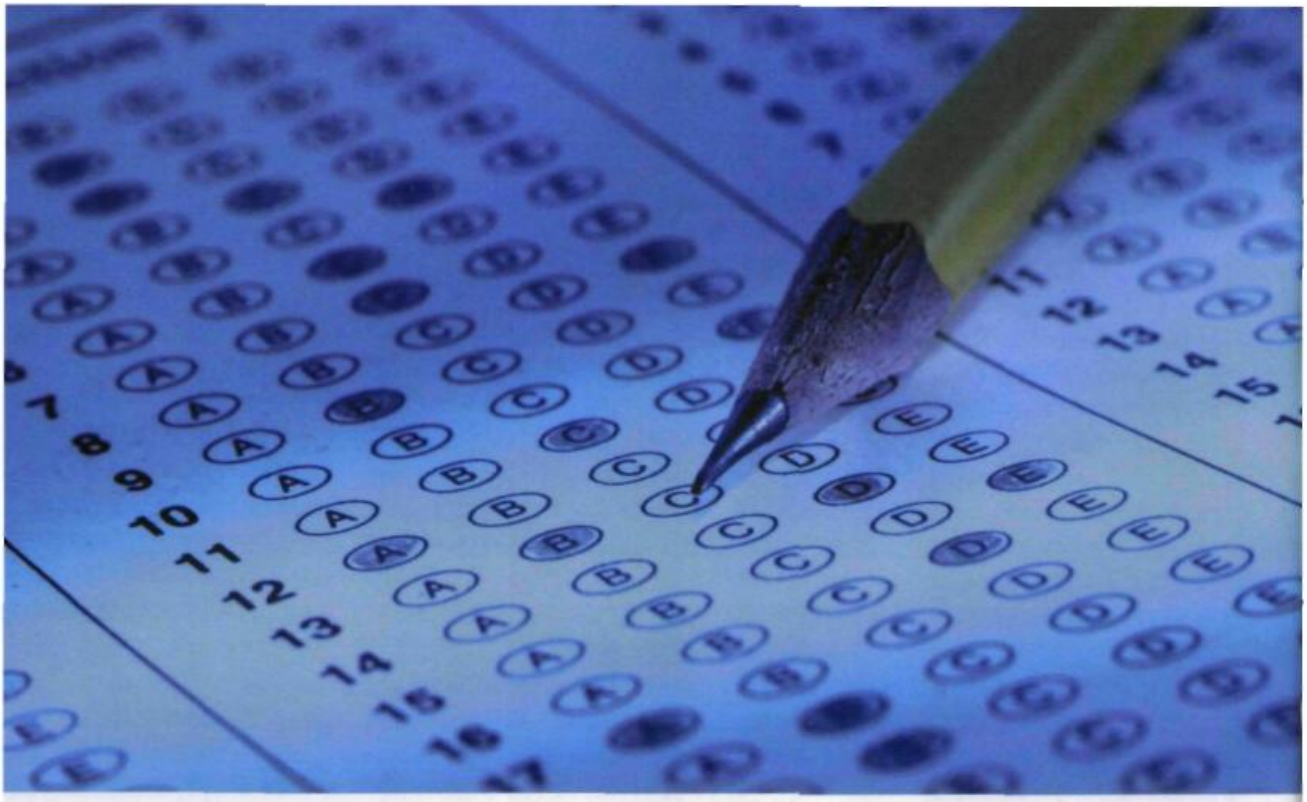

YGS ŞİFRESİ

Beş değişik seçenek kombinasyonunu çembersel kaydırarak kişiye özel sınav kitapçığı üretmek isteyen ÖSYM şıkları yeterince rasgele karmayınca olanlar oldu. Peki, bu sorun önlenebilir miydi? ÖSYM'nin temel hatası neydi? Soruyu kaydırmadan cevaplamak için biraz kriptoloji bilmek şart.

Albert Levi

Hollandalı şifre ve dilbilimci Auguste Kerckhoffs 1883'te yazdığı bir makalede "Sisteminizin güvenliğini kolayca değişemeyen hiçbir şeye bağlamayın. Şifreleme yönteminiz düşmanın eline geçebilir. Güvenliğinizi kolayca değiştirilebilen ve herhangi bir yerde yazılı olmayan anahtarla sağlayın" demişti. Literatüre Kerckhoffs Kuralları olarak giren bir dizi tavsiyeden alıntı olan bu altın öğüt, yüz yılı aşkın süredir şifrebilimcilerin aklının bir köşesinde sürekli durur, ama buna her zaman uyulamayabilir. II. Dünya Savaşı'nın müttefik kuvvetlerin zaferiyle sonuçlanmasındaki en büyük etkenlerden biri olan Alman Enigma şifrelerinin Bletchley Park'ta Turing, Gordon Welchman ve ekibi tarafından kırılmasında casuslar tarafından ele geçirilen kod kitapçıklarının etkisi büyüktü. Enigma ve türevi olan elektromekanik cihazlar, yapılarından dolayı ancak iç kablolamaları değiştirilerek anahtarları değiştirilebilen türdendi. Yani anahtar değiştirmek o kadar da kolay değildi.

1970'lerden itibaren yarıiletken ve bilgisayar teknolojisindeki ilerlemeye bağlı olarak ezberlenebilecek kadar kısa anahtarların kullanılabildiği, anahtar değişiminin yönteme müdahaleyi gerektirmediği standart şifreleme yöntemleri ortaya çıktı. Anahtar boyunun kısa olması nedeniyle artık güvenli kabul edilmeyen DES (Data Encryption Standard-Veri Şifreleme Standardı) ile başlayan bu eğilim günümüzde ağabeyi AES (Advanced Encryption Standard-İleri Şifreleme Standardı) ile devam ediyor. Hem donanım hem de yazılım ortamında gerçeklenebilen bu şifreleme yöntemleri aslında gündelik hayatımızın da içinde. Kredi kartıyla alışveriş yaparken, güvenli bir web sitesinde gezinirken, hatta cep telefonuyla konuşurken bile farkında olmadan şifreleme yapıyoruz veya karşı taraftan gelen şifreyi çözüyoruz.

Şifre terminolojisi

Peki, şifreleme ne demek? Hem sözlüklere hem de bilimsel literatüre göre şifre sözcüğü iki anlama geliyor. Bunlardan

biri gizli haberleşmeye yarayan işaret, kod; diğeri kontrollü erişime tabi olan fiziksel ve sanal kapıları açmaya yarayan gizli bilgi. İkinci anlam -İngilizce deyimiyle "password"- gündelik yaşamın tam ortasında. E-postalarımızı okurken, Facebook'a girerken kullandığımız şifre veya şifreler bu kategoride. İlk anlam, belki de ikinci anlamla karıştırmamak için, bilimsel literatürde daha çok "kripto" olarak geçiyor. Bu işin bilimi kriptografi olarak adlandırılıyor. Kriptografi, veriyi düzgün bir şekilde gizleme ve sadece yetkili kişilerin açmasını sağlayan tarafına odaklanıyor. Anahtara sahip olmadan şifreli veriyi açma, daha yaygın deyimiyle kırma çabalarına kriptanaliz deniyor. Kriptolojiyse kriptografi ve kriptanalizin ortak adı.

Kriptografi ile kriptanaliz birbirlerinden ayrılamayan düşman kardeşler gibidir. Her yeni kriptografik sistem önce dünya çapında kriptanalistlerin çalışma konusu olur. Kırılabilenler tarihin tozlu raflarında yerini alırken, testlerden başarıyla geçenler güvenlik uygulamalarında kullanılır. AES'in standartlaşması da bu süreçlerden sonra gerçekleşti.

Amerikan Standart ve Teknoloji Enstitüsü'nün (NIST–National Institute of Standards and Technology) 1997'de açtığı blok şifreleme algoritması yarışmasına yirminin üzerinde aday algoritma katıldı. Birkaç tur eleme sonunda 2000'de Rijndael algoritması yarışmayı kazandı. Bu algoritma bir yıl sonra AES olarak standartlaştırıldı. Eleme ve seçme sürecinde tüm dünyadaki kriptanalistler aday algoritmaları bilinen tüm kriptanaliz yöntemleriyle kırmaya çalıştı; bu çalışmalar değişik ortamlarda yayınlandı. Elbette tek kriter güvenlik değildi ama NIST güvenlik değerlendirmesini büyük ölçüde açık literatürde yer alan bu çalışmalara dayanarak yaptı.

Julian Assange, Wikileaks belgelerini 256 bit AES algoritmasıyla şifrelemişti.

Şifreleme ve rastsallık

Akla gelen ilk çözüm, tüm anahtarları deneme yoluyla kriptanaliz yapmak olsa da anahtar boylarının uzunluğu nedeniyle pratikte bu pek mümkün değil. AES örneğiyle devam edelim, AES'in üç standart anahtar uzunluğu var: 128 bit, 192 bit ve 256 bit. En kısa anahtar boyu olan 128 bit kullanılsa bile olası anahtar sayısı 38 basamaklı. Bu kadar çok anahtarı günümüz bilgisayar teknolojisini kullanarak makul bir sürede deneme-yanılma yoluyla test etmek imkânsız. Çünkü saniyede bir trilyon deneme bile yapsak bu tür bir analiz ortalama 54 katrilyon yüzyıl alır. Bu nedenle kriptanalistler daha karmaşık istatistiksel yöntemlerle şifreli metin, düz metin ve anahtar arasındaki ilişkileri anlamaya ve yorumlamaya çalışır.

Bir şifreleme algoritmasının sahip olması gereken olmazsa olmaz özelliklerden biri de açık metin ile şifreli metin arasında anlamlı ilişkilerin kurulamamasıdır. Kriptanalist ele geçirdiği birkaç açık ve şifreli metin ikilisinde bazı bağıntılar keşfederse bunu şifreli diğer metinlere de uygulayabilir; kısmen veya tamamen şifreli metinleri anahtara sahip olmadan çözebilir. Bu nedenle, bir şifreleme algoritması tasarlanırken açık metin ile şifreli metin arasındaki ilişkilerin doğrusallığı bozulmaya ve anahtara bağımlı hale getirilmeye çalışılır. Böylece anahtara sahip olmayan biri, açık ve şifreli metin arasında genellenebilecek bağıntılar bulamaz. Başka bir deyişle açık metinle şifreli metin arasındaki ilişki mümkün olduğunca rastsallaştırılır.

Konu rastsallaştırmadan açılmışken şunu da belirtelim. Rastsallaştırma, kriptografiden bağımsız olarak, gündelik yaşamda da karşımıza çıkar. Örneğin bir iskambil destesinin karılmasındaki amaç, kart sırasını rastsallaştırmaktır. Hatta bu amaçla geliştirilen profesyonel makineler, bilgisayar programları var. Burada dikkat edilmesi gereken en önemli nokta, rastsallıkta tekrarın önlenmesidir. Aynı rastsal dizi birkaç destede bir tekrarlarsa bu durum amaçla ters düşer. Yani rastsallaştırma aynı zamanda farklılaştırmayı da beraberinde getirmeli. Aynı malzemeyi bile

SİMETRİK BLOK ŞİFRELEME MODELİ

DES, AES gibi algoritmaların da bulunduğu simetrik blok şifreleme modelinde alıcı ve gönderen tarafın paylaştığı bir anahtar ile işlem yapılır. Gönderen şifreleme modülünü ve anahtarı kullanarak şifreli mesajı oluşturur; alıcı şifre çözme modülü ve anahtarı kullanarak açık mesajı elde eder.

Çizimler: Benbey

SSPL / Getty Images Turkey

Şifre kırıcılar Bletchley Park'ta çalışıyor, Ekim 1942.

kullansanız rastsal dizilişler birbirlerinden farklı olmalı.

YGS'de ne oldu?

Önce durum tespiti yapalım. 2011 YGS öncesindeki durum (yani iptal edilen KPSS sürecinde yaşananlar) doğru cevapların dışarıya sızdırılabileceğini, sızdırılmasa bile sınav sırasında dışarıdan içeriye cevapların iletilebileceği olasılığını ortaya koydu. ÖSYM yönetimi önlem olarak sınav kitapçığı ve dolayısıyla cevap anahtarı tipi sayısını arttırma yolunu seçti. Bu farklı setler, soruların yerlerinin yanısıra cevap şıklarının da yerleri değiştirilerek oluşturuldu.

Şifre iddialarına yol açan durumun adını net bir şekilde koyabilmek için ÖSYM'nin internet sitesinden de dağıttığı basına verilen soru kitapçığı ve cevap anahtarını, master soru kitapçığını ve adaylara sınav sırasında verilen soru kitapçıklarından bir seti inceledim. Matematik ve fen testi sorularını analiz ederek soru kitapçıklarındaki şıkların nasıl bir algoritmayla oluşturulduğunu anlamaya çalıştım. Yani bir tür tersine mühendislik çalışması yaptım. Çok fazla teknik terim kullanmadan kaba hatlarıyla özetlemeye çalıştığım yöntem, olası yöntemlerden sadece biri. Tersine mühendisliğin doğası gereği kullanılan esas yöntem ve algoritma detaylarda biraz farklı olabilir ama varılan sonuç aynı.

Beş şıklı bir soruda şıkların yerleriyle oynanarak 120 değişik kombinasyon elde edilebilir. Ancak ÖSYM her bir soru için, her biri doğru cevabın değişik bir şıkta olduğu beş değişik kombinasyon kullanmış. Üstelik bu kombinasyonlar birbirleriyle bağıntılı. Her bir soru için önce master kitapçıktaki şık sırası karılarak bir rastsallaştırılmış master oluşturulmuş. Daha sonra da her bir kitapçıkta söz konusu sorunun cevabı ilgili cevap anahtarına göre hangi şık olması gerekiyorsa rastsallaştırılmış masterdeki şıklar kaydırılarak doğru cevabın istenen şıkka gelmesi sağlanmış. Bu kaydırma sırasında da sondaki şık başa geçmiş, yani çembersel bir kaydırma yapılmış.

Bu yöntem, her ne kadar bir soru için 120 değişik kombinasyondan birbiriyle bağlantılı beşi kullanılmış olsa da, soru kitapçıklarındaki şıklara bakarak doğru cevabı bulmaya yarayan bir kodlama veya şifreleme barındırmaz. Ancak bunun ön koşulu, rastsallaştırılmış master için oluşturulan karışık şıkların her bir soru için birbirinden bağımsız ve şık sırasıyla doğru cevap arasında bir doğrusallık yaratmayacak şekilde oluşturulması gerekliliğidir. Başka bir deyişle, örneğin masterde doğru cevabın (c) şıkkında olduğu tüm sorularda şıkların kendi içindeki dizilişlerinin soru bazında farklılık göstermesi mümkün olmalı. Ayrıca doğru cevabın diğer şıklarda olduğu sorulardaki dizilişler de kullanılarak diziliş ile doğru cevap arasındaki doğrusal ilişki bozulmalı. Çünkü aksi halde masterdeki sıralamayı bilen veya tahmin edebilen biri, diziliş kurallarını da biliyorsa herhangi bir soru kitapçığındaki soruda doğru cevabın hangisi olduğunu anlayabilir.

ÖSYM'nin temel algoritmik hatası da bu noktada olmuş. Matematik ve fen sorularını incelendiğimde cevap şıklarının rastsallaştırılmasında kullanılan yöntemin sorunun master soru kitapçığındaki doğru cevabının hangi şıkta olduğuyla yoğun biçimde ilgili olduğunu saptadım. Yani herhangi bir soru kitapçığında yer alan bir sorunun şıklarının masterdeki aynı sorunun şıklarına göre nasıl dizildiği saptanırsa, yüzde 5 oranındaki birkaç istisna soru haricinde, doğru cevabın hangi şıkta olduğu saptanabilir. Tabii bunun için master soru kitapçığındaki şık dizilişlerinin bilinmesi de gerekir ki, sıralanabilir şıkların çoğunda bu sıra küçükten büyüğe gidiyor. Örneğin matematik testindeki 40 sorunun 37'sinde şıklar sıralanabiliyor ve bunların 30'unda sıra küçükten büyüğe.

Kriptografik anlamda şifrelemenin en temel şartlarından birinin açık metinle şifreli metin arasındaki doğrusal ilişkinin bozularak rastsallaştırılması olduğunu söylemiştik. YGS durumundaysa bir sorunun masterdeki doğru şıkkıyla soru kitapçıklarındaki şık dizilişi arasındaki doğrusal ilişki yüzde 95 oranında saklı kalmış. Buradan da ÖSYM'nin amacının doğru cevabın hangi şıkta olduğunu öğrencilerden gizlemeyi hedefleyen güçlü bir şifre tasarımı yapmak olmadığını çıkarmak mümkün. Aksine bu durum, masterdeki şıkların nasıl bir transformasyonla

karıldığını ve masterdeki şık sırasını bilen veya tahmin eden birinin doğru cevabın hangi şıkta olduğunu anlayabileceği kötü bir şifrelemeye (daha doğru bir teknik ifadeyle kodlamaya) dönüşmüş.

Bu durum nasıl önlenirdi?

Spesifik bir soru için 120 cevap şıkkı kombinasyonunun hepsinin değişik soru kitapçıklarında kullanılması gerekmiyor. Hatta her bir soru için tek bir kombinasyonu çembersel olarak kaydırarak da kullanmak mümkün. Böylece 120/5=24 değişik şık kombinasyonu elde ederiz. Bu 24 değişik kombinasyonun hangisinin hangi sorularda kullanılacağını sorunun doğru cevabının bulunduğu şıktan bağımsız ve rastsal bir şekilde belirlersek sorun büyük ölçüde çözülür. Buradaki tek risk, rastsallığın doğasından dolayı istenmeden oluşabilecek doğrusallık yaratan durumlardır. Örneğin 24 kombinasyondan bir veya birkaçı rastlantı sonucu tek bir doğru şık için kullanılıyor olabilir. Bu durum benzeri şifre tartışmalarını doğurabilir. Bu tür durumların yaratılmaması için birtakım çapraz kontrollerin de yapılması gerekir.

Yukarıdaki çözüm aslında karmaşık olanı. Daha basit bir çözüm, rastsallaştırılmış masteri hiç kullanmamak, şıkları esas masterin şıklarını çembersel olarak kaydırarak çeşitlendirmek olurdu. Örneğin, bir sorunun masterdeki doğru cevabı (b) şıkkıysa ve bir soru kitapçığında doğru cevabın (e) şıkkına gelmesi gerekiyorsa, masterdeki şıklar çembersel olarak üç kez sağa kaydırılır. Masterdeki sıra bilinse bile sadece cevabın kaç kez kaydırıldığı anlaşılabilir, ama masterdeki doğru cevap bilinmeden soru kitapçığındaki doğru şık da anlaşılamaz.

Kaynaklar

Nechvatal J. ve ark., "Report on the Development of the Advanced Encryption Standard (AES)", *NIST Technical Report*, 2 Ekim 2000 (http://csrc.nist.gov/archive/aes round2/r2report.pdf)

Menezes A. J., van Oorschot P., Vanstone S. A., *Handbook of Applied Cryptography*, CRC Press, 1997.

Kruh L., Deavours C., "The Commercial Enigma: Beginnings of Machine Cryptography", *Cryptologia*, 26(1), 2002.

http://www.osym.gov.tr

Doç. Dr. Albert Levi
Sabancı Üniversitesi • levi@sabanciuniv.edu

Boğaziçi Üniversitesi Bilgisayar Mühendisliği Bölümü'nden 1991'de lisans, 1993'te yüksek lisans ve 1999'da doktora derecelerini aldı. 1999-2002 arasında ABD Oregon Eyalet Üniversitesi Elektrik ve Bilgisayar Mühendisliği Bölümü, Bilgi Güvenliği Laboratuvarı'nda doktora sonrası araştırmalarda bulundu ve aynı bölümde ziyaretçi öğretim üyesi olarak ders verdi. 2002'den bu yana Sabancı Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Bilimi ve Mühendisliği Programı'nda öğretim üyesi olarak görev yapıyor. Araştırma alanları bilgi ve ağ güvenliği.

YGS'DE ŞIKLAR NASIL OLUŞTURULDU?

Soru kitapçıklarındaki şıkların nasıl oluşturulduğunu bir örnek üzerinde açıklayalım. Örneğin matematik testinde master soru kitapçığının 17. sorusunun şıkları aşağıdaki gibiydi ve doğru cevap 7 idi.

A) 5 B) 6 C) 7 D) 8 E) 9

Rastsallaştırılmış masterde aynı sorunun şıkları yandaki transformasyona göre karıldı. Herhangi bir soru kitapçığında bu sorunun doğru cevabının C) şıkkında olması gereklyorsa yandaki sıra aynen kullanıldı:

A) 8 B) 9 (C) 7 D) 6 E) 5

Eğer herhangi bir soru kitapçığında bu sorunun doğru cevabının D) şıkkında olması gerekiyorsa bu sıra yanda gösterildiği gibi çembersel olarak sağa bir kere kaydırıldı ve şık sırası

A) 5 B) 8 C) 9 (D) 7 E) 6 şeklinde değişti.

Eğer herhangi bir soru kitapçığında bu sorunun doğru cevabının E) şıkkında olması gerekiyorsa rastsallaştırılmış masterdeki sıra yanda gösterildiği gibi çembersel olarak sağa iki kere (veya eşdeğer olarak sola üç kere) kaydırıldı ve şık sırası

A) 6 B) 5 C) 8 D)9 (E) 7 şeklinde değişti.

Eğer herhangi bir soru kitapçığında bu sorunun doğru cevabının A) şıkkında olması gerekiyorsa rastsallaştırılmış masterdeki sıra yanda gösterildiği gibi çembersel olarak sağa üç kere (veya eşdeğer olarak sola 2 kere) kaydırıldı ve şık sırası

(A) 7 B) 6 C) 5 D)8 E) 9 şeklinde değişti.

Eğer herhangi bir soru kitapçığında bu sorunun doğru cevabının B) şıkkında olması gerekiyorsa rastsallaştırılmış masterdeki sıra yanda gösterildiği gibi çembersel olarak sağa dört kere (veya eşdeğer olarak sola 1 kere) kaydırıldı ve şık sırası

A) 9 (B) 7 C) 6 D)5 E) 8 şeklinde değişti.