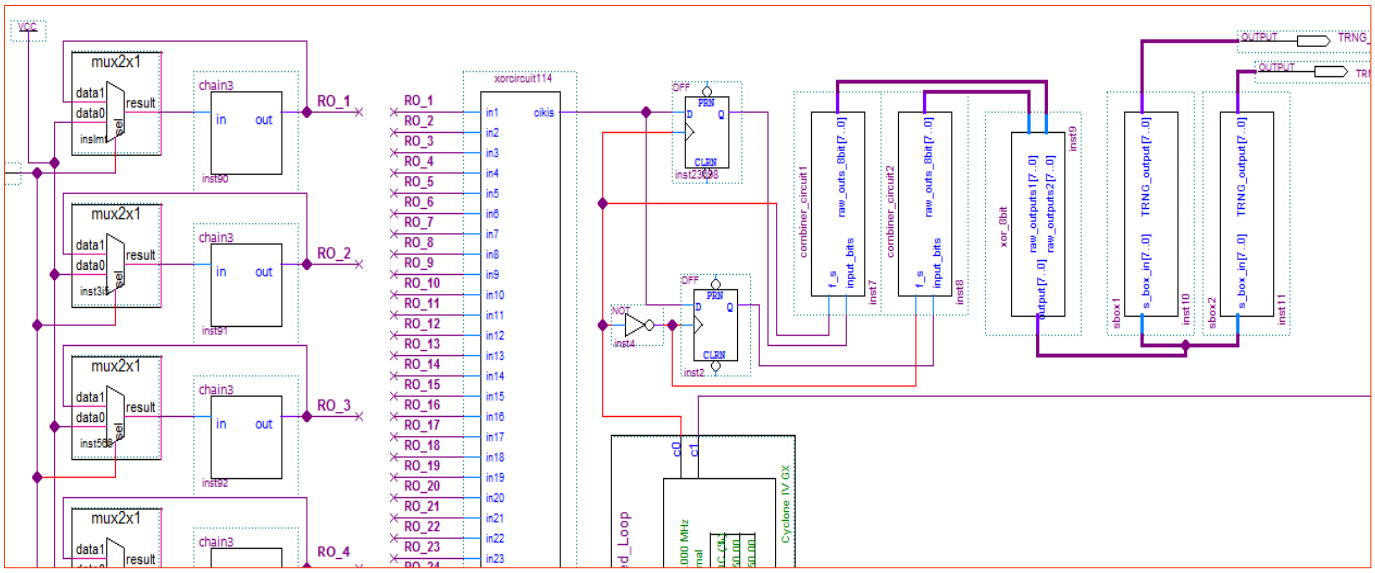
Çalışmanın bu bölümünde gerçek rasgele üretilmiş sayı dizilerinin rasgelelik varsayımını doğrulamak için kullanılabilecek test tekniklerine, bu testlerin uygulama yöntemlerine ve uygun parametre seçimlerine yer verilmiştir. Bu kapsamda bias, korelasyon, entropi, karmaşıklık ve gerçek rasgelelik olmak üzere beş farklı test tekniği ele alınmıştır. Test teknikleri her biri 524.288 bitten oluşan toplamda 100 farklı rasgele sayı dizisine uygulanmıştır. Teste tabi sayı dizileri, GRSÜ mimarisinden gerçek zamanlı olarak elde edilmiş ve ilgili test formatına uygun biçime dönüştürülmüştür. Test tekniklerinin farklı anlamlılık seviyelerinin seçimine bağlı olarak rasgelelik dağılımı üzerindeki istatistiki etkisi, analiz edilmiş ve yine bu testlerin uygulanmasına ilişkin bazı öneriler sunulmuştur.

Bias Analizi

Kriptografik amaçlar için kullanılacak ideal bir rasgele sayı dizisinde, bu diziyi oluşturan bit düzeyindeki elemanların oluşma olasılığı açısından beklenen değeri ($E(X)$ birbirine eşit yani $1/2$ olmalıdır. Herhangi bir sayı dizisi için bu ideal istatistiki dağılım ölçüsünden sapma miktarı bias olarak adlandırılır. Rasgele sayı dizisinin oluşma olasılığının ideal değerden sapma miktarı, bu sayı dizilerinin rasgele olmadığına karar verebilmek için tek başına yeterli bir kanıt sunabilmektedir.



Şekil 5. GRSÜ' nin temel tasarım bileşenleri



Şekil 6. GRSÜ' nin Quartus ortamında davranışsal modellemesi

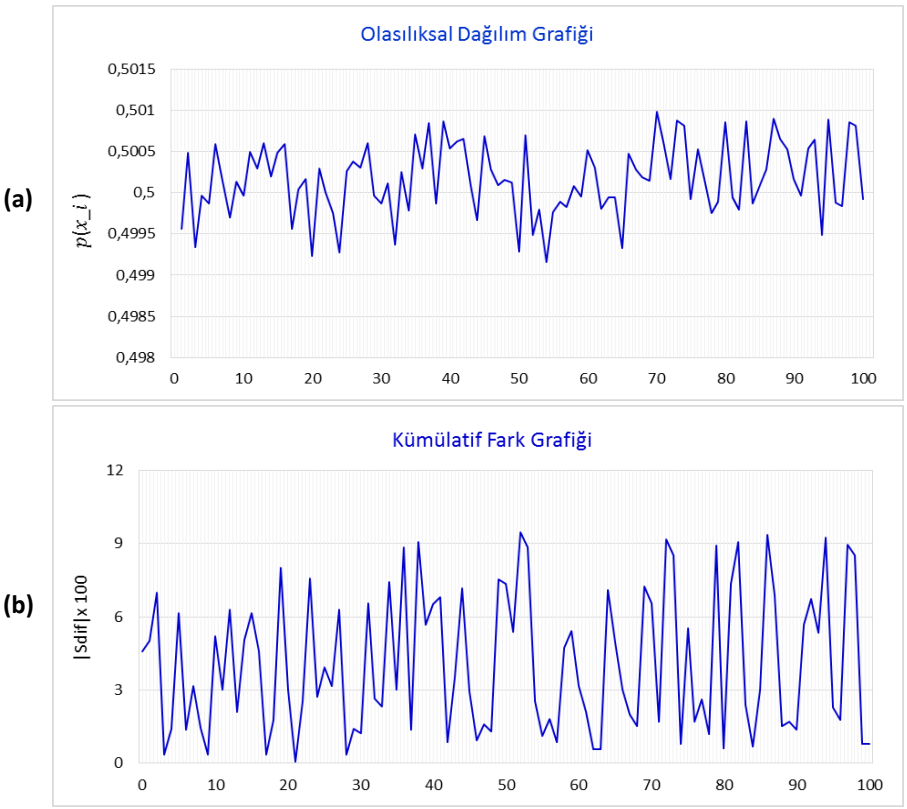
Bit düzeyindeki rasgele sayıların olasılıksal dağılımlarını incelemek için frekans (monobit) testi kullanılmıştır. Yarı normal dağılımın referans alındığı testin temel odak noktası, n bitlik rasgele sayı dizisinde 1' lerin veya 0' ların frekansının $n/2$' lik beklenen ideal değere yakınlığını tespit etmektir [18]. $X = x_0x_1x_2 \dots x_{n-1}$ n bitlik rasgele sayı dizisi olmak üzere Monobit testinin matematiksel tanımı Denklem 6-8' deki gibidir. Denklem 6 ve 7' de X_{dif} ve S_{obs} sırasıyla rasgele sayı dizisinin elamanları için hesaplanan kümilatif fark ve bu fark değerine karşılık hesaplanan görece p_{value} değeridir. $erfc$ tamamlayıcı hata fonksiyonu olmak üzere gerçek p_{value} değeri, diğer bir deyişle test istatistiği Denklem 8' deki gibi hesaplanır. Hipotez tabanlı testlerin sonuçlarını yorumlamak için test istatistiğini temsil eden p_{value} değeri dikkate alınır. Hipotez tabanlı bir testte H_0 hipotezinin kabul edilebilmesi için $[0-1]$ aralığında değişken değerler alabilen p_{value}' nin anlamlılık seviyesine (a) eşit veya daha büyük olması gerekir. p_{value}' nin anlamlılık seviyesinden düşük olduğu aksi durumlar, H_0' nın reddedilmesi veya alternatif hipotezin H_a' nın kabulü ile ilgili güçlü kanıtlar sunar. Frekans testi için istatistiki önem seviyesi çalışma kapsamında 0.01 olarak seçilmiştir. Testin başarılı kabul edilebilmesi için $p_{value} \geq a$ şartının mutlaka sağlanması gerekir. Bias analizi için test edilen 100 farklı rasgele sayı dizisinin olasılıksal dağılımları ve frekans testi sonuçları Şekil 7' deki gibidir.

$$X_{dif} = y_1 + y_2 + y_3 \dots + y_n \quad \text{ve} \quad y_i = 2x_i - 1 \tag{6}$$

$$X_{obs} = \frac{|X_{dif}|}{\sqrt{n}} \tag{7}$$

$$p_{value} = erfc(z) = erfc\left(\frac{X_{obs}}{\sqrt{2}}\right) = \frac{2}{\pi}\int_{z}^{\infty} e^{-u^2} du \tag{8}$$

Şekil 7 (a)' daki sonuçlar incelendiğinde test edilen sayı dizilerinde, dizi elemanlarının oluşma olasılıklarının $1/2$' lik ideal değere yakın varyasyonlarla oluştuğu görülebilir. Bu sonuçlar bir diğer yönüyle Frekans testinin başarılı olabilmesi için ihtiyaç duyulan ve Şekil 7 (b)' de verilen dizi elemanları arasındaki kümülatif fark değerleri ile de uyumludur. Test edilen her bir dizinin 524.288 bitten oluştuğu düşünüldüğünde, frekans testinde $p_{value} \geq a$ şartının yerine gelebilmesi için dizi elemanları arasındaki kümülatif fark için



Şekil 7. Test edilen sayı dizilerinin (a) bit düzeyinde olasılıksal dağılımları ve (b) bu olasılık değerleri için hesaplanmış kümülatif fark değerleri

sınır değer, $erfc(a)^{-1} = erfc(0.01)^{-1}$ için $|S_{dif}| = 1821$' dir. Kümülatif farkın sınır değerin üzerinde gerçekleştiği aksi durumlarda, $p_{value} \geq a$ şartı sağlanamadığı için frekans testi başarısız kabul edilir. Şekil 7' de verilen sonuçlar incelendiğinde, test edilen sayı dizilerinde kümülatif farkın ve dolayısıyla biasın, kümülatif fark için hesaplanan sınır değerin altında meydana geldiği görülebilir. Dolayısıyla test sonuçları başarılıdır ve aynı sonuçlar için test edilen sayı dizilerinde biasın, rasgelelik dağılımını bozmayacak şekilde kontrollü bir biçimde meydana geldiği söylenebilir.

Otokorelasyon Analizi

İkinci aşamada ise test edilen rasgele sayı dizilerinde ardışık gözlemler arasındaki belirgin bir doğrusal ilişkinin var olup olmadığını tespit etmek için otokorelasyon testi uygulanmıştır. Aynı zamanda biasın varlığına da işaret eden rasgele sayılar arasındaki güçlü bir ilişkinin varlığı, bu sayılar için rasgelelik varsayımının ihlal edildiğini ve tahmin edilebileceği anlamına gelir. Çalışma içerisinde rasgele sayı dizilerinin kendi içerisinde herhangi bir istatistiki bağımlılık içermediğini göstermek için matematiksel tanımı Denklem 9 ve 10' da verilen otokorelasyon testi kullanılmıştır. Test tekniği herhangi $x_0x_1x_2 \ldots x_{n-1}$ rasgele sayı dizisi ile bu dizinin $1 \leq d \leq \lfloor n/2 \rfloor$ aralığındaki farklı d tamsayıları için kaydırılmasıyla elde edilmiş alt dizileri (döngüsel olmayan) arasındaki otokorelasyon katsayısını (X_5) hesaplar. Denklem 9' da $\oplus$ işareti XOR işlemini temsil etmektedir [19].

$$A(d) = \sum_{i=0}^{n-d-1} b_i \oplus b_{i+d} \tag{9}$$

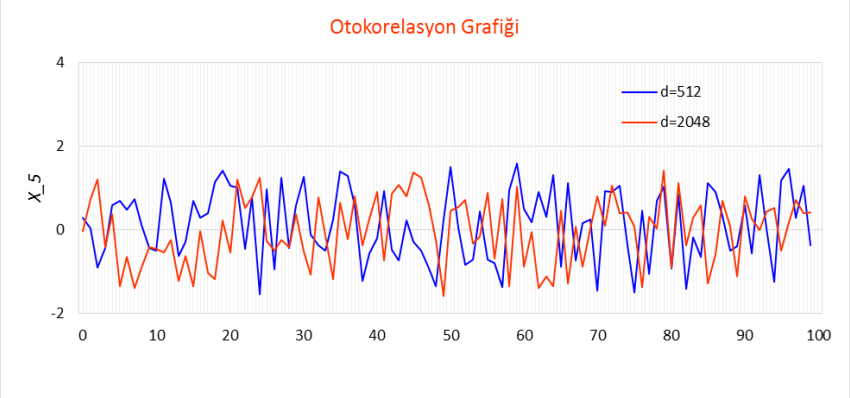
$$X_5 = \frac{2[A(d) - (n-d)/2]}{\sqrt{(n-d)}} \tag{10}$$

Eşitlikte Z_0, ortalaması (μ) 0 ve varyansı (σ^2) 1 olan ve genel karakteristiği standart dağılımla aynı olan rasgele değişkenin olasılık yoğunluk fonksiyonudur. a önem derecesi olmak üzere, $Z_{a/2}$ değerleri Z standart normal dağılım tablosu üzerinde yer almaktadır. Test tekniği uygulanırken H_0 hipotezinin reddedilme olasılığını temsil eden anlamlılık a değerleri sırasıyla 0.1 ve 0.05 olarak seçilmiştir. Her iki a değeri için H_0 hipotezinin reddedilmeme olasılığı diğer bir deyişle testin önem derecesi sırasıyla %90 $(1 - a = 0.9)$ ve %95 $(1 - a = 0.95)$' tir. $a = 0.1$ ve $a = 0.05$ için Z tablosu üzerinde tanımlı test istatistiğini temsil eden eşik değerler $(Z_{a/2})$ sırasıyla 1.649 ve 1.96' dır. Bu değerler, Z tablosunda $Z_{a/2}$' ye karşılık gelen satır ve sütun değerlerinin toplamı alınarak bulunur. Örneğin bazı önem dereceleri için standart Z tablosundan elde edilmiş sınır değerler, Tablo 4' te verilmiştir.

Bir hilesiz bozuk paranın havaya atılmasıyla tura yüzü için '0', yazı için ise '1' üretildiği 100 bitlik bir gerçek rasgele sayı dizisini ele alalım. Gerçek rasgele olayların çoğu zaman

Tablo 4. Farklı a değerleri için Z tablosu üzerinde tanımlı X_5 değerleri

a	0.1 (%90)	0.05 (%95)	0.025 (%97,5)	0.01 (% 99)	0.005 (%99.5)
X_5	1.6449	1.9600	2.24	2.575	2.81



Şekil 8. Test edilen rasgele sayı dizileri için otokorelasyon test sonuçları

düşük seviyeli rasgelelik içermesinin bir sonucu olarak, sayı dizisinin ilk 99 biti '0' olarak meydana gelse bile, sonraki üretilen bit değerinin '1' veya '0' olma ihtimali yine $1/2$ olmalıdır. Bu varsayım, ancak ve ancak herhangi bir andaki paranın havaya atılması olayının, bir önceki veya bir sonraki olayın sonucu üzerinde etkisinin olmamasıyla mümkündür. Yani havaya atılan paranın sonuçlarının bir seri halinde tura gelmiş olması gerçeği, bir sonraki atışta tekrar tura geleceği sonucunu olası kılmamalıdır. Otokorelasyon, olaylar arasında istatistiksel bağımsızlık olarak da anlamlandırılan bu etkinin varlığını ve yönünü matematiksel olarak belirlemeye yarayan bir test yöntemidir [20].

$d = 512$ *ve* $d = 2048$ kaydırma değerleri için rasgele sayı dizilerinin otokorelasyon test sonuçları Şekil 8' de verilmiştir. Test tekniğinde rasgele sayılar için bağımsızlık varsayımının yapılabilmesi için $0 \leq X_5 \leq Z_{a/2}$ şartının yerine gelmesi gerekir. Dolayısıyla her iki önem derecesinde test edilen sayı dizilerinde bağımsızlık varsayımını doğrulamak için korelasyon katsayısı, sırasıyla $-1.6449 \leq X_5 \leq 1.6449$ ve $-1.96 \leq X_5 \leq 1.96$ aralığında olmalıdır.

Şekil 8' de verilen sonuçları incelendiğinde her iki önem derecesi için test sonuçlarının otokorelasyon testi açısından kabul edilebilir sınır değerler içerisinde oluştuğu görülmektedir. Test tekniğinin başarılı kabul edildiği bu durumda dizi elemanları arasındaki bağıntıyı düzgün bir şekilde tanımlayan lineer bir modelin bulunmadığı söylenebilir. Özetle test edilen sayı dizilerinin kendi içerisinde tahmin edilebilirliğe yol açan herhangi bir istatistiki bağımlılık içermediği ve GRSÜ' nin bu sayı dizilerini sırasıyla %90 ve %95 güven aralığında bağımsız olarak ürettiği söylenebilir.

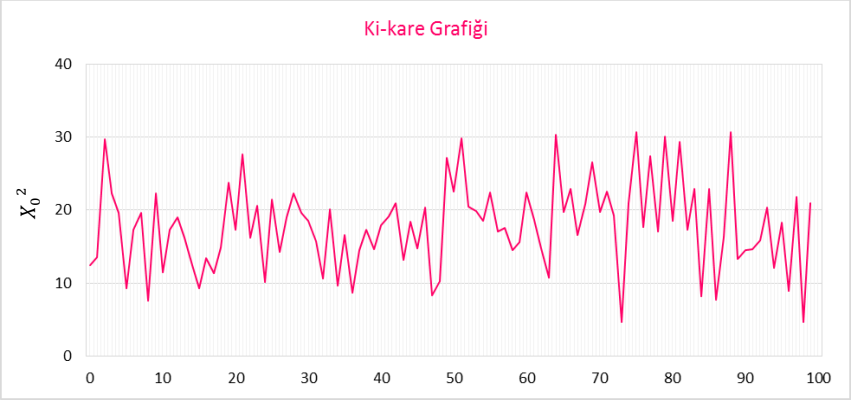
Ki-kare İyi Uyum Analizi

Çalışma içerisinde rasgele sayı dizilerinin ideal istatistiki dağılıma teorik olarak uyup uymadığını tespit edebilmek için ki-kare uyum testi kullanılmıştır. Ki-kare uyum testinin matematiksel tanımı Denklem 11' de verilmiştir. Denklemde k sınıf/grup sayıdır. O_i ve E_i sırasıyla $i.$ sınıf için gözlemlenen beklenen frekans değerleridir. Uyum testi, sayı dizisinde alt sınıflara ayrılmış rasgele değişkenlerin beklenen ve gözlemlenen frekansları arasındaki farkın anlamlı olup olmadığı temeline dayanır. Dolayısıyla boyutu n olan ideal bir rasgele sayı dizisinde her bir grubun bu dizi içerisindeki beklenen frekans dağılımı n/k olmalıdır. Bir dağılım içerisinde her bir sınıf $(1, \dots k)$ için gözlemlenen ve beklenen frekans değerleri arısındaki farkın kareleri toplamı, ki-kare $({X_0}^2)$ test istatistiğidir. Test istatistiğinin dağılımı aynı zamanda $k-1$ serbestlik derecesi ile karakterize edilen ki-kare dağılımıdır. Test sonuçları rasgele sayıların düzgün bir dağılımla oluşup olmadığına ve bu sayıların istatistiki açıdan rasgele olup olmadığına karar verebilmek için yeterli kanıt sunabilmektedir [13], [24].

$$ {X_0}^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} < X^2_{[a,k-1]} \qquad (11) $$

Denklem 11' deki eşitlikten yola çıkarak test istatistiğinin ki-kare dağılımı göstermesi için k ile temsil edilen sınıf sayısı, 5' ten büyük olmalıdır. Bu nedenle test sonuçlarının güvenilirliği için teste tabi bit düzeyindeki rasgele sayı dizileri kendi içerisinde 4 bitlik ardışıl olarak gruplanarak heksadesimal (hexadecimal) seviyeye dönüştürülmüştür. Dolayısıyla Denklem 11' de k grup sayısı 16' dır. Test edilen sayı dizilerinin heksadesimal seviyeye dönüştürülmesiyle beraber her biri 524.288 bitten oluşan sayı dizilerinin boyutu 131.072 (524.288/4)' ye indirgenmiştir. Bu tanımlı aralıktaki sayı dizilerinde heksadesimal olarak temsil edilen her bir sınıfın beklenen frekans değeri 8.192 $(131.072/16)$ dir. $a = 0.01$ ve $a = 0.005$ önem derecesi için test edilen 100 farklı rasgele sayı dizisinin ki-kare salınımları Şekil 9' daki gibidir. Ayrıca $k-1$ serbestlik derecesinde farklı a değerlerine karşılık gelen ki-kare $({X_0}^2)$ dağılımı ve test edilen üç farklı sayı dizisinin frekansları sırasıyla Tablo 5 ve 6' da verilmiştir.

Tablo 5' te $k-1$ serbestik derecesi için a=0.01 ve $a = 0.005$ önem derecelerine karşılık standart ki-kare dağılım tablosundan elde edilmiş $X^2_{[a,k-1]}$ sınır değerleri sırasıyla 30.578 ve 32.801' dir. Test tekniğinin başarımı için rasgele sayı dizileri için hesaplanan ki-kare test istatistiğinin



Şekil 9. Rasgele sayılar için ki-kare testi sonuçları

bu sınır değerlerin altında salınım göstermesi, diğer bir deyişle $X_0^2 < X^2_{[a,k-1]}$ şartının yerine gelmesi gerekir. Şekil 9' da verilen sonuçlar incelendiğinde, test edilen sayı dizileri için hesaplanan ki-kare istatistiğinin sınır değerlerin altında salınım gösterdiği görülmektedir. Aynı sonuçlar test edilen sayı dizilerinde gözlemlenen frekans değerleri ile beklenen frekans değerleri arasında, olasılık dağılımına uyum açısından bir tutarlılığın olduğunu göstermektedir. Bu durumda sayı dizilerinin tanımlı aralık içerisinde düzgün dağılımla oluştuğu ve bu sayıların rasgelelik hipotezi kabul edilir veya en azından reddedilmemiş olur.

Tablo 5. $k-1$ çin farklı a değerleri için X_0^2 değerleri

	Olasılık (a)					
SD	**0.90**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
15	8.574	22.307	24.996	27.488	30.578	32.801

Tablo 6. Üç farklı rasgele seçilmiş sayı dizisinin frekansları

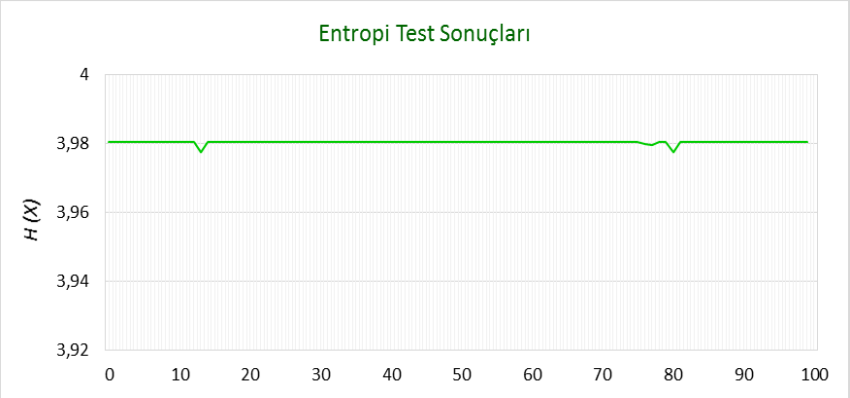
Hex	E_i	**Örnek 1 (6. test dosyası)**	**Örnek 2 (37. test dosyası)**	**Örnek 3 (99. test dosyası)**
		O_i	O_i	O_i
0	8192	8069	8103	8185
1		8193	8101	8189
2		8250	8197	8191
3		8320	8161	8240
4		8138	8220	8143
5		8188	8294	8262
6		8258	8184	8131
7		8238	8153	8134
8		8194	8142	8111
9		8248	8151	8273
A		8087	8313	8210
B		8150	8171	8129
C		8283	8256	8192
D		8152	8239	8218
E		8192	8283	8222
F		8112	8104	8242

Entropi Testi

Entropi ile ilgili farklı kavramsal ve matematiksel tanımlamalar olmasına rağmen en genel tanım Claude Elwood Shannon tarafından yapılmıştır. Matematiksel tanımı Denklem 1' de verilen entropi, bilgi teorisinde bir rasgele değişkenin tahmin edilemezlik ölçüsü olarak kabul edilir. Denklem 1' de $p(x_i)$, $X[n] = x_0x_1x_2 \dots x_{n-1}$ $x \in \{0,1\}$ olmak üzere dizideki x_i. rasgele değişkenin oluşma olasılığını temsil etmektedir. Rasgele sayı dizisi içerisindeki her bir elemanın oluşma olasılığının eşit olduğu durumun teorik sonucu olarak entropi maksimum değerini alır yani 1' e eşit olur. Belirsizliğin olmadığı, çıkışların kesin ve bilinir olduğu durumlar için ise entropi minimumdur yani 0' dır. Rasgele sayıların tahmin edilemezliğini teorik olarak garanti edebilmek için, $X[n]$ şeklindeki n-bitlik rastgele bir vektörün entropisi, mümkün olduğunca n' e yakın olmalıdır [12], [15].

Düzgün dağılımla ve her bir elemanın bağımsız üretildiği bir rasgele sayı dizisinde bir bitin oluşma olasılığı $1/2$' dir. "00", "01", "10", "11" şeklinde iki bitlik kodlanmış bir dizide art arda herhangi iki bitin üretilme olasılığı ise $1/4$' tür. Çünkü bağımsız iki olayın gerçekleşme olasılığı, bu olayların kendi olasılıklarının çarpımına eşittir. Dolayısıyla bağımsız ve düzgün dağılımla oluşmuş n bitlik bir rasgele sayı dizisinin doğru tahmin edilmesi olasılığı, toplamda $\frac{1}{2^n}$' ye eşit veya yakın olmalıdır. Her bir bit için oluşma olasılığı eşit durumların olası sonuçları ve taşıdığı bilgi miktarı da göz önünde bulundurulduğunda sayı dizisinin entropisi de n' yakın veya eşit olduğu görülebilir. Saldırganın her bir bitin olası sonuçlarını tahmini için eşit derecede zorlukla karşılaştığı bu durumda, belirsizlik yani entropi maksimumdur. Bu durumdaki bir sayı dizisinde, bu diziyi oluşturan öncül veya ardıl bitlerin $1/2$' den daha yüksek bir olasılıkla tahmin edilmemesini matematiksel olarak garanti etmiş oluruz. Test edilen rasgele sayı dizileri için entropi testi sonuçları, Şekil 10' daki gibidir.

Herhangi bir rasgele değişkenin belirsizliğini matematiksel olarak ölçebilen entropi, değişkenin/bilginin kodlama yöntemiyle doğrudan ilgilidir. Şekil 10' da verilen test sonuçları için heksadesimal olarak dönüştürülmüş rasgele sayı dizileri kullanılmıştır. Dolayısıyla test edilen heksadesimal seviyedeki sayı dizilerinde maksimum entropi



Şekil 10. Test edilen sayı dizileri için entropi dağılımı

değeri, bu dizileri oluşturan 0-15 arasındaki her bir sayı başına 4 bittir. Entropinin maksimum değere ulaşabilmesi ancak rasgele olduğu varsayılan olay ya da değişkenlerin tanımlı bir aralıkta eşit oluşma olasılığıyla mümkündür. Bu olasılığın; değişkenin belirli çıkışları üzerinde yoğunlaştığı veya eşit oluşmadığı düşük olasılıklı durumlar, tahmin edilebilirlik açısından daha fazla bilgi içerirler. Oysaki, bağımsızlık ve eşit oluşma olasılığı gibi iyi istatistiksel özelliklerine ek olarak rasgele sayı dizilerinin entropisinin yüksek olması istenir. Dolayısıyla bu gibi durumlarda entropi düşüktür. Şekil 10' da verilen test sonuçları incelendiğinde test edilen sayı dizileri için entropinin maksimum değere yakın bir dağılım (spektrum) sergilediği görülmektedir. İdeal değere yakın entropi sonuçları test edilen rasgele dizilerinin düzgün bir dağılımla oluştuğunu ve sayı dizisinde bit başına entropinin yüksek olduğunu doğrulamaktadır. Entropi sonuçları, bu yönüyle Şekil 9' da verilen ki-kare test sonuçlarıyla da uyumludur. İyi istatistiksel özelliklerine ek olarak test edilen rasgele sayılar için entropinin yüksek olması bu sayıların kriptoanaliz senaryolarına karşı dirençli olduğu çıkarımı yapılabilir.

Dieharder İstatistiki Rasgelelik Analizi

Rasgele sayı dizilerinin olasılıksal özelliklerinin tanımlanabilmesine imkan tanıyan istatistiksel analize dayalı farklı test paketleri bulunmaktadır. Bu test paketlerinden biri olan Dieharder [24] testi, diğer test yöntemlerinde olduğu gibi sayı dizilerinin ampirik dağılımlarını analiz ederek, bu dizilerin rasgele olmadığına işaret eden örüntüleri tarar. NIST SP 800-22 [25] test paketinde yer alan test tekniklerinin bir bölümünü içeren Dieharder testi, Tablo 7' de yer alan 31 farklı alt test grubunu içermektedir. RSÜ' lerin geniş bir yelpazede değerlendirilmesine olanak sağlayan bu durum, aynı zamanda test çeşitliliği açısından üretecin birçok yönünü analiz edebilecek kapsamlı bir bakış açısı sunabilmektedir.

Test metodolojisinde her bir kriter için rasgele sayı dizileri için test istatistiği olarak da kabul edilen $p-değer$ parametresi hesaplanır. Bu parametre aynı zamanda, rasgelelik açısından H_0 hipotezinin doğru olduğu varsayımına karşı olan kanıtların nicel bir ölçüsüdür. Farklı dağılım ölçütlerinin dikkate alındığı test kriterleri için spesifik olarak $p-value$ değerleri hesaplanabilmektedir. Ancak genel olarak rasgele sayılar için hesaplanan test istatistiğinin rasgelelik varsayımı altındaki beklenen dağılımla karşılaştırılması esas alınır. Test edilen sayı dizilerinin Dieharder testlerini geçebilmesi için p-değer parametresinin $[0+a/2, 1-a/2]$ aralığında olması gerekir. Diğer bir deyişle $a = 0.01$ önem derecesi için her bir test kriterinde $0.005 \leq p-değer \leq 0.995$ şartı sağlanmalıdır. Test edilen örnek bir rasgele sayı dizisi için Dieharder test sonuçları Tablo 7' deki gibidir. Çalışma kapsamında test edilen sayı dizilerinin tamamı, Dieharder testlerini başarıyla geçmiştir.

Tablo 7. Rasgele sayılar için Dieharder test sonuçları

	Test Adı	$p-değeri$	Sonuç
1	diehard birthdays test	0.73858321	Geçti
2	diehard operm5 test	0.12481644	Geçti
3	diehard 32x32 binary rank test	0.01401876	Geçti
4	diehard 6x8 binary rank test	0.25343627	Geçti
5	diehard bit stream test	0.64233008	Geçti
6	diehard opso test	0.81128183	Geçti
7	diehard oqso test	0.42451618	Geçti
8	diehard dna test	0.72141254	Geçti
9	diehard count the 1s (stream)	0.97591478	Geçti
10	diehard count the 1s test (byte)	0.92267082	Geçti
11	diehard parking lot test	0.96749733	Geçti
12	diehard min. distance (2d circle) test	0.97488440	Geçti
13	diehard 3d sphere (min. distance)	0.62540617	Geçti
14	diehard squeeze test	0.92231336	Geçti
15	diehard sums test	0.35664329	Geçti
16	diehard runs test	0.23238026	Geçti
17	diehard craps test	0.38974780	Geçti
18	marsaglia and tsang gcd test	0.49064950	Geçti
19	sts monobit test	0.77660845	Geçti
20	sts runs test	0.91894503	Geçti
21	sts serial test (generalized)	0.98904112	Geçti
22	rgb bit distribution test	0.83630253	Geçti
23	rgb generalized min. distance test	0.70698375	Geçti
24	rgb permutations test	0.81118202	Geçti
25	rgb lagged sum test	0.94342790	Geçti
26	rgb kolmogorov-smirnov test	0.32679982	Geçti
27	Dab byte distribution test	0.38705684	Geçti
28	dab discrete cosine transform test	0.83667580	Geçti
29	dab fill tree test	0.43147552	Geçti
30	dab fill tree 2 test	0.28552909	Geçti
31	dab fill monobit 2 test	0.87813217	Geçti

Çalışmanın bu bölümünde kriptografik rasgeleliğin değerlendirilmesi amacıyla ele alınan bağımsız testler ile bu testlere ait sonuçların, literatürde yaygın kabul görmüş test gruplarının sonuçlarıyla uyumluluğu analiz edilmiştir. Test istatistikleri için farklı dağılım ve metrik değerlerin dikkate alındığı bağımsız testlerin sonuçları, Tablo 7' deki Dieharder test paketine ait sonuçlar ile uyumlu ve anlamlı benzerlikler göstermektedir. Bu tutarlı durumun bir sonucu olarak çalışma içerisinde tercih edilen bağımsız test teknikleriyle pratikte güvenilir ve geçerli sonuçların elde edilebilirliğini doğrulamaktadır. Standart test paketlerine ek olarak bağımsız test stratejilerinin özel senaryoların test edilebilmesi, rasgeleliğin anlaşılabilirliği, uygulama bağımsızlığı ve özelleştirilebilirlik gibi özellikleri rasgelelik açısından daha geniş kapsamlı bir değerlendirme yapılmasına olanak sağlar.

Standart Sapma ve Gerçek Rasgelelik Analizi

Herhangi bir GRSÜ' nin temel güvenlik değerlendirmesi, R1 ve R2 gereksinimleri ile karakterize edilmiştir. R1 gereksinimi kapsamında; sınırsız hesaplama kaynağına sahip olsa bile saldırganların rasgele sayıları tahmin etmek için başvurabileceği en iyi yöntem kaba kuvvet olmalıdır. Rasgele sayıların tahmini için tüm olası kombinasyonların denenmesini zorunlu kılan bu saldırıları etkisiz kılabilmek için GRSÜ' nin iyi istatistiksel özelliklere sahip olması şarttır. Fakat GRSÜ' lerin zayıf istatistiki rasgelelik özellikleri, kriptografik amaçlar için kullanımını sınırlayan önemli bir faktördür. Bu durumun üstesinden gelmek ve GRSÜ' leri kriptografik açıdan güçlü kılmak için son işlem teknikleri kullanılmaktadır. Bu kapsamda çalışma içerisinde rasgelelik analizi için tercih edilen ve detayları Bölüm 2.1' de verilen GRSÜ mimarisinde son işlem aşamasında sürekli ve ayrık zamanlı kaos tabanlı s-box' lar kullanılmıştır. Zira bu bölümde sunulan bağımsız test teknikleri kullanılarak elde edilen istatistiki analiz sonuçları GRSÜ mimarisinin R1 gereksinimini başarıyla yerine getirdiğini doğrulamaktadır.

İstatistiki test tekniklerinin (entropi testi de dahil olmak üzere) rasgele değişkenlerin gözleme dayalı sonuçlarına uygulanması, güvenirlik açısından yanıltıcı sonuçların elde edilmesine neden olabilir. Zira bu yöntemlerle test edilecek sayı dizilerinin basit yazılımsal yöntemlerle üretilmesi de mümkündür. Dolayısıyla istatistiki testler rasgele sayıların elde edilme yöntemleri, diğer bir deyişle kaynağın güvenirliği konusunda geçerli kanıtlar sunmazlar. İstatistiki test sonuçlarından yola çıkarak bu sayı dizilerinin kriptografik açıdan güvenilir bir kaynaktan elde edilip edilmediği hakkında yorum yapmak mümkün değildir. Oysaki kriptografik uygulamaların temel güvenlik varsayımı, kullanılan rasgele sayıların istatistiki rasgelelik yeterliliklerinin yanı sıra bu sayıların fiziksel bir gerçekliğe bağlı olarak üretilmesi ile de güçlü bir şekilde ilişkilidir [7],[9]-[10].

Fiziksel gerçekliğe dayalı bir kaynaktan beslenen GRSÜ' ler, tahmin edilemezlik açısından istatistiki rasgelelik testlerinden daha güçlü özelliklere sahip olabilirler. Öyle ki, bu üreteçlerin tek yönlü bir fonksiyon gibi çalışmasını sağlayan gürültü kaynaklarının stokastik davranışı, tahmin edilemezlik açısından R2 ve bu gereksinimle bağlantılı R3 ve R4 gereksinimlerini tek başına yerine getirebilmektedir. Gürültü kaynaklarının non-deterministiklik sağlayan bu özelliği, GRSÜ' lerde istatistiki rasgeleliğin, yani R1 gereksiniminin önemli bir tamamlayıcı özelliğidir. Dolayısıyla bu üreteçlerde gürültü kaynağının non-deterministik davranışının matematiksel araçlarla modellenmesi, üretecin ve rasgele sayıların güvenirliği açısından önemlidir. Çalışmanın bu bölümünde kullanılan GRSÜ' nin ve rasgele sayıların R2 gereksinimiyle bağlantılı tahmin edilemezlik özelliklerini, entropi varsayımı açısından doğrulamak için standart sapma analizi kullanılmıştır.

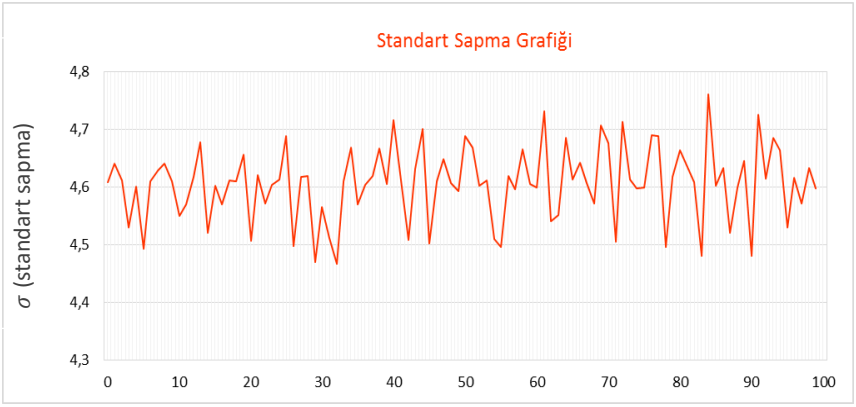
Standart sapma analizi için ihtiyaç duyulan sayı dizileri, [13]' teki GRSÜ mimarisinin aynı başlangıç şartlarında 100 defa yeniden başlatılmasıyla elde edilmiştir. Heksadesimal seviyeye dönüştürülen bu bölümdeki sayı dizilerinin tamamı, doğrudan gürültü kaynağından elde edilmiş saf gerçek rasgele (son işlem uygulanmamış) çıkışlardan oluşur. Bu sayı dizilerinin standart sapma (σ) değerleri Denklem 12' deki gibi hesaplanmıştır. $p(x_i)$, saf rasgele sayı dizisi içerisinde x_i. elamanın oluşma olasılığı olmak üzere varyans (Var) ve rasgele değişkenlerin ağırlıklı toplamlarını temsil eden $E(X)$ değerleri ise sırasıyla Denklem 13 ve 14' teki gibi hesaplanmıştır. Test edilen sayı dizileri için hesaplanan standart sapma değerlerinin değişimi Şekil 11' deki gibidir.

$$\sigma = \sqrt{Var\ (X)} \tag{12}$$

$$Var\ (X) = \sum_{i=1}^{k} [x_i - E(X)]^2 . p(x_i) \tag{13}$$

$$\mu = E(X) = \sum_{i=1}^{k} x_i . p(x_i) \tag{14}$$

Tohum gibi başlangıç girdilerine ihtiyaç duyan deterministik RSÜ' ler aynı zamanda tipik matematiksel bir bağıntıdır. Deterministikliğin doğası gereği bu üreteçler başlangıç şartları farklı olsa da aynı başlangıç girdileri için çıkışında hep aynı sayı dizilerini üretirler. Bu durumda, Denklem 12-14' teki bağıntıya göre test edilecek sayı dizilerinin hepsi aynı standart sapma değeriyle oluşması kaçınılmazdır. Fakat her bir rasgele sayı dizisinin farklı bir standart sapma değeriyle oluştuğunu gösteren Şekil 11' deki sonuçlar, test edilen rasgele sayı dizilerinin kriptografik açıdan güvenilir bir kaynaktan elde edildiğini doğrular niteliktedir. Özellikle GRSÜ' nin tek yönlü bir fonksiyon gibi davranmasını sağlayan gürültü kaynağının non-deterministik (gerçek rasgele) davranışı, test edilen sayı dizilerinin farklı standart sapma değerleriyle oluşması noktasında belirleyici bir etkiye sahiptir. Kaynağın bu davranışı, aynı zamanda test edilen rasgele sayı dizileri için R2 ile bağlantılı tekrar üretilemezlik ve tahmin edilemezlik gibi önemli bir karakteristik gereksinimin yerine getirdiğini göstermektedir.



Şekil 11. Test edilen saf rasgele sayı dizilerinin standart sapma dağılımı. Test aşamasında sayı dizilerinin fiziksel bir gerçeklikten elde edildiğini doğrulamak için doğrudan gürültü kaynağından örneklenmiş saf (son işlem uygulanmamış) halleri kullanılmıştır.

Sonuç Ve Öneriler

Rasgele sayı dizileri, katı güvenlik gereksinimlerine ihtiyaç duyan günümüz modern kriptografik protokollerinin önemli girdileridir. Dolayısıyla bu sayıların güvenlikle bağlantılı temel gereksinimleri göz ardı edilemez. Bu gereksinimler doğrultusunda rasgele sayılar iyi istatistiksel yeterliliklerinin yanı sıra, tahmin edilemez ve tekrar üretilemez olması istenir. Bu nedenle kriptografik rasgelelik kavramının iyi anlaşılması ve nicel istatistiksel yöntem ve araçlarla doğru analiz edilmelidir. Çünkü bu protokollerin temel güvenlik varsayımı kullanılan rasgele sayıların istatistiki kalitesi ve tahmin edilemezlik özellikleri ile doğrudan bağlantılıdır.

En genel tanımıyla kriptografide gerçek rasgelelik, güvenilir şekilde yeniden üretilemeyen, birbirinden bağımsız, düzgün dağılımla oluşmuş ve tahmin edilemez değerler üreten olasılıklı bir sürecin sonucunu ifade eden önemli bir metrik kavramdır. Bu kavramın/sürecin doğru anlaşılması ve yorumlanması noktasında olasılık teorisi ve istatistik, güçlü ve kullanışlı bir dizi araç ve yöntem sunmaktadır. Literatürde rasgelelik değerlendirilmesiyle ilgili güçlü argümanlar sunabilen hipotez tabanlı test yöntemleri bulunmaktadır. Tek başına yeterli olmamakla birlikte bu testler RSÜ' lerin başlangıç değerlendirmesinde önemli bir yere sahiptir. RSÜ' ler açısından tahmin edilemezlik önemli bir güvenlik gereksinimidir. Özellikle GRSÜ' ler için literatürde yaygın kullanılan test paketleri, bu temel gereksinim ile alakalı kapsamlı bir değerlendirme yapmazlar. Dolayısıyla testlerin yanı sıra bu üreteçlerde entropi analiziyle birlikte, fiziksel gerçekliğe dayalı stokastik analizin yapılması daha geçerli sonuçların elde edilmesini sağlayacaktır.

Çalışma kapsamında kriptografik RSÜ' ler için rasgelelik kavramı ve bu kavramla bağlantılı istatistiki gereksinimlerin doğrulanması için bazı bağımsız test yöntemlerine yer verilmiştir. Bu testler aynı zamanda, FPGA tabanlı bir GRSÜ' den gerçek zamanlı olarak elde edilmiş sayı dizilerine uygulanmıştır. Rasgeleliğin derinlemesine ve çok yönlü bir bakış açısıyla analizi için bu testlerin uygulanma biçimleri ve görece dağılımları, doğrudan kriptografik gereksinimler özelinde incelenmiştir. Klasik paradigmayı aşmayı sağlayan bağımsız yöntemler, bilindik test paketleriyle benzer derecede geçerli sonuçlar sağlayarak, rasgelelik analizinde daha geniş bir bakış açısı ve kapsamlı bir değerlendirme imkanı sağlamaktadır. Çalışma bu yönüyle, rasgelelik kavramının daha derinlemesine anlaşılmasına katkı sağlarken, güvenlik protokollerinin ve kriptografik RSÜ' lerin tasarımında sağlam bir temel oluşturabilecek farklı bir metodoloji önermektedir. Çalışma bir diğer yönüyle, kriptografik düzlemde yapılacak akademik çalışmalara ve uygulayıcılara fikir ve uygulama bazında kaynaklık edeceğine inanmaktayız. Gelecek çalışmalarımızda sunulan test yöntemlerinin farklı alternatif test yöntemlerini de kapsayacak genişletilmesini ve bir web platformu üzerinden kullanıcılara sunulduğu gelişmiş bir test platformunun hayata geçirilmesini amaçlamaktayız.

KAYNAKLAR

[1]. C. K. Koc, Cryptographic Engineering, Springer, Signals and Communication Theory, Berlin. DOI: 10.1007/978-0-387-71817-0_2, 2009.

[2]. A. M. Garipcan, E. Erdem, "Hardware implementation of chaotic zigzag map based bitwise dynamical PRNG on FPGA", *Informacije MIDEM*, vol. 50(4), pp. 243-254, 2020

[3]. I. Cicek, A. E. Pusane, G. Dündür, "Random number generation using field programmable analog array implementation of logistic map", *In 2013 21st Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE, 2013

[4]. G.C. Bilginer, "Kriptoloji", *Tubitak Bilim ve Teknik Dergisi*, cilt 55(658), ss. 16-3, 2022.

[5]. F. Özkaynak, "Kriptolojik Rasgele Sayı Üreteçleri", *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, cilt 8(2), ss. 37-45, 2015

[6]. K. Marton, A. Suciu, I. Ignat, "Randomness in digital cryptography: A survey", *Rom. J. Inf. Sci. Technol*, vol. 13(3), pp. 219-240. 2010

[7]. V. Fischer, M. Deutschmann, S. Lattacher, …., G. Battum, "Report on Selected TRNG and PUF Principles", *HECTOR Project Technichal Report D2.1, UJM* (Université Jean Monnet), 2016

[8]. A. J. Acosta, T. Addabbo, E. Tena-Sánchez, "Embedded electronic circuits for cryptography, hardware security and true random number generation: an overview", *International Journal of Circuit Theory and Applications*, vol. 45(2), pp. 145-169, 2017

[9]. W. Schindler, W. Killmann, "Evaluation criteria for true (physical) random number generators used in cryptographic applications" *In International Workshop on Cryptographic Hardware and Embedded Systems* (pp. 431-449). Springer, Berlin, Heidelberg, 2022

[10]. K. Wold, Security properties of a class of true random number generators in programmable logic. PhD Thesis, Faculty of Computer Science and Media Technology, Gjøvik University College, 2011

[11]. E. Avaroğlu, T. Tuncer, A. B. Özer, B. Ergen, M. Türk, "A novel chaos-based post-processing for TRNG" *Nonlinear Dynamics*, vol. 81(1), pp. 189-199, 2015

[12]. H. Demirhan, N. Bitirim, "Statistical testing of cryptographic randomness" *İstatistikçiler Dergisi: İstatistik ve Aktüerya*, cilt 9(1), ss. 1-11, 2016

[13]. A. M. Garipcan, E. Erdem, "A gigabit TRNG with novel lightweight post-processing method for cryptographic applications" *The European Physical Journal Plus*, vol.137(4), pp. 1-26, 2022

[14]. I. Vattulainen, "New tests of random numbers for simulations in physical systems" arXiv preprint cond-mat/9411062, 1994

[15]. O. Bahadır, H. Türkmençalıkoğlu, "Bilgi Kuramında Shannon Entropisi ve Uygulamaları", *Avrupa Bilim ve Teknoloji Dergisi*, cilt (32), ss. 491-497, 2021

[16]. O. Kocak, F. Sulak, A. Doğanaksoy, U.G.U.Z Muhiddin, "Modifications of Knuth randomness tests for integer and binary sequences" *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, vol. 67(2), pp. 64-81, 2018

[17]. M. Bar-Hillel, W. A. Wagenaar, "The perception of randomness", *Advances in applied mathematics*, vol. 12(4), pp. 428-454, 1991

[18]. A. M. Garipcan, E. Erdem, "Design, FPGA implementation and statistical analysis of a high-speed and low-area TRNG based on an AES s-box post-processing technique" *ISA transactions*, vol. 117, pp. 160-171, 2021

[19]. A. J. Menezes, P.C. Van Oorschot, , S. A. Vanstone, "Handbook of applied cryptography", CRC press, 2018

[20]. C. Wheelan, "Çıplak İstatistik", Buzdağı Yayınevi, Eryaman-Ankara, 2022

[21]. L. F. Rojas-Muñoz, S. Sánchez-Solano, M. C. Martínez-Rodríguez, P. Brox, "True Random Number Generation Capability of a Ring Oscillator PUF for Reconfigurable Devices" Electronics, vol. 11(23), 4028, 2022

[22]. F. Özkaynak, "Construction of robust substitution boxes based on chaotic systems" Neural Computing and Applications, vol. 31(8), pp. 3317-3326, 2019.

[23]. F. Ozkaynak, F. "A novel random number generator based on fractional order chaotic Chua system", *Elektronika ir Elektrotechnika*, vol. 26(1), pp. 52-57, 2020

[24]. R. G. Brown, D. Eddelbuettel, D. Bauer, "Dieharder" Duke University Physics Department Durham, NC, 27708-0305, 2018

[25]. L. E. Bassham, A. L. Rukhin, J. Soto, J. R. Nechvatal, M. E. Smid, E. B. Barker, ... S. Vo, "Sp 800-22 rev. 1a. a statistical test suite for random and pseudorandom number generators for cryptographic applications" National Institute of Standards and Technology, 2010

Etik kurul onayı ve çıkar çatışması beyanı

Hazırlanan makalede etik kurul izni alınmasına gerek yoktur. Ayrıca hazırlanan makalede herhangi bir kişi/kurum ile çıkar çatışması bulunmamaktadır.

Yazar Katkıları

Yazarlar makaleye eşit derecede katkıda bulunmuştur.

Teşekkür

Yazarlar, değerlendirme sürecinde bilgi birikimleri ve yapıcı geri bildirimleriyle makalenin bilimsel kalitesinin iyileştirilmesine, katkıda bulunan anonim değerlendiricilere ve editörlere içtenlikle teşekkür eder.