

# GENETİK ALGORİTMALARDA TEK VE ÇOK NOKTALI ÇAPRAZLAMANIN SÖZDE RASSAL POPULASYONLARA ETKİSİ

Pınar SANAC

Ali KARCI

Bilgisayar Mühendisliği Bölümü  
Mühendislik Fakültesi  
Fırat Üniversitesi  
23119 Elazığ

## ÖZET

Genetik algoritmalarda genetik operatörler çok büyük önem taşımaktadır. En genel haliyle çaprazlama, mutasyon ve seçim operatörleri genetik algoritmanın temel parçalarıdır. Genetik operatörlerin algoritma üzerindeki etkileri daha önceki çalışmalarda ispatlanmıştır. Özellikle çaprazlama operatörü algoritma üzerinde en fazla oranla kullanılan operatörlerden birisidir. Aynı şekilde başlangıç popülasyonu oluşturma yöntemleri üzerinde de son zamanlarda çalışmalar gittikçe artmıştır. Bu makalede, geleneksel şema teorisine dayalı olarak çaprazlama operatörünün özel bir başlangıç popülasyonu oluşturma yöntemi olan sözde rassal başlangıç popülasyonları üzerindeki etkisi araştırılmıştır. Çaprazlamanın çeşitliliği ne derece sağladığı ve hangi durumlarda şemayı bozmadığı incelenmiştir ve bu durumda algoritmanın iyiye doğru gidip gitmediği gözlemlenmiştir. Çaprazlama yöntemlerinden olan n noktalı çaprazlama metodu kullanılmıştır. Buna bağlı olarak uygunluk değerlerinin gösterdiği sıçramalar gözlemlenmiştir. Teorik açıklamalar bir problem üzerinde denenmiş ve elde edilen sonuçlar karşılaştırılmıştır.

**Anahtar Kelimeler:** Sözde Rassal Popülasyon, Başlangıç popülasyonu, n noktalı çaprazlama.

## 1. GİRİŞ

Gelişigüzel arama metodlarından biri olan genetik algoritmalarda, rassal değerlerle yola çıkarak problemin çözümünü bulmaya çalışılır[1]. Bu aramanın belirli kurallara dayalı olarak yapılmasıyla ve rassallığın azaltılmasıyla bu metodlar daha iyiye götürülmeye çalışılır. Bu sayede daha kısa sürede daha iyi sonuçlar bulmak hedeflenir.

Geleneksel olarak rassal başlangıç popülasyonlarıyla arama yapmaya başlayan genetik algoritmalarda başlangıç popülasyonu sözde rassal dizilerden faydalanılarak oluşturulmuştur[2]. Bu sayede çeşitliliğin daha iyi sağlanması beklenmektedir.

Kullanılan genetik operatörler de genetik algoritma için oldukça önemlidirler. Özellikle çaprazlama operatörü genetik algoritmada mutasyondan daha yüksek oranda kullanılır ve sonuca yakınsamada etkili olması beklenir.

Daha önceki çalışmalarda çaprazlama operatörünün rassal popülasyonlar üzerindeki etkileri teorik olarak incelenmiştir[6].

Bu çalışmada çaprazlamanın sözde rassal başlangıç popülasyonları üzerindeki etkileri araştırılacaktır. Buna göre makalenin ikinci bölümünde sözde rassal popülasyonlar, 3. bölümde çaprazlamanın rassal popülasyon üzerindeki etkisi, 4. bölümde sözde rassal popülasyon için çaprazlamanın beklenen etkisi anlatılmıştır. 5. bölümde deneysel sonuçlar verilmiş ve 6. bölümde bu sonuçlar değerlendirilmiştir.

## 2. SÖZDE RASSAL POPULASYON

Başlangıç popülasyonu genetik algoritmanın ilk adımını oluşturduğu için, başlangıçta bireylerin çözüm uzayına dağılımı önemlidir. Rassal dağıldığında bireyler birbirlerinin çok yakın noktalarda toplanabilirler ve bu noktalar da çözümden çok uzakta olabilir ve doğru sonuca ya çok fazla iterasyon sonunda ulaşılır veya da hiç ulaşamayabilir.

Sonuca daha erken ulaşabilmek için başlangıç popülasyonları çeşitli yöntemlere göre oluşturulmaktadır.[2,3,4,5] Bunlardan bir tanesi de popülasyon oluşturulurken sözde rassal dizilerden faydalanmaktır.

### 2.1. Sözde Rassal Popülasyonları Oluşturma yöntemleri:

Eğer gerçel kodlama yapılacaksa, bir kromozomdaki gen sayısı kadar asal sayı bulunur ve her sayı bu asal tabanlarda yazılır ve daha sonra tersi alınır. Hesaplanacak sayı  $\phi_p(n)$  olarak gösterilirse;

$$\phi_p(n) = a_0 * p^{-1} + a_1 * p^{-2} + \dots + a_m * p^{-m-1}$$

Burada p asal tabanı, m ise basamak sayısını gösterir,  $a_i$ ' ler katsayılarıdır ve  $0 \leq a_i < p$  dir. Buna göre ikilik tabanda  $\phi_p(n)$  değerleri Tablo 1'deki gibi hesaplanır:

Tablo 1' de elde edilen sayılar popülasyonu oluşturmak için kullanılır. Bu noktadan sonra popülasyon, bu sayılar yardımıyla 2 farklı yöntem kullanılarak oluşturulabilir. Bunlar; Halton ve Hammersley dizileridir[2,3]. Bu çalışmada Halton sözde rassal dizileri kullanılmıştır.

**Tablo 1:** İki tabanında  $\phi_p(n)$  değerleri

ondalık	ikilik	ikilik	ondalık
n = 1	= 1	$\phi_2(n) = 0.1$	= 0.5
2	10	0.01	0.25
3	11	0.11	0.75
4	100	0.001	0.125
5	101	0.101	0.625
.	.	.	.
.	.	.	.
.	.	.	.

Halton dizileriyle kromozom oluşturulursa, popülasyondaki i. kromozom için ;

$$x_i = (\phi_{p1}(i), \phi_{p2}(i), \dots, \phi_{pn}(i)) \quad i=1, 2, \dots, N$$

(N : popülasyon boyutu) şeklinde oluşturulur. Her kromozomun bu şekilde oluşturulmasıyla Halton sözde rassal popülasyonu meydana gelir. Halton popülasyon, çözümleri uzaya belirli aralıklarla dağıtır. Bu nedenle sözde rassal popülasyon oluşturma yöntemleri içerisinde oldukça sık kullanılır[3].

### 3. ÇAPRAZLAMANIN RASSAL POPULASYONA ETKİSİ

Herhangi bir kodlama alfabelerinden oluşabilecek kromozom gruplarına “şema” denir. Örneğin ikili kodlama kullanılıyorsa, alfabe  $A = \{0, 1\}$ ’ den oluşur. Bu alfabeğe göre oluşabilecek şemalardan bazıları  $0*11*$ ,  $1**00$ ,  $1110*$ , şeklinde gösterilebilir. Burada \*, o genin 0 veya 1 olmasının önemli olmadığını gösterir. Yani \*’ lı genler 1 de olsalar 0 da olsalar aynı şemaya dahil olacaklardır demektir. \*’ lar dışındaki 1 veya 0’ lara “tanımlayıcı genler” denir. Bir şemada tanımlayıcı genlerin sayısı o şemanın mertebesini verir.

Buna göre Holland, bir şemanın çaprazlama sonrasında hayatını sürdürme olasılığını araştırmıştır. Goldbeg, t anında verilen H şemasındaki birey sayısından yola çıkarak t+1 anında elde edilecek olan H şemasındaki birey sayısını hesaplamıştır[6].  $m_t$ ; t anında H şemasına ait birey sayısı,  $\bar{f}_t$ , h şemasının t anındaki ortalama uygunluk değeri ve  $\bar{f}_{t+1}$ , t anında popülasyonun ortalama uygunluk değeri olmak üzere, t+q anında H şemasına ait bireylerin sayısının şöyle olması beklenir:

$$m_{t+1}(H) \geq m_t(H) \cdot \frac{\bar{f}_t(H)}{\bar{f}_t} \cdot P_y(H)$$

Burada  $P_y(H)$ , H şemasının çaprazlama sonucunda yaşamını devam ettirme olasılığıdır. t+1 anında H

şemasına ait birey sayısının daha fazla olması beklenir, çünkü şema bir sonraki nesilde yaşamını sürdürebilir, hem de başka şemaların çaprazlanmasından H şemasına ait birey meydana gelebilir.

#### 3.1. N Noktalı Çaprazlamanın Etkisi

Sabit uzunluklu bir evrimsel algoritma için, eğer L uzunluğunda kodlama yapılacaksa ve kullanılan alfabenin eleman sayısı C ise,  $C^L$  tane katar oluşturmak mümkündür. Eğer k. mertebeden bir şema  $H_k$  ile gösterilirse, bu şema  $C^{L-k}$  tane katar içerebilecektir.

Çaprazlama olayı, R rassal değişkeniyle tanımlansın. İkili kodlama düşünüldüğünde  $H_k$  için  $2^k$  tane çaprazlama olayının olması mümkündür ( $0 \leq R \leq 2^k - 1$ ). Her R çaprazlama olayı k uzunluğunda bir bit maskeleye sunulabilir. Eğer o bitin maskesi 1 ise, bu demektir ki j. noktadaki aleller yer değiştirecek, 0 ise yer değiştirmeyecektir. Bütün n noktalı çaprazlamada bit maskeleye kullanılabilir. Bir nesilde bütün çaprazlama olaylarının toplam olasılığı 1 olmalıdır.

$$\sum_R P(R) = 1$$

Bir  $H_k$  şeması için çaprazlama olayı düşünülürse, seçilecek ebeveynlerin biri  $H_k$ ’ ya ait olsun, diğer birey ise keyfi seçilsin. Bu iki bireyin çaprazlanması sonucunda oluşan evlatlardan bir tanesi  $H_k$ ’ ya ait ise,  $H_k$  şeması yaşamını sürdürür, hiçbir  $H_k$ ’ ya ait değilse,  $H_k$  bozulur. Bu durumda  $P_y$ ,  $H_k$ ’ nin yaşamını sürdürme olasılığı,  $P_b$  de  $H_k$ ’ nin bozulma olasılığı olmak üzere;

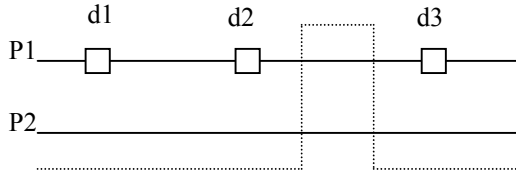
$$P_y(H_k) = \sum_R P(R) \cdot P(H_{k_y}/R) = \sum_R P(R) \cdot P_y(H_k/R)$$

$$P_b(H_k) = \sum_R P(R) \cdot P(H_{k_b}/R) = \sum_R P(R) \cdot P_b(H_k/R)$$

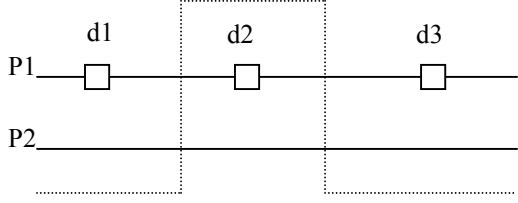
$$P_y(H_k) + P_b(H_k) = 1$$

#### 3.2. N Noktalı Çaprazlama için Yaşamını Sürdürme Teorisi

N noktalı çaprazlamada, çaprazlama olayı için n tane kesim noktası belirtilir. Bu çaprazlama olaylarını tek ve çift noktalı çaprazlama olarak sınıflandırmak uygundur. Çift olayda  $H_k$ ’ nin tanımlama pozisyonları (genleri) arasına çift sayıda kesim noktası gelir(Şekil 1). Bu durumda tanımlayıcı alellerin hiçbir yer değiştirmeyecektir ya da tümü yer değiştirecektir. Bunun tersine tek çaprazlama olayında; tanımlama pozisyonları arasına gelen kesim noktaları tek sayıdadır(Şekil 2). Bu durumda da bu tanımlayıcı alellerin bir bölümü yer değişir. Bu durumda şemanın bozulma ihtimali yüksektir.



Şekil 1. Çift Çaprazlama Olayı



Şekil 2. Tek Çaprazlama Olayı

$$P_{\text{çift}}(H_k) = \sum_{R_{\text{çift}}} P(R)$$

$$P_{\text{tek}}(H_k) = \sum_{R_{\text{tek}}} P(R)$$

Tek ve çift çaprazlama olayları bütün çaprazlama kümesini oluşturduğundan toplamaları 1' dir.

$$P_{\text{çift}}(H_k) + P_{\text{tek}}(H_k) = 1$$

Bu durumda  $H_k$  şemasının bozulma ve yaşama olasılıkları sırasıyla şöyledir:

$$P_d(H_k) = \sum_{R_{\text{çift}}} P(R) \cdot P_d(H_k/R) + \sum_{R_{\text{tek}}} P(R) \cdot P_d(H_k/R)$$

$$P_s(H_k) = \sum_{R_{\text{çift}}} P(R) \cdot P_s(H_k/R) + \sum_{R_{\text{tek}}} P(R) \cdot P_s(H_k/R)$$

Eğer R çift ise;  $P_d(H_k/R) = 0$  ; eğer R tek ise;  $P_s(H_k/R) = 1$  ' dir. Buna göre k, tanımlama uzunluğu olmak üzere, k=2 için çift çaprazlama olayının olma olasılığı n noktalı çaprazlama için şöyledir;

$$P_{\text{çift}}(H_2, L, L_1, n) = \sum_{x=0}^{\lfloor n/2 \rfloor} \binom{n}{2x} \left(\frac{L_1}{L}\right)^{2x} \left(\frac{L-L_1}{L}\right)^{n-2x}$$

Burada  $L_1/L$  oranı kesim noktasının iki tanımlama noktası arasına yerleşme olasılığını,  $(L-L_1)/L$  ise kesim noktasının iki tanımlama noktası dışında olma olasılığını gösterir.  $2x$  indeksi daima çifttir, 0' dan n' e kadar bütün çift sayıları içerir. Kombinasyon ise, n tane kesim noktasından seçilebilen çift sayı yollarının sayısını hesaplar. Bu sonuç sadece ikinci dereceden şemalar için geçerlidir. Bu ilişkiyi daha yüksek şemalar için genelleştirmek mümkündür, çünkü çift sayıların toplamı yine bir çift sayıdır. Bu denklem en genel haliyle şöyle gösterilebilir:

$$P_{\text{çift}}(H_k, L, L_1, \dots, L_{k-1}, n) = \sum_{x=0}^{\lfloor n/2 \rfloor} \binom{n}{2x} \left(\frac{L_1}{L}\right)^{2x} \left(\frac{L-L_1}{L}\right)^{n-2x} \cdot P_{\text{çift}}(H_{k-1}, L_1, \dots, L_{k-1}, 2x)$$

#### 4. SÖZDE RASSAL POPULASYONLAR ÜZERİNDE ÇAPRAZLAMANIN ETKİSİ

3. bölümdeki rassal populasyon için tek ve çift çaprazlama olaylarının olma olasılığı sözde rassal populasyon için de geçerlidir. Buna göre eğer çift çaprazlama olayı gerçekleşmişse, şema bozulmaz. Tek çaprazlama olayı gerçekleşmişse, bu durumda şema, bozulur veya hayatını sürdürmeye devam eder. Buna göre 3. bölümdeki denklemler sözde rassal populasyonlar için de geçerlidir.

Bu durumda çaprazlamanın genetik algoritmayı iyiye götürmesi gereklidir.

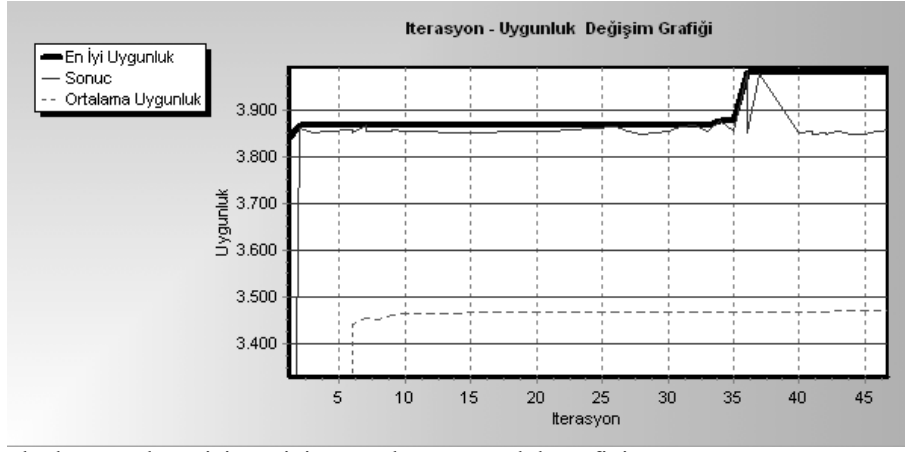
#### 5. DENEYSEL SONUÇLAR

Deneysel çalışmalar yapılırken genetik algoritma kullanılarak program yapılmıştır. Programda, sık kullanılan

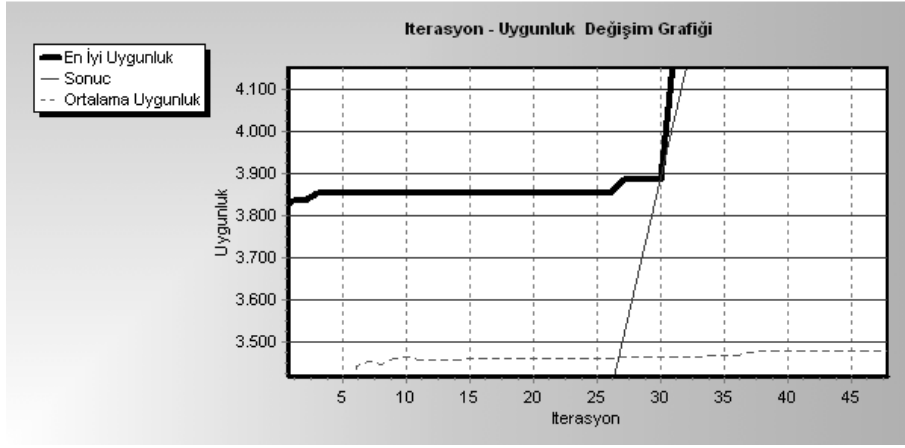
$$f(x) = \sum_{i=1}^N (x_i^2 - 10 \cos(2\pi x_i) + 10)$$

trigonometrik fonksiyon kullanılmıştır[4]. Probleme değişkenlerin tanım aralığı [-5,5] olarak alınmış ve fonksiyonun maksimumu araştırılmıştır. Problem boyutu 128 olarak seçilmiştir. Programda populasyon boyutu 60 olarak alınmıştır. Gerçek kodlama yapılmıştır. Başlangıç populasyonu Halton Sözde Rassal Populasyon şeklinde oluşturulmuştur. Programda klasik genetik algoritmadan farklı olarak, mutasyon yapılmamıştır. Tek noktalı ve iki noktalı çaprazlama kullanılmış, bunlar uygunluk değerlerine göre değerlendirilmiştir. Seçim operatörü ise sadece ebeveynlerin seçiminde kullanılmıştır. Bir sonraki nesil için sadece çaprazlama sonucu oluşan bireyler kullanılır; eski bireyler bir sonraki nesle taşınmaz. Elitizm kullanılmamıştır. Şemalar kontrol edilmemiştir, çünkü gerçek kodlamada şemaya hakim olmak, genler verilen aralıkta sonsuz değer alabileceğinden zordur. Bu nedenle, popülasyondaki ortalama uygunluk değeri ve en iyi uygunluk değerlerine bakılmıştır ve buna dayalı olarak şemalar hakkında yorum yapılmıştır.

Buna göre tek ve iki noktalı çaprazlamanın uygunluk-iterasyon grafik değişimi aşağıda verilmiştir.



Şekil 3. Tek noktali çaprazlama için en iyi ve ortalama uygunluk grafiđi



Şekil 4. Çift noktali çaprazlama için en iyi ve ortalama uygunluk grafiđi.

Şekil 3 ve Şekil 4' te tek ve çok noktali çaprazlama için her jenerasyondaki en iyi uygunluk ve ortalama uygunluk değeri verilmiştir. Buna göre tek ve çok noktali çaprazlamada populasyonun ortalama uygunluk değeri çok fazla farklılık görülmemektedir. Çift noktali daha inişli çıkışlı bir grafik mevcutken, tek noktali çaprazlamada ortalama uygunluk değeri daha düzgün bir grafik oluşturmuştur.

En iyi uygunluk değeri bakılırsa, iki çaprazlama yöntemi arasında açık bir fark vardır. Tek noktali çaprazlamada 30. iterasyona kadar en iyi uygunluk değeri hemen hemen değişmemiş, 30. iterasyondan sonra sıçrama gerçekleştirmiş ve daha sonra tekrar birbirine benzer değerlerle 50. iterasyona kadar gitmiştir. Çok noktali çaprazlamada ise, en iyi uygunluk birkaç sıçrama yapmış, fakat 30. iterasyondan sonra sürekli artış geçmiştir. Çok noktali çaprazlama olayından sonra populasyonda genel olarak çok fazla bir değişim olmasa da en iyi birey sürekli daha iyiye doğru gitmiştir.

**Tablo 2.** Çaprazlama metodlarının Karşılaştırılması

	En iyi Uygunluk	Ortalama Uygunluk	İterasyon No
Tek Noktali Çaprazlama	4123.540	3217.12	40
Çok Noktali Çaprazlama	4644.322	3122.0	65

Tablo 2' de görüldüğü gibi en iyi uygunluğa tek noktali çaprazlamada 40. iterasyonda ulaşılırken, çok noktali çaprazlamada 65. iterasyonda ulaşılmıştır. Fakat elde edilen uygunluk değeri arasında oldukça büyük bir fark vardır.

## 6. SONUÇ

Elde edilen grafikler ve tabloya göre, çaprazlama olayının genetik algoritmayı iyiye götürdüğünü söyleyebiliriz. Sözde rassal populasyon başlangıçta çeşitliliği iyi sağladığı için çaprazlamayla büyük sıçramalar gerçekleşmemiştir. Fonksiyon tutarlı bir yol izlemiştir. N noktali çaprazlama metodu, sözde rassal başlangıç populasyonu üzerinde olumlu etkiler yapmıştır ve fonksiyon hiçbir zaman daha kötüye gitmemiştir.

Eğer şemalar bozulsaydı, bu durumda yeni şemalar oluşmuş olacak veya kötü şemalar elde edilmiş olacaktı ve bunun sonucu olarak da fonksiyon kötüye gidecekti veya algoritma uzun iterasyonlar boyunca devam edecekti. O zaman tek ve çok noktalı çaprazlamanın sözde rassal populasyon için şemayı bozma olasılığının düşük olduğu sonucu çıkarılabilir.

#### **KAYNAKLAR:**

- [1] D: E. Goldberg, Genetic Algorithm in Search, Optimization and Machine Learning ,Addison-Wesley, 1989
- [2] J. H. Halton. On the e.ciency of certain quasi-random sequences ofpoints in evaluation multi-dimensional integrals. Numer. Math., 2:84–90, 1960
- [3] P. SANAÇ, A. KARCI, Genetik Algoritmalarda Sanki Rassal Populasyonların arşılaştırılması, Ya/Em'2004 Yöneylem Araştırması/Endüstri MühendisliğiXXIV. Ulusal Kongresi 16-18 Haziran 2004, Çukurova Üniversitesi, Adana
- [4] Yiu-Wing Leung, An Orthogonal Genetic Algorithm with Quantization for Global Numercal Optimization, , IEEE Trans. on Evolutionary Computation, vol:5, no:1, 2001.
- [5] Y. J. Cao, Q. H. Wu, D. W. Shimmin, Study Of Initial Population In Evolutionary Programming
- [6] William M.Spears, Evolutionary Algorithms The Role of Mutation and Recombination, 2000