

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
FİZİK ANABİLİM DALI**

**BİYOLOJİK TOPLULUKLARIN VE GELİŞEN AĞ
YAPILARININ DOLAŞIK DOĞA MODELİ İLE
İNCELENMESİ**

**Hazırlayan
Kamil ARGİN**

**Danışman
Prof. Dr. Osman CANKO**

Doktora Tezi

**Eylül 2015
KAYSERİ**

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
FİZİK ANABİLİM DALI**

**BİYOLOJİK TOPLULUKLARIN VE GELİŞEN AĞ
YAPILARININ DOLAŞIK DOĞA MODELİ İLE
İNCELENMESİ**

**Hazırlayan
Kamil ARGİN**

**Danışman
Prof. Dr. Osman CANKO**

Doktora Tezi

**Bu çalışma TÜBİTAK tarafından 111T735 nolu proje ile
desteklenmiştir**

**Eylül 2015
KAYSERİ**

Bu alıřmadaki tm bilgilerin, akademik ve etik kurallara uygun bir řekilde elde edildiđini beyan ederim. Aynı zamanda bu kural ve davranıřların gerektirdiđi gibi, bu alıřmanın znde olmayan tm materyal ve sonuları tam olarak aktardıđımı ve referans gsterdiđimi belirtirim.

Kamil ARGİN



“Biyolojik Toplulukların ve Gelişen Ağ Yapılarının Dolaşık Doğa Modeli ile İncelenmesi” adlı Doktora tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi’ne uygun olarak hazırlanmıştır.

Tezi Hazırlayan
Kamil ARGİN



Danışman
Prof. Dr. Osman CANKO



Fizik ABD Başkanı
Prof.Dr. Ayhan GÜLDESTE

Prof. Dr. Osman CANKO danışmanlığında **Kamil ARGİN** tarafından hazırlanan “**Biyolojik Toplulukların ve Gelişen Ağ Yapılarının Dolaşık Doğa Modeli ile İncelenmesi**” adlı bu çalışma, jürimiz tarafından Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fizik Anabilim Dalında **Doktora** tezi olarak kabul edilmiştir.

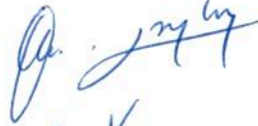
28.09/2015

JÜRİ :

Başkan : Prof. Dr. Ahmet ERDİNÇ



Üye : Prof. Dr. Osman CANKO



Üye : Doç. Dr. Gökhan KOÇAK



Üye : Doç. Dr. Mehmet ŞAHİN



Üye : Doç. Dr. Asım SOYLU



ONAY :

Bu tezin kabulü Enstitü Yönetim Kurulunun 06/10/2015 tarih ve 2015/4428 sayılı kararı ile onaylanmıştır.

06/10/2015



Enstitü Müdürü
Prof. Dr. Kâzım KEŞLİOĞLU

TEŞEKKÜR

Öncelikle tez konumun belirlenmesinde yardımcı olan, çalışmalarımın başından sonuna kadarki bütün aşamalarda yardımlarını esirgemeyen, değerli fikir ve tecrübeleri ile bana büyük destek sağlayan saygıdeğer hocam ve tez danışmanım Prof. Dr. Osman CANKO'ya ve tez izleme komitesi üyeleri değerli hocalarım Prof. Dr. Ahmet ERDİNÇ'e ve Doç. Dr. Gökhan KOÇAK'a en içten dileklerimle teşekkürlerimi sunarım.

Tez çalışması süresince 111T735 nolu proje ile sağladıkları maddi destekten dolayı Türkiye Bilimsel ve Teknolojik Araştırma Kurumu'na (TÜBİTAK) teşekkür ederim.

Tezin L^AT_EX ile yazımında desteğini gördüğüm değerli hocam Doç. Dr. Ferhat TAŞKIN'a teşekkür ederim.

Tezin yazım aşamasındaki katkılarından dolayı değerli arkadaşım Onur KAVCI'ya teşekkür ederim.

Tez çalışması boyunca bana destek olan, kardeşim gibi sevdiğim çok kıymetli arkadaşım Erkan SAĞIR'a teşekkür ederim.

Son olarak, tüm bu yorucu çalışmalarım süresince her zaman sabır ve anlayış gösteren, maddi ve manevi desteklerini hiçbir zaman esirgemeyen babam Mehmet Metin ARGİN'a, annem Aynur ARGİN'a ve kardeşlerim Yunus Emre ARGİN ve Yiğitcan ARGİN'a en içten dileklerimle teşekkürlerimi sunarım.

BİYOLOJİK TOPLULUKLARIN VE GELİŞEN AĞ YAPILARININ DOLAŞIK DOĞA MODELİ İLE İNCELENMESİ

Kamil ARGİN
Erciyes Üniversitesi, Fen Bilimleri Enstitüsü
Doktora Tezi, Eylül 2015
Tez Danışmanı: Prof. Dr. Osman CANKO

ÖZET

Ekosistemdeki biyolojik toplulukların evrimsel süreci, birey tabanlı eş evrimleşme modeli olan Dolaşık Doğa Modeli ile incelendi. Eş-evrimleşme olayını gözlemlemek için, canlı türlerin bir arada yaşadığı fiziksel çevredeki değişiklikler ihmal edildi. Basitlik için genotip ve fenotip arasında ayrım yapılmadı. Bütün türler genotip uzayında düşünüldü. Simülasyon sonuçlarına göre, evrimsel süreçte topluluktaki çeşitlilik ve toplam popülasyon artmaktadır. Çeşitliliğin ve toplam popülasyonun artışından dolayı topluluğun kompleksliliği artmaktadır. Bu komplekslik artışı dinamik ağ yapısıyla incelendi. Ağ yapısını analiz edebilmek için, bazı ağ parametrelerinin zamanla değişimi incelendi. Elde edilen sonuçlar gerçek dünyadaki evrimsel ağ sistemleri ile karşılaştırıldı ve uyumluluk gözlemlendi. Evrimsel süreç içerisinde, biyolojik çeşitliliğin kökeninin araştırılması için önemli olan soy ağacı çıkarımı yapıldı. Elde edilen soy ağacı, literatürde geniş çapta kullanılan En Fazla Olabilirlik ve Komşu Kurma metotları ile tahmin edilen soy ağaçları ile karşılaştırıldı. Metotların yaptıkları tahminlerin doğrulukları nümerik olarak analiz edildi. Topluluk yapısı, soy ağacındaki türlerin ağ yapısı için elde edildi. Topluluktaki grup sayısının gerçek dünyadaki bazı yiyecek ağları ile uyumlu olduğu görüldü.

Anahtar Kelimeler: Eş-Evrimleşme, Kompleks Ağ, Ağ Parametreleri, Soy Ağacı, En Fazla Olabilirlik Metodu, Komşu Kurma Metodu.

INVESTIGATION OF BIOLOGICAL COMMUNITIES AND THEIR DEVELOPING NETWORK STRUCTURE WITH TANGLED NATURE MODEL

Kamil ARGIN

Erciyes University, Graduate School of Natural and Applied Sciences

Ph.D. Thesis, August 2015

Thesis Supervisor: Prof. Dr. Osman CANKO

ABSTRACT

The evolutionary process of biological communities in ecosystem was be investigated with Tangled Nature Model which is an individual based co-evolutionary model. Changing in the physical environment in which co-exist the living species was be neglected to observe co-evolution phenomenon. The distinction between genotype and the phenotye was not be made for simplicity. All of the species were be thought in the genotype space. According to simulation results, diversity and total population increase in evolutionary process. The complexity of the community is increase because of the increasing of the diversity and the total population. This increasing of the complexity is analysed with dynamic network structure. Some network parameters changing with time are examined in order to analysed the network structure. The obtained results are compared with the evolutionary network system in the real world and compatibility is observed. In the evolutionary process, inference of the phylogenetic tree which is important to investigate the origin of biological diversity is made. The obtained phylogenetic tree is compared with phylogenetic trees which is estimated with Maximum Likelihood and Neighbor-Joining methods which is widely used in literature. The accuracy of estimates which is made these methods is analysed numerically. The community structure is obtained for the network structure of species in the phylogenetic tree. It is observed that the number of group in the community is compatible with some food webs in the real world.

Keywords: Co-Evolution, Complex Network, Network Parameters, Phylogenetic Tree, Maximum Likelihood Method, Neighbor-Joining Method.

İÇİNDEKİLER

BİYOLOJİK TOPLULUKLARIN VE GELİŞEN AĞ YAPILARININ DOLAŞIK DOĞA MODELİ İLE İNCELENMESİ

BİLİMSEL ETİĞE UYGUNLUK SAYFASI	i
YÖNERGEYE UYGUNLUK SAYFASI	ii
KABUL VE ONAY	iii
TEŞEKKÜR	iv
ÖZET	v
ABSTRACT	vi
İÇİNDEKİLER	vii
TABLolar LİSTESİ	x
ŞEKİLLER LİSTESİ	xi
KISALTMALAR VE SİMGELER	xvi
GİRİŞ	1

1. BÖLÜM

EVİRİMSEL EKOLOJİ VE AĞ SİSTEMLERİ

1.1. Evrimsel Değişim Mekanizmaları	11
1.1.1. Mutasyon	13
1.1.2. Gen Akışı	14
1.1.3. Genetik Sürüklenme	14
1.1.4. Doğal Seçilim	15
1.2. Eş- Evrimleşme	15
1.3. Biyolojik Çeşitlilik	16
1.4. Sınıflandırma	21

1.5. Evrimsel Modeller	21
1.5.1. NK Modeli	22
1.5.2. Kendini Organize Eden Evrim Modeli	23
1.5.3. Hiyerarşik Yiyecek Ağı Modeli	24
1.6. Ağ Sistemlerinin Matematiksel Temelleri	25
1.6.1. Bitişiklik Matrisi	26
1.6.2. Yönelimli ve Yönelimsiz Ağlar	28
1.6.3. Ağırlıklı ve Ağırlıksız Ağlar	31
1.7. Ağ Analizinde Sıklıkla Kullanılan Parametreler	32
1.7.1. Derece Dağılımı	33
1.7.2. Ortalama Yol Uzunluğu ve Çap	36
1.7.3. Kümelenme Katsayısı	39
1.7.4. Modülerite	43
1.8. Ağ Modelleri	46
1.8.1. Rastgele Ağ Modeli	46
1.8.2. Küçük Dünya Ağları	52
1.8.3. Ölçeksiz Ağlar	57

2. BÖLÜM SOY AĞACI

2.1. En Fazla Olabilirlik (ML) Metodu	65
2.2. Komşu Kurma (NJ) Metodu	68

3. BÖLÜM DOLAŞIK DOĞA MODELİ (DDM)

3.1. Genel Özellikleri	73
3.2. Matematiksel İşleyişi	74
3.2.1. Organizmalar ve Genotip Uzayı	74
3.2.2. Ağırlık Fonksiyonu ve Üreme İhtimaliyeti	76
3.2.3. Zaman Adımı ve Nesilsel Zaman	86

4. BÖLÜM SİMÜLASYON SONUÇLARI

4.1. Tek Bir Simülasyon Sonucu	90
4.1.1. Toplam Popülasyon ve Çeşitlilik	93
4.1.2. Dinamik Ağ Yapısı	94
4.1.3. Soy Ağacı	100
4.2. Çok Sayıda Simülasyonun Ortalama Sonuçları	110
4.2.1. Toplam Popülasyon ve Çeşitlilik	110
4.2.2. Dinamik Ağ Yapısı	112

5. BÖLÜM

TARTIŞMA, SONUÇ VE ÖNERİLER

KAYNAKLAR	121
ÖZGEÇMİŞ	133

TABLOLAR LİSTESİ

Tablo 1.1.	Jeolojik zaman ölçeği. Ref.[80]	17
Tablo 1.2.	Bazı ağlardaki köşeler ve kenarlar.	26
Tablo 1.3.	Şekil 1.14(a)'daki olası yollar.	37
Tablo 1.4.	Şekil 1.14(b)'deki olası yollar.	37
Tablo 1.5.	Gerçek dünya ağlarının sahip oldukları L ortalama yol uzunluklarının rastgele ağ modeli ile elde edilen L_r ortalama yol uzunlukları ile karşılaştırılması.	50
Tablo 1.6.	Gerçek dünya ağlarının sahip oldukları C kümelenme katsayılarının rastgele ağ modeli ile elde edilen C_r kümelenme katsayıları ile karşılaştırılması.	51
Tablo 1.7.	Küçük dünya ağları için Watts ve Strogatz tarafından verilen bazı deneysel örnekler. Ref.[41]	57
Tablo 4.1.	Bazı gerçek dünya ağlarının modülerite değerleri.	99
Tablo 4.2.	Şekil 4.9 a'da verilen soy ağacındaki 45 ana türün genom dizileri. . .	103
Tablo 4.3.	TOPD/FMTS yazılımı ile DDM ile elde edilen soy ağacının, ML ve NJ metotlarının tahmin ettikleri ağaçlar ile karşılaştırılması.	107
Tablo 4.4.	Gerçek dünyadaki yiyecek ağları için topluluklarda bulunan grup sayıları. Ref.[152]	110

ŞEKİLLER LİSTESİ

Şekil 1.1.	DNA ve onu oluşturan yapıların gösterimi. a) DNA molekülünü oluşturan nükleotit birimlerinin içeriği. b) DNA molekülünün tek bir ipliğini oluşturan nükleotit dizileri. 5 sayısı nükleotidin fosfatlı kısmını, 3 sayısı ise fosfatsız kısmını göstermektedir. c) Nükleotit dizilerindeki bazların hidrojen bağları ile bağlanması sonucu oluşan DNA çift ipliği. d) DNA molekülünün ters kutuplu çift iplikten oluşan sarmallı gerçek yapısı. Ref.[53]	12
Şekil 1.2.	Fanerozoyik zaman boyunca, deniz canlıları için çeşitliliğin milyon yıl ölçeğindeki zamana göre değişimi. Kesikli çizgi lineer, noktalı çizgi eksponansiyel fit eğrileridir. Ref.[82]	18
Şekil 1.3.	Fanerozoyik zaman boyunca, deniz canlıları için milyon yıl başına yok olmanın (ana şekil) ve milyon yıl başına çıkmanın (küçük şekil) milyon yıl ölçeğindeki zamana göre değişimi. Ref.[82]	19
Şekil 1.4.	Raup tarafından yapılan, Fanerozoyik zamanın 79 jeolojik devresi için yok olan tür sayısının dağılımı. Ref.[88]	20
Şekil 1.5.	6 tane köşeden oluşan bir ağ.	27
Şekil 1.6.	6 tane köşeden oluşan bir ağ sistemindeki çoklu ve öz-kenar durumları.	28
Şekil 1.7.	6 tane köşeden oluşan yönelimli bir ağ.	29
Şekil 1.8.	6 tane köşeden oluşan çift yönelimli bir ağ.	30
Şekil 1.9.	6 tane köşeden oluşan yönelimsiz bir ağ.	30
Şekil 1.10.	6 tane köşeden oluşan yönelimsiz ağırlıklı bir ağ. Kenarların üzerindeki sayılar ilgili kenarın ağırlığını belirtmektedir.	31
Şekil 1.11.	5 tane köşeden oluşan yönelimsiz ağırlıksız basit bir ağ.	33
Şekil 1.12.	Şekil 1.11'deki ağın derece dağılımı.	34
Şekil 1.13.	25 köşeden oluşan yönelimsiz karmaşık bir ağ ve derece dağılımı. . .	35
Şekil 1.14.	Yönelimsiz ve yönelimli ağlardaki yollar.	36

Şekil 1.15.	Köşeler arasındaki kümelenmenin ve ya geçişkenliğin şematik gösterimi.	40
Şekil 1.16.	Yönelimli ağlarda oluşabilecek üçgenler. Kümelenmenin olacağı köşe i köşesi olarak belirlenmiş bu yüzden siyah renkle gösterilmiştir.	42
Şekil 1.17.	Metodun adımlarının örnek bir ağ üzerinde gösterimi.	45
Şekil 1.18.	Farklı p ihtimaliyetlerine göre bağlanmış $N = 10$ tane köşeye sahip rastgele ağ örnekleri.	47
Şekil 1.19.	$N = 10000$ ve $p = 0.0015$ değerleri alınarak oluşturulmuş bir rastgele ağ için Poisson derece dağılımı.	49
Şekil 1.20.	$N = 5000$ tane köşeden oluşan bir rastgele ağ için ortalama yol uzunluğunun ağın ortalama derecesine göre değişimi.	50
Şekil 1.21.	$N = 5000$ köşeye sahip bir rastgele ağdaki kümelenme katsayısının ortalama derece ifadesine göre değişimi.	51
Şekil 1.22.	Milgram'ın deneyinde alıcıya ulaşan 44 mektubun, alıcıya ulaşmak için gerekli kişi sayısına göre dağılımı. Ref.[40]	53
Şekil 1.23.	Watts ve Strogatz tarafından elde edilen, düzenli örgü ile rastgele ağ arasında artan p değerine göre kenarların rastgele yeniden bağlanmasının şematik gösterimi. Ref.[41]	54
Şekil 1.24.	Watts ve Strogatz tarafından $N = 1000$ ve $k = 10$ için elde edilen ortalama yol uzunluğunun ve kümelenme katsayısının p ihtimaliyetine göre değişimi. Ref.[41]	55
Şekil 1.25.	Barrat ve Weigt tarafından $N = 1000$ ve $\langle d \rangle = k = 6$ değerlerine göre farklı p değerleri için elde edilen derece dağılımı. Ref.[126]	56
Şekil 1.26.	$N = 10$, $m_0 = 2$ ve $m = 2$ değerleri için ölçeksiz ağın $t = 8$ zaman adımı evrilmesi.	58
Şekil 1.27.	Barabási et al. tarafından $N = t + m_0 = 300000$ tane köşe için elde edilen derece dağılımı. <i>Daire</i> için $m_0 = m = 1$, <i>Kare</i> için $m_0 = m = 3$, <i>Elmas</i> için $m_0 = m = 5$ ve <i>Üçgen</i> için $m_0 = m = 7$ değerleri alınmıştır. Ref.[127]	60

Şekil 1.28.	Albert ve Barabási tarafından $\langle k \rangle = 4$ için elde edilen ortalama yol uzunluğu ifadesinin ağdaki köşe sayısı N 'ye göre değişimi. <i>Daire</i> Barabási - Albert modeli ve <i>Kare</i> ise rastgele ağ modeli ile elde edilen değerlerdir. Kesikli çizgi Denklem (1.47)'ye göre, sürekli çizgiler ise Denklem (1.37)'ye göre elde edilmiştir. Ref.[101]	61
Şekil 1.29.	Albert ve Barabási tarafından $\langle k \rangle = 4$ için elde edilen kümelenme katsayısı ifadesinin ağdaki köşe sayısı N 'ye göre değişimi. <i>Daire</i> Barabási - Albert modelindeki ve <i>Düz çizgi</i> ise rastgele ağ modelindeki değerlerdir. Ref.[101]	61
Şekil 2.1.	Soy ağacının temel gösterimi.	63
Şekil 2.2.	Ağaç topolojilerinin gösterimi.	64
Şekil 2.3.	ML metodunun anlatımı sırasında kullanılacak ağaç örneği. A, C, G ve T bazları 1, 2, 3 ve 4 değerleri ile gösterilmiştir. 0, 5 ve 6 değerleri ise ağacın dallanma noktalarıdır. v 'ler ise dal uzunlukları olup nesil zamana karşılık gelmektedir.	66
Şekil 2.4.	Saitou ve Nei tarafından verilen 8 türden oluşan yıldız benzeri bir ağaç. Ref.[138]	69
Şekil 2.5.	En yakın iki türün ayrılarak bir küme oluşturması. Ref.[138]	69
Şekil 2.6.	Saitou ve Nei tarafından verilen 8 türden oluşan köksüz bir soy ağacı. Ref.[138]	71
Şekil 3.1.	$L = 4$ boyutlu hiperküp ile temsil edilen genotip uzayı örneği.	75
Şekil 3.2.	Üreme ihtimaliyetinin ağırlık fonksiyonuna göre değişimi. $H \rightarrow -\infty$ iken $P_{\text{üreme}} \rightarrow 0$ ve $H \rightarrow +\infty$ iken $P_{\text{üreme}} \rightarrow 1$ değerlerine eksponansiyel olarak değişmektedir.	80
Şekil 3.3.	Eşeysiz üremede yavruların oluşumu.	81
Şekil 3.4.	Eşeyli üremede yavrunun oluşumu.	82
Şekil 3.5.	Eşeysiz üreme durumunda bir nesil zaman için akış diyagramı. $t = N(t)/P_{\text{öldür}}$ adet zaman adımı tekrarlanınca bir nesil zaman tamamlanmış olur.	87
Şekil 4.1.	Genotip uzayının doluluğunun zaman ile değişimi.	91

Şekil 4.2.	Şekil 4.1'deki simülasyon için, (a) Toplam popülasyonun ve (b) Çeşitliliğin zamanla değişimi.	93
Şekil 4.3.	Şekil 4.1'deki simülasyon için köşelerin derece dağılımları. Kırmızı çizgiler 3 parametrelili Gaussian fit eğrileridir.	95
Şekil 4.4.	Ortalama yol uzunluğunun zamanla değişimi.	97
Şekil 4.5.	Çapın zamanla değişimi.	97
Şekil 4.6.	Ortalama kümelenme katsayısının zamanla değişimi.	98
Şekil 4.7.	Modüleritenin zamanla değişimi.	99
Şekil 4.8.	Modülerite neticesinde toplulukda oluşan grupların sayısının zamanla değişimi.	100
Şekil 4.9.	a İlk q-ESS periyodunun sonunda yaşayan 187 türün gerçek soy ağacı gösterimi. Düşey eksen 712889241 numaralı başlangıç türünden meydana gelen türlerin numaralarını göstermektedir. Ancak bu numaralar keyfi olarak yerleştirilmiş olup, artan yada azalan bir sıra söz konusu değildir. İlk q-ESS periyodu boyunca yaşayan 45 türün ağaçda sahip olduğu dallar dört farklı renk (kırmızı, yeşil, pembe ve mavi) ile gösterilmiştir. b ve c Soy ağacında ilk türleşmelerin gerçekleştiği dönemin büyültülmüş halleri.	102
Şekil 4.10.	Şekil 4.9 a 'daki 45 ana tür için ML metodunun tahmin ettiği soy ağacı. Dallanma noktalarında verilen sayılar, ilgili dallanma noktalarının üretilen 100 farklı ağaçtan kaç tanesinde aynı noktada olduğunu belirtmektedir. Türler, Şekil 4.9 a 'daki gerçek soy ağacında sahip oldukları renklere göre dairelerle gösterilmiştir.	105
Şekil 4.11.	Şekil 4.9 a 'daki 45 ana tür için, NJ metodunun tahmin ettiği soy ağacı. Dallanma noktalarında verilen sayılar, ilgili dallanma noktalarının üretilen 100 farklı ağaçtan kaç tanesinde aynı noktada olduğunu belirtmektedir. Türler, Şekil 4.9 a 'daki gerçek soy ağacında sahip oldukları renklere göre dairelerle gösterilmiştir.	106

- Şekil 4.12. Şekil 4.9 **a**'daki 45 ana türün topluluk yapısı. Toplulukda kesikli dairelerle gösterilen dört grup bulunmaktadır. Türler Şekil 4.9 **a**'da sahip oldukları renklere göre farklı büyüklükteki dairelerle gösterilmiştir. Dairelerin büyüklükleri, türlerin sahip oldukları $n_i(t)$ birey sayıları ile $50 \times \log(10 \times n_i(t))$ bağıntısına göre belirlenmiştir. 109
- Şekil 4.13. 200 simülasyonun ortalaması için, (a) Toplam popülasyonun ve (b) Çeşitliliğin zamanla değişimi. 111
- Şekil 4.14. 50 simülasyon üzerinden ortalaması alınan birinci ve ikinci q-ESS periyotlarının derece dağılımları. Siyah kesikli ve mavi sürekli çizgiler sırası ile birinci ve ikinci q-ESS periyotlarının ortalama derece dağılımlarıdır. Kırmızı yıldız semboller, birinci q-ESS periyodunun ortalama derece dağılımının iç detaylarını göstermektedir. 112
- Şekil 4.15. 300 simülasyon üzerinden ortalamaları alınan ağ parametrelerinin zamanla değişimi. **a)** Ortalama Yol Uzunluğu **b)** Çap **c)** Kümelenme Katsayısı **d)** Modülerite. 113

KISALTMALAR VE SİMGELER

<u>Sembol</u>	<u>Anlamı</u>
DDM	Dolaşık Doğa Modeli
DNA	Deoksiribo Nükleik Asit
S^α	Birey
S^a	Tür
L	Genom uzunluğu
K	Türün genom zincirindeki etkileşen genlerin sayısı
$f(S^\alpha)$	Bireyin uygunluk değeri
$P_{\text{üreme}}$	Üreme ihtimaliyeti
$P_{\text{öldürme}}$	Öldürme ihtimaliyeti
A	Ağın bitişiklik matrisi
d	Derece
P(d)	d derecesinin frekansı
L	Ortalama yol uzunluğu
L_r	Rasgele ağlardaki ortalama yol uzunluğu
D	Çap
C	Kümelenme katsayısı
C_r	Rasgele ağlardaki kümelenme katsayısı
Q	Modülerite
ML	En Fazla Olabilirlik (Maximum Likelihood)
NJ	Komşu Kurma (Neighbor - Joining)

δ_{ij}	Kronecker Delta fonksiyonu
H_a	Ağırlık fonksiyonu
c	Etkileşme şiddeti
$N(t)$	t anındaki toplam popülasyon sayısı
t	Nesilsel zaman
J_{ab}	Türler arasındaki etkileşme
n_b	Genotip uzayındaki dolu konumların birey sayısı
μ	Taşıma kapasitesi
I_a	Etkileşme terimi
ρ_b	Bireylerin yoğunluğu
θ	Etkileşme oranı
T	Bağılantılılık matrisi
P_{mut}	Mutasyon ihtimaliyeti

GİRİŞ

Etrafımızda birbirleriyle etkileşen veya birbirleriyle ilişkili olan cisim ya da varlıklardan oluşan birçok *sistem* vardır. Bu sistemlerden, kolaylıkla analiz edilip hakkında oldukça kolay bilgi sahibi olunan *basit sistem*'ler, sıklıkla uzay-zaman kavramı içerisinde basit davranış kuralları ile açıklanabilir. Bu sistemlerde hiçbir rastgelelik yoktur ve sistemin sonraki davranışı kolaylıkla öngörülebilir. İdeal kristal yapılar ya da ideal gaz sistemleri basit sistemlere örnektir. Sistem birçok alt birimlerden oluşmasına rağmen, alt birimlerin ortaya çıkardığı davranışlar basittir. Bunun nedeni, ideal kristal yapının yüksek düzene sahip olması ve bu düzenliliğin neticesinde sistemin fiziksel ve matematiksel açıdan incelenebilmesidir. Benzer olarak ideal gaz sistemlerinde de, istatistiksel olarak homojen bir yapıya sahip olan sistem basınç, hacim ve sıcaklık gibi birkaç parametre ile tanımlanabildiği için yine basit sistemdir. Son otuz yıl içerisinde sistemler hakkında ortaya çıkan yeni bir görüşe göre; bütün fizik kurallarının geçerli olduğu ancak şu an teori ve model kurmak için kullanılan standart yaklaşımlarla çok iyi anlaşılmayan bir çok sistemin *kompleks* olduğu belirtilmiştir. *Kompleks sistemler*, çok fazla sayıda karşılıklı etkileşen alt birimlerden oluşur ve sistemin bütün davranışı bu alt birimlerin davranışları ile şekillenir. Bu yüzden kompleks sistemleri bir bütün olarak incelemek istediğimizde sistem anlaşılmaz olacaktır. Basit sistemlerin aksine kompleks sistemler, kolaylıkla analiz edilemeyen, gelecekteki durumu hakkında herhangi bir öngöründe bulunulamayan ve rastgeleliğin baskın olduğu sistemlerdir. Bu yüzden böyle sistemler tarihsel zamanın varlığını kabul eden teorilerle açıklanamaz.

Süperpozisyon prensibi gereği bir sistem için iki farklı çözüm varsa sistem iki kısma ayrılabilir. Bu kısımların çözümlerinin ayrı ayrı yapıp toplanması ile sistemin tam çözümüne ulaşılabilir. Süperpozisyon prensibi lineer sistemler için geçerlidir. Bunun nedeni, lineer sistemlerin değişkenlerinden bir tanesi üzerinde yapılacak değişimin diğer değişkenleri etkilemeyecek olmasıdır. Bu yüzden lineer sistemlerin çıktıları girdileri ile orantılıdır. Ancak lineer olmayan (nonlinear) sistemler, çıktılarının girdileri ile orantılı

olmamasından yani sistemin deęişkenlerinden bir tanesi üzerinde yapılacak küçük bir farklılığın sistem üzerinde büyük farklılık meydana getirebilme olasılığından dolayı süperpozisyon prensibine uymazlar. Kompleks sistemler, lineer olmayan (nonlinear) sistemler için bir örnek teşkil eder.

Zamanla bazı evrilme kuralları çerçevesinde sistemlerin durumlarında deęişiklik meydana gelir. Dinamik sistemler olarak adlandırılan bu sistemlerdeki evrilme kuralları, sistemin gelecekte nasıl bir durumda bulunacağını tanımlayan diferansiyel denklemlerdir. Bu sebeple kompleks sistemler, alt birimlerden bazıları üzerine deęişiklik yapıldığı taktirde kendini, yapılan deęişikliğe adapte ederek varlığını devam ettirebilen dinamik sistemlerdir [1].

Kompleks sistemler, birçok disiplinde çalışılmış ve çalışılmaya devam eden ilginç konulardan biridir [2]. Farklı disiplinlerde yapılan çalışmaların ortak paydası, kompleks sistemlerin elemanları arasındaki etkileşimler tarafından inşa edilen ve incelenen sistemlerin doğası hakkında bilgiler sunan yüksek boyutlu *engebeli peyzaj (rugged landscape)*'in varlığıdır. Engeli peyzaj, camsı sistemlerin termodinamiği [3–6] üzerine çalışan bir *fizikçi* için spin camları, evrim hakkında araştırma yapan bir *biyolog* için türlerin uygunluk (*fitness*) peyzajı [7–9] ve bazı optimizasyon problemleri [10–12] hakkında çalışan bir *bilgisayar mühendisi* için optimizasyon problemindeki izin verilen konfigürasyonlar takımı olarak yorumlanmaktadır.

Fizikte spin camları, homojenliğin olmadığı en basit düzensiz malzemelerdir [13]. Fiziksel olarak rastgele dağıtılmış manyetik momentler, aralarındaki ferromanyetik ve antiferromanyetik etkileşmelerin rekabeti neticesinde düşük sıcaklık durumunda yani taban durumunda manyetik olarak düzensizlerdir. Manyetik momentler arasındaki etkileşmelerin ferromanyetikten antiferromanyetiğe deęişmesi ya da bu durumun tam zıttı bağlardan bağlara deęiştiği için etkileşmeler ne tamamen ferromanyetik ne de tamamen antiferromanyetik olarak alınabilir. İstatistik mekanikte sistemlerin fiziksel olarak kararlılığını karakterize eden serbest enerji kavramı, nonlinear dinamik sistemler için deęişik yaklaşımlarla incelenmiştir. Bu yaklaşımlar için anahtar unsur, dinamik sistemlerin termodinamik formalizmidir [14]. Sistemin içinde bulunduğu durumlar, enerji peyzajında çukurlarla temsil edilir. Bu çukurlardan en derin olanı en düşük serbest

enerjiye sahip olanıdır ve sistemin içinde bulunduğu en kararlı durumdur. Bu nedenle, spin camları kompleks sistemler olarak incelenmiş ve faz geçişlerinin doğası üzerine yapılan çalışmalar oldukça ilginç bulunmuştur [15–19]. Biyolojide spin camlarında gözlemlenen bu faz geçişlerinin bir örneği ise, amino asit polimerleri olan proteinlerin doğal yapısında bulunan bazı alt birimlerin yüksek sıcaklık ve radyasyon gibi dış etkilere maruz kalarak denatürasyona (doğal yapının bozulması) uğramasıdır [20]. Bu sistem, bir çok amino asit grubunun toplu bir davranışını sergilediği için kompleks sistem olarak adlandırılır. Amino asit grupları arasındaki etkileşimleri ideal sistemlerdeki gibi ihmal edemeyeceğimiz için bu sistemi incelemek oldukça karışıktır.

Kompleks sistemlere verilebilecek bir diğer önemli örnek ise ekosistemdir. Ekosistem, canlı (biyotik) organizmaların çevrelerinde bulunan hava, su ve toprak gibi cansız (abiyotik) bileşenlerle etkileşim içerisinde olduğu bir topluluktur. Biyotik organizmalar, izole edilmiş yani bir başına bırakılmış ortamda varlıklarını sürdüremezler. Organizmaların ve ekosistem fonksiyonlarının bir bütün olarak varlıklarının devam edebilmesi için organizmaların birbirleri ve çevreleri ile etkileşimleri temel bir durumdur. Bu nedenle ekosistem, kompleks bir etkileşme ağı ile tanımlanır. Kompleks bir sistem olan ekosistemin alt birimleri olarak tanımlanan biyotik organizmalar, birey veya tür seviyesinde düşünülebilir. Organizmalar arasındaki biyolojik etkileşimler ise hem aynı türlerin bireyleri arasında, hem de farklı türlerin bireyleri arasında olabilir. Biyolojik etkileşimler genel olarak, *doğrudan* ve *dolaylı* etkileşimler olmak üzere iki temel başlık altında kategorize edilebilir [21]. Doğrudan etkileşimler, bir bireyin bir başka bireyle olan ilişkisinde fayda veya zararı birbirlerinden gördükleri etkileşimlerdir. Dolaylı etkileşimler ise, bireyler arasındaki etkileşimlerin herhangi bir taşıyıcı ya da üçüncü bir birey üzerinden yapıldığı etkileşimlerdir.

Bir bireyin *uygunluk (fitness)* değeri, o bireyin çevresine ne kadar iyi uyum sağladığının (adaptasyon) bir ölçüsüdür. Matematiksel olarak uygunluk değeri, bir bireyin sahip olması beklenen yetişkinlik çağına ulaşmış çocuklarının sayısı ile ilişkilendirilebilir. Belirli bir türün bireylerinin sahip olduğu uygunluk değerinin belirlenmesinde diğer türlerin önemli bir role sahip olduğu düşünülebilir. Çünkü ekosistemi şekillendiren etkenlerden bir tanesi, bireyler arasındaki doğrudan ya da dolaylı etkileşimlerdir. Tanımlaması oldukça kolay olsa da, bir bireyin uygunluk değerini belirlemek oldukça

zordur. Bunun nedeni uygunluk değerini belirlerken, sadece aynı genotipe sahip bireylerin sayısı yeterli değildir. Aynı zamanda topluluktaki aynı genotipe sahip bireylerin *sıklığı* (*frequency*) yani topluluktaki toplam bireyler içerisinde belirli bir genotipe sahip bireylerin sayısı, belirli bir bölgede bulunan topluluk içerisinde belirli bir genotipe sahip bireylerin *yoğunluğu* (*density*) gibi konular da uygunluk değerini belirlerken anahtar rollere sahiptirler [22]. Bu faktörler doğal olarak toplulukta bulunan diğer bireylerin varlığı ve bu bireyler arasındaki etkileşmelerle alakalıdır. Bireyler arasındaki etkileşmelerin, topluluğun genelinde nasıl bir etki oluşturduğunu görmek için uygunluk peyzajına bakılır. Uygunluk peyzajı, bireylerin sahip oldukları uygunluk değeri üzerinden genotip uzayının modellenmesidir. Bireylerin uygunluk değerinin kolayca belirlenememesinden dolayı uygunluk peyzajı da kolay bir şekilde bulunamayacaktır. Bu sebeple doğru uygunluk peyzajının genel özellikleri bilinemeyecektir [23]. Uygunluk peyzajı, spin camlarındaki enerji peyzajına benzer bir mantığa sahiptir. Aralarındaki fark, enerji peyzajı enerji çukurları ile temsil edilirken uygunluk peyzajı uygunluk tepeleri ile temsil edilmesidir. Enerji peyzajında en derin enerji çukuru en kararlı durumu gösterirken, uygunluk peyzajında çevreye en iyi uyum sağlayan türler en yüksek uygunluk tepesine sahip olacaktır. Kısacası enerji peyzajındaki enerji değerleri, uygunluk peyzajındaki negatif uygunluk değerleri olarak düşünülür [24].

Karmaşıklık derecelerine göre çeşitli uygunluk peyzajları vardır. Bunlardan en basit olanı Fujiyama peyzajıdır [25]. Fujiyama peyzajında tek bir maksimum uygunluk değeri vardır. Peyzajın herhangi bir yerinde bulunan organizmalar, ortama sağladıkları uyumu artırarak her zaman bu maksimum uygunluk tepesine ulaşırlar. İlk kez Derrida [26] tarafından, düzensiz bir sistem olan ve oldukça karmaşık enerji peyzajına sahip olduğu bilinen spin camları rastgele bir enerji modeli olarak incelenmiştir. Kauffman [27] tarafından geliştirilen NK peyzajı da bir başka karmaşık uygunluk peyzajıdır. N uzunluğundaki genom zincirinde, genler arasındaki etkileşme parametresi olan K'nın değerine bağlı olarak peyzajın karmaşık olup olmayacağına karar verilir. K'nın artması karmaşıklığı arttırarak peyzajda birçok maksimum uygunluk değerinin olmasına neden olurken $K = 0$ değerinde peyzaj Fujiyama peyzajı olacak ve tek bir maksimum uygunluk değeri olacaktır. K'nın limit değeri olan $K = N - 1$ durumunda ise, NK peyzajı Derrida'nın sıfır sıcaklıktaki rastgele enerji modeline eşit olacaktır [28]. Genetik evrim ve uyum sağlama

(adaptasyon) için geliştirilen NK peyzajında, her bir tür komşuları içerisindeki en yüksek uygunluk değerine sahip türe mutasyon yapacaktır. Böylelikle peyzajda meydana gelen değişiklikler gözlenebilecektir.

Yaşamları boyunca doğada bulunan her bir birey, gerek ait oldukları türün diğer bireyleri ile gerekse farklı türlere ait bireylerle etkileşim içerisinde. Özellikle rekabet halinde olan (*competitors*) veya aralarında av-avcı (*prey-predator*) ilişkisi olan türler, birbirlerinin uygunluk değerlerine etki ederek evrimsel değişikliğe zorlayabilirler [23, 29]. Bireylerin bulundukları ortama uyum sağlayıp hayatta kalabilmek ve nesillerini devam ettirebilmek için yapacakları bu değişiklik, *eş-evrimleşme* olarak adlandırılır. Daha açık bir ifade ile eş-evrimleşme, bir popülasyondaki bireylerin sahip oldukları özelliklerden birinde, bir başka popülasyondaki bireylerin sahip oldukları bir özelliğe karşı gerek savunma gerekse avlanma için meydana gelen evrimsel bir değişikliktir [30].

Eş-evrimleşme olayı çeşitli modellerle incelenmiştir. Neumann ve Morgenstern [31] tarafından geliştirilen oyun teorisi (*Game Theory*), Smith [32] tarafından biyolojik topluluklara uyarlanmış ve eş-evrimleşme *Evrimsel Oyun Teorisi* olarak incelenmiştir. Evrimsel oyun teorisinde, biyolojik topluluktaki bireyler *oyuncular*, bireylerin genotipleri *stratejiler* ve bireylerin sahip oldukları uygunluk değerleri de *ödenen bedel* olarak düşünülmüştür. İyi bir model kurmanın dışında evrimsel oyun teorisinin esas amacı, evrimsel olarak kararlı stratejiler (ESS) olarak adlandırılan uygunluk peyzajında yüksek uygunluk değerlerine sahip bireyleri bulmaktır. Eğer bir topluluktaki bireyler çevrelerine uyum sağlamışsa, uygunluk değerlerini arttıracakları başka bir strateji bulunmayacaktır. Kauffman ve Levin [33], uygunluk peyzajındaki maksimum uygunluk değerini bulmak için *uyumlu adımlar (adaptive walks)* denilen bir teori kurmuşlardır. Model, genom uzayındaki konumların uygunluk değerlerinin belirlenmesi için, konumlardan biriyle başlayarak, tek bir mutasyonla ulaşabileceği en yakın komşularının uygunluk değerlerine bakar. Eğer kendi uygunluk değerinden daha büyük bir uygunluk değerine sahip bir konum bulursa o konuma geçerek uyum sağlar. Bu yürüyüş sırasında birçok *yerel (local) maksimum* bulunur. Ancak uygunluk peyzajında tek bir *evrensel (global) maksimum* vardır. Modelin amacı, bu evrensel maksimumu bulmaktır. Kauffman'ın NK modeline oldukça benzer bir model Bak ve Sneppen modelidir [34]. Bu modelde NK modelindeki K değeri $K = 2$ şeklindedir. Bir çembere N tane tür yerleştirilir ve her bir türe 0 ile 1

arasında rastgele uygunluk değerleri verilir. Daha sonra en düşük uygunluk değerine sahip tür seçilir ve bu türün uygunluk değeri, rastgele yeni bir uygunluk değeri ile değiştirilir. Bu değişim sonunda bu türün etkileştiği sağındaki ve solundaki ($K = 2$) en yakın iki komşusu da yeni bir uygunluk değerine sahip olacaktır. Bu uzun hesaplamalar sonunda, türlerin hayatta kalabilmek için ihtiyaç duydukları en düşük uygunluk değerleri bulunur. Amaral ve Meyer [35] tarafından kurulan bir model ise av-avcı (prey-predator) ilişkisine dayanır. Modelde ekosistem, organize bir *yiyecek ağı* olarak düşünülür. Türlerin sahip oldukları uygunluk değerleriyle ilgilenmeden fosil kayıtlarında görülen yok olmaların incelenmesi amaçlamaktadır. Ekosistem farklı tabakalar halinde tanımlanır. Av ve avcılar bu tabakalardaki boşluklara rastgele yerleştirilirler. Yukarı tabakalara çıkıldıkça, en üst tabakada hiç bir avcısı olmayan avcılara, aşağı tabakalara inildikçe de en alt tabakada hiç bir avı olmayan avlara ulaşılır. Komşu tabakalarda bulunan av ve avcılar arasında ise bir etkileşme tanımlanmıştır. Eğer bir avcının yiyebileceği avlar tükenmiş ise, bu durum avcının öleceği anlamına gelmektedir. Bu işlem en alt tabakadan en üst tabakaya doğru tekrarlanır.

Bu tez çalışmasının bel kemiğini oluşturan eş-evrimleşme modeli ise, Christensen et al. [36] tarafından kurulan Dolaşık Doğa Modelidir (DDM). DDM, fiziksel çevredeki değişikliği ihmal eden birey seviyesinde bir eş-evrimleşme modelidir. Birey seviyesinde düşünülmesinin nedeni, tür seviyesinde tanımlanan eş-evrimleşme modellerinde incelemenin zor olduğu türlerin ortaya çıkışı problemini inceleyebilmektir. Çünkü ekosistem tür seviyesinde incelenirse, belirli bir türe ait birey öldüğünde bütün bir tür ölmüş olacaktır. Ancak birey seviyesinde incelenen bir ekosistemde sadece belirli bir birey ölmüş olacak ve türün geriye kalan bireyleri yaşamlarına devam ederek yeni türlerin ortaya çıkmasına neden olabilecektir. Bu durum, ekosistemin gerçek doğası ile daha çok uyumludur. Modelde, bireylerin ait oldukları türler genom uzayındaki konumlarla temsil edilir. Her bir konumda birden fazla birey bulunabilir. Bireylerin sahip oldukları uygunluk değeri ise, genom uzayındaki konumlar içinde bu bireyin ait olduğu konumun dışındaki diğer konumların doluluğuna bağlıdır. DDM'nin diğer modellerden bir farkı, bireyler arasındaki tek bir çeşit etkileşimi değil, olabilecek farklı etkileşimleri barındırmasıdır.

Ekolojik sistemlerde bulunan biyolojik topluluklar arasındaki etkileşimlerden dolayı, gerek evrimsel gerekse de ekolojik sistemler, kompleks topolojili dinamik ağlar olarak

düşünülebilirler. Bu düşünce ile birlikte, ekolojik ve evrimsel biyolojide etkileşmelerin hayati öneme sahip olmasından dolayı, son zamanlarda biyolojik sistemlerin kompleks ağlar olarak incelenmesinde yoğun ilgi vardır [20]. Topluluktaki bireylerin çevrelerine sağlayacakları uyumun, ağ etkileşimleri altında şekillenen hem çevresel değişikliklere hem de doğal seçilime bağlı olduğu söylenebilir. Bu nedenle ekolojik ağlar kompleks olmasına rağmen, komplekslik ve ekolojik denge arasındaki ilişkinin anlaşılmasını sağlayarak, ekolojik etkileşmeler altında yatan mekanizmayı aydınlatır [37]. Kompleks ağ yapıları için öncü bir çalışma Erdős ve Rényi [38, 39] tarafından geliştirilen *rastgele ağ* modeli olmuştur. Bu modele göre, kompleks ağda bulunan noktalar birbirlerine eşit ihtimaliyetle bağlanırlar ve böylece rastgele bir ağ oluşturulmuş olur. Bu model kompleks ağ çalışmaları için kılavuz olma niteliği taşır. Ancak kompleks sistemler üzerine artan ilgiden dolayı, gerçek dünyada bulunan kompleks sistemlerdeki ağ yapılarının hepsinin rastgele olup olmadığı tartışılmaktadır. Bu tartışma ile ilgili ilk düşünce Milgram [40] tarafından ortaya atılan *küçük-dünya (small-world) ağı* düşüncesi olmuştur. Milgram yaptığı incelemede, Amerika'daki insanlar arasındaki en uzak mesafenin 6 kişi olduğu sonucuna varmıştır. Küçük-dünya ağları ile ilgili bir diğer çalışma Watts ve Strogatz [41] tarafından yapılmıştır. Bu çalışmada, kompleks olan bir çok biyolojik, teknolojik ve sosyal ağların ne tamamen düzenli ne de tamamen rastgele olmadığı düşüncesine varılmıştır. Model ilk başta tamamen düzenli bir ağ ile başlar ve daha sonra ağdaki bağlantıları rastgele yeniden bağlayarak yeni bir ağ yapısı oluşturur. Elde ettikleri sonuçlar, Milgram'ın sonuçları ile uyumludur. Barabási ve Albert [42], kompleks topolojili dinamik ağlar için yeni bir model kurmuşlardır. *Ölçeksiz (scale-free) ağ* olarak adlandırılan bu modelde ise, ağda bulunan bireylerin bağlantılı oldukları komşu sayılarının dağılımının bir kuvvet-kanunu (power-law) davranışı sergileyeceğini belirtir. Bu modelin iki temel unsuru vardır. Bunlardan ilki, ağa yeni bireylerin eklenmesi ile ağın sürekli *genişlemesi* ve ikincisi eklenen bu yeni bireylerin ağda en avantajlı konum hangisi ise ona *tercihli eklenmesidir*.

Dinamik kompleks sistemler incelenirken sistemin doğasının anlaşılması için kompleks ağ yapılarının incelenmesi oldukça önemli bir konudur. Özellikle kompleks bir sistem olan ekosistemde, türlerde gerçekleşecek evrimsel değişiklikler neticesinde biyolojik toplulukların genel yapısındaki değişiklikler, dinamik ağ yapıları incelenerek

açıklanabilir. En iyi bilgilerimiz dahilinde, biyolojik sistemlerin tamamı için geçerli olan kapsamlı bir dinamik kompleks ağ çalışması yapılmamıştır.

Bu tez çalışmasının amacı, ekolojik çevrede evrimleşen biyolojik toplulukları DDM ile oluşturmak ve bu evrimleşen biyolojik topluluklar tarafından kurulan dinamik ağ yapılarını, ağların yerel ve evrensel istatistik özellikleri ile açıklamaktır. Ayrıca, biyolojik topluluklar için yapılacak simülasyon sonucunda türlerin gerçek yaşam ağacı elde edilip şu an literatürde kullanılan ağaç metotları test edilecektir.

Bu giriş bilgilerinden sonra Bölüm 1’de, evrimsel ekoloji ile ilgili temel bilgiler verilecektir. Evrim olayının gerçekleşmesi için gerekli olan değişim mekanizmaları üzerinde durulacaktır. Bu mekanizmalar neticesinde gerçekleşen eş-evrimleşme ve fosil kayıtlarında gözlemlenen biyolojik çeşitlilik ile ilgili oldukça temel bir anlatım yapılacaktır. Jeolojik zamanlarda görülen çeşitlilik ve kitlesel yok olma olayları için, fosil kayıtlarında gözlemlenen değerleri elde etmeye çalışan bazı evrim modelleri anlatılacaktır. Türlerin sınıflandırılması ve soy ağaçları hakkında terimsel düzeyde bilgiler sunulacaktır. Ağ yapısı oluşturan ekosistem incelemeleri için, ağ yapılarının matematiksel temeli, ağ yapılarının analiz edilmesinde kullanılan derece dağılımı (degree distribution), ortalama yol uzunluğu (average path length), çap (diameter), kümelenme katsayısı (clustering coefficient) ve modülerite (modularity) gibi ağ parametreleri yine bu bölüm çerçevesinde anlatılacaktır. Ayrıca gerçek dünya ağlarını incelemek için kurulan literatürdeki bir takım önemli ağ modelleri anlatılacaktır.

Bölüm 2’de, biyolojik türlerin sahip oldukları soy ağaçları için terminolojik bilgiler verilecektir. Literatürde soy ağacı tahminlerinde kullanılan çeşitli metotlar içerisinde iki metodun genel prensipleri üzerinde durulup matematiksel işleyişleri anlatılacaktır.

Bölüm 3’de, bu tez çalışmasının iskeletini oluşturan DDM hakkında genel bilgiler verildikten sonra modelin matematiksel işleyişi anlatılacaktır. Modelin dinamiğini oluşturan, biyolojik topluluklardaki bireylerin etkileşmesi, bireylerin üremesi, mutasyona uğraması ve ölmesi gibi temel parametreleri hakkında bilgi verilecektir. Simülasyonun zaman birimi olan nesil zamanı açıklanarak, her bir nesil zamanı için yapılan işlemlerin akış diyagramı verilecektir.

Bölüm 4’de, bu tez çalışmasında yapılan simülasyon sonuçları sunulacaktır. Öncelikle tek bir simülasyon sonucu yorumlanacak ve sonrasında biyolojik toplulukların genel doğası hakkında anlamlı sonuçlar elde edebilmek için gerekli olan çok sayıda simülasyonların ortalama hesapları verilecektir. Tek bir simülasyon yapılırken, daha geniş bir toplulukla çalışma düşüncesi ile daha büyük bir genom uzayında çalışılacak, ortalama alınırken ise simülasyon süresi düşünülerek daha küçük bir genom uzayında çalışılacaktır. Her iki hesap için de, topluluktaki çeşitlilik ve popülasyon dalgalanması ve dinamik ağın parametreleri için sergilenen genel davranışlar verilerek literatürdeki sonuçlarla karşılaştırma yapılacaktır. Bunların dışında, türlerin belirli bir nesilsel zaman aralığında sahip oldukları soy ağacı verilecek ve bu ağaç, Bölüm 2’de anlatılan ağaç metotları ile karşılaştırılarak bu metotların tahminlerinin doğruluğu üzerinde durulacaktır.

Son bölümde ise, yapılan çalışmalar kısaca özetlenerek, elde edilen sonuçlar üzerine tartışma yapılacaktır.

1. BÖLÜM

EVİRİMSEL EKOLOJİ VE AĞ SİSTEMLERİ

Yaklaşık 4.5 milyar yıl yaşında olan Dünyamızda hemen hemen herkesin dikkatini çeken konu, Dünya üzerinde bulunan türlerin son durumlarına nasıl ulaştıklarıdır. Bu konu ile ilgili biyolojik evrimin ilk yazılı bildirisi 19.yy'ın başlarında Lamarck [43] tarafından yapılmıştır. Lamarck'ın evrim teorisi iki temel düşünceyi içerir. *i)* Bireyler kullanmadıkları veya ihtiyaç duymadıkları karakteristik özelliklerini kaybederken, kullandıkları karakteristik özellikleri geliştirirler. *ii)* Bireyler kazandıkları karakteristik özellikleri kalıtsal olarak sonraki nesillere aktarırlar. Lamarck, türlerin nesillerinin tükenmediğini belirtmekle birlikte, türlerin yavaş yavaş diğer türlere dönüştüğünü ileri sürerek yaşam formlarının zaman geçtikçe daha karmaşık bir yapıya dönüştüğünü belirtmiştir [44]. 19.yy'ın ortalarında Darwin [45] yayınladığı *Türlerin Kökeni* adlı kitabında herkesin dikkatini çeken bu konu için *doğal seçim* adı verilen bir süreç yardımıyla evrim teorisi önermiştir. Darwin'in bu teorisi üç temel faktöre dayanır [46]. *i)* Bir popülasyondaki farklı bireyler, farklı morfolojiye, farklı fizyolojiye ve farklı davranışlara sahiptir. Bu faktör fenotipik farklılıklar olarak da adlandırılabilir. *ii)* Farklı çevrelerde bulunan farklı fenotiplere sahip bireyler, farklı hayatta kalma ve çoğalma oranlarına sahiptir. Bu faktör ise uygunluk değerlerindeki farklılığa dikkat çekmektedir. *iii)* Sahip olunan özellikler açısından nesiller arasında bir korelasyon vardır. Bu son faktör ise, uygunluk değerinin kalıtsal bir durum olduğunu vurgulamaktadır. Bu üç faktör geçerli olduğunda, popülasyon evrimsel değişikliğe uğrayabilir. Darwin, yaşam formlarındaki karmaşık yapı artışını basitçe, canlıların nesilden nesile çevrelerine daha iyi uyum sağlamalarının bir sonucu olarak açıklamıştır [45]. 19.yy'ın sonlarında Darwin'in doğal seçim düşüncesini benimseyen Wallace [47] ve Weismann [48], Lamarck'ın düşüncesine karşı çalışmalar yapmış ve eşeyli üremenin her nesilde değişken bir birey topluluğu

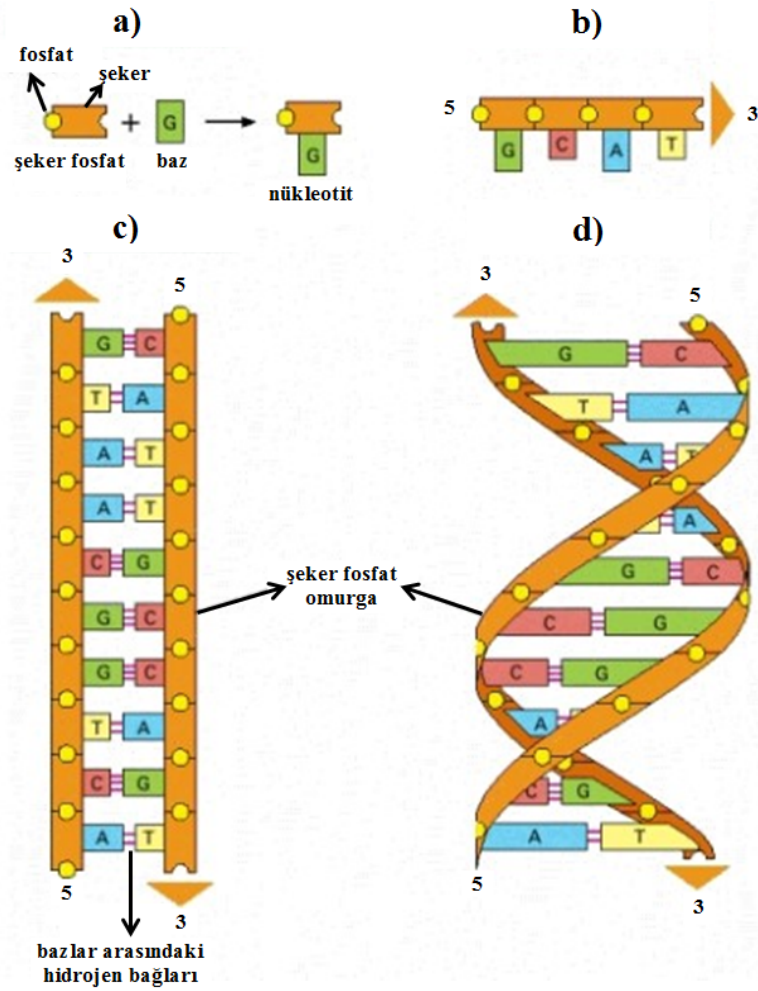
oluşturacağını ve doğal seçilimin bu farklılaşan topluluk üzerinde rol oynayarak evrimsel değişikliğe neden olduğunu öne sürmüşlerdir [49]. Darwin'in evrim teorisinin ortaya çıkmasından sonra Mendel [50], bezelyelerin kalıtımı üzerine çalışarak, sahip olunan bazı özelliklerin nesilden nesile aktarıldığını kaydedip kalıtım kuramının bütün canlılar için geçerli olduğunu belirlemiştir. Böylelikle Darwin kuramı ve Mendel'in kalıtım çalışması bütünleştirilerek Darwin teorisinin genişletilmiş hali olan *Neo-Darwinizm* düşüncesi ortaya çıkmıştır.

1.1. Evrimsel Değişim Mekanizmaları

Mendel'in bütün canlılar için geçerli olan kalıtım kuramını [50] keşfetmesinden sonra, Tamm et al. [51] kalıttan sorumlu nitrojen bazları üzerine bir eşleme metodu uygulayarak DNA (Desoxyribo Nucleic Acid) molekülünü tanımladı. Metotta, pürin ve pirimidin olarak sınıflandırılan nitrojen bazlarının birbirleri ile karşılıklı olarak eşleştiği bulundu. Tamm et al. tarafından yapılan bu çalışmadan sonra Watson ve Crick [52], DNA molekülünün "çift sarmal" yapıda olduğunu buldular.

Her canlı organizma canlılık işlevleri ve biyolojik gelişimleri için gerekli olan, genetik karakteristikleri taşıyan bir DNA'ya sahiptir. Kompleks bir yapıya sahip olan DNA, nükleotit olarak isimlendirilen birimlerin diziliminden oluşan bir polimer zincirdir. Nükleotit birimler bir şeker, bir fosfat ve bir bazdan oluşurlar [53]. Adenin (A), Sitozin (C), Guanin (G) ve Timin (T) bazları, nükleotit birimlerinin baz kısmını oluşturan pürin ve pirimidin türevleridir. Bu dört adet bazdan A ve G birer pürin, T ve C ise birer pirimidin türevidir. Birbirine sıkıca sarılı bir çift molekülden oluşan iki uzun ipliksi yapı, birbirine sarmaşık gibi sarılarak bir DNA çift sarmalı oluştururlar. Sarmallarda bulunan bazlar arasındaki hidrojen bağları sayesinde, DNA sağlam bir yapıya ulaşır.

Şekil 1.1'de, DNA molekülünü oluşturan bütün yapılar ve bu yapılar ile DNA molekülünün adım adım kuruluşu gösterilmiştir. İlk olarak, tek bir fosfat grubunun eklendiği 5 karbon şeker ile nitrojen içerikli bir bazın bileşiminden oluşan bir nükleotit yapısı oluşturulmuştur. Baz, şeker fosfat kanalına kovalent bağ ile bağlanmaktadır [53]. Daha sonra bu nükleotit birimleri birbirine, bir köşesi fosfatlı diğer köşesi fosfatsız olacak şekilde kimsayal bir kutup boyunca bağlanarak DNA molekülünün tek bir ipliği



Şekil 1.1. DNA ve onu oluşturan yapıların gösterimi. **a)** DNA molekülünü oluşturan nükleotit birimlerinin içeriği. **b)** DNA molekülünün tek bir ipliğini oluşturan nükleotit dizileri. 5 sayısı nükleotidin fosfatlı kısmını, 3 sayısı ise fosfatsız kısmını göstermektedir. **c)** Nükleotit dizilerindeki bazların hidrojen bağları ile bağlanması sonucu oluşan DNA çift ipliği. **d)** DNA molekülünün ters kutuplu çift iplikten oluşan sarmallı gerçek yapısı. Ref.[53]

oluşturulmuştur. Son olarak, bazları birbirine hidrojen bağı ile bağlanmış olan ters kutuplu iki DNA ipliği ile DNA molekülünün çift sarmallı yapısı gösterilmiştir. A - T bazları arasında ikili, C - G bazları arasında üçlü hidrojen bağları bulunmaktadır.

Organizmalar üreme gerçekleştirirken, DNA dizilerini kopyalar ve sahip oldukları karakteristik özellikleri yavrusuna aktarır. DNA dizilerinin kopyalanması sırasında, genetik birimler üzerinde bir takım değişimler oluşabilmektedir. Bu değişimler evrimsel olayların ortaya çıkmasına neden olmaktadır.

Evrimsel olaylar, bir canlı popülasyonunun genetik kompozisyonunun bir takım evrimsel mekanizmalar ile zamanla değişmesi ile gerçekleşirler. Bu mekanizmaları oluşturan alt birimlerin incelenmesi, evrim sürecinin anlaşılması için büyük öneme sahiptir. Bu kesimde, bu evrimsel değişim mekanizmaları üzerinde oldukça kısa bir anlatım yapılacaktır.

1.1.1. Mutasyon

Mutasyon, bir organizmanın DNA dizilimindeki nükleotit birimlerinin, DNA onarımı, mayoz bölünme, radyasyon ve bazı kimyasal ilaçlar gibi sebeplerden dolayı bozulması ile meydana gelen değişimlerdir. Mutasyona sahip bir organizma mutant olarak adlandırılır. Mutasyonlar genel olarak, *i*) Vücut hücreleri ve *ii*) Eşey (üreme) hücrelerinde gerçekleşen mutasyonlar olmak üzere ikiye ayrılır [54]. Bedensel (somatik) mutasyonlar, doku hücrelerinde gerçekleştiği için kalıtsal değildir ve nesilden nesile aktarılmaz. Germ hattı mutasyonları olarak da adlandırılan eşey (üreme) hücrelerinde görülen mutasyonlar kalıtsaldır ve sonraki nesillere aktarılır.

DNA dizilerinin kopyalanması sırasında dört tip mutasyon gerçekleşebilmektedir [55]. Gen seviyesinde tanımlanan birinci tip mutasyon *baz değişimidir*. DNA dizilerindeki tek bir nükleotit biriminde bulunan bazın başka bir baz ile değişmesi neticesinde oluşur. Baz değişimi, her ne kadar bir gendeki tek bir dizi için gerçekleşmiş olsa da, bu tip bir mutasyon protein üretimi için farklı etkiler oluşturabilir. Bu yüzden baz değişim mutasyonu üç gruba ayrılır. Birinci grup, kalıtsal bilginin okunması sırasında protein molekülüne yanlış bir amino asidin eklenmesine neden olur. İkinci grup, protein molekülünün üretilmesinin erken bitirilmesine neden olur. Üçüncü grupta ise, baz değişiminden etkilenmemiş bir şekilde aynı amino asit grubu üretilir. İkinci tip mutasyon *nükleotit eklenmesidir*. DNA dizisine ekstra nükleotit eklenmesi sonucu meydana gelir. Bunun neticesinde DNA dizilerinde kaymalar oluşacaktır. Üçüncü tip mutasyon *nükleotit çıkarılmasıdır*. Bu tip mutasyonlarda, kopyalanan DNA dizisinden nükleotit çıkarılması ile yine dizilerde kaymalar meydana gelecektir. Dördüncü tip mutasyon *çerçeve kaymasıdır*. Bu tip mutasyonlar, kalıtsal bilginin okunması sırasında bir proteinin kodon ya da üçlü nükleotit grubu olarak adlandırılan amino asit dizilerinin değişmesine neden olmaktadır. Bu nedenle bireylerin DNA dizilerinde gerçekleşecek mutasyonlar

genetik çeşitlilik için önemli bir role sahip olacaktır.

Mutasyonlar türler arasında farklı oranlarda gerçekleşmektedir. Örneğin, bir RNA virüsünün DNA diziliminde tipik olarak 10 bin baz çifti bulunmaktadır. Bu virüs için mutasyon oranı, her nesilde baz çifti başına yaklaşık olarak 10^{-4} şeklindedir [56]. İnsan DNA'sında ise yaklaşık olarak 3 milyar baz çifti bulunmaktadır. İnsan da gözlenen mutasyon oranı ise, her nesilde baz çifti başına yaklaşık olarak 10^{-8} şeklindedir [57].

Oluşan mutasyonların organizmalar üzerinde her zaman negatif ya da pozitif etkilere sahip olduğu söylenemez. DNA diziliminde 120 milyon baz çifti bulunan *Drosophila melanogaster* (sirke sineği) üzerinde yapılan çalışmalar, gen dizilimleri tarafından oluşturulan bir protein mutasyonunun yaklaşık olarak % 70'inin zararlı etkiye sahip olduğunu ve geri kalan kısmının ise, nötr veya zayıf faydalı etki gösterdiğini ortaya koymaktadır [58].

1.1.2. Gen Akışı

Gen akışı, bir popülasyondaki genlerin başka bir popülasyona başarılı bir şekilde taşınmasıdır [59]. Başka bir ifade ile, bir popülasyondaki bireylerin başka bir popülasyona göç etmesidir. Bu sayede gen akışının gerçekleştiği popülasyonun gen havuzunda, farklı DNA dizilimlerine sahip yeni genetik çeşitler ortaya çıkabilmektedir. Popülasyonlar arasında karşılıklı olarak yeterli seviyede gen akışı gerçekleştiği takdirde, popülasyonların genetik farklılıkları azalacaktır [59]. Gen akışını türleri bir arada tutan bir yapışkan gibi düşünen Mayr [60], gen akışından dolayı verilen bir popülasyon için nesil başına genetik değişikliğin en az 10^{-3} ya da 10^{-2} gibi bir değer olarak tahmin etmiştir. Gen akışı ile ilgili yapılan bir diğer çalışmada [61] gen akışı, hayvanlarda ve bitkilerde ayrı ayrı düşünülerek doğadaki farklı popülasyonlar için gen akışı oranının ortak bir değere sahip olmasından ziyade, popülasyonlardaki bireylerin hareketliliği (mobility) ile orantılı olarak gen akışı oranının değişeceği belirtilmiştir.

1.1.3. Genetik Sürüklenme

Genetik sürüklenme, bir popülasyondaki genetik çeşitlilik sayısındaki değişim olarak tanımlanmaktadır [62]. Kısa zaman periyodunda, popülasyonda büyük değişimlerin

meydana gelmesine neden olabilen rasgele bir süreçtir. Popülasyondaki bireylerin hayatta kalma ya da üreme kabiliyetlerinden bağımsız olarak gerçekleşmektedir. Birkaç bireyden oluşan küçük bir popülasyonda genetik çeşitlilik az olduğu için, genetik sürüklenmenin etkisi büyük popülasyonlara nazaran daha fazladır [63, 64]. Bu etki, popülasyondaki bazı genetik çeşitlerin tamamen kaybolmasından dolayı gerçekleşen genetik çeşitlilikteki azalma şeklinde gözlenmektedir. Bu nedenle nadir görülen bir genetik çeşitliliğin yok olması, yaygın görülen genetik çeşitliliğe göre daha yüksek ihtimaldir. Bunun sonucunda genetik sürüklenme, organizmaları evrimsel değişikliğe zorlayacaktır.

1.1.4. Doğal Seçilim

Darwin'ın [45] popüler hale getirdiği, evrimsel değişim mekanizmalarının köşe taşı olan doğal seçim ifadesi, herhangi bir popülasyondaki bireylerin hayatta kalma veya üreme kabiliyetlerindeki farklılaşmalar olarak tanımlanmaktadır [65]. Fenotip olarak adlandırılan bireylerin görünür karakteristikleri üzerinde etkilidir [66]. Ancak fenotipik özelliklerin temeli bireylerin sahip oldukları genetik yapıları olduğu için, doğal seçim özellikle genetik sürüklenme üzerinde önemli bir role sahiptir. Buradan da anlaşılacağı üzere doğal seçim mekanizmasının işleyebilmesi için genetik çeşitliliğe ihtiyaç vardır. Bir popülasyondaki bireyler hayatta kalmak ve çoğalabilmek için, içinde bulundukları ortama uyum sağlamaları (adaptasyon) gerekmektedir. Doğal seçim, bireylerin gerçekleştirecekleri bu uyum sağlama olayı üzerinde etkili olan bir evrim mekanizmasıdır. Ortama olan uyumluluklarını artıran bireyler, bu karakteristik özelliklerini gelecek nesillere aktaracaktır. Ayrıca doğal seçim değeri birçok nesil için aynı değere sahip olursa, ortama uyumlu olan bireylerin sayısı gittikçe artacaktır. Karıştırılmaması gereken önemli husus, doğal seçilimin kendisinin bizzat evrim olayı olmadığıdır.

1.2. Eş- Evrimleşme

Ekosistemde her bir canlı organizma, benzer ya da farklı organizmalarla ve fiziksel çevrelerinin iklim gibi dinamik bileşenleri ile daima bir etkileşim içerisinde [67]. Organizmalar arasındaki karşılıklı etkileşimler, birbirlerinin evrimleşmelerini etkilemektedir. Bunun sonucunda, karşılıklı etkileşen organizmalardan birinde oluşacak evrimsel bir değişikliğin, diğer organizmanın da evrimleşmesine neden olması durumu

eş-evrimleşme olarak adlandırılmaktadır [68]. Eş-evrimleşme, mikroskobik seviyede bir protein molekülünde bulunan amino asit gruplarından, makroskobik seviyede aynı fiziksel çevrede yaşayan farklı türlere kadar, etkileşen taraflardan birinin diğeri üzerine seçim baskısı uygulaması ile çeşitli biyolojik seviyelerde gerçekleşebilmektedir.

Organizmalar arasındaki çeşitli etkileşmelere göre bazı eş-evrimleşme modelleri kurulmuştur. Valen'in Red Queen hipotezi [69] bunlardan birisidir. Bu hipotezin ana fikri, aralarında konak-parazit ilişkisi bulunan canlıların hiçbir zaman aynı kalmayacağı ve parazit canlının uyum sağlamak için gerçekleştireceği evrime karşılık olarak konak canlının da evrimleşeceği yönündedir. Dawkins ve Krebs [70], aralarında rekabetçi ilişki bulunan canlıların eş-evrimleşmesini türler arası ve tür içi olmak üzere iki sınıfa ayırmıştır. Her iki sınıfta da türlerin hayatta kalabilmek için bir silahlanma yarışına (Arms Races) girdikleri ve bu sebeple eş-evrimleşme geçirdikleri sonucuna varılmıştır.

Eş-evrimleşmenin canlı organizmalar arasında birçok örneği bulunmasına karşın, fosil kayıtlarında kesin olarak gözlenmemiştir. Bunun nedeni, eş-evrimleşmenin Dünya tarihindeki büyük ölçekli değişikliklere neden olmamasıdır [71]. Fosil kayıtlarındaki arılar ve çiçekler arasındaki etkileşimler gibi eş-evrimleşmenin klasik örnekleri bile bu durum için zayıf kanıtlardır. Çünkü her iki organizmanın evrimleşme zamanlarının örtüştüğü konusunda net bir bilgi bulunmamaktadır [72].

1.3. Biyolojik Çeşitlilik

Biyolojik çeşitliliğin biyologlar tarafından en sık yapılan tanımı, bir bölgedeki genlerin, türlerin ve ekosistemlerin toplamı şeklindedir [73]. Yapılan bu tanımdan, biyolojik çeşitlilik için üç faktör ortaya çıkmaktadır. Birinci faktör genetik çeşitliliktir. Dünyada yaşayan bütün canlılar, kalıtım materyali olan genlere sahiptir. Genler, her bir canlının fiziksel ve biyokimyasal karakteristiklerini tanımlayan bilgileri içerir. Böylelikle bir tür içerisinde farklı genetik karakteristikte bireylerin bulunması genetik çeşitliliği ortaya çıkaracaktır. İkinci faktör tür çeşitliliğidir. Tür çeşitliliği, verilen bir topluluk içerisinde bulunan farklı türlerin sayısını vermektedir. Üçüncü faktör ise ekosistem çeşitliliğidir. Ekosistem, canlı organizmaların çevrelerinde bulunan cansız organizmalarla etkileşim içerisinde olduğu topluluktur. Verilen bir bölgedeki bu grupların sayısı ekosistemin

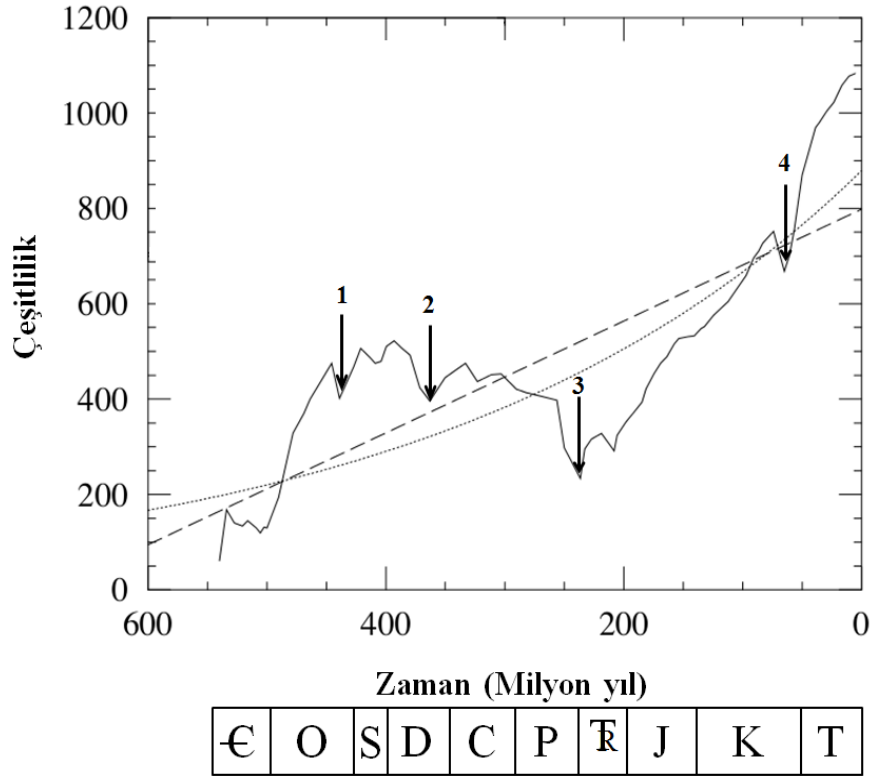
ne kadar çeşitli olduğunu belirlemektedir. Bu üç faktör, biyolojik çeşitliliğin temelini oluşturmaktadır. Biyolojik çeşitliliğin tarihi üzerine çeşitli çalışmalar yapılmıştır [74, 75]. Ancak bu çalışmaların sonuçlarının dikkatle incelenmesi gerekmektedir. Çünkü biyolojik çeşitlilik hakkındaki geçmiş zaman bilgileri, eksiksizlik, doğruluk ve korunma kalitesi gibi tartışmaların bulunduğu fosil kayıtlarından elde edilmektedir. [76]. Fosil kayıtlarından elde edilen bilgilerdeki ana sorun ise, fosilleri birbirinden ayıran zaman aralıklarının çok uzun olması ve verilerin farklı bölgelerdeki birbirinden bağımsız kaynaklardan elde edilmesidir. Bu sebeple benzer çeşitlenme yapılarına, birbirinden bağımsız veri setlerinde karşılaşılabilmektedir. Ancak bütün fosillerin aynı jeolojik faktörlerden etkilenip etkilenmediği tartışılmaktadır. Bu durum için soyoluşsal çalışmalarda, fosil kayıtlarında görünen katmanlar morfolojik olarak analiz edilerek birbirinden bağımsız fosillerin kardeş grupları arasındaki ilişkiler tanımlanmıştır [77–79]. Dünya tarihi boyunca gözlemlenmiş bu tarz veriler arasındaki zamanlamanın ve ilişkilerin tanımlanması için de jeolojik zaman ölçeği kullanılmıştır. Kullanılan bu jeolojik zaman ölçeği Tablo 1.1’de [80] gösterilmiştir.

Tablo 1.1. Jeolojik zaman ölçeği. Ref.[80]

Zaman	Devir	Devre	Çağ	Milyon yıl önce	Sembol
Kriptonozoyik	Arkeozoyik			3600	
	Proterozoyik			2500	
Fanerozozyik	Palezozyik	Kambiriyen		543	Є
		Ordovizyen		500	O
		Silüriyen		439	S
		Devoniyen		409	D
		Karbonifer		354	C
		Permiyen		290	P
	Mezozozyik	Triyas		251	Ƨ
		Jura		206	J
		Kretase		144	K
	Senozozyik	Tersiye	Paleyosen	65	T
			Eosen	55.6	
			Oligosen	33.5	
			Miyosen	23.8	
			Pliyosen	5.2	
		Kuaterner	Pleyistosen	1.8	
			Holosen	0.01	

Biyolojik çeşitlilik üzerine yapılan çalışmalarda ortak kabullenim, çeşitliliğin jeolojik

zaman boyunca arttığı yönündedir. Sepkoski [81] tarafından deniz canlıları üzerinde yapılan bir çalışmada bu durum açıkça görülmektedir. Bu çalışmanın sonucu Şekil 1.2'de [82] gösterilmiştir.

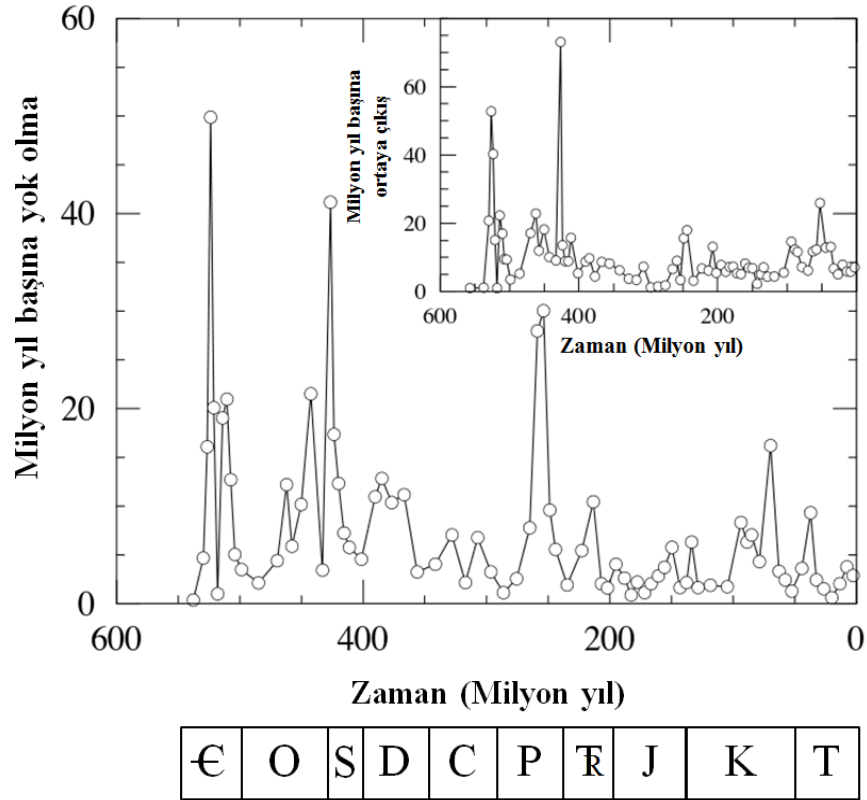


Şekil 1.2. Fanerozoik zaman boyunca, deniz canlıları için çeşitliliğin milyon yıl ölçeğindeki zamana göre değişimi. Kesikli çizgi lineer, noktalı çizgi eksponansiyel fit eğrileridir. Ref.[82]

Şekil 1.2'de Kambiriyen devresinin başlangıcında çeşitlilik sayısında hızlı bir artışın olduğu görülmektedir. Bu hızlı artış, birçok hayvan sınıfının ortaya çıktığı "Kambiriyen patlaması" olarak bilinir [83, 84]. Çeşitlilikteki bu artış Ordovizyen devresi boyunca devam etmektedir. Bu devrenin sonuna doğru 1 numaralı ok ile gösterilen zamanda, çeşitlilik ilk kez ciddi bir azalma sergilemektedir. Gerçekleşen bu azalma, Dünya'nın sahip olduğu oldukça kararlı çevresinin bozulmaya uğradığı şeklinde ifade edilmiştir [85]. Silüriyen devresinde çeşitlilikte tekrar bir artış gerçekleşmektedir. Ancak 2 numaralı ok ile gösterilen Devoniyen devresinin sonunda çeşitlilikte ikinci kez büyük bir azalma görülmektedir. Karbonifer devresinde çeşitlilik hemen hemen sabit kalmakta iken, Permian devresi boyunca hem jeolojik hem de iklimsel etkilerden dolayı 3 numaralı ok ile gösterilen Triyas devresinin başına kadar çeşitlilikte sürekli azalma gerçekleşmektedir.

Triyas devresinin ilk zamanlarında çeşitlilikde yeniden artış başlamış ve Mezozoyik devri boyunca devam etmektedir. Ancak 4 numaralı ok ile gösterilen Mezozoyik devrinin son devresi olan Kretase devresinin sonunda çeşitlilik tekrar azalmaktadır.

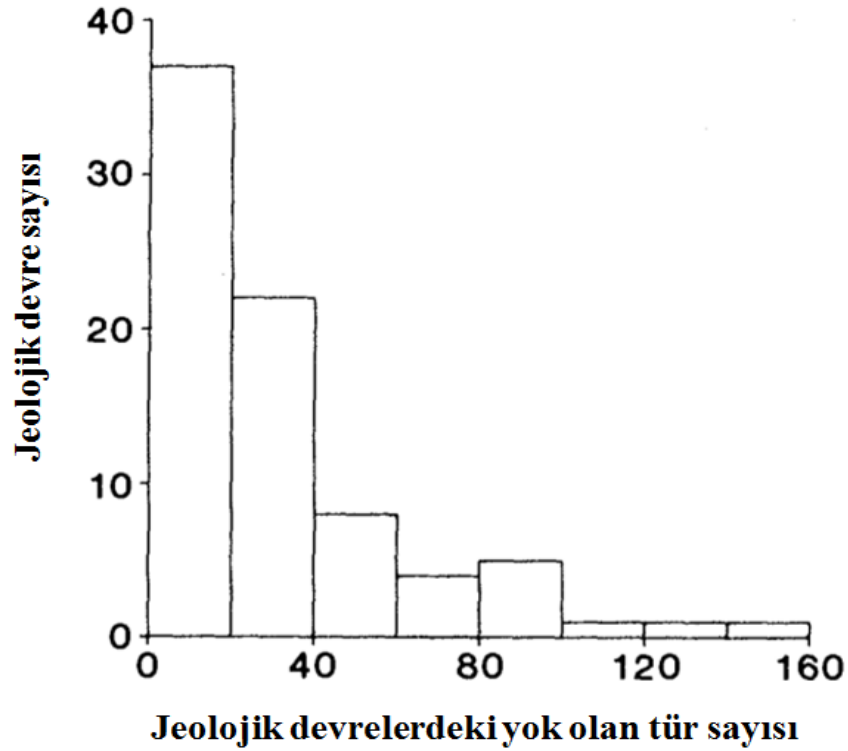
Yine Sepkoski [81] tarafından deniz canlıları üzerinde, yok olma ve ortaya çıkma olayları incelenmiştir. Bu çalışmanın sonucu Şekil 1.3'de gösterilmiştir.



Şekil 1.3. Fanerozoik zaman boyunca, deniz canlıları için milyon yıl başına yok olmanın (ana şekil) ve milyon yıl başına ortaya çıkmanın (küçük şekil) milyon yıl ölçeğindeki zamana göre değişimi. Ref.[82]

Şekil 1.3'de, Şekil 1.2'deki çeşitlilikde azalmaların gerçekleştiği 1, 2, 3 ve 4 nolu oklarla gösterilen jeolojik zamanlarda, kitlesel yok olmalar görülmektedir. Kambriyen devresinin ortalarında yok olma oranında büyük bir artış olmasına rağmen, bu zamandaki çeşitlilik çok büyük olmadığı için kitlesel bir yok olmadan söz edilememektedir. Deniz canlılarının ilk kitlesel yok olmalarını, Şekil 1.2'de 1 numaralı ok ile gösterilen Ordovizyen devresinin sonunda yaşadıkları görülmektedir. Diğer taraftan Şekil 1.3'da ana şeklin içinde verilen küçük şekile bakıldığında, ortaya çıkma oranlarında gözlenen piklerin yok olma oranlarında gözlenen piklerle uyumlu olduğu görülmektedir.

Canlı organizmaların yok olmalarının, yaşadıkları çevrelerinde değişikliğe neden olan çeşitli jeolojik olaylardan kaynaklandığı düşünülmektedir. Organizmalar çevrelerinde meydana gelen bu değişikliğe uyum sağlayamadıkları takdirde yok olmaktadır. Alvarez et al. [86] kitlesel yok olmaları iklimsel değişikliklere bağlayarak, bir türün yok olmasının kitlesel yok olma zincirini başlatacağını ileri sürmüştür. Alvarez et al. [87] bu hipotezlerini genişleterek kuyruklu yıldızların kitlesel yok olmalarını incelemişlerdir. Sepkoski'nin [81] deniz canlılarının fosil kayıtlarındaki yok olma oranı için yaptığı çalışmanın verilerini kullanan Raup [88], Fanerozoik zaman boyunca 79 jeolojik devredeki yok olan tür sayısının dağılımını incelemiştir.



Şekil 1.4. Raup tarafından yapılan, Fanerozoik zamanın 79 jeolojik devresi için yok olan tür sayısının dağılımı. Ref.[88]

Şekil 1.4’de, Fanerozoik zamanın 79 jeolojik devresi için yok olan tür sayısının dağılımı verilmiştir. 79 jeolojik devrenin çoğunda ölen tür sayısı düşüktür. Ayrıca ölen tür sayısının, jeolojik devre sayısı ile ters orantılı olarak değiştiği görülmektedir. Solé et al. [89], Raup’un elde ettiği bu dağılımın üzerine eksponansiyel bir eğri fit etmişlerdir. Bu eğrinin, logaritmik eksenlerde $P(s) \sim s^{-\tau}$ şeklinde kuvvet kanunu dağılımı sergilediğini belirtmişlerdir. Burada s , jeolojik devrelerdeki yok olan tür sayısı ve $P(s)$, jeolojik

devre sayısıdır. Dağılımın üstelini ise $\tau = -2$ olarak elde etmişlerdir. Böylelikle fosil kayıtlarındaki yok olma oranının eksponansiyel olarak azaldığını ifade etmişlerdir.

1.4. Sınıflandırma

Kompleks sistemler hiyerarşik bir yapı sergilemektedirler [90]. Ekosistem de kompleks bir sistem olduğu için hiyerarşik bir düzenleme oldukça kullanışlı olacaktır. Biyologlar, paylaşılan karakteristiklere göre organizmaları sınıflandırmak için kapsamlı bir hiyerarşik yapı kullanmaktadırlar. Bu yapıda yukarıya doğru çıkıldıkça birimler gittikçe azalırken, aşağıya doğru inildikçe sınıflandırılan bütün varlıkları içeren birçok birim olacaktır. Linnaeus [91], bu sistemi kuran ve her bir birime isim veren ilk kişidir. Linnaeus, bütün organizmaların bir türün parçası olarak tanımlamıştır. Türleri grupladığı birime cins adını vermiştir. Her bir cinsde en az bir tür bulunmak üzere üst limiti yoktur. Ancak her bir tür sadece bir cinse üye olabilmektedir. Hiyerarşik yapıda yukarı doğru çıktıkça diğer birimleri aile, takım, sınıf, şube ve alem olarak isimlendirmiştir. Tür ve cins arasındaki durum diğer birimler için de geçerlidir. Linnaeus'un yaptığı bu çalışma, bugün biyologların kullandığı isimlendirmeye önemli katkı sağlamıştır.

Son yıllarda daha uzun ve daha fazla sayıda DNA dizileri ile inceleme yapılabilmesi, sınıflandırma için evrimsel sistematiklerin geliştirilmesini sağlamıştır [92]. Evrimsel sınıflandırma olarak adlandırılan bu sistematiklerde, yaşayan ve fosil organizmaların DNA dizileri üzerinde incelemeler yapılmaktadır. Bu organizmaların soyoluşsal ilişkileri üzerinden sınıflandırmalar gerçekleştirilerek soy ağacı çıkarılmaktadır.

1.5. Evrimsel Modeller

Evrimsel bir modelinin kuruluşu sırasında model,

- 1) Üreme
- 2) Kalıtım
- 3) Mutasyon
- 4) Doğal seçim

şeklindeki dört temel evrim unsurunu içermelidir. Biyolojik topluluktaki organizmaların üremeleri gerekmektedir. Bunun nedeni, organizmaların nesillerinin devamlılığını

sağlamak istemesidir. Eğer bir organizma üreme kabiliyetine sahipse, sahip olduğu karakteristik özellikleri kalıtım yolu ile gelecek nesillere aktarabilecektir. Eğer çevre koşullarında herhangi bir değişiklik olmaz ise, gelecek nesil çevre şartlarında yaşayabilmek için en az ailesi kadar başarılı olacaktır. Çünkü doğal seçilim, en başarısız organizmayı topluluktan kaldıracak başka bir ifade ile en başarılı olan organizmayı destekleyecektir. Doğal seçilim olayının işleyebilmesi için, toplulukta farklı türlerin bulunmasına ihtiyaç vardır. Çeşitliliğin gerçekleşmesi için gerekli olan en temel kaynak ise üreme sürecinde gerçekleşecek mutasyondur. Bu kesimde, bazı evrim modelleri anlatılacaktır.

1.5.1. NK Modeli

NK Modeli, Kauffman [8] tarafından geliştirilen, türlerin sahip oldukları uygunluk değerleri üzerine kurulan standart bir eş-evrimleşme modelidir. Modelde her bir tür vektörel olarak

$$\mathbf{S}^\alpha = \{S^\alpha_1, S^\alpha_2, \dots, S^\alpha_L\} \quad (1.1)$$

şeklinde L uzunluklu genom zinciri ile gösterilmektedir. Burada S^α_i 'ler ± 1 değerlerinden birini almakta ve $\mathbf{S}^\alpha$ türünün i . genini temsil etmektedir. Her bir genom zinciri L boyutlu hiperküp olan genom uzayında belirli bir konuma sahiptir. Bu konumlar hiperkübün köşeleridir. Modelin adındaki N, türün sahip olduğu genotipindeki genlerin sayısını yani L 'yi ve K ise, türün kendi genom zincirindeki genler arasındaki etkileşmelerin sayısıdır. Model bir türün uygunluk değerini

$$f(\mathbf{S}^\alpha) = \frac{1}{L} \sum_{i=1}^L f_i(S^\alpha_i, S^\alpha_{i1}, \dots, S^\alpha_{iK}, S^{\alpha 1}_{i1}, \dots, S^{\alpha 1}_{iG}, \dots, S^{\alpha m}_{i1}, \dots, S^{\alpha m}_{iG}) \quad (1.2)$$

ifadesi ile tanımlamaktadır. Burada $S^\alpha_{i1}, \dots, S^\alpha_{iK}$, $\mathbf{S}^\alpha$ türünün kendi genom zincirindeki S^α_i geninin K tane gen ile iç etkileşim içerisinde olduğunu göstermektedir. $S^{\alpha m}_{i1}, \dots, S^{\alpha m}_{iG}$ ise, $\mathbf{S}^\alpha$ türünün genom zincirindeki S^α_i geninin, topluluktaki diğer m tane türün her birinin G tane genleri ile dış etkileşim içerisinde olduğunu göstermektedir. Bu şekilde her bir tür için uygunluk değeri hesabı yapılarak, bir uygun peyzajı elde edilmektedir. Uygunluk peyzajında en yüksek uygunluk değerine sahip olan tür, en yüksek uygunluk tepesine sahip olan türdür. Bu çerçevede modelin temel düşüncesi iki ihtimale dayanır. Birincisi, türler mutasyon geçirerek uygunluk değerini artıracak ve

farklı konumlara hareket edecek. İkincisi, türler mutasyon geçirmeyerek olduğu konumda kalacaktır.

Her bir tür için $K = 0$ olduğunda, uygunluk peyzajında tek bir uygunluk tepesi bulunmaktadır. Bu peyzaj Fujiyama tipi uygunluk peyzajı olarak adlandırılır. Peyzaja bu isim Japonya'daki Fuji dağından esinlenerek verilmiştir. Fuji dağı ülkenin en yüksek dağıdır. Bölgede geniş bir düzlükte bulunan tek bir tepedir. Dağın yamaçları düzgün bir koni şeklinde yükselerek tepede son bulur. $K = L - 1$ durumunda ise, uygunluk peyzajında çok sayıda uygunluk tepeleri bulunacaktır. Bu durum uygunluk peyzajının engebeli olmasına neden olacaktır. Modelin simülasyon hesabının işleyişi ise,

1- Simülasyonun başlangıcında her bir S_i geninin rasgele seçilen diğer K tane genle etkileşmesi

2- Daha sonra 2^{K+1} tane olası gen diziliminin her birindeki genlere $[0, 1]$ aralığındaki düzgün dağılımdan rasgele seçilen değerlerin her bir genin uygunluk değeri olarak f_i değerine atanması

3- Son olarak olası gen dizilimlerinin her biri için f_i değerlerinin aritmetik ortalaması alınarak, genom uzayındaki her bir türün uygunluk değerinin bulunması

olarak verilmektedir. Modelden iki olası dinamik ortaya çıkmaktadır. Birincisi, genom uzayında karşılıklı olarak fayda sağlayan iki tür her ikisinin de memnun olduğu bir yerel maksimum konumuna ulaşabilir. Bu konumun uygunluk değeri, uygunluk peyzajındaki en düşük değer ile en yüksek değer arasında bir değerdir. İkincisi, bu iki tür temkinli davranıp bir global maksimum konumuna ulaşabilir. Bu konumun uygunluk değeri, uygunluk peyzajının en yüksek değeridir.

Kauffman, bu modeli ile yok olma oranları dağılımının, fosil kayıtlarında gözlemlenen $P(s) \sim s^\tau = s^{-2}$ şeklindeki kuvvet kanunu dağılımı ile aynı olduğunu belirtmiştir.

1.5.2. Kendini Organize Eden Evrim Modeli

Bak ve Sneppen [34] tarafından geliştirilen model, NK modelinin basitleştirilmiş hali olarak düşünülebilir. Her iki modelde de birlikte varolan türlerin hayatta kalmak için mücadele ettiği görülmekte ancak konuya farklı açılardan bakmaktadırlar. Bak ve

Sneppen'in modeli, türler arasındaki etkileşmelerden dolayı ekosistemin kendini organize eden bir sistem olarak düşünülebileceği görüşünden yararlanmaktadır. Bu durumda model, kuvvet serisi dağılımına sahip yok olma olaylarının olduğu bir durum içerisinde evrimleşir.

Model basit olarak, D tane türün tek boyutta bir çizgi üzerine periyodik sınırlama şartı yani çizgideki son türün ilk türe bağlanması ile düzenlenir. Her bir $\mathbf{S}^\alpha$ ($\alpha = 1, 2, \dots, D$) türü, $\mathbf{S}^{\alpha+1}$ ve $\mathbf{S}^{\alpha-1}$ olmak üzere iki en yakın komşu türü ile etkileşir ve her bir $\mathbf{S}^\alpha$ türünün $[0, 1]$ aralığında düzgün dağılımdan rasgele seçilen bir eşik değeri bulunmaktadır. Bu durumda NK modelindeki m değeri 2 olacaktır. Burada önemli olan nokta, eğer $\mathbf{S}^\alpha$ türü mutasyon geçirerek konumunu değiştirmek isterse, yeni konumu bu en yakın komşu türlerinden birinin konumu olacaktır. Eşik değeri düşük olan türler mutasyona daha yatkın olan türlerdir. Yüksek eşik değerine sahip olan türler, bulundukları konumu koruma eğilimindedirler.

Modelin dinamik işleyişinde,

1- Her zaman adımında, en düşük eşik değerine sahip $\mathbf{S}^\alpha$ türü seçilir ve yeni bir tür ile değiştirilerek yeni bir eşik değeri belirlenir

2- Her bir $\mathbf{S}^\alpha$ türü, sadece $\mathbf{S}^{\alpha+1}$ ve $\mathbf{S}^{\alpha-1}$ gibi iki tür ile etkileşir

şeklinde iki temel kural bulunmaktadır. Bu kurallar simülasyonun her zaman adımında yeniden işletilir.

Model yok olma oranları için $P(s) \sim s^\tau$ şeklinde kuvvet kanunu dağılımına sahiptir. Dağılımın üsteli $\tau = -0.9 \pm 0.1$ 'dir. Bu değer, fosil kayıtlarından elde edilen yok olma oranları dağılımının üsteli $\tau = -2$ 'den farklıdır.

1.5.3. Hiyerarşik Yiyecek Ağı Modeli

Amaral ve Meyer [35] tarafından önerilen modelde, ekosistem hiyerarşik bir yiyecek ağı olarak gösterilmiştir. Amaral ve Meyer bu çalışmalarında, uygunluk değeri ya da türler arasındaki rekabet kavramlarını ele almamışlardır. Fosil kayıtlarında gözlemlenen kuvvet serisi dağılımını iyileştirebilen basit bir modelin varlığını ispat etmeye çalışmışlardır.

Model, ekosistemi yiyecek ağının besleyici seviyelerini gösteren farklı tabakalardan şekillenen bir sistem gibi düşünülmektedir. Avcılar ve avlar, modeldeki her bir tabakadaki boşluklara rasgele yerleştirilir. En yukarıdaki tabaka hiçbir avcısı olmayan saf avcılardan, en alttaki tabaka ise hiçbir avı olmayan saf avlardan oluşmaktadır. Bütün bu tabakalardaki türler arasında ise tabakaları birbirine bağlayan etkileşimler bulunmaktadır.

Modelin dinamik işleyişi,

1- Her zaman adımında rasgele yeni bir tür oluşturulur. Verilen bir L tabakasında varolan her tür, sabit bir $P_{\text{üreme}}$ üreme ihtimaliyeti ile potansiyel olarak yeni bir türün oluşmasını sağlar. Yeni tür ya aynı tabakadaki ya da komşu tabakalardan birindeki rasgele seçilen bir boşluğu dolduracaktır. Eğer yeni tür $L \neq 0$ tabakasındaki bir boşluğu doldurmuş ise, avcı etkileşimlere sahip olacağından dolayı hemen alt tabakadaki k tane dolu konumlara bağlanır

2- Daha sonra, $L = 0$ tabakasından $P_{\text{öldürme}}$ ihtimaliyeti ile rasgele seçilen bir tür yok edilir. Bu olay sonunda eğer $L = 1$ tabakasındaki bir türün bütün avları yok olmuş ise bu tür de yok olacaktır. Bu sorgulama süreci en üst tabakaya doğru tekrar edilir

şeklinde iki temel kurala bağlıdır.

Bu basit model $k > 2$ durumunda yok olma oranı için $P(s) \sim s^\tau$ şeklinde kuvvet kanunu dağılımına sahip ve dağılımın üsteli $\tau = -1.97 \pm 0.05$ 'dir. Bu değer, fosil kayıtlarında gözlemlenen yok olma dağılımının üsteli $\tau = -2$ 'ye çok yakın bir değerdir.

1.6. Ağ Sistemlerinin Matematiksel Temelleri

En basit tanımı ile bir ağ, birbirlerine çizgilerle bağlanmış noktaların bir araya gelmesi ile oluşmuş bir yapıdır. Teriminolojik olarak ağ tanımındaki noktalar *köşe (vertex)* ve noktaları birbirine bağlayan çizgiler de *kenar (edge)* olarak adlandırılmaktadır. Bu köşe ve kenarlar *bilgisayar mühendisi* için *nokta (node)* ve *bağlantı (link)*, bir *fizikçi* için *konum (site)* ve *bağ (bond)* ve bir *ekolojici* için *tür (species)* ve *etkileşim (interaction)* olarak adlandırılırlar. Bu çalışmanın temeli biyolojik topluluklar üzerine kurulduğu için ağdaki köşe ve kenarlar sırasıyla tür ve etkileşim olarak belirtilecektir.

Tablo 1.2'de ağlardaki köşe ve kenarların ne anlama geldiğini belirtmek için bazı ağ

Tablo 1.2. Bazı ağlardaki köşeler ve kenarlar.

Ağ	Köşe	Kenar
İnternet	Bilgisayar	Kablolu veya kablosuz bağlantı
Yiyecek ağı	Tür	Avlanma
Enerji nakli	Üretim istasyonları	Taşıma hatları
Ulaşım	Ülke ya da şehir	Ulaşım araçları ya da yollar
Atıf	Makale ya da patent	Yapılan atıflar
Arkadaşlık	Kişi	Dostluk ilişkisi
Sinirsel	Sinir hücreleri	Sinaps

örnekleri verilmiştir. Bunlardan yiyecek ağına bakıldığında, köşeler topluluktaki av veya avcı türlere karşılık gelirken, kenarlar av ile avcı arasındaki ilişkiyi tanımlamaktadır. Başka bir ifade ile, hangi avcı türün hangi av türü ile beslendiği ağ yapısı olarak incelenebilmektedir. Bu kesimde ağın genel yapısını çizen bazı temel tanımlar verilecektir.

1.6.1. Bitişiklik Matrisi

Ağlar oluşturulurken ilk olarak ağda bulunan köşelere olası karışıklığı önlemek için birer etiket verilir. N tane köşeden oluşan bir ağ sisteminde köşeler 1’den N ’ye kadar tam sayılarla etiketlenir. Sadece köşelerin ayırt edilebilmesi istendiği için hangi köşeye hangi sayının verildiğinin bir önemi yoktur. Bu noktadan hareketle ağ sistemlerini matematiksel olarak göstermenin farklı yolları vardır [93]. Eğer herhangi bir i köşesi ile herhangi bir j köşesi arasında bir kenar varsa, bu (i,j) şeklinde gösterilir. Bu şekildeki gösterim *kenar listesi* olarak adlandırılır ve $G(V,E)$ ile gösterilir. Burada V İngilizce köşe anlamına gelen *Vertex* kelimesinin baş harfi ve E İngilizce kenar anlamına gelen *Edge* kelimesinin baş harfidir. Bu gösterim genellikle ağları bilgisayar ortamında depolayabilmek için kullanılır. Daha açık bir gösterim için yaygın olarak *bitişiklik (adjacency) matrisi* kullanılır.

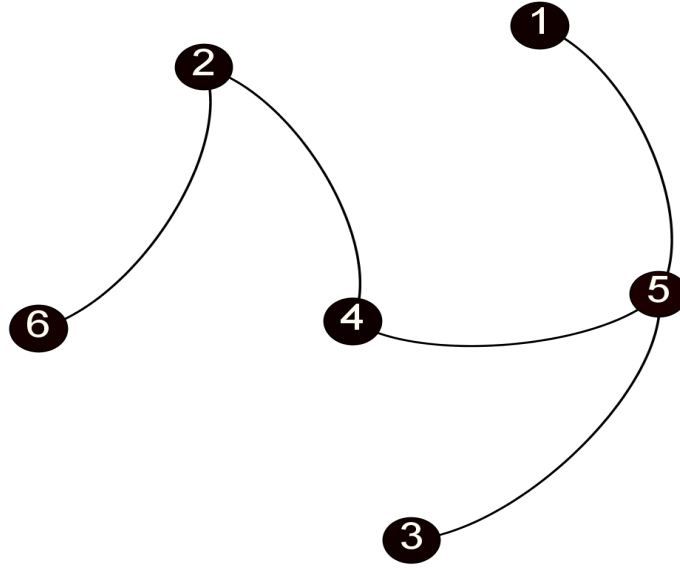
A ile gösterilen bitişiklik matrisi her bir A_{ij} matris elamanının,

$$A_{ij} = \begin{cases} 1 & \text{eğer } i \text{ ve } j \text{ köşeleri arasında bir kenar varsa} \\ 0 & \text{diğer durumda} \end{cases} \quad (1.3a)$$

$$(1.3b)$$

şeklinde aldığı değerlerle doldurulur.

Şekil 1.5’de 6 tane köşe ve 5 tane kenardan oluşan bir ağ sistemi verilmiştir. Köşelerin



Şekil 1.5. 6 tane köşeden oluşan bir ağ.

numaraları tamamen keyfi bir şekilde verilmiş ve herhangi bir anlamı bulunmamaktadır.

Bu ağın kenar listesi ile gösterimi,

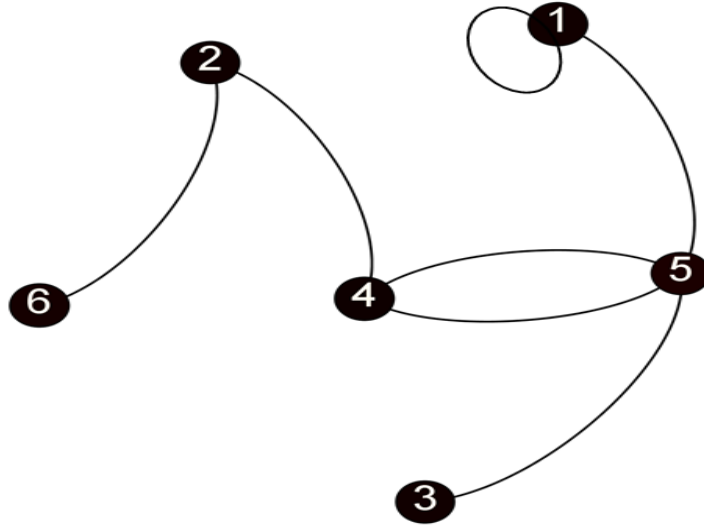
$$G(V, E) = \{(1, 5), (2, 4), (2, 6), (3, 5), (4, 5)\} \quad (1.4)$$

şeklindedir. $G(V, E)$ kenar listesinin eleman sayısı ağda bulunan kenarların sayısını vermektedir. Bitişiklik matrisi ile gösterimi ise,

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (1.5)$$

şeklinde olacaktır. Bitişiklik matrisinde dikkat çeken iki önemli husus vardır. Birincisi matrisin köşegen elemanları $A_{ii} = 0$ dır. Bunun anlamı, köşelerin kendileri üzerine kapanan herhangi bir kenar olmamasıdır. İkincisi ise matrisin simetrik yani $A_{ij} = A_{ji}$ olmasıdır.

Ağ sistemlerinde herhangi iki köşe arasında *basit ağ* olarak adlandırılan Şekil 1.5’de olduğu gibi sadece tek bir kenar olabileceği gibi bazen *çoklu kenar* olarak adlandırılan birden fazla kenarların olduğu durumlarda olabilir. Bunun yanında köşelerin kendileri ile yaptıkları *öz-kenar* durumu da söz konusu olabilir. Şekil 1.6’de bu durumlar için bir örnek verilmiştir. Şekil 1.6’da ise üç farklı kenar durumunun varlığı söz konusudur.



Şekil 1.6. 6 tane köşeden oluşan bir ağ sistemindeki çoklu ve öz-kenar durumları.

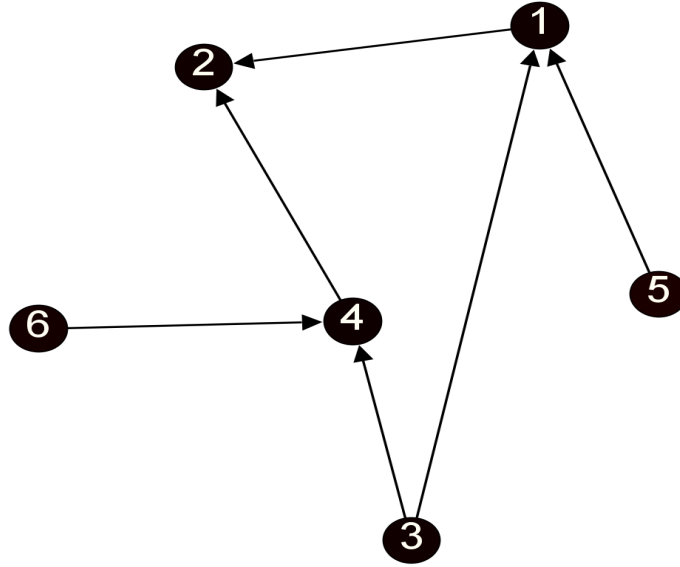
Genel olarak tek kenarlar bulunmasına rağmen 4 ve 5 numaralı köşeler arasında iki tane kenar yani çoklu kenar durumu mevcuttur. 1 numaralı köşede ise öz-kenar durumu görülmektedir. Bu şekilde çoklu ve öz-kenar durumlarının da söz konusu olduğu ağlar *basit olmayan ağlar* olarak adlandırılır. Böyle bir ağın bitişiklik matrisi ile gösterimi ise,

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (1.6)$$

şeklinde olacaktır. Matris simetrik bir matristir. Eğer i ve j köşeleri arasında bir kenar varsa $A_{ij} = A_{ji} = 1$ olacaktır. 4 ve 5 numaralı köşeler arasında 2 kenar ya da 2 kenar bağlantı noktası olduğu için bitişiklik matrisinin A_{45} ve A_{54} elemanları 2'ye eşittir. Aynı durum öz-kenar durumuna sahip 1 numaralı köşe için de geçerlidir. Öz-kenar durumunda köşe 2 kenar bağlantı noktasına sahiptir. Bu nedenle A_{11} matris elemanı 2'ye eşittir.

1.6.2. Yönelimli ve Yönelimsiz Ağlar

Yönelimli ağlar, ağda bulunan kenarların belirli bir yönde oldukları yani iki köşe arasındaki kenarın hangi köşeden *çıkıp* hangi köşeye *girdiği* anlaşılabilen ağlardır. Şekil 1.7'de yönelimli bir ağ örneği gösterilmiştir. Kenarlara yönelim kazandırmak için gösterimde ok işareti kullanılmıştır. Örneğin 1 numaralı köşeden 2 numaralı köşeye bir



Şekil 1.7. 6 tane köşeden oluşan yönelimli bir ağ.

kenar varken, 2 numaralı köşeden hiçbir köşeye bir kenar yoktur.

$$A_{ij} = \begin{cases} 1 & \text{eğer } i \text{ köşesinden } j \text{ köşesine bir kenar varsa} \\ 0 & \text{diğer durumda} \end{cases} \quad (1.7a)$$

$$(1.7b)$$

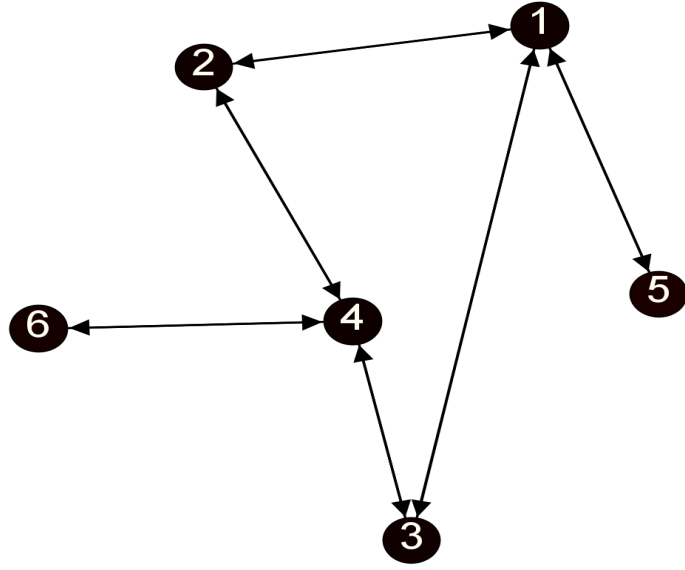
Denklem (1.7) ile elemanları oluşturulan bitişiklik matrisi ise,

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (1.8)$$

şeklinde olacaktır. Görüldüğü üzere bitişiklik matrisi simetrik olmayan bir matristir. Yönelimli ağlar için bu genel bir durumdur [93]. Kenarların yöneliminden bahsedildiği için Şekil 1.7’de ki ağda mevcut olan her bir yönelimli kenarın ters yönelimi de söz konusu olabilirdi. Daha açık bir ifade ile, köşeler birbirleri ile çift yönlü kenar oluşturabilirdi.

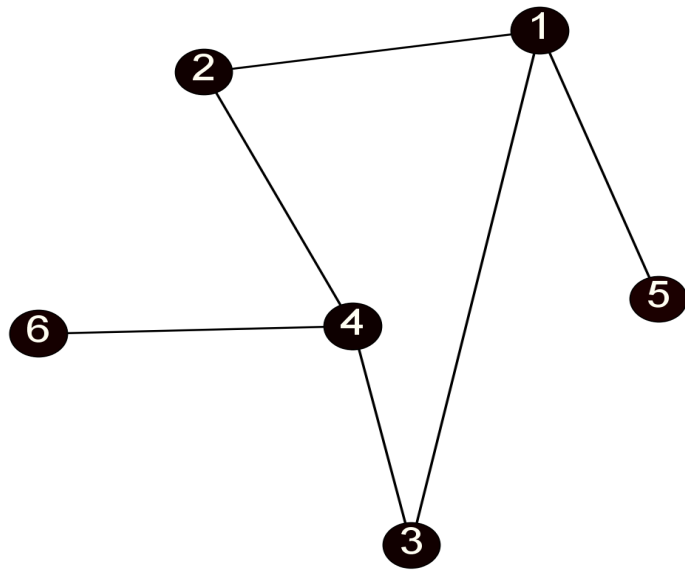
Şekil 1.8’de çift yönelimli bir ağ örneği gösterilmiştir. Böyle bir ağın bitişiklik matrisi Denklem (1.7)’e göre,

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (1.9)$$



Şekil 1.8. 6 tane köşeden oluşan çift yönelimli bir ağ.

şeklinde simetrik bir matris olacaktır. Bu simetriden dolayı $A_{ij} = A_{ji}$ olacaktır. i ve j köşeleri arasında her iki yönde de kenarın olması, köşeler arasındaki kenarın yönelimden bağımsız olması anlamına gelecektir. Bu şekilde ki ağlar *Yönelimsiz Ağlar* olacaktır. Buradan anlaşılacağı üzere, yönelimli ağlarda kenarların yönelimi sadece tek bir yönde gerçekleşmekte [94] aksi bir durumda ağ yönelimsiz ağ olarak incelenmektedir. Yönelimsiz ağlarda önemli olan köşeler arasında kenarın olup olmamasıdır.



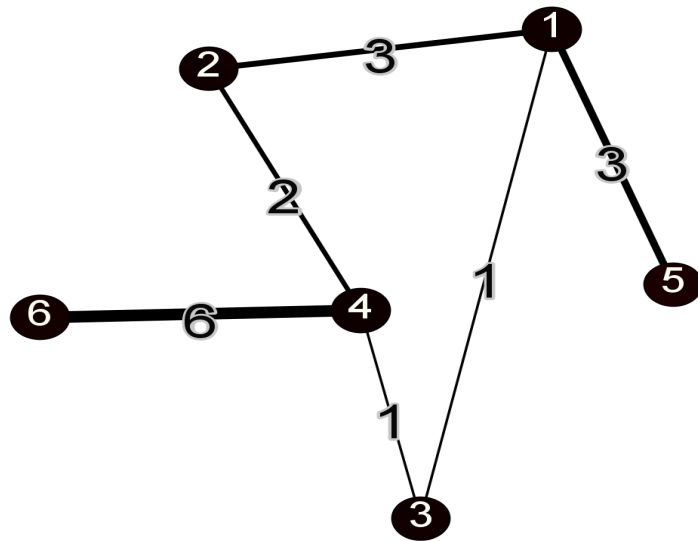
Şekil 1.9. 6 tane köşeden oluşan yönelimsiz bir ağ.

Şekil 1.9'da yönelimsiz bir ağ verilmiştir. Bu yönelimsiz ağın bitişiklik matrisi Denklem

(1.3)'e göre oluşturulduğunda Denklem (1.9)'nin aynısı olacaktır. Ayrıca önemli olan başka bir husus ise, Şekil 1.7'deki yönelimli ağda Şekil 1.6'daki gibi bir öz-kenar durumu söz konusu olsaydı, matrisin köşegen elemanı 2 değil 1 olacaktı. Bunun nedeni ise Şekil 1.6 yönelimsiz bir ağ olduğu için aslında 2 tane yönelimli kenar vardır. Bu yüzden Şekil 1.6'da bitişiklik matrisinin köşegen elemanı 2 alındı. Ancak yönelimli ağda bu kenar 1 tane olacağı için köşegen elemanı 1 olarak alınır.

1.6.3. Ağırlıklı ve Ağırlıksız Ağlar

Birçok ağ yapıları genellikle ikili sistem halinde çalışılmıştır [95]. Başka bir ifade ile, ağdaki köşeler arasında kenar ya *vardır* ya da *yoktur*. Bu yüzden bitişiklik matrislerinin elamanları kenar var ise 1 yok ise 0 mantığı ile oluşturulmaktadır. Ancak bazı durumlarda bu düşünce yeterli olmayıp kenarlara bir *ağırlık* verilmesi gerekir. Bu ihtiyacın sebebi, köşeler arasındaki kenarların varlığına ilaveten köşelerin birbirleri ile ne derece bağlantılı olduğunu inceleyebilmektir. Biraz daha anlaşılabilir olması için, incelenen ağ bir arkadaşlık ağı olarak düşünüldüğünde, eğer ağ ağırlıksız olsaydı herkes herkesle aynı derecede arkadaş olacaktı. Ancak bu durum gerçek hayatta böyle değildir. Bu yüzden, *kimin hangi arkadaşı ile ne kadar sıkı ya da ne kadar zayıf bir ilişki kurduğunu* bilebilmek için ağdaki kenarların ağırlıklı olmasına ihtiyaç duyulur.



Şekil 1.10. 6 tane köşeden oluşan yönelimsiz ağırlıklı bir ağ. Kenarların üzerindeki sayılar ilgili kenarın ağırlığını belirtmektedir.

Şekil 1.10'da yönelimsiz ağırlıklı bir ağ verilmektedir. Köşelerin birbirleri ile yaptıkları bağlantıların önemini belirtmek için kenarlara ağırlık verilmiştir. 4 ve 6 köşeleri arasındaki bağlantı 1-3 ve 3-4 köşeleri arasındaki bağlantılardan 6 kat, 2-4 köşeleri arasındaki bağlantılardan 3 kat ve 1-2 ile 1-5 köşeleri arasındaki bağlantıdan 2 kat daha ağır veya daha yoğundur denilebilir. Ağırlıklı ağın bitişiklik matrisi ise, kenarların ağırlıklarına göre oluşturulur. Kenarın ağırlık değeri bitişiklik matrisinin ilgili elemanı olarak alınır ve

$$\mathbf{A} = \begin{pmatrix} 0 & 3 & 1 & 0 & 3 & 0 \\ 3 & 0 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 6 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 \end{pmatrix} \quad (1.10)$$

olarak yazılır. Ağ yönelimsiz olduğu için bitişiklik matrisi simetriktir. Ağdaki ağırlık değerleri pozitif olabileceği gibi negatif de olabilir. Bu, kişilerin bazı insanlarla dostça ilişki kurabildiği gibi bazı insanlarla düşmanca ilişki kurabileceği anlamına gelmektedir. Dolayısıyla bu kararı köşeler arasındaki ilişkinin doğası belirleyecektir. Ağırlıklı ağlar yönelimsiz olduğu gibi yönelimli de olabilir.

Ağırlıksız ağlar ise kenarların herhangi bir ağırlığa sahip olmadığı anlamına gelmez. Bu durum, ağda bulunan bütün kenarların sahip oldukları ağırlık değerlerinin birbirine eşit olması ile açıklanır. Eğer Şekil 1.10'daki ağ ağırlıksız olmuş olsaydı, Şekil 1.9'deki yönelimsiz ağ olarak alınacak ve bitişiklik matrisi Denklem (1.9)'daki gibi olacaktı.

1.7. Ağ Analizinde Sıklıkla Kullanılan Parametreler

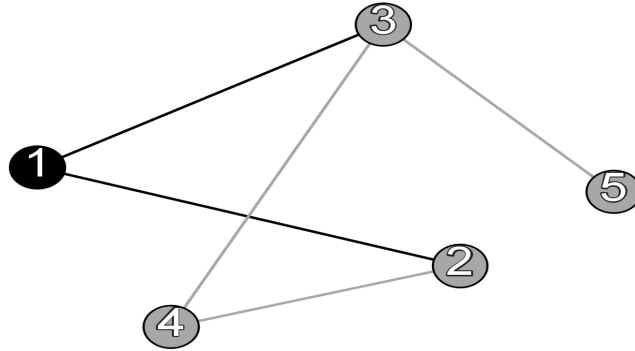
Bir ağın istatistiğinin yapılabilmesi için ağın yapısı hakkında bilgi sahibi olmak gerekir [96]. Ağın yapısı hakkında bilgi ise bitişiklik matrisinden alınmaktadır. Kesim 1.6'de ağın sahip olabileceği genel kimliği üzerinde durulmuştu. Ancak ağın analiz edilebilmesi için bitişiklik matrisinin detaylıca incelenmesi gerekmektedir. Bu çerçevede ağı çeşitli açılardan ele alarak, ağın karakteristik yapısını açıklamada kullanılan birtakım parametreler vardır. Bu kesimde bu parametrelerden bazıları incelenecektir.

1.7.1. Derece Dağılımı

Derece, ağdaki bir köşenin sahip olduğu kenar sayısı olarak tanımlanır ve ağ üzerinde gerçekleşen herhangi bir bilgi iletiminde köşelerin bu bilgiye sahip olma olasılığını gösterdiği için önemli bir parametredir [96–101]. Köşelerin sahip oldukları dereceler belirlenirken bitişiklik matrisinden faydalanılır ve

$$d_i = \sum_{j \in N} A_{ij} \quad (1.11)$$

olarak ifade edilir. d_i , i ile j köşesi arasındaki kenar sayısının toplamını yani i köşesinin derecesini belirtir. A_{ij} ise N tane köşeden oluşan ağın bitişiklik matrisi elemanıdır. $\sum$ ise i köşesinin kendisi de dahil olmak üzere ağdaki tüm köşelerle olan kenar durumunu sorgulamaktadır. Denklem (1.11) yönelimsiz bir ağ için tanımlanan derece ifadesidir.



Şekil 1.11. 5 tane köşeden oluşan yönelimsiz ağırlıksız basit bir ağ.

Şekil 1.11’de yönelimsiz, ağırlıksız bir basit ağ örneğinde, 1 numaralı köşenin derecesi üzerinde durulmuştur. Bu yüzden bu köşe ve kenarları siyah, diğer köşeler ve kenarları gri renkte gösterilmiştir. 1 numaralı köşenin sadece 2 ve 3 numaralı köşelerle kenarı olduğu görülmektedir. Bu yüzden $d_1 = 2$ dir.

Eğer ağ yönelimli olursa, derece ifadesi iki kısımdan oluşacaktır. Birincisi i köşesine giren kenarların derecesi d_i^g , ikincisi i köşesinden çıkan kenarların derecesi d_i^c olarak ifade edilir.

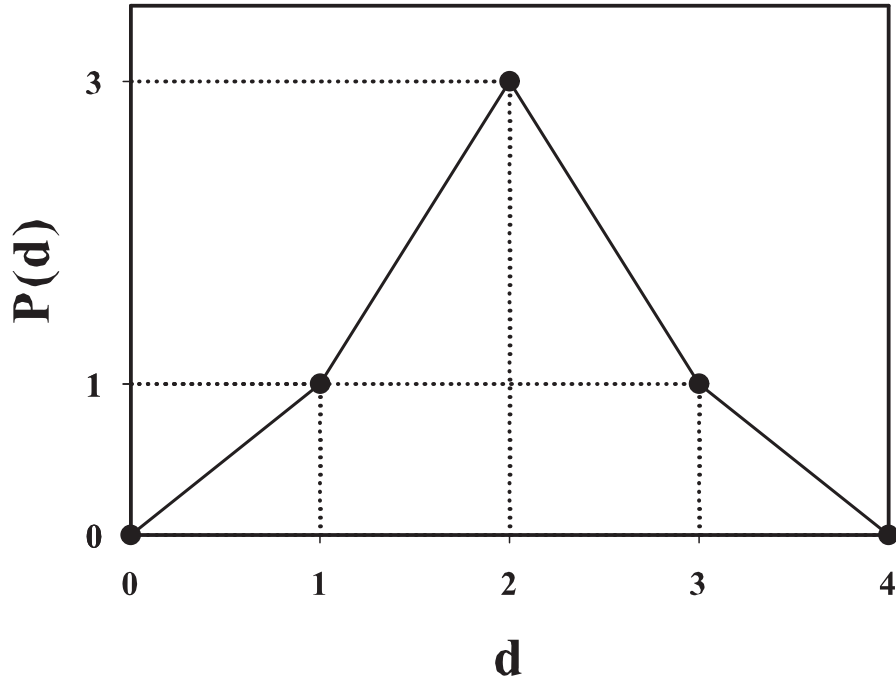
$$d_i^g = \sum_{j \in N} A_{ji} \quad \text{ve} \quad d_i^c = \sum_{j \in N} A_{ij} \quad (1.12)$$

i köşesinin toplam derecesi ise,

$$d_i^t = d_i^g + d_i^e \quad (1.13)$$

şeklinde olacaktır.

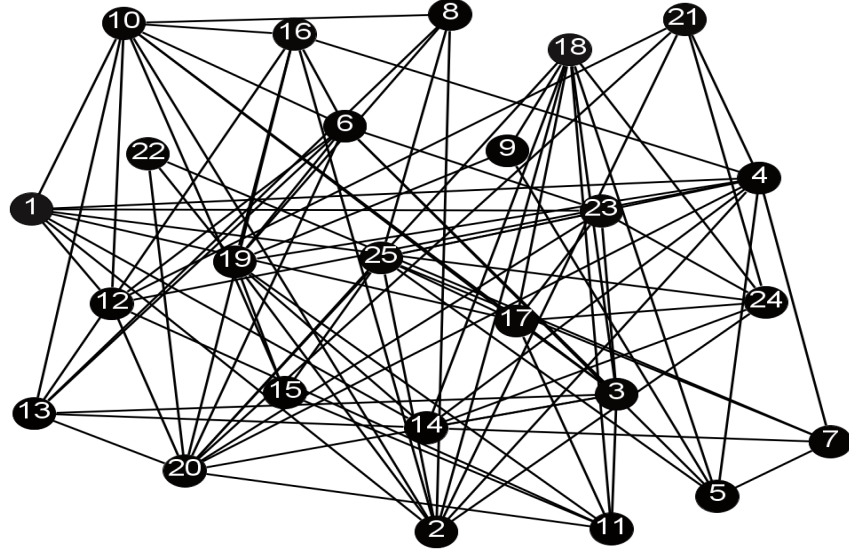
Ağın en temel karakteristik incelemesi, tüm ağın sahip olduğu köşe derecelerinin *derece dağılımı* (*degree distribution*) ile yapılır [97]. $P(d)$ ile gösterilen derece dağılımı, ağdan rastgele seçilen bir köşenin derecesinin d değerine eşit olma ihtimaliyeti olarak tanımlanır. Başka bir ifade ile, ağda d derecesine sahip kaç tane köşe olduğunu belirler. Bu yüzden derece dağılımı, köşelerin yerel özelliklerinden yararlanarak ağın genel özelliklerini üretir. Eğer ağda d derecesine sahip en az bir köşe varsa $P(d)$ pozitif yoksa $P(d)$ sıfır olacaktır. O halde derece dağılımı her zaman $P(d) \geq 0$ dir. Eğer ağ yönelimli ise derece dağılımı, sadece köşelere giren kenarların derece dağılımı $P(d)^g$ ve sadece köşelerden çıkan kenarların derece dağılımı $P(d)^e$ olmak üzere iki ayrı şekilde incelenir [97].



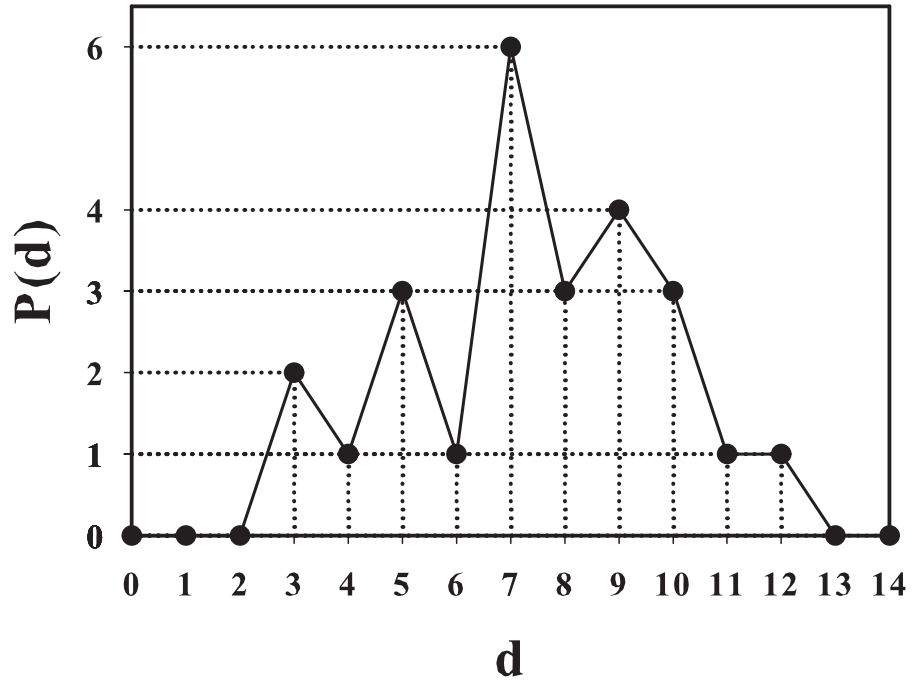
Şekil 1.12. Şekil 1.11'deki ağın derece dağılımı.

Şekil 1.11'deki gibi bir ağın derece dağılımı Şekil 1.12'de verilmiştir. Görüleceği gibi, ağın yapısında hangi dereceleri barındırdığı ve bu derecelerin nasıl bir istatistiğe sahip olduğu derece dağılımından açıkça anlaşılabilir. Ağın ortalama derecesinin yani ağdaki en yüksek ihtimaliyetli d derecesinin $\langle d \rangle = 2$ olduğu görülmektedir. Aslında Şekil 1.11'e bakıldığında derece dağılımına gerek kalmadan aynı sonuca ulaşılabilir.

Ancak incelenen ağlar her zaman Şekil 1.11'deki gibi az sayıda köşe ve kenar içeren basit bir ağ olmayabilir. Derece dağılımının önemi daha büyük ve karmaşık ağlarda daha net anlaşılmaktadır.



(a) Ağ yapısı



(b) Derece dağılımı

Şekil 1.13. 25 köşeden oluşan yönelimsiz karmaşık bir ağ ve derece dağılımı.

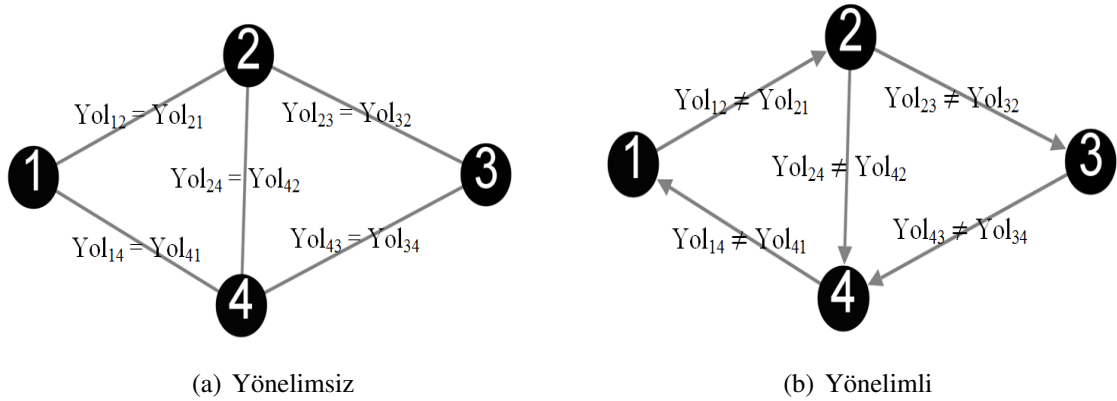
Şekil 1.13'de derece dağılımının önemini belirtmek için karmaşık bir ağ örneği

verilmiştir. Şekil 1.13(a)'da 25 köşeden oluşan yönelimsiz bir ağ verilmiştir. Ağa ilk bakıldığında karmaşık yapısından dolayı pek birşey anlaşılmaz. Ancak Şekil 1.13(b)'deki derece dağılımına bakıldığında, ağdaki köşelerden d derecesine sahip olan ağların oranı rahatça görülebilmekte ve $\langle d \rangle = 7$ olduğu anlaşılmaktadır.

Derece dağılımları farklı ağlar için farklı tipte dağılımlar sergiler. Bu dağılımlar ağın karakteristik yapısını ortaya koyar. Ancak bu kesimde derece ve derece dağılımının temeli anlatıldığı için bu dağılım tipleri burada verilmeyecektir. Ağ modellerinin incelendiği kesimde açıkça anlatılacaktır.

1.7.2. Ortalama Yol Uzunluğu ve Çap

Ağdaki herhangi iki köşe arasındaki ortalama en kısa mesafeyi veren *ortalama yol uzunluğu* (*average path length*), herhangi bir bilginin iletimi ağdaki yollar üzerinden sağlandığı için ağ incelemelerinde ağın genel özelliklerini belirlemek için kullanılan önemli parametrelerden biridir [93, 94, 96, 97, 102–104]. Ağdaki bir *yol*, kenarlarla birbirlerine bağlanmış köşeler dizisi olarak tanımlanır [93] ve ağdaki bir köşeden diğer köşeye ulaşmak için gerekli olan minimum mesafe veya kenar uzunluğu olarak hesaplanır.



Şekil 1.14. Yönelimsiz ve yönelimli ağlardaki yollar.

Şekil 1.14'de yönelimsiz ve yönelimli ağlar için ağdaki yollar gösterilmiştir. Ağlar ağırlıksız yani bütün kenarların ağırlığı 1 olarak alınmıştır. Şekil 1.14(a)'da ağ yönelimsiz olduğu için herhangi iki köşe arasındaki yol birbirine eşittir. Ancak Şekil 1.14(b)'de köşelerin yönelimlerinden dolayı bu durum geçerli olmayacaktır. Çünkü yönelimli ağda bilgi iletimi veya etkileşme tek yönde olmaktadır.

Tablo 1.3. Şekil 1.14(a)'daki olası yollar.

Köşeler	Olası yollar	Uzunluklar	Köşeler	Olası yollar	Uzunluklar
1 - 2	$1 \rightarrow 2$	1	3 - 1	$3 \rightarrow 2 \rightarrow 1$	2
	$1 \rightarrow 4 \rightarrow 2$	2		$3 \rightarrow 4 \rightarrow 1$	2
	$1 \rightarrow 4 \rightarrow 3 \rightarrow 2$	3		$3 \rightarrow 2 \rightarrow 4 \rightarrow 1$	3
1 - 3				$3 \rightarrow 4 \rightarrow 2 \rightarrow 1$	3
	$1 \rightarrow 2 \rightarrow 3$	2	3 - 2	$3 \rightarrow 2$	1
	$1 \rightarrow 4 \rightarrow 3$	2		$3 \rightarrow 4 \rightarrow 2$	2
	$1 \rightarrow 2 \rightarrow 4 \rightarrow 3$	3		$3 \rightarrow 4 \rightarrow 1 \rightarrow 2$	3
1 - 4	$1 \rightarrow 2 \rightarrow 4 \rightarrow 3$	3	3 - 4		
	$1 \rightarrow 4$	1		$3 \rightarrow 4$	1
	$1 \rightarrow 2 \rightarrow 4$	2		$3 \rightarrow 2 \rightarrow 4$	2
2 - 1	$1 \rightarrow 2 \rightarrow 3 \rightarrow 4$	3	4 - 1	$3 \rightarrow 2 \rightarrow 1 \rightarrow 4$	3
	$2 \rightarrow 1$	1		$4 \rightarrow 1$	1
	$2 \rightarrow 4 \rightarrow 1$	2		$4 \rightarrow 2 \rightarrow 1$	2
2 - 3	$2 \rightarrow 3 \rightarrow 4 \rightarrow 1$	3	4 - 2	$4 \rightarrow 3 \rightarrow 2 \rightarrow 1$	3
	$2 \rightarrow 3$	1		$4 \rightarrow 2$	1
	$2 \rightarrow 4 \rightarrow 3$	2		$4 \rightarrow 1 \rightarrow 2$	2
2 - 4	$2 \rightarrow 1 \rightarrow 4 \rightarrow 3$	3	4 - 3	$4 \rightarrow 3 \rightarrow 2$	2
	$2 \rightarrow 4$	1		$4 \rightarrow 3$	1
	$2 \rightarrow 1 \rightarrow 4$	2		$4 \rightarrow 2 \rightarrow 3$	2
	$2 \rightarrow 3 \rightarrow 4$	2		$4 \rightarrow 1 \rightarrow 2 \rightarrow 3$	3

Tablo 1.4. Şekil 1.14(b)'deki olası yollar.

Köşeler	Olası yollar	Uzunluklar	Köşeler	Olası yollar	Uzunluklar
1 - 2	$1 \rightarrow 2$	1	3 - 1	$3 \rightarrow 4 \rightarrow 1$	2
1 - 3	$1 \rightarrow 2 \rightarrow 3$	2	3 - 2	$3 \rightarrow 4 \rightarrow 1 \rightarrow 2$	3
1 - 4	$1 \rightarrow 2 \rightarrow 4$	2	3 - 4	$3 \rightarrow 4$	1
	$1 \rightarrow 2 \rightarrow 3 \rightarrow 4$	3			
2 - 1	$2 \rightarrow 4 \rightarrow 1$	2	4 - 1	$4 \rightarrow 1$	1
	$2 \rightarrow 3 \rightarrow 4 \rightarrow 1$	3			
2 - 3	$2 \rightarrow 3$	1	4 - 2	$4 \rightarrow 1 \rightarrow 2$	2
2 - 4	$2 \rightarrow 4$	1	4 - 3	$4 \rightarrow 1 \rightarrow 2 \rightarrow 3$	3
	$2 \rightarrow 3 \rightarrow 4$	2			

Tablo 1.3 ve Tablo1.4'de sırasıyla Şekil 1.14(a) ve Şekil 1.14(b)'deki yönelimsiz ve yönelimli ağlar için bütün köşeler arasındaki olası bütün yollar verilmiştir. Yolların doğrudan iki köşe arasında olabileceği gibi, başka köşe veya köşeler üzerinden de olabileceği gösterilmiştir.

Örneğin Tablo 1.3'de belirtildiği gibi 1 ve 2 numaralı köşeler arasında doğrudan bir Yol_{12} yolu olabileceği gibi 4 numaralı köşe üzerinden geçerek $Yol_{14} \rightarrow Yol_{42}$ ve birde 4 ile 3 numaralı köşelerden geçerek $Yol_{14} \rightarrow Yol_{43} \rightarrow Yol_{32}$ yolları da olabilir. 1 ve 2 numaralı

köşeler arasındaki bu olası 3 yolun l_{12} *uzunlukları* ise bitişiklik matrisi ile belirlenir ve

Yol₁₂ için

$$l_{12} = A_{12} = 1 \quad (1.14)$$

Yol₁₄ → Yol₄₂ için

$$l_{12} = l_{14} + l_{42} = A_{14} + A_{42} = 1 + 1 = 2 \quad (1.15)$$

ve Yol₁₄ → Yol₄₃ → Yol₃₂ için

$$l_{12} = l_{14} + l_{43} + l_{32} = A_{14} + A_{43} + A_{32} = 1 + 1 + 1 = 3 \quad (1.16)$$

olarak bulunur. $A_{ij}(i, j = 1, 2, 3, 4)$ 'ler bitişiklik matrisi elemanlarıdır. Görüldüğü gibi 1 ve 2 numaralı köşeler arasındaki yollardan en kısa uzunluğa sahip olan yol doğrudan 1 ve 2 numaralı köşeler arasında olan Yol₁₂'dir. Ortalama yol uzunluğu hesaplamasında belirleyici faktör köşeler arasındaki *en kısa uzunluktur*. Bu yüzden hem Tablo 1.3'de hem de Tablo 1.4'de köşeler arasındaki en kısa uzunluğa sahip olan yollar kırmızı renkle gösterilmiş ve numaralandırılmıştır. Ancak Tablo 1.3'e bakıldığında l_{13} ve l_{31} için aynı en kısa uzunluk değerine sahip iki farklı yol vardır. Ortalama yol uzunluğunu hesaplarken aynı en kısa uzunluk değerine sahip yollardan sadece bir tanesi alınmaktadır. Her iki ağda da bu şekilde en kısa uzunluğa sahip 12 tane yol bulunmaktadır.

Gerekli en kısa mesafelerin tayin edilmesinden sonra ağdaki herhangi bir i köşesinden diğer bütün j köşelerine olan ortalama en kısa yol uzunluğu

$$L_i = \frac{\sum_{j=1}^{N-1} l_{ij}}{n} \quad (1.17)$$

şeklinde belirlenir. Burada L_i , i köşesinin ağdaki diğer j köşeleri ile arasındaki ortalama yol uzunluğu, l_{ij} ağlardaki i ve j köşeleri arasındaki en kısa yol uzunluğu, n bu köşeler arasındaki en kısa yolların sayısı ve N ise ağdaki toplam köşe sayısıdır. Köşelerin kendilerine olan en kısa yol uzunluğu l_{ii} , ortalama yol uzunluğu hesabı ile ilişkili olmadığı için ihmal edilir [93]. Bu yüzden toplam ifadesinde j , $N-1$ tane farklı değer alabilir. Eğer i köşesinin ağda hiçbir köşe ile bir kenarı yoksa bu köşenin ağdan ayrık olduğu ve herhangi bir yola sahip olmadığı söylenebilir [41]. Bu durumda Denklem (1.17)'de $n = 0$ olacağından dolayı, i köşesinin ortalama yol uzunluğu sonsuz olacaktır. Bu durumdan

kurtulmak için, hiç bir kenara sahip olmayan yani izole olmuş köşelerin ağdan çıkarılması gerekmektedir [93].

Böylelikle bütün bir ağın ortalama yol uzunluğu ise

$$L = \frac{\sum_{i \neq j} l_{ij}}{n_t} \quad (1.18)$$

şeklinde olacaktır. n_t ağdaki en kısa yolların toplam sayısıdır. Yönelimsiz ağlar için $n_t = N(N - 1)$ 'dir.

Ağın *çapı (diameter)* ise ağdaki en kısa uzunluklar içerisinde en büyük değer olarak tanımlanır ve ağın büyüklüğünü yani ağdaki en uzak iki köşe arasındaki mesafeyi belirler.

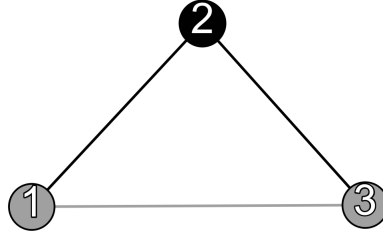
$$D = \max(l_{ij}) \quad (1.19)$$

Ortalama yol uzunluğu çap değerinden daha büyük olamayacağı için, ortalama yol uzunluğunun sınırı çap ile belirlenir. Tablo 1.3'e bakıldığında, Şekil 1.14(a)'daki ağ için çapın eşit uzunluklara sahip l_{13} veya l_{31} olduğu; Tablo 1.4'den ise Şekil 1.14(b)'deki ağ için yine eşit uzunluklara sahip l_{32} veya l_{43} uzunlukları olduğu görülmektedir. Her iki ağın ortalama yol uzunlukları bu değerlerden küçük veya eşit olmalıdır.

1.7.3. Kümelenme Katsayısı

Kümelenme katsayısı (clustering coefficient), ağdaki üçlü köşelerin sayısını yani köşelerin birbirleri ile olan geçişkenliğini tanımlamak için kullanılan ve bunun sonucunda ağın yerel özellikleri açıklanabilen önemli ağ parametrelerinden birisidir [41,93,97,105–109].

Şekil 1.15'de üç köşenin birbirleri ile yaptıkları kenarlar gösterilmiştir. 1 ile 2 ve 2 ile 3 arasında olmak üzere iki tane kenar vardır. 2 numaralı köşe, 1 ve 3 numaralı köşeler için ortak köşe olduğu için köşe ve kenarları siyah renkle, 1 ile 3 numaralı köşeler arasında bir kenar olmadığı için köşeler ve olası bir kenar gri renkle gösterilmiştir. Matematiksel olarak, a , b ve c gibi üç ifade arasında $a = b$ ve $b = c$ gibi bir ilişki varsa $a = c$ olması gerekmektedir. Bu düşünce ile, eğer 1 ile 2 ve 2 ile 3 arasında bir kenar varsa 1 ile 3 arasında da bir kenar olma olasılığının olduğu ve 1 ile 3 arasında da bir kenar varsa bu üç köşenin bir *üçgen* oluşturduğu söylenebilir. Sosyal ağlarda bu durum *arkadaşımın arkadaşı benim de arkadaşımdır* olarak ifade edilir. Bununla birlikte



Şekil 1.15. Köşeler arasındaki kümelenmenin ve ya geçişkenliğin şematik gösterimi.

kümelenme katsayısı, *ağda varolan üçgen sayısının, varolan ve varolması muhtemel olan üçgen sayısına oranı* olarak tanımlanır [109]. Kümelenme katsayısı yönelimsiz ağlarda olduğu gibi yönelimli ağlarda da incelenebilir. Yönelimsiz bir ağdaki i köşesinin kümelenme katsayısı

$$\begin{aligned}
 c_i &= \frac{\text{Varolan üçgen sayısı}}{\text{Varolan ve varolması muhtemel olan üçgen sayısı}} \\
 &= \frac{\frac{1}{2} \sum_{j \neq i} \sum_{h \neq (i,j)} A_{ij} A_{ih} A_{jh}}{\frac{1}{2} d_i (d_i - 1)} \\
 &= \frac{\sum_{j \neq i} \sum_{h \neq (i,j)} A_{ij} A_{ih} A_{jh}}{d_i (d_i - 1)}
 \end{aligned} \tag{1.20}$$

şeklinde elde edilir. A_{ij} , A_{ih} ve A_{jh} 'ler bitişiklik matrisi elemanlarıdır. d_i , i köşesinin derecesini belirtmektedir. Toplam ifadesi ise, i 'nin iki ayrı j ve h komşuları üzerinden alınmaktadır. Denklem (1.20)'nin ikinci eşitliğinin hem pay hem de paydasındaki $\frac{1}{2}$ sabiti, ağ yönelimsiz olduğu için her bir kenar iki defa sayılmakta ve bu kenarlar aynı olduğu için bir tanesinin alındığını belirtmektedir. Denklem (1.20)'ye göre c_i değerinin 0 ile 1 arasında bir değer alacağı anlaşılmaktadır. $c_i = 0$ olması, i köşesinin komşuları arasında hiçbir bağlantının olmadığı ve $c_i = 1$ olması ise i köşesinin komşularının hepsinin birbirleri ile bağlantılı olduğu anlamına gelmektedir. Köşelerin kümelenme katsayısı c_i , ağın yerel özelliklerini oldukça net bir şekilde ortaya koymaktadır.

Ağın genel özellikleri hakkında bilgi sahibi olmak için ise,

$$C = \frac{1}{N} \sum_{i=1}^N c_i \quad (1.21)$$

şeklinde ağdaki tüm i köşeleri için bulunan kümelenme katsayılarının ortalaması alınmaktadır [109]. Burada C , ağın *ortalama kümelenme katsayısı* ve N de ağdaki toplam köşe sayısıdır. Denklem (1.20)'nin payda kısmına bakıldığında derecesi $d_i = 0$ ve $d_i = 1$ olan köşeler için bu denklem çalışmayacaktır. Bu durumu çözmek için, derecesi 2'den küçük olan köşelerin yani bağlantısız veya tek bağlantılı olduğundan üçgen oluşturamayacak köşelerin c_i değerleri sıfır alınır ve ağdan çıkarılarak

$$C^* = \frac{1}{1 - \frac{N_f}{N}} C = \frac{1}{N - N_f} \sum_{i=1}^N c_i \quad (1.22)$$

şeklinde yeni bir kümelenme katsayısı hesaplanır [108]. N_f , derecesi 2'den küçük olan köşelerin sayısıdır. Birçok gerçek dünya ağlarında kümelenme katsayısının değeri 0.1 ile 0.5 arasında değişmektedir [105].

Şekil 1.15'de 1 ve 3 numaralı köşeler arasında da bir kenar olduğu kabul edilirse 1 numaralı köşenin kümelenme katsayısı,

$$c_1 = \frac{\sum_{j \neq i} \sum_{h \neq (i,j)} A_{ij} A_{ih} A_{jh}}{d_i(d_i - 1)}$$

$$c_1 = \frac{A_{12}A_{13}A_{23} + A_{13}A_{12}A_{32}}{2(2 - 1)} = \frac{1.1.1 + 1.1.1}{2.1} \quad (1.23)$$

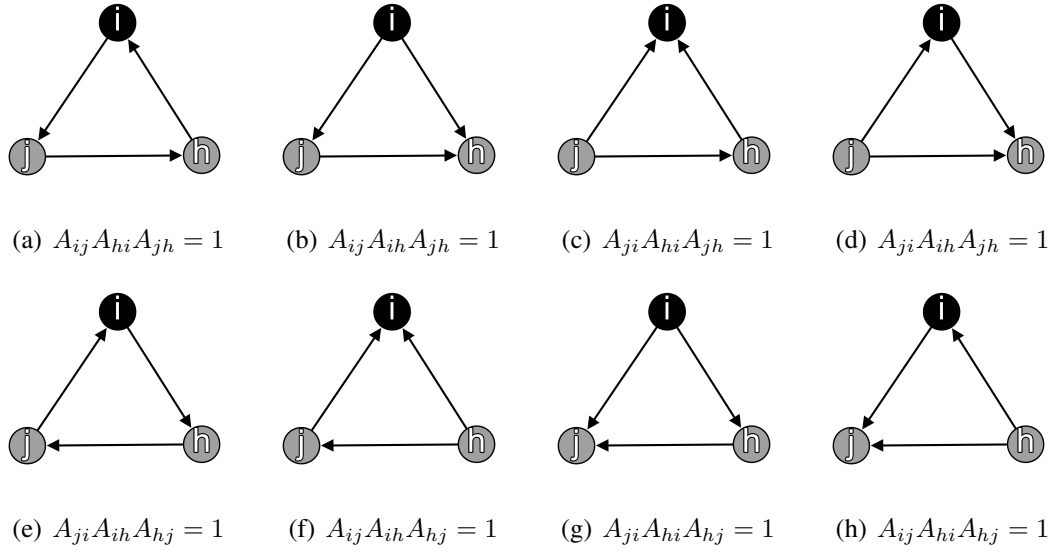
$$c_1 = 1$$

şeklinde elde edilir. Benzer şekilde $c_2 = 1$ ve $c_3 = 1$ olarak elde edilir. Ağın genel kümelenme katsayısı ise

$$C^* = \frac{1}{N - N_f} \sum_{i=1}^N c_i = \frac{1}{3 - 0} (c_1 + c_2 + c_3) = \frac{3}{3} = 1 \quad (1.24)$$

olarak elde edilir. Denklem (1.24), ağdaki bütün köşelerin birbirleri ile bağlantılı olduğunu belirtmektedir.

Yönelimli ağlarda ise ağdaki her bir köşe, bağlantılı olduğu köşelerle 8 farklı üçgen oluşturabilir [109].



Şekil 1.16. Yönelimli ağlarda oluşabilecek üçgenler. Kümelenmenin olacağı köşe i köşesi olarak belirlenmiş bu yüzden siyah renkle gösterilmiştir.

Herhangi bir i köşesi üzerinde oluşabilecek bir kümelenme için bu üçgenler Şekil 1.16'da verilmiştir. Ağ yönelimli olduğu için i köşesinin derecesi d_i^g giren ve d_i^c çıkan dereceler olmak üzere iki kısımdan oluşacağı Kesim 1.7.1'de belirtilmiş ve i köşesinin toplam derecesi Denklem (1.13)'de verilmiştir. Ağın yönelimli olması i köşesinin komşuları ile karşılıklı bir bağlantı içinde olmasına engel değildir ve eğer i köşesi böyle bir bağlantıya sahipse bu bağlantıların sayısı

$$d_i^{\leftrightarrow} = \sum_{j \neq i} A_{ij}A_{ji} \quad (1.25)$$

şeklinde olacaktır. Burada $\leftrightarrow$, i köşesinin komşuları ile karşılıklı etkileştiğini göstermektedir. Şekil 1.16'da varolan bütün üçgenleri üretecek

$$\text{Varolan üçgen sayısı} = (A_{ij} + A_{ji})(A_{ih} + A_{hi})(A_{jh} + A_{hj}) \quad (1.26)$$

şeklinde genel bir ifade tanımlanabilir. Böylelikle yönelimli ağlarda bir i köşesinin kümelenme katsayısı

$$c_i = \frac{\text{Varolan üçgen sayısı}}{\text{Varolan ve varolması muhtemel olan üçgen sayısı}} \quad (1.27)$$

$$= \frac{\frac{1}{2} \sum_{j \neq i} \sum_{h \neq (i,j)} (A_{ij} + A_{ji})(A_{ih} + A_{hi})(A_{jh} + A_{hj})}{d_i^t(d_i^t - 1) - 2d_i^{\leftrightarrow}}$$

ifadesi ile elde edilecektir [109]. Şekil 1.16'da (a) ile (e), (b) ile (f), (c) ile (g) ve (d) ile (h) arasında, i köşesinin komşuları arasındaki farklı yönelimlerden dolayı bir

simetri söz konusudur. Bu yüzden aslında 8 yerine (a), (b), (c) ve (d) olmak üzere 4 farklı üçgen vardır. Denklem (1.27)'nin pay kısmındaki $\frac{1}{2}$ bu simetriyi belirtmekte ve sadece farklı üçgenlerin alındığını ifade etmektedir. $d_i^t(d_i^t - 1)$, varolan ve varolması mümkün olan bütün üçgenlerin sayısını vermektedir. Ancak i köşesinin komşu ya da komşuları ile karşılıklı bağlantı durumu söz konusu olduğunda bu ifade fazla üçgenleri toplayacaktır. Bu nedenle doğru değerin bulunması için, hem i köşesi hem de karşılıklı bağlantı yaptığı komşusu için 2 tane fazla üçgen sayılacağından dolayı bu ifadeden $2d_i^{\leftrightarrow}$ değeri çıkarılmıştır. Tüm ağın kümelenme katsayısı ise yine Denklem (1.21) veya (1.22) ile hesaplanabilir.

1.7.4. Modülerite

Ağ yapılarında genellikle bir kaç köşeden daha fazla köşelerin bir arada bulunduğu, *topluluk* olarak adlandırılan alt ağ yapıları mevcuttur. Aynı topluluklarda bulunan köşeler arasında oldukça fazla sayıda kenarlar olmasına karşın, farklı topluluklarda bulunan köşeler arasında daha az sayıda bir etkileşme vardır. Ağın karakteristik yapısını belirlerken bu toplulukların incelenmesi ağ hakkında önemli bilgiler vermektedir. İncelenen bir ağın bu şekilde topluluklara nasıl ayrılabilirliğini belirlemek için çeşitli algoritmalar kurulmuştur. Bu algoritmalar, topluluklar arasındaki kenarları belirleyip onları ağdan kaldıran *bölücü algoritmalar* [105, 110, 111], kendi kendini tekrar ederek aynı köşe veya topluluğu birleştiren *büyütücü algoritma* [112] ve hedeflenen bir fonksiyonun maksimize edilmesine dayanan *en uygun hale getirici algoritmalar* [113–116]. Bu algoritmalarından elde edilen topluluk ayrımlarının kalitesi ise Newman ve Girvan [110, 117] tarafından tanımlanan ve *modülerite (modularity)* olarak adlandırılan bir kalite fonksiyonu ile belirlenir. Newman'ın [118] ağırlıklı ağlara da uyguladığı bu modülerite değeri, "Ağda aynı topluluk içerisinde bulunan köşeler arasındaki varolan toplam kenar sayıları ve, derecelerinin sabit kalması koşuluyla, köşelerin rastgele bağlanmasıyla kurulacak yeni ağdaki köşeler arasında olması beklenen toplam kenar sayısı arasındaki fark" olarak ifade edilir ve matematiksel olarak

$$Q = \left(\frac{1}{2m} \sum_{i,j} A_{ij} \delta(c_i, c_j) \right) - \left(\frac{1}{2m} \sum_{i,j} \frac{d_i d_j}{2m} \delta(c_i, c_j) \right) = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{d_i d_j}{2m} \right) \delta(c_i, c_j) \quad (1.28)$$

şeklinde verilir. A_{ij} , i ve j köşeleri arasındaki kenarı gösteren bitişiklik matrisi

elamanıdır. d_i ve d_j sırasıyla i ve j köşelerinin dereceleri. c_i ve c_j , i ve j köşelerinin ait oldukları toplulukları gösterirken $\delta(c_i, c_j)$ ise $c_i = c_j$ olduğunda 1 olarak alınır diğer durumda 0'dır. m ağdaki toplam kenar sayısını gösterir ve $m = \frac{1}{2} \sum_{ij} A_{ij}$ dir. Denklem (1.28) ile , bir ağın topluluklara ne kadar iyi ayrılıp ayrılmadığı başarılı bir şekilde belirlenebilir. Q değeri 0 ile 1 aralığında değişmektedir. $Q = 0$ değerini alması, verilen bir ağda aynı topluluk içinde bulunan köşeler arasındaki bağlantı sayısının, rastgele bağlanma durumundakinden daha iyi olmadığı anlamına gelmektedir. $Q = 1$ durumunda ise, ağın topluluklara en iyi şekilde ayrıldığı anlaşılmaktadır. Modülerite ifadesinin pratikteki tipik değeri 0.3 ile 0.7 aralığındadır [110, 118]. Newman ve Girvan [110] bu düşünce ile Zachary'nin karate kulübü ağı için başarılı bir topluluk yapısı elde etmiştir. Zachary'nin karate kulübü ağı, 1970'lerde bir Amerikan üniversitesindeki karate kulübü üyelerinin birbirleriyle olan ilişkileri üzerine kurulmuş bir arkadaşlık ağıdır. Ağ kurulduktan kısa bir süre sonra, kulüpteki üyeler arasında çıkan anlaşmazlıktan dolayı ağ 2 gruba ayrılmıştır. Newman ve Girvan bu ağ üzerinde topluluk incelemesi yaptıktan sonra, gerçek hayatta 2 gruba ayrılan bu ağı aynı şekilde 2 gruba ayırmıştır. Bu inceleme yapıldıktan Clauset et al.'nın [114] hızlı bir modülerite hesabı için yaptıkları çalışmada ise başarısız sonuçlara rastlanmaktadır. Ayrıca Newman'ın [117], modülerite matrisinin özvektör hesabına bağlı bir algoritma hesabında ise bazen hatalı sonuçlara ulaşılmaktadır. Newman'ın [119], gerek bilgisayarda üretilen gerekse gerçek dünyada bulunan ağlar ile karşılaştırma yaparak en uygun modülerite değerini bulan bir metot geliştirmesine rağmen, çok fazla sayıda kombinasyon hesaplanmasını gerektirdiğinden kesin bir modülerite değeri bulmak oldukça zordur. Blondel et al. [113], modülerite değerinin maksimize edilmesi için hızlı ve güvenilir bir metot geliştirmişlerdir. Diğer metotlarla karşılaştırıldığında geliştirilen bu metot, büyük ağlar için oldukça kısa zaman içinde oldukça yüksek modülerite değerini bulmaktadır. N tane köşeden oluşan bir ağ için metodun çalışması ise şu şekildedir:

Adım 1 Ağdaki bütün köşelerin her biri farklı birer topluluk olarak düşünülür. Bu nedenle başlangıçta ağda ne kadar köşe varsa o kadar topluluk vardır.

Adım 2 Her bir i köşesi kendi topluluğundan kaldırılıp, komşusu olan j köşesinin topluluğu içine katılması durumunda modülerite değerinde pozitif bir artışın olup olmadığına bakılır ve en büyük artışın olduğu komşu topluluğa yerleştirilir. Eğer eklendiği komşularından hiç birinde pozitif bir modülerite artışı yoksa kendi topluluğunda

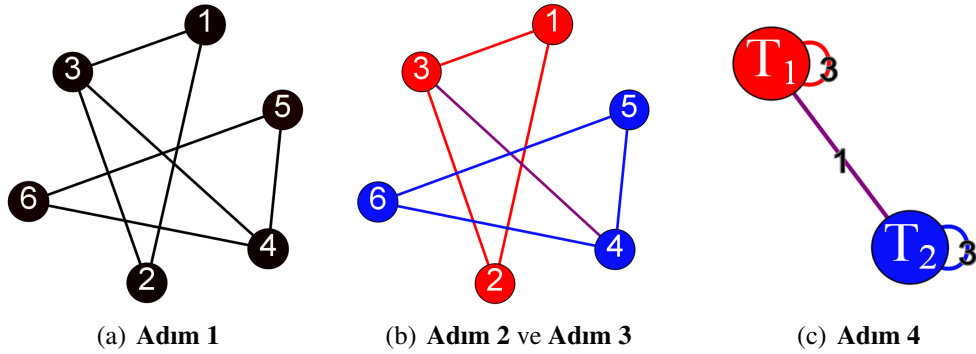
bırakılır. Modüleritedeki artış ise

$$\Delta Q = \left[\frac{(\sum_{in}) + (2k_{i,in})}{2m} - \left(\frac{(\sum_{tot}) + (k_i)}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right] \quad (1.29)$$

ifadesi ile belirlenir. Burada $\sum_{in}$, i köşesinin ekleneceği topluluğun içindeki kenarların toplam sayısı, $k_{i,in}$, i köşesinin ekleneceği toplulukta bulunan köşelere yaptığı kenarların sayısı, $\sum_{tot}$, i köşesinin ekleneceği toplulukta bulunan köşelere bağlanmış olan kenarların toplam sayısı, k_i , i köşesine bağlanan kenarların sayısı ve m ise ağda bulunan kenarların toplam sayısıdır.

Adım 3 Bir önceki adım ağdaki bütün köşeler için tekrarlanır.

Adım 4 Sürecin sonunda, köşelerin bir arada olduğu çeşitli topluluklar oluşmaktadır. Bu toplulukların kendileri de birer köşe gibi düşünülerek yeni bir ağ oluşturulur ve ilk 3 adımdaki süreç tekrar edilir. Böylece uygun toplulukların da birleştirilmesiyle en uygun modülerite değerine ulaşılır.



Şekil 1.17. Metodun adımlarının örnek bir ağ üzerinde gösterimi.

Şekil 1.17’de metodun daha iyi anlaşılması için, adımlar örnek bir ağ üzerinde teker teker gösterilmiştir. Şekil 1.17(a)’da metodun ilk adımı olduğu için ağda 6 tane topluluk mevcuttur. Şekil 1.17(b)’de ikinci ve üçüncü adımların uygulanması ile en yüksek modülerite değerine sahip olan ağda 1, 2 ve 3 numaralı köşeler ile 4, 5 ve 6 numaralı köşelerden oluşan iki ayrı topluluğun olduğu ve topluluk sayısının Şekil 1.17(a)’ya göre azaldığı görülmektedir. Aynı toplulukta olan köşeler ve bu köşeler arasındaki kenarlar aynı renkle gösterilirken, farklı topluluklarda olan köşeler arasındaki kenarlar farklı renkle gösterilmiştir. Şekil 1.17(c)’de ise aynı toplulukta olan köşeler bir araya toplanmıştır. 1, 2 ve 3 numaralı köşeler T_1 topluluğunda, 4, 5 ve 6 numaralı köşeler ise T_2 topluluğundadır.

Görüldüğü gibi kendi başlarına birer köşe gibi düşünülen T_1 ve T_2 topluluklarında öz-kenar durumu mevcuttur. Her bir köşenin öz-kenarı, o topluluk içerisinde toplam kaç tane kenar olduğunu belirtmektedir. Bu durumda T_1 ve T_2 toplulukları içerilerinde toplam 3 tane kenar barındırmaktadır. T_1 ve T_2 toplulukları arasında ise sadece 3 ve 4 numaralı köşeler arasındaki 1 tane kenar ile bağlanmaktadır. Topluluklar içerisinde, topluluklar arasına kıyasla daha fazla sayıda kenar olması modülerite kavramı ile uyumluluk sergilemektedir. Bu durumda dördüncü adımın uygulanması ile, T_1 ve T_2 topluluklarının birleştirilmesi modüleritede artış sağlamadığından ağ iki topluluğa ayrılmış olur.

1.8. Ağ Modelleri

Bu kesimde, Kesim 1.6 ve 1.7’de matematiksel temelleri ve karakteristik özellikleri incelenen ağ sistemleri için; her biri kendine has motivasyona, ağ kurma yöntemine sahip olan ve literatürde sıkça üzerinde durulan ağ modelleri anlatılacaktır.

1.8.1. Rastgele Ağ Modeli

Rastgele ağlar için sistematik çalışmaların temeli, rastgele bağlantıların artan sayısının bir fonksiyonu olarak olasılıkçı metotlar şeklinde çalışmayı amaçlayan Erdős ve Rényi [38, 39] tarafından atılmıştır. Rastgele ağ ifadesi, ağdaki farklı köşeler arasındaki kenarların rastgele, düzensiz bir şekilde belirlendiği anlamına gelmektedir. Erdős ve Rényi yaptıkları ilk çalışmada, N tane köşeye ve K tane kenara sahip rastgele bir ağ üretmeyi amaçlamışlardır. Bunun için N tane izole yani hiçbir kenara sahip olmayan köşe ile başlamışlar ve köşeleri birbirlerine bağlayan kenarları, $N(N - 1)/2$ tane olasılık içerisinde rastgele seçerek kenar sayısı K olana kadar devam etmişlerdir. Bu incelemede, hiçbir köşenin çoklu kenar yapmasına müsaade edilmemiştir [38]. Diğer bir çalışmada ise, N tane izole köşeyi birbirine $0 < p < 1$ aralığındaki bir p ihtimaliyeti ile bağlamışlardır [39]. Ağda olması beklenen kenar sayısı ise

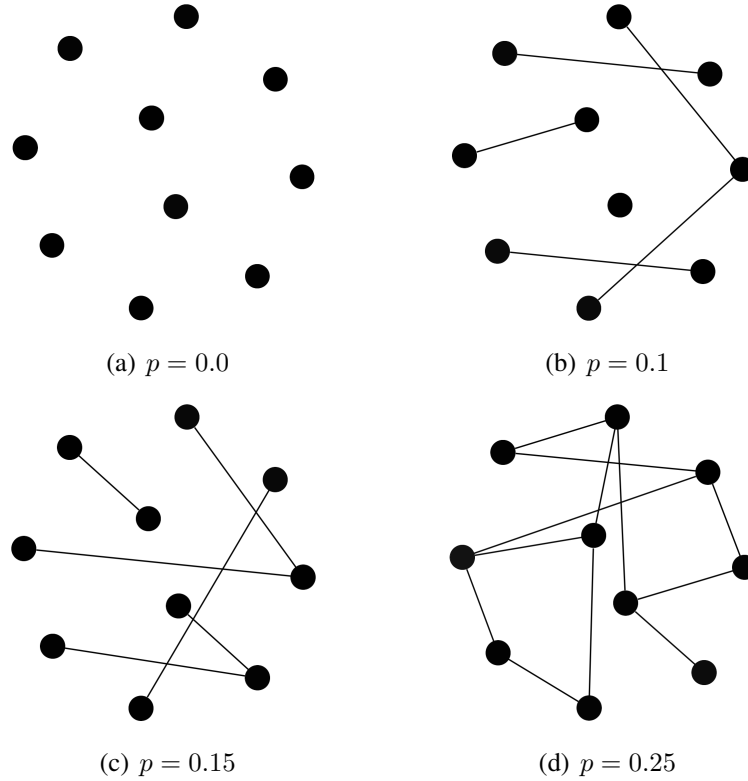
$$E(K) = p \frac{N(N - 1)}{2} \quad (1.30)$$

şeklinde dir. Dolayısıyla maksimum kenar sayısı $p = 1$ durumunda $N(N - 1)/2$ ’dir. Bu yöntem ile ağda farklı kenar sayılarına sahip topluluklar tanımlanmaktadır. Böylelikle N

tane köşe ve K tane kenara sahip ağın elde edilme ihtimaliyeti

$$P(G(N, K)) = p^K (1 - p)^{N(N-1)/2 - K} \quad (1.31)$$

olacaktır. [97, 101]. Denklem (1.31), ağda p^K ihtimaliyet ile kenarların bulunması ve $(1 - p)^{N(N-1)/2 - K}$ ihtimaliyet ile de kenarların bulunmaması olarak yorumlanabilir.



Şekil 1.18. Farklı p ihtimaliyetlerine göre bağlanmış $N = 10$ tane köşeye sahip rastgele ağ örnekleri.

Şekil 1.18’de $N = 10$ tane köşeden oluşan ve farklı p bağlanma ihtimallerine göre kurulmuş ağ örnekleri verilmiştir. Şekil 1.18(a)’da $p = 0.0$ yani bütün köşeler izole ve ağda hiçbir kenar bulunmamaktadır. Bağlanma ihtimalinin $p = 0.1$ olduğu Şekil 1.18(b)’de izole köşe ve ayrık toplulukların varlığı söz konusudur. Şekil 1.18(c)’de bağlanma ihtimaliyeti artırılmış ve $p = 0.15$ olarak alınmıştır. Bu durumda ağda izole köşeler olmamasına rağmen ağın hala ayrık topluluklardan oluştuğu görülmektedir. Son olarak Şekil 1.18(d)’de $p = 0.25$ alındığında ağda ne izole köşeler ne de ayrık topluluklar bulunmaktadır.

Erdős ve Rényi tarafından kurulan rastgele ağların yapısal özellikleri p bağlanma ihtimaliyetine göre değiştiği için bu ihtimaliyetin $p_c = \frac{1}{N}$ gibi belirli bir kritik değeri

olmalıdır. Bu ihtiyacın sebebi, ağın tamamen bağlantılı yani ayrık toplulukların olup olmadığını belirlemektir. Ağın tamamen bağlantılı olup olmadığı ise

$$\lim_{N \rightarrow \infty} P_{N,p} = \begin{cases} 0 & , \text{ eğer } \frac{p}{p_c} \rightarrow 0 \text{ ise} \\ 1 & , \text{ eğer } \frac{p}{p_c} \rightarrow \infty \text{ ise} \end{cases} \quad (1.32a)$$

$$(1.32b)$$

şeklinde bir $P_{N,p}$ ihtimaliyetinin $N \rightarrow \infty$ limit değeri ile belirlenir. Eğer p değerindeki artış p_c değerindeki artıştan daha az ise Denklem (1.32a)'da görüldüğü gibi $P_{N,p} = 0$ olacak ve ağda ayrık topluluklar olacaktır. Eğer p değerindeki artış p_c değerindeki artıştan daha fazla ise bu defa Denklem (1.32b)'ye göre $P_{N,p} = 1$ olacak ve ağ tamamen bağlantılı olacaktır. Ağın ortalama derecesi ise

$$\langle d \rangle = p(N - 1) \cong pN \quad \{N \rightarrow \infty \text{ iken}\} \quad (1.33)$$

ifadesi ile verilmektedir [101].

İlk olarak Erdős ve Rényi [38] rastgele ağlar için maksimum ve minimum derece dağılımlarını çalıştıktan sonra bütün bir derece dağılımı Bollobás [120] tarafından çalışılmıştır. Bağlanma ihtimaliyeti p olan bir ağda bir i köşesinin d_i derecesine sahip olma ihtimaliyeti

$$P(d_i) = \binom{N-1}{K} p^K (1-p)^{N-1-K} \quad (1.34)$$

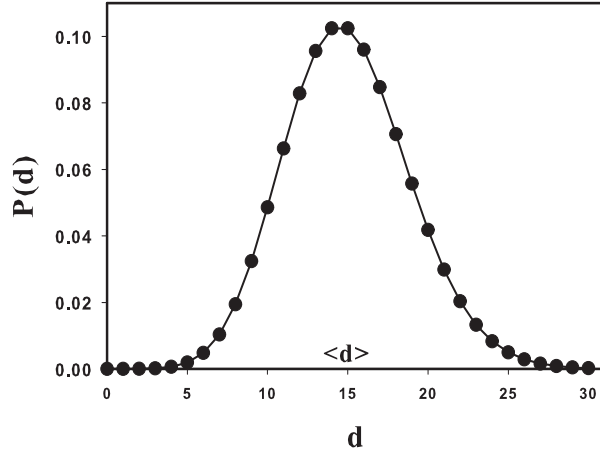
şeklinde bir Binom dağılımı ile verilmektedir. Burada $\binom{N-1}{K}$ ifadesi, i köşesinin K tane kenarı ağdaki diğer $N - 1$ tane köşeden hangileri ile kenar yapabileceğini, p^K K tane kenar yapılma ihtimalini ve p^{N-1-K} ise K tane kenar yapılmama ihtimalini belirtmektedir. Eğer N çok büyük değerler alırsa derece dağılımı

$$P(d_i) = e^{-\langle d \rangle} \frac{\langle d \rangle^d}{d!} \quad (1.35)$$

şeklinde bir Poisson dağılımı olacaktır.

Şekil 1.19'de Denklem (1.35)'ün sergilediği örnek bir dağılım verilmiştir. Denklem (1.33)'e göre $N \gg 1$ olduğu için $\langle d \rangle = pN = 15$ olduğu görülmektedir. Buradan anlaşılacağı üzere, ağda $P(d) = 0.1024$ ihtimaliyeti ile bulunma ihtimaliyeti en yüksek olan derece $d = 15$ derecesidir.

Bir ağın çapı, ağın ortalama yol uzunluğu için bir üst sınır belirler. Bunun için ilk önce rastgele ağlardaki çap tanımlanmalıdır [101]. Ayrık toplulukların bulunduğu bir ağın çapı



Şekil 1.19. $N = 10000$ ve $p = 0.0015$ değerleri alınarak oluşturulmuş bir rastgele ağ için Poisson derece dağılımı.

sonsuz olacaktır. Bu durumdan kurtulmak için, ağdaki en büyük topluluk ele alınarak çap belirlenir. Bu düşünce ile Chung ve Lu [121], aynı N ve p değerlerine sahip rastgele ağlar için çap ifadesinin hemen hemen aynı olduğunu belirtmiş ve

$$D_r = \frac{\ln(N)}{\ln(pN)} = \frac{\ln(N)}{\ln(\langle d \rangle)} \quad (1.36)$$

şeklinde tanımlamışlardır. r indisi, rastgele ağı belirtmektedir. Bu noktada birkaç önemli sonuç ortaya çıkmaktadır. Bunlar

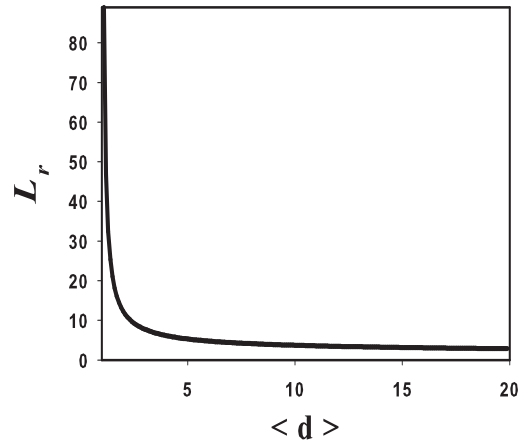
- a)** Eğer $\langle d \rangle = pN < 1$ ise, ağda izole olarak düşünülebilecek en fazla $d = 2$ derecesine sahip olan ayrık topluluklar bulunacaktır. Ağın çapı ise 2 olacaktır.
- b)** Eğer $\langle d \rangle = pN > 1$ ise, ağda $d > 2$ derecesine sahip olan ayrık topluluklar bulunacaktır. Ağın çapı ise en büyük d derecesine sahip topluluğun çapına eşit olacaktır.
- c)** Eğer $\langle d \rangle = pN \geq \ln(N)$ ise, ağda ayrık topluluklar bulunmayacak yani ağ tamamen bağlantılı olacaktır. Ağın çapı ise Denklem (1.36) ile elde edilecektir.

Ağdaki bilgi akışının incelenmesini sağlayan ve çap ile sınırları çizilen ortalama yol uzunluğu yine çap gibi hesaplanır ve

$$L_r \sim \frac{\ln(N)}{\ln(\langle d \rangle)} \quad (1.37)$$

şeklindedir. r indisi, rastgele ağı belirtmektedir.

Şekil 1.20'da Denklem (1.37) ile elde edilen ortalama yol uzunluğunun ağın ortalama derecesine göre değişim grafiği verilmiştir. Küçük $\langle d \rangle$ değerlerinde ağ üzerindeki



Şekil 1.20. $N = 5000$ tane köşeden oluşan bir rastgele ağ için ortalama yol uzunluğunun ağın ortalama derecesine göre değişimi.

bilgi akışı için uzun mesafeli yol almak gerekirken, $\langle d \rangle$ değerinin artması ile ağdaki herhangi iki köşe arasındaki ortalama yol uzunluğu kılalacaktır. Ortalama yol uzunluğundaki bu azalma logaritmik bir şekilde gerçekleşmektedir.

Tablo 1.5. Gerçek dünya ağlarının sahip oldukları L ortalama yol uzunluklarının rastgele ağ modeli ile elde edilen L_r ortalama yol uzunlukları ile karşılaştırılması.

Ağ	N	$\langle d \rangle$	L	L_r
Yiyecek Ağı [122]	182	26.05	1.89	1.77
Sinir Ağı [41]	282	14	2.65	2.25
Enerji Nakli Hattı [41]	4941	2.7	18.7	12.4

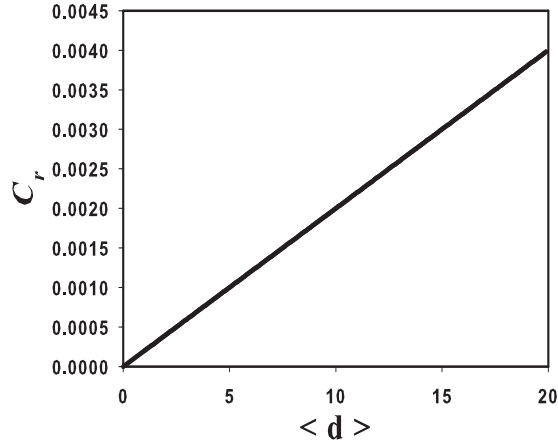
Tablo 1.5’de bazı gerçek dünya ağları için verilen ortalama yol uzunluğu değerlerine bakıldığında yüksek $\langle d \rangle$ değerleri için L ve L_r değerlerinin birbirine çok yakın olduğu görülmektedir. Diğer taraftan $\langle d \rangle$ değeri azaldıkça hem L hem de L_r değerinde artış meydana gelmektedir.

Rastgele ağlarda herhangi bir köşenin en yakın iki komşusu ile bağlantılı olma ihtimali, ağdan rastgele seçilen iki köşenin bağlantılı olma ihtimaline eşittir [101, 109, 123]. Bu düşünce ile rastgele ağlar için kümelenme katsayısı

$$C_r = \frac{\langle d \rangle}{(N - 1)} = \frac{p(N - 1)}{(N - 1)} = p \quad (1.38)$$

şeklinde olacaktır. r indisi, rastgele ağı belirtmektedir. Açıkça görüldüğü gibi rastgele

ağların kümelenme katsayısı C_r , ağın ortalama derecesinin ağdaki toplam köşe sayısına oranı olarak ifade edilmektedir. Bu ise Denklem (1.33) gereği ağdaki kenarların bağlanma ihtimaliyetine eşit olacaktır.



Şekil 1.21. $N = 5000$ köşeye sahip bir rastgele ağdaki kümelenme katsayısının ortalama derece ifadesine göre değişimi.

Şekil 1.21’de kümelenme katsayısının ağın ortalama derece değeri ile değişim grafiği verilmiştir. Denklem (1.38)’den de anlaşılacağı gibi C_r ’nin $< d >$ ile lineer bir şekilde değiştiği görülmektedir. Bu durum, ağdaki köşelerin bağlanma ihtimali ne kadar fazla ise ağın kümelenme katsayısının da o kadar fazla olacağını belirtir. Ancak rastgele ağlardaki kümelenme katsayısı, gerçek dünya ağlarındaki kümelenme katsayısı ile karşılaştırıldığında oldukça küçük olduğu görülmektedir [123]. Bu yüzden rastgele ağ modelinde incelenen kümelenme özelliği, gerçek dünya ağları ile uyumamaktadır [101].

Tablo 1.6. Gerçek dünya ağlarının sahip oldukları C kümelenme katsayılarının rastgele ağ modeli ile elde edilen C_r kümelenme katsayıları ile karşılaştırılması.

Ağ	N	$< d >$	C	C_r
Enerji Nakli Hattı [41]	4941	2.7	0.08	0.00054
İnternet [124]	6374	3.8	0.24	0.0006
Yiyecek Ağı [125]	134	8.7	0.22	0.065
Sinir Ağı [41]	282	14	0.28	0.049

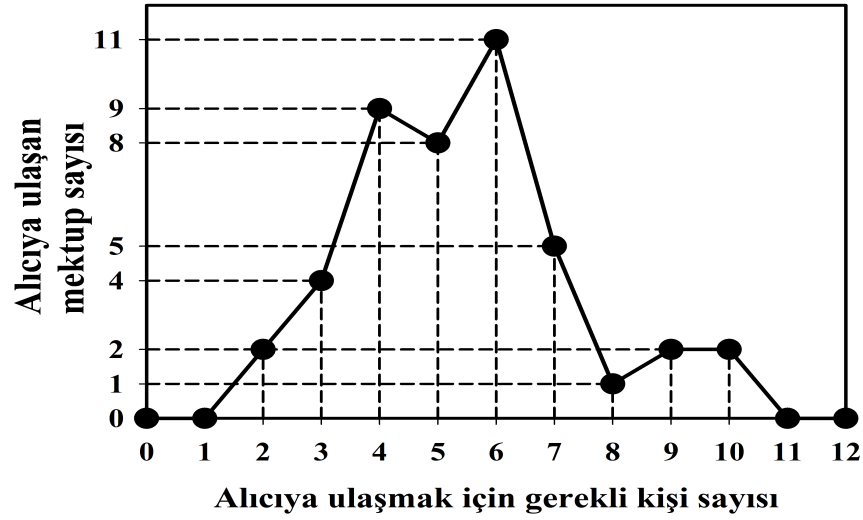
Tablo 1.6’de N tane köşeden oluşan ve ortalama derecesi $< d >$ olan bazı gerçek dünya ağları için kümelenme katsayısı değerleri verilmiştir. Verilen bütün ağlarda, rastgele ağlar

gerçek dünya ağları ile aynı köşe sayısına ve aynı ortalama dereceye sahip olmasına rağmen bütün ağlar için $C \gg C_r$ olduğu açıkça görülmektedir.

1.8.2. Küçük Dünya Ağları

Küçük dünya ağları ile ilgili en gerçekçi çalışma Milgram [40] tarafından yapılmıştır. Milgram yaptığı bu deneyde, Amerika'da birbirlerini tanımayan iki insan arasındaki en kısa yakınlık ilişkisi zincirini bulmayı amaçlamıştır. Deneyde Nebraska'dan seçilen kişilere verilen bir mektubun, Massachusetts'deki tanımadıkları bir kişiye iletmeleri istenmiştir. Bu iki şehrin seçilme nedeni, Amerika'da birbirine en uzak iki şehir olmasıdır. Başlangıçta, Nebraska'dan rastgele seçilen kişilere gönderilen davet yazısında, mektubun gönderileceği alıcı kişi hakkında çalıştığı yer, yaşadığı yer, gittiği okul gibi oldukça temel bilgiler verilmektedir. Ayrıca davet yazısına kendi isimlerini yazacakları bir isim listesi de eklenmiştir. Seçilen kişilere davet yazısında, mektubun gönderileceği alıcı kişiyi kişisel olarak tanıyıp tanımadığı sorulmaktadır. Eğer tanıyorsa isim listesine kendi ismini ve alıcının ismini yazarak mektupla birlikte doğrudan alıcı kişiye göndermesi istenmiştir. Çoğu durumda mektubu gönderen kişi alıcı kişiyi kişisel olarak tanımamaktadır. Bu durumda, kendi ismini ve alıcıyı daha iyi tanıyabileceğini düşündüğü arkadaşının ismini isim listesine yazarak mektupla birlikte arkadaşına göndermesi istenmiştir. Her iki durumda da sürecin ilerleyişinin takip edilebilmesi için, gönderen kişinin kendi ismini ve alıcının ya da arkadaşının ismini yazdığı isim listesini Harvard'daki araştırmacılara da göndermesi istenmiştir. Mektup alıcıya ulaşana kadar bu şekilde devam edilecektir. Mektup alıcıya ulaştığında, araştırmacılar isim listelerini inceleyerek alıcıya kaç kişi üzerinden ulaşıldığını belirleyeceklerdir. Nebraska'da 160 farklı kişi ile başlanan bu deneyde 44 kişi mektubu alıcıya ulaştırabilmiştir. Çoğu durumda da, alıcıya çok kısa bir mesafe kala süreç durmuş ve mektup iletilenmemiştir.

Şekil 1.22'de görüldüğü gibi mektupları alıcıya ulaştırmak için gerekli kişi sayısı 2 ile 10 arasında değişmekle beraber ortalama değer 6 kişidir. Bu ortalama değer 200 milyon nüfuslu Amerika'da birbirini tanımayan insanlar arasındaki yakınlık derecesinin en büyük değerini vermektedir [40]. Küçük dünya ağları için bu deney "6 ayrılık derecesi (six degrees of separation)" olarak ifade edilmektedir.



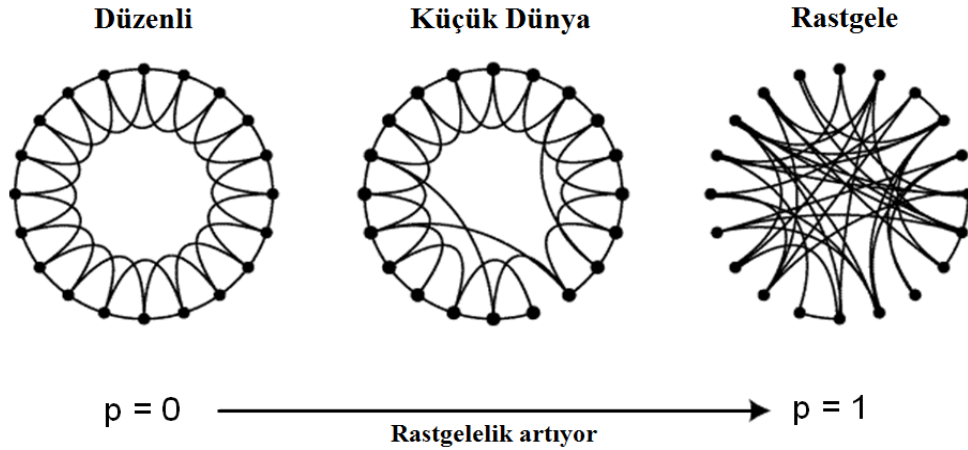
Şekil 1.22. Milgram'ın deneyinde alıcıya ulaşan 44 mektubun, alıcıya ulaşmak için gerekli kişi sayısına göre dağılımı. Ref.[40]

Milgram'ın bu deneyinden sonra Watts ve Strogatz [41] küçük dünya ağları için, düzenli örgüler kadar yüksek kümelenme katsayısına ve rastgele ağlar kadar düşük ortalama yol uzunluğuna sahip teorik bir model geliştirmişlerdir. Modelin işleyişi,

Adım 1 N tane köşe ve her bir köşenin yönelimsiz olarak k tane en yakın komşuya (saat yönünde $k/2$ tane) sahip olduğu düzenli bir örgü ile başlanır. Düzenli örgüde hiçbir ayrık topluluk bulunmamakla beraber örgü olarak adlandırılan ağda her zaman $N \gg k \gg \ln(N)$ yani tamamen bağlantılı olma durumu söz konusudur.

Adım 2 Örgüdeki her bir kenar $0 < p < 1$ aralığında değişen bir p ihtimaliyetiyle ve öz ve çoklu kenar durumları da ihmal edilerek rastgele değiştirilir. şeklinde verilmektedir.

Şekil 1.23'da Watts ve Strogatz modelinin işleyişi için basit şematik bir gösterim verilmiştir. Başlangıçta ağ, $p = 0$ iken $N = 20$ ve $k = 4$ değerlerine sahip, yönelimsiz kenarlardan oluşan düzenli bir örgü yapısına sahiptir. Bu durumda iken, ağda bir köşe ve bu köşenin saat yönündeki en yakın komşusu ile yaptığı kenar seçilir. Seçilen bu kenar p ihtimaliyeti ile ağdaki başka bir köşeye öz ve çoklu kenarlara müsaade edilmeden yeniden bağlanır. Bu süreç tüm köşeler için saat yönünde tekrarlanır. Daha sonra bir köşe ve bu köşenin yine saat yönündeki ikinci en yakın komşusu ile yaptığı kenar seçilerek en yakın komşu için yapılan p ihtimaliyeti ile yeniden bağlama süreci aynı şekilde tüm köşeler için uygulanır. Her bir köşe saat yönünde $k/2$ tane en yakın komşuya sahip



Şekil 1.23. Watts ve Strogatz tarafından elde edilen, düzenli örgü ile rastgele ağ arasında artan p değerine göre kenarların rastgele yeniden bağlanması şematik gösterimi. Ref.[41]

olacağı için yeniden bağlama süreci her bir köşenin sahip olduğu kenarlar için en fazla $k/2$ defa gerçekleşecektir. Modelin bu işleyişi, $p = 0$ olan tamamen düzenli ağlar ile $p = 1$ olan tamamen rastgele ağlar arasında farklı p değerleri için gerçekleştirilebilir. Önemli olan nokta, ağdaki toplam köşe ve kenar sayılarının daima aynı kalmasıdır. p değerindeki artış ağdaki düzensizliğin yani rastgeleliğin artması anlamına gelmektedir. Watts ve Strogatz'ın küçük dünya ağları $0 < p < 1$ aralığında bir p ihtimaliyetine sahip olup, ne tamamen düzenli ne de tamamen rastgele bir ağıdır. Ortalama yol uzunluğu ve kümelenme katsayısı ifadeleri $p \rightarrow 0$ iken

$$L \sim \frac{N}{2k} \gg 1 \quad (1.39)$$

ve

$$C = \frac{3(k-2)}{4(k-1)} \approx \frac{3}{4} \quad (1.40)$$

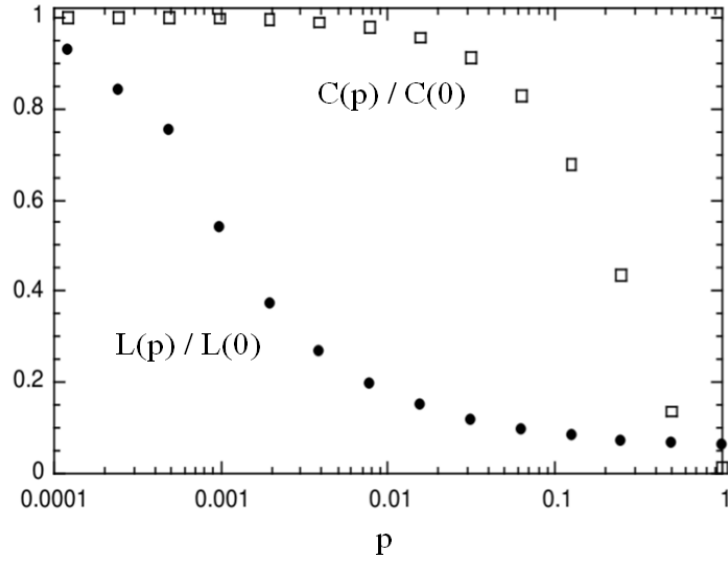
olarak ve $p \rightarrow 1$ iken

$$L \approx L_r \sim \frac{\ln(N)}{\ln(k)} \quad (1.41)$$

ve

$$C \approx C_r \sim \frac{k}{N} \ll 1 \quad (1.42)$$

olarak ifade edilmektedir [41]. Denklem (1.39)'a bakıldığında L , N ile lineer olarak artarken Denklem (1.41)'de logaritmik olarak artmaktadır. Aynı zamanda Denklem (1.40)'daki C değeri Denklem (1.42)'dekinden daha büyük olduğu görülmektedir.



Şekil 1.24. Watts ve Strogatz tarafından $N = 1000$ ve $k = 10$ için elde edilen ortalama yol uzunluğunun ve kümelenme katsayısının p ihtimaliyetine göre değişimi. Ref.[41]

Şekil 1.24’de ağın genel özelliği olan ortalama yol uzunluğu $L(p)$ ve yerel özellik olan kümelenme katsayısı $C(p)$ ’nin bağlanma ihtimaliyeti p ’ye göre değişimi verilmiştir. Her iki değerde $p = 0$ daki değerlerine bölünerek normalize edilmiştir. $p = 0$ ’daki düzenli örgü için hem $L(p)/L(0)$ hem de $C(p)/C(0)$ ifadeleri en yüksek değerlerine sahiptir. $p = 1$ ’de ise rastgele ağ için hem $L(p)/L(0)$ hem de $C(p)/C(0)$ ifadeleri en düşük değerlerine sahiptir. Ancak $0 < p < 1$ aralığındaki bazı p değerlerinde $L(p)/L(0)$ ve $C(p)/C(0)$ değerleri arasında geniş bir aralığın olduğu görülmektedir. Bu değerlerdeki $L(p)$ değerlerinin, düzenli ağlardan oldukça düşük neredeyse rastgele ağlardaki $L(1)$ değerine eşit olduğu görülmektedir. Diğer taraftan $C(p)$ değerlerinin ise rastgele ağlardan oldukça büyük neredeyse düzenli ağlardaki $C(0)$ değerine eşit olduğu görülmektedir. Sonuç olarak bu p değerlerinde, ağ ne tamamen düzenli ne de tamamen rastgeledir. Bu şekildeki ağlar küçük dünya ağları olarak adlandırılmaktadır.

Küçük dünya ağları için derece dağılımı çalışması ise Barrat ve Weigt [126] tarafından incelenmiştir. Başlangıçta $p = 0$ yani ağ düzenli bir halka örgü şeklinde iken her bir köşenin derecesi $d = k$ dır. Sadece saat yönündeki komşuluklar düşünüldüğünde her bir köşenin derecesi $d = k/2$ dir. Ancak $p \neq 0$ iken ağda rastgele yeniden bağlanmalar ortaya çıkacağı için her köşenin sahip olduğu derece değişecek ancak ortalama derece

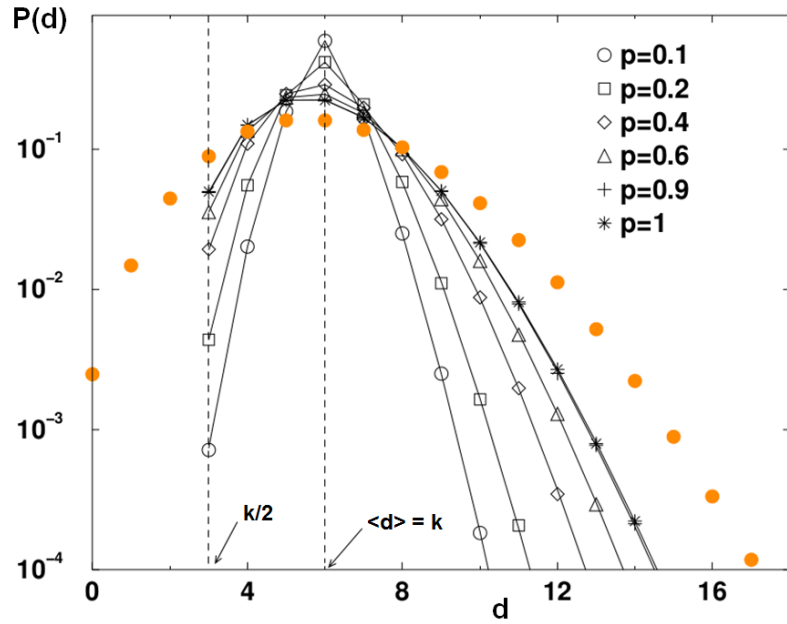
yine $\langle d \rangle = k$ olacaktır. Yeni bağlanmalar ortaya çıktıkça ağdaki bir i köşesinin yeni derecesi

$$d_i = \frac{k}{2} + n_i \quad (1.43)$$

olacaktır. Buradaki n_i , i köşesinin saat yönündeki komşuları ile olan bağlantısının değişmediği yani $1 - p$ ihtimaliyetle yeniden bağlanmanın olmadığı kenar sayısı $n_i^1 \leq k/2$ ve N ağdaki köşe sayısı olmak üzere p/N ihtimaliyetle yeniden bağlanan kenar sayısı $n_i^2 = n_i - n_i^1$ olarak iki kısımdan oluşmaktadır. Böylelikle ağın derece dağılımı yani bir köşenin d derecesine sahip olma ihtimali $P(d)$

$$P(d) = \sum_{n=0}^{\min(d-\frac{k}{2}, \frac{k}{2})} \binom{\frac{k}{2}}{n} (1-p)^n p^{\frac{k}{2}-n} \frac{(kp)^{d-\frac{k}{2}-n}}{(d-\frac{k}{2}-n)!} e^{(-p\frac{k}{2})}, \quad d \geq \frac{k}{2} \quad (1.44)$$

şeklinde verilmiştir [126].



Şekil 1.25. Barrat ve Weigt tarafından $N = 1000$ ve $\langle d \rangle = k = 6$ değerlerine göre farklı p değerleri için elde edilen derece dağılımı. Ref.[126]

Şekil 1.25’de küçük dünya ağları için Denklem (1.44)’e göre farklı p değerlerindeki derece dağılımları farklı sembollerle gösterilmiştir. Turuncu renkli daireler ise rastgele ağlar için $\langle d \rangle = k = 6$ değerine göre Denklem (1.35) ile elde edilmiştir. Şekil itibari ile küçük dünya ağlarındaki derece dağılımlarının rastgele ağlardaki derece dağılımına kısmen benzediği görülmektedir. Bütün dağılımlar $\langle d \rangle = k = 6$ değerinde bir pike

sahiptirler. Küçük dünya ağlarında p değerinin artması ile dağılımların daha geniş aralığa sahip olduğu anlaşılmaktadır.

Tablo 1.7. Küçük dünya ağları için Watts ve Strogatz tarafından verilen bazı deneysel örnekler. Ref.[41]

Ağ	L	L_r	C	C_r
Film aktörleri	3.65	2.99	0.079	0.00027
Enerji nakil hattı	18.7	12.4	0.080	0.005
Sinir Ağı	2.65	2.25	0.28	0.05

Tablo 1.7’de, ortalama yol uzunluğu ve kümelenme katsayısı için Watts ve Strogatz tarafından yapılan küçük dünya ağlarının rastgele ağlarla karşılaştırılması verilmiştir. Film aktörleri ağında $N = 225226$ ve $\langle d \rangle = 61$, enerji nakil hattı ağında $N = 4941$ ve $\langle d \rangle = 2.67$ ve sinir ağında $N = 282$ ve $\langle d \rangle = 14$ dir. Küçük dünya ağlarının karakteristiği olarak $L \geq L_r$ ve $C \gg C_r$ olduğu görülmektedir.

1.8.3. Ölçeksiz Ağlar

İlk olarak Barabási ve Albert [42] tarafından geliştirilen modelde, rastgele ve küçük dünya ağlarının aksine birçok gerçek dünya ağlarının açık bir sistem olduğu ve ağın, yeni köşelerin eklenmesiyle *büyümesi* gerektiği ifade edilmektedir. Rastgele ve küçük dünya ağlarında iki köşe arasındaki bağlanma ihtimaliyeti köşelerin derecelerinden bağımsızdır. Ancak bu durumun aksine Barabási ve Albert, birçok gerçek dünya ağlarında iki köşe arasındaki bağlantıların *tercihli eklendiğini* yani köşelerin sahip oldukları dereceler bu köşelere yapılacak bağlantılar için önemli bir role sahip olduğunu belirtmişlerdir. Bu iki unsur Barabási ve Albert modelinin temelini oluşturmaktadır. Bu unsurlar yardımı ile modelin işleyişi

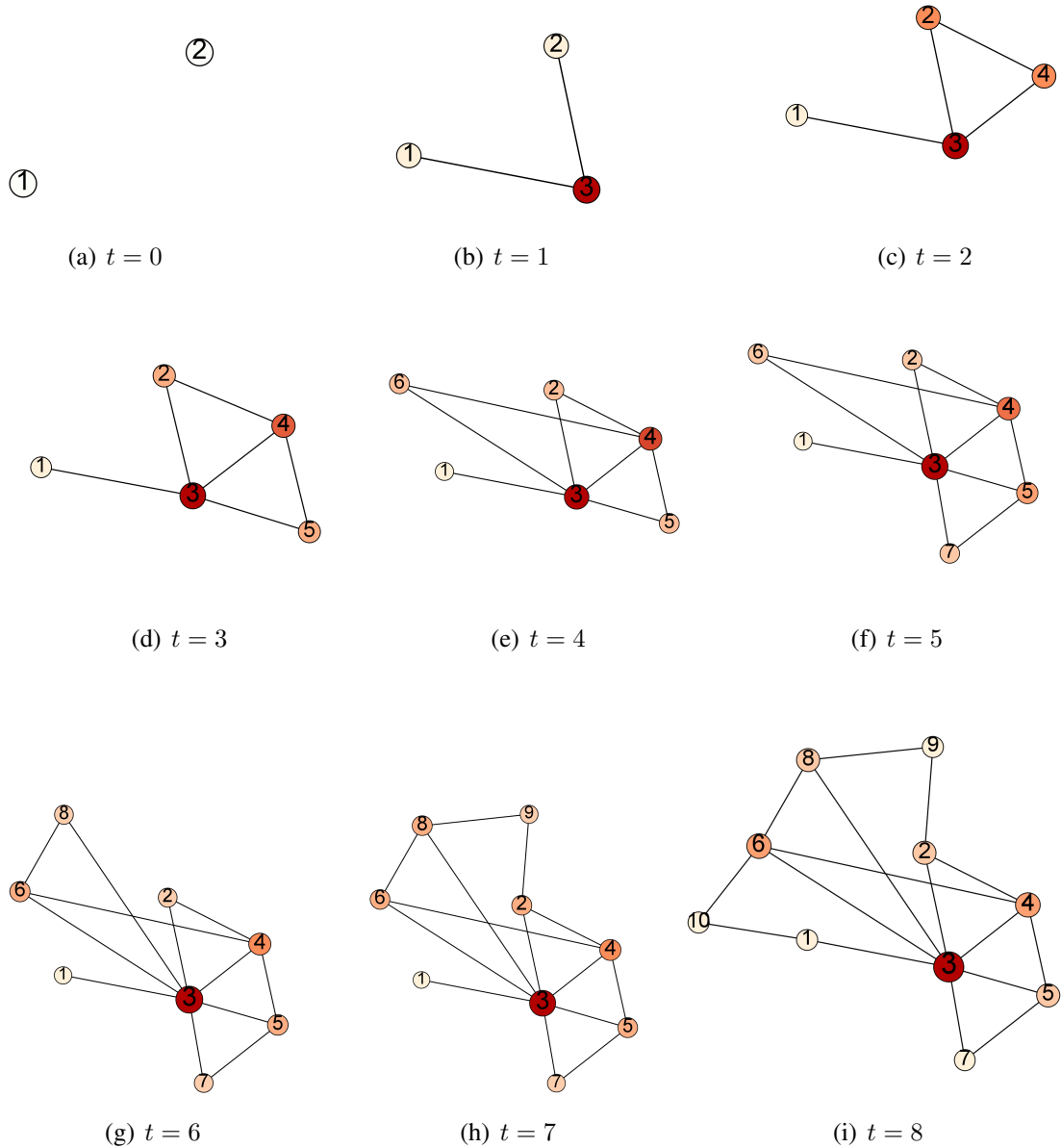
Adım 1 Başlangıçta çok küçük bir m_0 köşe ve sıfır kenar sayısı ile başla ve her zaman adımında $m \leq m_0$ kenar sayısına sahip olduğu için m farklı köşe ile bağlantı kurabilecek yeni bir köşe ekle (Bu ağın *büyüme* unsurudur.)

Adım 2 Daha sonra bu yeni köşeyi

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (1.45)$$

ihtimaliyeti ile k_i derecesine bağı olarak ağdaki bir i köşesine bağla (Bu ise ağın *tercihli eklenme* unsurudur.)

olarak verilmektedir [42]. t zaman adımı sonunda ağdaki toplam köşe sayısı $N = t + m_0$ ve toplam kenar sayısı ise mt olacaktır. Böylelikle ağ zamanla ölçeksiz olarak büyümektedir [42, 101].



Şekil 1.26. $N = 10$, $m_0 = 2$ ve $m = 2$ değerleri için ölçeksiz ağın $t = 8$ zaman adımında evrilmesi.

Şekil 1.26'de ölçekten bağımsız ağın $t = 8$ zaman adımındaki evrilmesinin şematik bir

örneği verilmiştir. Her bir zaman adımında eklenen yeni köşe, ağda varolan bağlanacağı köşenin derecesine bağlı olarak Denklem (1.45)'deki ihtimaliyete göre $m = 2$ kenar ile eklenmiştir. Şekil 1.26(a)'da $t = 0$ anında ağda hiçbir bağlantıları olmayan $m_0 = 2$ köşe ile başlanmıştır. Şekil 1.26(b)'de $t = 1$ anında $m = 2$ kenarı olan 3 numaralı köşe mecburen 1 ve 2 numaralı köşelere bağlanmıştır. Bu anda ağdaki toplam köşe sayısı $N = t + m = 3$ ve toplam kenar sayısı $m \times t = 2 \times 1 = 2$ dir. Her bir zaman adımında yeni köşeler bu şekilde eklenmeye devam edilerek Şekil 1.26(i)'deki son zaman adımı $t = 8$ anında 10 numaralı köşe 1 ve 6 numaralı köşelere bağlanmıştır. Bu son zaman adımında ağdaki toplam köşe sayısı $N = 10$ ve toplam kenar sayısı 16 dır. Ağdaki köşeler sahip oldukları kenar sayılarına göre farklı renklerde gösterilmiştir. Kırmızı renkli olan 3 numaralı köşe, $k_3 = 7$ kenar sayısı ile ağdaki en fazla kenar sayısına sahip olan köşedir. 9 ve 10 numaralı köşeler dışında ağdaki bütün köşeler 3 numaralı köşe ile bağlantılıdır. Bu nedenle 3 numaralı köşe ağın merkezi konumundadır. Diğer taraftan 1, 7, 9 ve 10 numaralı köşeler ise 2'şer kenar sayısı ile ağdaki en düşük kenar sayısına sahip köşelerdir.

Barabási et al. [127], birçok gerçek dünya ağlarının derece dağılımlarının, rastgele ve küçük dünya ağlarında olduğu gibi bir Poisson dağılımı değil de γ kritik üsteline sahip bir kuvvet serisi şeklinde olduğunu belirtmiş ve

$$P(k) = 2m^{\frac{1}{\beta}} k^{-\gamma} = 2m^2 k^{-3} \quad (1.46)$$

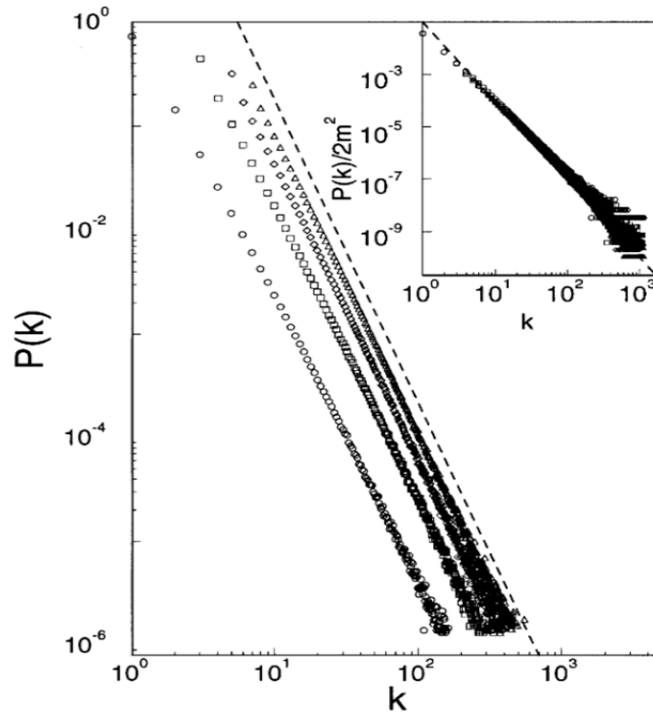
olarak tanımlamışlardır. Görüldüğü üzere γ üsteli, yeni eklenen köşelerin sahip oldukları kenar sayısı m 'den bağımsız olarak $\gamma = 3$ değerine sahiptir.

Şekil 1.27'de $N = t + m_0 = 300000$ köşeden oluşan bir ağ için farklı m_0 ve m değerlerindeki derece dağılımları verilmiştir. Dağılım, ölçeksiz olarak büyüyen ağın dağılımıdır. Dağılımın kritik üsteli, kesikli çizgi ile gösterilen doğrunun eğimi olan $\gamma = 2.9 \pm 1$ 'dir. Şeklin içindeki küçük şekil ise derece dağılımı $P(k)$ 'nın, aynı m değerlerinde $2m^2$ 'ye bölünerek yeniden boyutlandırılmış halidir. Bu grafikte kesikli çizginin eğimi olan kritik üstel ise $\gamma = 3$ dür.

Ölçekten bağımsız ağların ortalama yol uzunluğu ifadesi için analitik bir ifade Bollobás ve Riordan [128] tarafından

$$L \sim \frac{\ln N}{\ln \ln N} \quad (1.47)$$

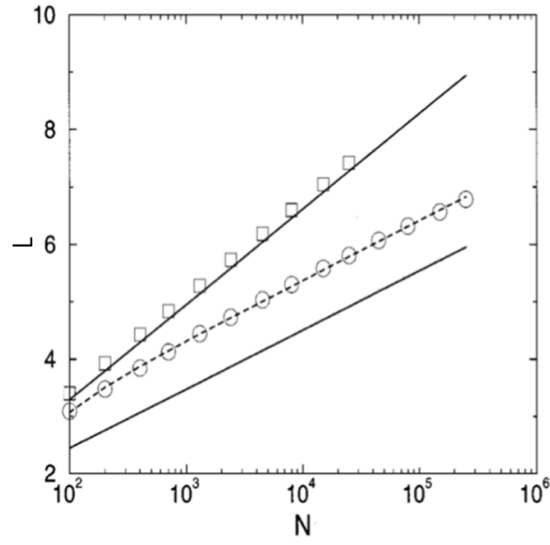
şeklinde elde edilmiştir.



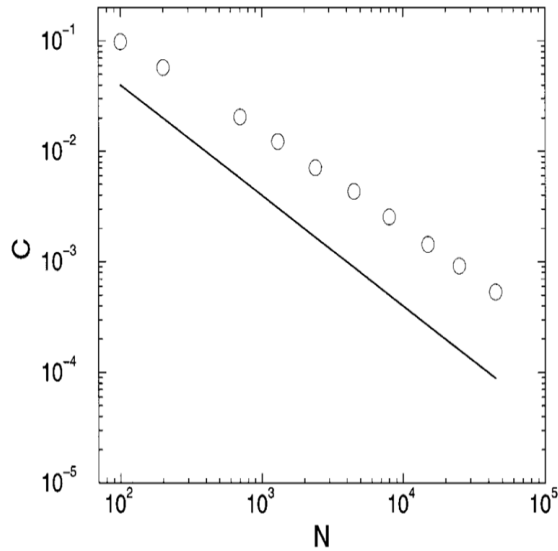
Şekil 1.27. Barabási et al. tarafından $N = t + m_0 = 300000$ tane köşe için elde edilen derece dağılımı. *Daire* için $m_0 = m = 1$, *Kare* için $m_0 = m = 3$, *Elmas* için $m_0 = m = 5$ ve *Üçgen* için $m_0 = m = 7$ değerleri alınmıştır. Ref.[127]

Şekil 1.28’de aynı N ve $\langle k \rangle$ değerleri için Barabási - Albert modeli ve rastgele ağ modeli ile elde edilen ortalama yol uzunluğu ifadesinin karşılaştırılması yapılmıştır. Herhangi bir N değerinde Barabási - Albert modeli ile elde edilen ortalama yol uzunluğu ifadesinin rastgele ağ modeli ile elde edilenden daha küçük olduğu görülmektedir. İki model arasındaki bu fark, derece dağılımının Barabási - Albert modeli için kuvvet serisi rastgele ağ modeli için Poisson dağılımına sahip olmasından kaynaklanmaktadır [101]. Bu durumda, ölçeksiz ağdaki merkezi yoğun bağlantılı köşeler nedeniyle diğer köşeler arasındaki yolların, bu köşeler üzerinden kısalacağı anlaşılabılır.

Şekil 1.29’de aynı N ve $\langle k \rangle$ değerleri için Barabási - Albert modeli ve rastgele ağ modelindeki kümelenme katsayısı ifadesinin karşılaştırılması yapılmıştır. Barabási - Albert modelinde kümelenme katsayısı için rastgele ağ modelinde olduğu gibi analitik bir öngörü bulunmamaktadır. Rastgele ağ modelinde $C_r = \langle k \rangle N^{-1}$ iken, Barabási - Albert modelinde $C \sim N^{-0.75}$ dir.



Şekil 1.28. Albert ve Barabási tarafından $\langle k \rangle = 4$ için elde edilen ortalama yol uzunluğu ifadesinin ağdaki köşe sayısı N 'ye göre değişimi. *Daire* Barabási - Albert modeli ve *Kare* ise rastgele ağ modeli ile elde edilen değerlerdir. Kesikli çizgi Denklem (1.47)'ye göre, sürekli çizgiler ise Denklem (1.37)'ye göre elde edilmiştir. Ref.[101]



Şekil 1.29. Albert ve Barabási tarafından $\langle k \rangle = 4$ için elde edilen kümelenme katsayısı ifadesinin ağdaki köşe sayısı N 'ye göre değişimi. *Daire* Barabási - Albert modelindeki ve *Düz çizgi* ise rastgele ağ modelindeki değerlerdir. Ref.[101]

2. BÖLÜM

SOY AĞACI

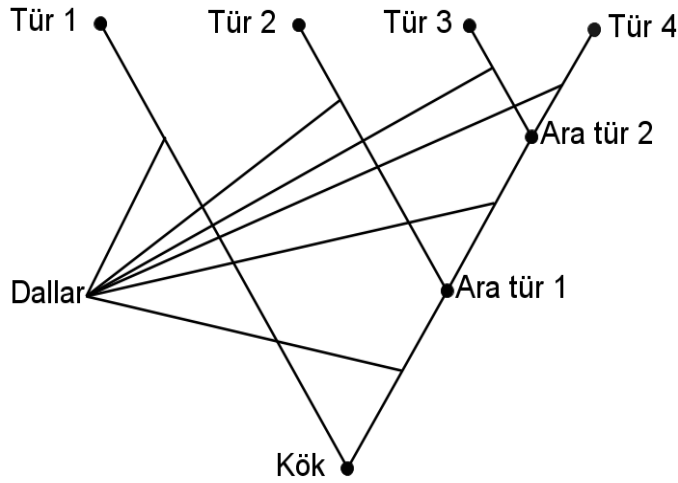
Biyolojinin temel prensiplerinden bir tanesi kalıttır. Kalıtımın keşfedilmesi, evrimsel sistemlerde büyük değişikliklerin yapılmasına neden olmuştur [129–132]. Evrimsel sistemlerin temeli, canlı organizma grupları arasındaki ilişkiler neticesinde ortaya çıkan hiyerarşik bir düzene dayanır ve evrimsel süreci yorumlamak için oldukça önemlidirler. Bu nedenle evrimsel sistemler, canlı organizmalar için hem tür hemde tür üzerindeki seviyeler için çalışılmıştır [133]. Tür seviyesinde çalışmalar yapan Mayr [134] ve Simpson [131], türlerin sınıflandırılmasına önemli katkılar yapmışlardır. Mayr ve Simpson'ın çalışmalarının geniş çapta kabul gören sonuçlarından bir tanesi, türlerin hiyerarşik düzendeki diğer kategorik birimlerden (cins, aile, takım, sınıf, şube ve alem) temel olarak farklı olduğudur. Diğer bir sonuç ise, türlerin evrimin bir sonucu olmadığı aslında evrimsel süreçteki esas fonksiyonun türler olduğudur. Türler, Mayr'ın çalışmasında gen havuzu, Simpson'ın çalışmasında ise zaman boyunca genişleyen nesiller olarak kabul edilmiştir. Ortak ata varlığına inanan Hennig [132], tür seviyesinden daha yüksek sınıfların varolduğu sonucuna varmıştır. Hennig'e göre tür seviyesi üzerindeki evrim üreticileri, ata türlerin ve bu ata türlerin soylarının oluşturduğu gruplardır. Bu gruplar türlerin sınıflarını belirlemektedir. Hennig bu düşüncesiyle sınıflandırmadaki evrim anlayışını değiştirmiştir.

Bilim insanları evrimsel çerçevede biyolojik çeşitliliği anlamak için, türlerin evrimsel tarihinin anlaşılmasına yardımcı olan soy ağacı incelemeleriyle sınıflandırma yapmaktadırlar. Türlerin sınıflandırılmasını sağlayan soy ağacı, *anatomik* ve *moleküler* olmak üzere iki temel teknikle yapılmaktadır.

Soy ağacı çıkarımı için geleneksel bir yöntem olan anatomik teknik, türlerin morfolojik

benzerliklerine yani sahip oldukları fiziksel karakteristiklere göre yapılmaktadır. Türler üzerinde karakter analizi yapılarak, sahip oldukları karakterler belirlenmektedir. Karakterlerin evrimsel olarak ortaya çıkış zamanlarına göre türlerin soy ağacı oluşturulmaktadır. Aynı karaktere sahip olan türlerin, soy ağacında ortak bir ataya sahip olduğu belirtilmektedir.

Moleküler teknik ise, bütün canlılarda bulunan protein, DNA ve RNA gibi genetik diziler kullanılarak yapılmaktadır. Bu dizilerde, baz değişimi olarak adlandırılan mutasyonlar neticesinde yeni türlerin meydana geldiği varsayılmaktadır. Türlerin genetik olarak birbirlerine benzerliklerini belirleyebilmek için, sahip oldukları aynı diziler (protein dizileri, DNA dizileri veya RNA dizileri) karşılaştırılmaktadır. Karşılaştırma sonucu, en düşük genetik farklılığa yani en az sayıda farklı bazlara sahip türlerin, soy ağacında ortak bir ataya sahip olduğu belirtilmektedir.

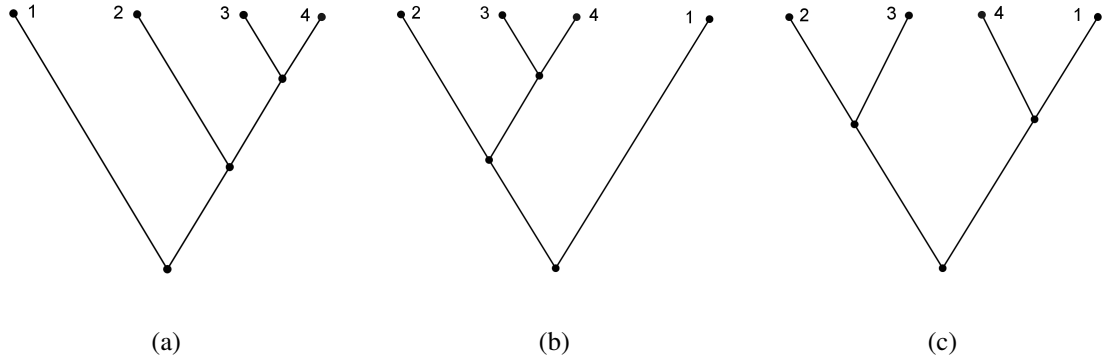


Şekil 2.1. Soy ağacının temel gösterimi.

Soy ağaçları, günümüzde varolan türlerin ve bu türlerin atalarının bir arada gösterildiği bir dallanma şemasıdır. Şekil 2.1'de soy ağacı terminolojisini anlatabilmek için dört türden oluşan örnek bir soy ağacı verilmiştir. Bugün gözlemlenebilen ve DNA(DeoksiriboNükleik Asit) dizileri bilinen türler *ağacın uç noktalarını* oluşturmaktadır. Bu türler Tür 1, Tür 2, Tür 3 ve Tür 4 olarak gösterilmiştir. *Ağacın dallanma noktalarını* oluşturan Ara tür 1 ve Ara tür 2, evrimsel süreç içerisinde ortaya çıkmış ve günümüzde varlıkları söz konusu olmayan ata türlerdir. Türlerin, hangi atadan ürediği *ağacın dalları* ile gösterilmekte ve bu dalların uzunlukları ise nesilsel zaman

olarak türlerin ne kadar yaşadıklarını belirtmektedir. Son olarak, gözlemlenen bütün türlerin ortak atası olan tür ise, *ağacın kökü* olarak tanımlanmaktadır.

Uç noktalar, dallar, dallanma noktaları ve kökten oluşan ağacın şekline *topoloji* denir. Evrimsel biyolojide, topolojilerdeki dallar sökülmeden bir topolojiden diğer topolojiye ulaşılamıyorsa, iki topolojinin birbirinden farklı olduğu söylenir.



Şekil 2.2. Ağaç topolojilerinin gösterimi.

Şekil 2.2’de eşit sayıda dallanma ve uç noktalara sahip üç ağaç topolojisi verilmiştir. Şekil 2.2(a) ve Şekil 2.2(b)’deki her iki topolojide de dallar sökülmeden birinci topolojiden ikinci topolojiye ya da tam tersi ikinci topolojiden birinci topolojiye ulaşılabildiği için aynı topolojiye sahiptirler. Ancak Şekil 2.2(c)’deki topolojide dallanma noktaları Şekil 2.2(a) ve Şekil 2.2(b)’deki dallanma noktalarından farklıdır. Bu topolojiye ulaşmak için hem Şekil 2.2(a)’da hemde Şekil 2.2(b)’de bazı dalların sökölüp farklı noktalara bağlanması gerekmektedir. Bu yüzden Şekil 2.2(c), Şekil 2.2(a) ve Şekil 2.2(b)’den farklı bir topolojiye sahiptir.

Son yıllarda gelişen teknoloji ile birlikte bilgisayar kapasitelerinde çok ciddi artışlar olmuştur. Bu durum gerçek soy ağaçlarına ulaşmak için gerekli olan DNA dizilerinin daha uzun ve daha fazla sayıda çalışılmasına olanak sağlamıştır. Bu avantajlar, en iyi soy ağacını elde etmek için birçok istatistik modelin de geliştirilmesini beraberinde getirmiştir. Bu modeller kullandıkları stratejilere göre iki ana gruba ayrılır [135]. İlk grup, gerçekleşmesi olası olan bütün soy ağaçlarını inceler. Daha sonra bu ağaçlar içerisinde en doğru ağacı seçmek için maksimum evrimleşme ihtimaliyeti gibi bazı kriterler kullanır. Maximum parsimony (MP) [136] ve maximum likelihood(ML) [137] metodları bu gruba aittir. İkinci grup ise, DNA dizileri arasındaki mesafeyi analiz ederek

en iyi ağacı kurmaya çalışır. Neighbor-joining (NJ) [138] metodu bu grubun en iyi bilinen metodudur. Bu istatistik modeller yardımı ile verilen bir DNA dizileri seti için bilinmeyen bir soy ağacının, gerçek ağaca ne kadar benzediği öngörülebilir.

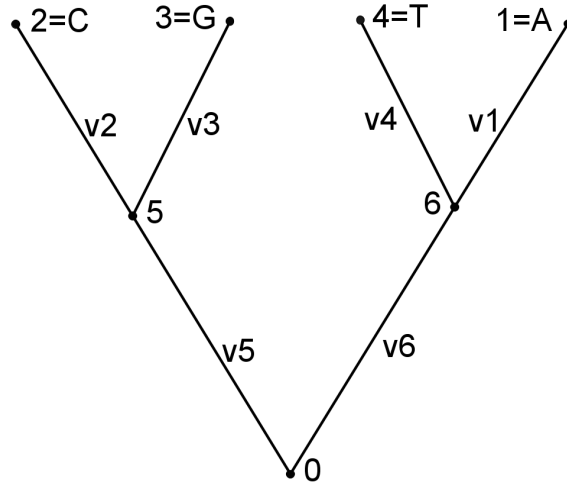
Bu tez çalışmasında her iki gruptan birer metot alınarak yani ML ve NJ metotları seçilerek analizler yapılacaktır.

2.1. En Fazla Olabilirlik (ML) Metodu

Felsenstein [137] tarafından geliştirilen ML (Maximum Likelihood) metodu, DNA dizilerindeki bazlarda görülen değişme ya da evrimleşme ihtimaliyetine göre en iyi soy ağacını bulmayı amaçlar. Ağacın topolojisi ve dal uzunlukları, metottaki en önemli faktörlerdir. Türlerin sayısı yeteri kadar çok olduğunda, olası topoloji sayısı da çok yüksek değerlere ulaşacaktır. Bu durum olası bütün ağaç topolojisini ve her bir topoloji için tanımlanması gereken en uygun dal uzunluklarını aramayı oldukça zorlaştıracaktır. Bu zorluğu ortadan kaldırmak için Felsenstein [139] bir yöntem geliştirmiştir. Bu yöntemde soy ağacı oluşturmak için başlangıçta sadece iki tür ile başlanmış ve daha sonra diğer türler ağaca başarılı bir şekilde eklenmiştir. Böylelikle olası topoloji sayısı sistematik olarak düşürülmüştür. k . türün ağaca eklenmesi ile en yüksek olabilirlik değerine sahip ağaç topolojisini bulmak için $2k - 5$ tane topolojinin incelenmesi gerekmektedir. Üçüncü ve daha sonraki türlerin eklenmesi ile türlerin ağaçtaki yerlerinin belirlenmesi gerektiğinden bu yöntemin de yine zorlukları bulunmaktadır.

Türlerin sahip oldukları DNA dizilerinin herhangi bir konumundaki i bazının t nesilsel zaman sonunda j bazına mutasyon verme ihtimaliyeti $P_{ij}(t)$ stokastik olmakla birlikte, başka bir konumdaki bazın mutasyon verme ihtimaliyetinden bağımsızdır. Burada i ve j sırasıyla A, C, G ve T bazlarına karşılık gelen 1, 2, 3 ve 4 değerlerini alabilmektedirler. Bu yüzden metot, DNA dizisindeki konumların mutasyona uğrama ihtimalini tek tek incelemektedir. $t = 0$ başlangıç durumunda iken DNA dizisindeki bazların frekansı π_i ile verilmektedir. π_i değeri sabit olmakla birlikte ağaçtan bağımsızdır.

Şekil 2.3’de ML metodunun anlatımı için kullanılacak 4 türden oluşan örnek bir ağaç verilmiştir. Metodun çalışma prensibinden dolayı ağaç, türlerin sahip oldukları DNA dizilerinin tek bir konumu için oluşturulmuştur. Bu yüzden dalların uç noktaları, türlerin



Şekil 2.3. ML metodunun anlatımı sırasında kullanılacak ağaç örneği. A, C, G ve T bazları 1, 2, 3 ve 4 değerleri ile gösterilmiştir. 0, 5 ve 6 değerleri ise ağacın dallanma noktalarıdır. v'ler ise dal uzunlukları olup nesil zamana karşılık gelmektedir.

DNA dizilerinin birer konumunu göstermektedir.

Eğer özel olarak 0, 5 ve 6 numaralı dallanma noktalarındaki bazlar biliniyorsa, böyle bir ağacın olabilirliği

$$L = \pi_{S_0} P_{S_0 S_5}(v_5) P_{S_5 S_2}(v_2) P_{S_5 S_3}(v_3) P_{S_0 S_6}(v_6) P_{S_6 S_1}(v_1) P_{S_6 S_4}(v_4) \quad (2.1)$$

ifadesi ile elde edilir. Burada π_{S_0} , S_0 bazının frekansı, S' ler ilgili konumdaki baz değerleri ve P 'ler de her bir daldaki bazların mutasyon verme ihtimaliyetidir.

Ancak uygulamada, S_0 , S_5 ve S_6 bazlarının değerleri bilinmediği için, bu bazlar 4 bazdan biri olacaktır. Bu durumda Şekil 2.3'deki gibi bir ağacın olabilirliği

$$L = \sum_{S_0=1}^4 \sum_{S_5=1}^4 \sum_{S_6=1}^4 \pi_{S_0} P_{S_0 S_5}(v_5) P_{S_5 S_2}(v_2) P_{S_5 S_3}(v_3) P_{S_0 S_6}(v_6) P_{S_6 S_1}(v_1) P_{S_6 S_4}(v_4) \quad (2.2)$$

şeklinde ifade edilecektir. Denklem (2.2) 64 terimden oluşan oldukça büyük bir ifadedir.

Tür sayısı artıkça bu ifadedeki terim sayısı da artacaktır.

Bu süreç Felsenstein tarafından [137] *şartlı olabilirlik* denen yöntemle yeniden ifade

edilmiştir. Bu yöntem, $L_S^{(k)}$ ile gösterilen şartlı olabilirlik ifadesi ile ağaçta S bazına sahip k noktalarının olabilirliği üzerinde durmaktadır. Eğer k noktası ağacın dallarının uç noktalarından birisi ise, bu noktanın sahip olduğu gerçek S_k bazının bu noktada olabilirliği $L_{S_k}^{(k)} = 1$ ve diğer bazların bu noktada olabilirliği $L_S^{(k)} = 0$ olacaktır. Eğer i ve j , k noktasının dallarının uç noktaları ise, 4 farklı S_k bazından birinin k noktasındaki olabilirliği için genel ifade

$$L_{S_k}^{(k)} = \sum_{S_k=1}^4 \left[\left(\sum_{S_i=1}^4 P_{S_k S_i}(vi) L_{S_i}^{(i)} \right) \left(\sum_{S_j=1}^4 P_{S_k S_j}(vj) L_{S_j}^{(j)} \right) \right] \quad (2.3)$$

şeklinde verilir. Her k noktası, S_k baz değerlerinin bulunması ile kendi atalarının dallarının uç noktaları olarak düşünülür. Bu süreç bir budama işlemi gibi ağacın köküne ulaşana kadar devam eder.

Örneğin Şekil 2.3'deki ağaçta 5 ve 6 konumlarının S_5 ve S_6 baz değerlerinin şartlı olabilirlikleri

$$L_{S_5}^{(5)} = \sum_{S_5=1}^4 (P_{S_5 S_2}(v2) L_{S_2}^{(2)}) (P_{S_5 S_3}(v3) L_{S_3}^{(3)}) \quad (2.4)$$

ve

$$L_{S_6}^{(6)} = \sum_{S_6=1}^4 (P_{S_6 S_1}(v1) L_{S_1}^{(1)}) (P_{S_6 S_4}(v4) L_{S_4}^{(4)}) \quad (2.5)$$

şeklinde dir. Burada $L_{S_1}^{(1)} = L_{S_2}^{(2)} = L_{S_3}^{(3)} = L_{S_4}^{(4)} = 1$ dir. S_5 ve S_6 baz değerleri belirlendikten sonra, π_{S_0} frekansına sahip ağacın kökü olan 0 konumunun S_0 baz değerinin şartlı olabilirliği ise

$$L_{S_0}^{(0)} = \sum_{S_0=1}^4 \pi_{S_0} (P_{S_0 S_5}(v5) L_{S_5}^{(5)}) (P_{S_0 S_6}(v6) L_{S_6}^{(6)}) \quad (2.6)$$

şeklinde elde edilir. Yine $L_{S_5}^{(5)} = L_{S_6}^{(6)} = 1$ dir. Ağacın toplam olabilirliği ise Denklem (2.4), (2.5) ve (2.6)'nın çarpımı ile Denklem (2.2) olarak elde edilir.

DNA dizilerinin herhangi bir konumundaki i bazının t nesilsel zaman sonunda j bazına mutasyon verme ihtimaliyeti $P_{ij}(t)$, Markov süreci özelliği taşımaktadır. Bu süreçte i bazının j bazına mutasyon verme ihtimaliyeti, i bazının geçmişteki değerinden bağımsız ve $P_{ij}(t)$ ihtimaliyeti Poisson dağılımına sahiptir. ML metodunun temelini oluşturan $P_{ij}(t)$ ihtimaliyeti

$$P_{ij}(t) = e^{-ut} \delta_{ij} + (1 - e^{-ut}) \pi_j \quad (2.7)$$

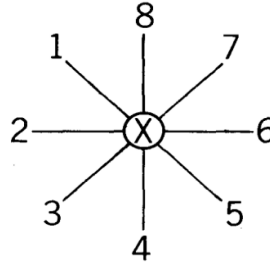
ifadesi ile verilmektedir. Burada u , birim zaman başına mutasyon oranıdır. δ_{ij} , Kronecker delta fonksiyonudur. t zaman sonunda $i = j$ durumunda $\delta_{ij} = 1$ değerini almakta ve e^{-ut} ihtimaliyeti ile mutasyon gerçekleşmemektedir. Mutasyonun gerçekleşmemesi, i bazının i bazı ile değiştirilmesi anlamına gelmektedir. $i \neq j$ durumunda ise $\delta_{ij} = 0$ olmakta ve $(1 - e^{-ut})$ ihtimaliyeti ile bir mutasyon gerçekleşmektedir. π_j , i bazının mutasyon verdiği j bazının frekansdır.

ML metodunda dal uzunlukları, olabilirliği en yüksek olan ağaç topolojisini bulmak için hayati öneme sahiptir. Dal uzunlukları doğrudan bütün ihtimaller denenerek bulunabilir. Ancak bu süreç çok uzun hesaplama süresi gerektirecektir. Metot bu sorunu çözmek ve en uygun dal uzunluklarını bulmak için makara prensibini kullanır. Bu çerçevede metodun kurduğu algoritma, her bir zaman adımında v_i dal uzunluklarını değiştirerek, olabilirliği en yüksek olan ağaç topolojisini bulacak şekilde işlemektedir. Bunun sonucunda her bir zaman adımında ağacın olabilirliği artacaktır. Olabilirlik değeri hiç bir zaman düşmeyeceği için bu süreç sonsuz bir döngü içerisine girmeyecektir. Böylelikle olabilirliği en yüksek ağaç topolojisinin dal uzunluklarının da elde edilmesiyle, en uygun soy ağacı tahmini tamamlanacaktır. ML metodu, en yüksek olabilirlik değerine sahip ağacı aradığı için uzun hesaplama süreleri gerektirmektedir.

2.2. Komşu Kurma (NJ) Metodu

Saitou ve Nei [138] tarafından geliştirilen NJ (Neighbor-Joining) metodunun amacı, türlerin DNA dizileri arasındaki mesafeyi kullanarak evrimsel olarak en yakın türleri bulmak ve en iyi ağacı kurmaktır. Türler arasındaki mesafe, DNA dizilerindeki farklı baz sayıları ile doğru orantılıdır. Başka bir ifade ile, iki türün sahip oldukları DNA dizilerindeki farklı bazların sayısı ne kadar az ise bu iki tür birbirine o kadar yakındır. Bu özelliğinden dolayı metot sadece en uygun ağaç topolojisini üretmekle kalmaz aynı zamanda ağacın dal uzunluklarını da belirler.

NJ metodu soy ağacını kurmaya, Şekil 2.4’de gösterildiği gibi yıldız benzeri, hiyerarşik bir düzenin olmadığı köksüz bir ağaçla başlar. 1 – 8 ile gösterilenler ağacın uç noktaları olan türlerdir. Bütün türler tek bir X noktasına bağlanmıştır. Bu durumda iken, türlerin DNA dizileri arasındaki uzaklıklarla $N \times N$ boyutlu bir D uzaklık matrisi kurar.

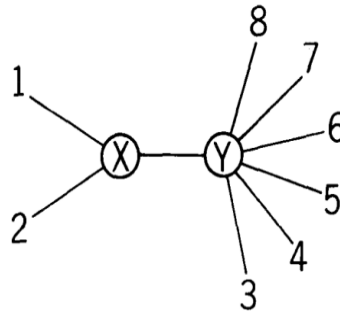


Şekil 2.4. Saitou ve Nei tarafından verilen 8 türden oluşan yıldız benzeri bir ağaç. Ref.[138]

Daha sonra, türlerin doğru komşuluklarını belirleyebilmek için D matrisini yeniden düzenleyerek elemanlarını

$$S_{ij} = \frac{1}{2(N-2)} \sum_{k \neq i,j}^N (D_{ik} + D_{jk}) + \frac{1}{2} D_{ij} + \frac{1}{N-2} \sum_{k \neq i,j}^N \sum_{\substack{l \neq i,j \\ l > k}}^N D_{kl} \quad (2.8)$$

ifadesi ile belirlediği ve bütün i ve j türleri arasındaki dal uzunluklarının toplamı olan yeni bir S uzaklık matrisi kurar. Burada N toplam tür sayısı ve D_{**} terimleri D uzaklık matrisinin elemanlarıdır. Belirlenen S matrisinin en düşük değerine karşılık gelen iki türün birbirleri ile komşu oldukları anlaşılır. Bu türler dışındaki diğer bütün türleri, X dallanma noktasından ayırarak yeni bir Y dallanma noktasına bağlar. Bu yöntem ile Şekil 2.4'de verilen ağaçta komşu iki tür olarak 1 ve 2 numaralı türler seçilmiş ve Şekil 2.5'de gösterildiği gibi diğer türlerden ayrılarak bir küme oluşturmaktadır.



Şekil 2.5. En yakın iki türün ayrılarak bir küme oluşturulması. Ref.[138]

Komşu olan 1 ve 2 numaralı türlerin X dallanma noktasına olan uzaklıkları ise

$$L_{1X} = \frac{D_{12} + \left(\sum_{i=3}^N D_{1i} \right) / (N-2) - \left(\sum_{i=3}^N D_{2i} \right) / (N-2)}{2} \quad (2.9)$$

ve

$$L_{2X} = \frac{D_{12} + \left(\sum_{i=3}^N D_{2i} \right) / (N-2) - \left(\sum_{i=3}^N D_{1i} \right) / (N-2)}{2} \quad (2.10)$$

ifadeleri ile elde edilmektedir. Bu aşamadan sonra 1 ve 2 numaralı türler tek bir tür gibi düşünülerek D uzaklık matrisinin boyutu $(N-1) \times (N-1)$ mertebesine düşürülür. 1 ve 2 numaralı türlerin atası konumunda olan X dallanma noktası ile diğer türler arasındaki mesafe

$$D_{Xj} = D_{(1-2)j} = \frac{D_{1j} + D_{2j} - D_{12}}{2} \quad \{j = 3, \dots, N\} \quad (2.11)$$

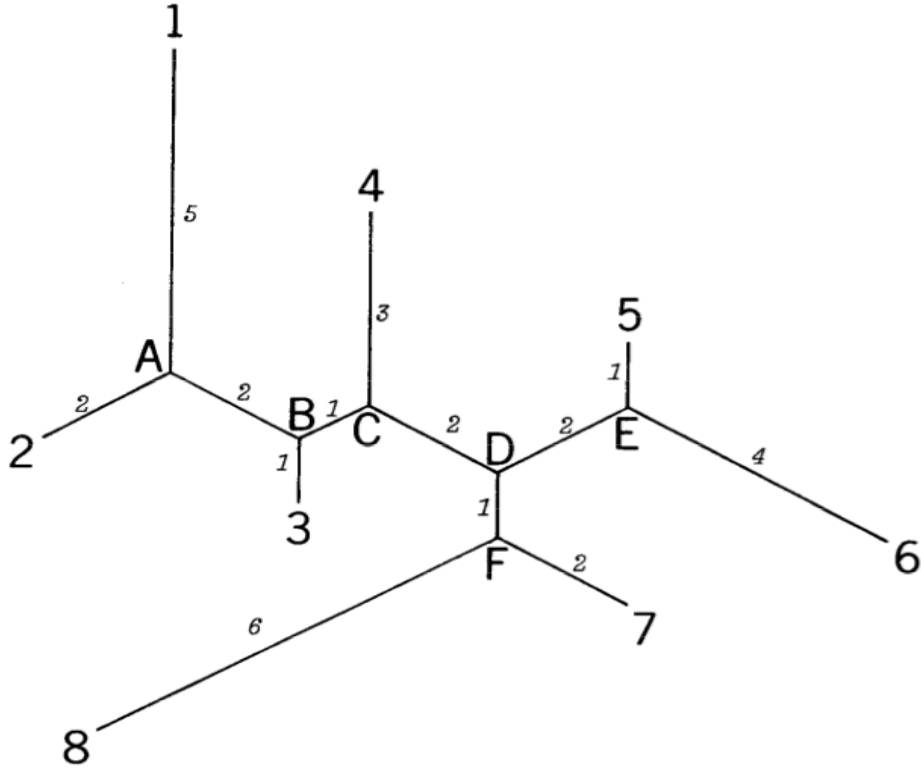
ifadesi ile belirlendikten sonra D uzaklık matrisi güncellenir. Bu sürecin $N-3$ defa tekrarlanması sonucunda soy ağacının son haline ulaşılır.

Saitou ve Nei'nin Şekil 2.4'deki 8 tür için belirledikleri D uzaklık matrisi [138]

$$D = \begin{pmatrix} 0 & 7 & 8 & 11 & 13 & 16 & 13 & 17 \\ 7 & 0 & 5 & 8 & 10 & 13 & 10 & 14 \\ 8 & 5 & 0 & 5 & 7 & 10 & 7 & 11 \\ 11 & 8 & 5 & 0 & 8 & 11 & 8 & 12 \\ 13 & 10 & 7 & 8 & 0 & 5 & 6 & 10 \\ 16 & 13 & 10 & 11 & 5 & 0 & 9 & 13 \\ 13 & 10 & 7 & 8 & 6 & 9 & 0 & 8 \\ 17 & 14 & 11 & 12 & 10 & 13 & 8 & 0 \end{pmatrix} \quad (2.12)$$

ile yapılan hesaplamalar neticesinde 8 tür için köksüz soy ağacı Şekil 2.6'deki gibi elde edilmektedir.

Şekil 2.6'da 8 türden oluşan köksüz bir soy ağacı gösterimi verilmiştir. $A-F$ ile gösterilen noktalar dallanma noktaları olan ara türlerdir. Ağacın dalları üzerinde gösterilen sayılar ise dal uzunlukları olup türler arasındaki evrimsel mesafeyi belirtmektedir. *Komşu türler* (*neighbors*), aralarındaki dal uzunlukları en düşük olan türlerdir. Bu bağlamda Şekil 2.6'daki ağaç topolojisine bakıldığında komşu türlerin, 1 – 2, 5 – 6 ve 7 – 8 olduğu görülmektedir. Ayrıca 3 numaralı türün 1 – 2 komşuluğuna komşu olduğu ve 4 numaralı türün ise (1 – 2) – 3 komşuluğuna komşu olduğu çıkarımları da yapılabilmektedir. Bunun



Şekil 2.6. Saitou ve Nei tarafından verilen 8 türden oluşan köksüz bir soy ağacı. Ref.[138]

yanında komşu türlerin aynı dallanma noktalarına da sahip olmaları gerektiği Şekil 2.6'da açıkça görülmektedir. Ağaçta bulunabilecek komşu sayısı, ağacın topolojisine bağlıdır. N çift sayı ve $N \geq 4$ olmak şartı ile her zaman için en düşük komşu sayısı 2 iken en yüksek komşu sayısı $N/2$ dir. Eğer N tek sayı ise en yüksek komşu sayısı $(N - 1)/2$ dir. Soy ağacını bölgesel matematiksel ilişkiler üzerine kurmasından dolayı metot hızlıdır.

3. BÖLÜM

DOLAŞIK DOĞA MODELİ (DDM)

Kompleks sistemler, sistemin işleyişi hakkında kolay bilgi sahibi olunamadığı, sistemi oluşturan alt birimler arasındaki etkileşimlerin doğasının açıkça anlaşılamadığı ve sistemin sahip olduğu özelliklerin zamanla nasıl değiştiğinin bilinmediği sistemlerdir. Bu sebeplerden dolayı, kompleks sistemlerin işleyişleri durdurulmadan incelenememektedir. Bu şekildeki sistemleri inceleyebilmek için, sistemin simülasyonu (benzetimi) yapılır. Simülasyon, gerçek sistemin bir kopyasını oluşturarak sistem hakkında öğrenilmek istenen bilgilerin bu kopya sistem üzerinde araştırılmasına olanak tanıyan bir bilgisayar tekniğidir. Yeni geliştirilecek bir sistemin performansını tahmin etmek için bir simülasyon yapılabileceği gibi, varolan sistem üzerinde yapılması düşünülen potansiyel değişikliklerin gerçekleştirilmeden önce nasıl bir etki oluşturacaklarını tahmin etmek için de yapılabilir. Daha açık bir ifade ile gerçek bir sistemin simülasyonu yapılarak,

- Sistemin alt birimlerinin kompleks etkileşimleri incelenebilir.
- Sistem üzerinde yapılacak değişiklikler neticesinde, sistemin nasıl bir davranış sergileyeceği gözlemlenebilir.
- Gerçek sistem durdurulmadan sistem hakkında bilgi sahibi olunabilir.
- Sistemin çalışmasını etkileyen değişkenlerin önemi hakkında çıkarımlar yapılabilir.

Kompleks sistemlerin incelenmesi için yapılacak simülasyon, sistemin kopyasını oluşturacak bir matematiksel modele gereksinim duymaktadır. Kurulacak model genellikle, sistemdeki olaylar hakkındaki varsayımlarla şekillenir ve sistemin alt birimleri arasındaki mantıksal ve matematiksel ilişkilerin incelenebilmesi için sistemin bilgisayar

ortamına aktarılmasını sağlar. Bir modelin kurulması sırasında sistemin, yapılan çalışmanın amacına uygun yönleri dikkate alınarak gereksiz detaylandırmadan kaçınılır.

Bu giriş bilgilerinden sonra bu bölümde, kompleks, dinamik ve stokastik bir sistem olan ekosistemin evrimsel sürecinin işleyişinin incelenmesi için kullanılan DDM üzerinde durulacaktır.

3.1. Genel Özellikleri

DDM, Christensen et al. [36] tarafından geliştirilen birey tabanlı eş-evrimleşme modelidir. Modelin tam olarak odaklandığı nokta, adından da anlaşılacağı üzere, ekosistemde bulunan organizmalar arasındaki etkileşimleri inceleyerek doğanın sahip olduğu karmaşık yapıyı analiz etmektir. Çünkü kompleks bir sistemin genel davranışı hakkında bilgi sahibi olabilmek için, sistemi oluşturan elemanların davranışları incelenmelidir. Modelde türlerin içerisinde yaşadığı fiziksel çevre ihmal edilmiş ve bütün ilgi organizmalar arasındaki etkileşimlere toplanmıştır. Bu düşüncenin birkaç nedeni vardır. Bunlardan biri, ekosistemin temel özelliklerinden biri olan herhangi bir coğrafik sınırlandırma ya da organizmaların izole edilme durumu söz konusu olmaksızın gerçekleşecek olan türleşmeye önem verilmesidir. Böylelikle evrimin önemli mekanizmalarından biri olan türleşme tamamen türlerin kendi aralarındaki etkileşimler neticesinde gerçekleşecektir. Bu olay, evrimin fiziksel çevreden bağımsız olarak sahip olduğu iç dinamiklerdendir. Türleşmenin dışında evrimin sahip olduğu iç dinamiklerden biri olan neslin tükenmesi olayı da, yine sadece fiziksel çevrede meydana gelen değişiklikler sebebi ile gerçekleşmez. Organizmalar arasındaki etkileşimler de neslin tükenmesine neden olabilir. Bu sebepten dolayı fiziksel çevrenin ihmal edilmesindeki diğer neden bu etkileşimlerin üzerinde durulmasıdır. Organizmalar için bu etkileşimler, hayatta kalmak ve neslin devamlılığını sağlamak için hayati bir öneme sahiptir. Fiziksel çevrenin koşulları her nasıl olursa olsun, organizmalar yaşam alanlarını paylaşırlar ve bu yaşam alanlarını sağlamlaştırma eğilimindedirler. Bu eğilimin sebebi, yaşam alanına adapte olmaları ve neslin devamlılığı için uygun koşulları oluşturmaları gereğindendir. Bunun için organizmalar birbirleri ile rekabet de edebilirler, farkında olmadan dolaylı yollarla birbirlerini de destekleyebilirler. DDM, bu sebeplerden dolayı organizmaların kendi aralarındaki etkileşimleri üzerine odaklanarak, eş-evrimleşmenin sabit bir çevrede

ekolojinin varolma kurallarını nasıl etkilediğini açıklamak için kurulmuştur.

DDM'nin organizmaları birey seviyesinde incelemesi, doğada meydana gelen doğal seçilim olayını açıklamak için önemli bir özelliktir. Çünkü doğal seçilim olayı, bir popülasyondaki gen sıklığına bağlıdır. Gen sıklığını değiştiren ise, hayatta kalma ihtimaliyetini arttıran ya da azaltan alellerdir. Bu nedenle doğal seçilimi tanımlamak için en avantajlı seviye birey seviyesidir. Türlerle ait bireylerin sahip oldukları genotiplerden bazılarında gerçekleşecek beklenmedik değişiklikler, dinamik özelliklerin ortaya çıkmasına neden olacaktır. Bu düşünce tarzı ile tür üzerindeki etkiler, birey seviyesindeki değişiklikler ile incelenebilir ve mikroevrim ile makroevrim arasında bir bağlantı kurulabilir. Türler arasındaki çeşitliliğin temel nedenleri genotipik ve fenotipik etkenlerdir. Bireylerin fenotiplerinde yani fiziksel görünüşlerindeki farklılıkların en temel nedeni ise, sahip oldukları genetik karakteristiklerdir. Bu sebeple DDM, basitlik için genotip ile fenotip arasında bir ayrım yapmamaktadır. Başka bir ifadeyle, bir türün bütün bireylerini aynı fenotipe sahip olarak kabul etmektedir. Böylelikle bütün bireyler genotip uzayında düşünülmektedir. Bireylerin sadece genotipik özellikleri incelendiği için, genotip uzayının çeşitliliği yani tür sayısı daha fazla olacaktır. Buna ilaveten DDM, üreme, mutasyon, seçilim ve çevrenin taşıma kapasitesi gibi evrim unsurları içerir. Anlaşılacağı üzere fiziksel çevreyi sabit kabul eden DDM, birebir evrimi açıklayan bir model değil evrimin basit bir durumunu inceleyen ve eş-evrimleşme mekanizmasının varlığını araştıran bir modeldir.

3.2. Matematiksel İşleyişi

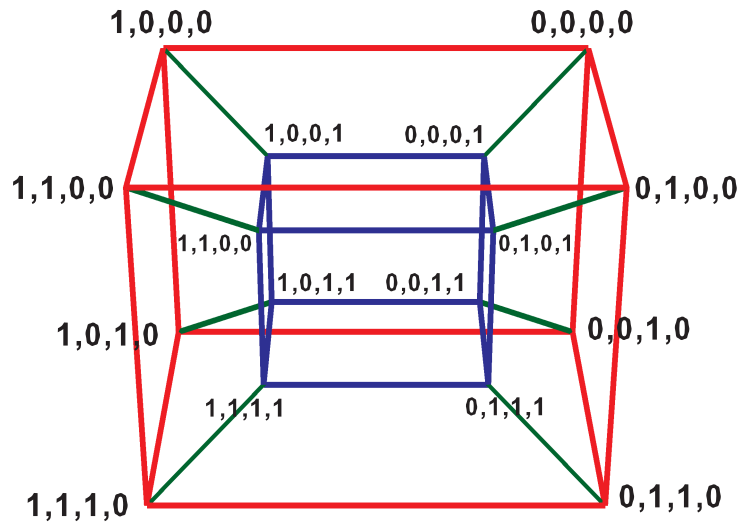
Bu kesim, modelin matematiksel olarak tanımlandığı ve işleyiş mekanizmasının ayrıntılı olarak incelendiği bir kesim olacaktır. Modelin yaptığı varsayımların net bir şekilde ortaya konması amaçlanmıştır.

3.2.1. Organizmalar ve Genotip Uzayı

Canlıların genom zincirleri, *lokus* adı verilen konumlardan oluşur. Bu lokuslarda, kalıtım birimi olan *genler* bulunmaktadır. Genler, pürin türevleri olan Adenin (A) ve Guanin (G), pirimidin türevleri olan Sitozin (C) ve Timin (T) nükleotit bazlarının diziliminden

oluşur. Bir karakterin kalıtımının incelenebilmesi için, genlerde birbirine zıt iki veya daha fazla durumun olması gerekmektedir. Örneğin Mendel'in bezelyeler üzerinde yaptığı kalıtım deneyinde [50], bezelyelerin şeklini belirleyen gende düzgün veya buruşuk olma durumları bulunmaktadır. Bu tarz genlere *alel gen* denir. Canlı birey, bu zıt durumların sadece birini gösterebilir.

DDM, ekosistemin alt birimlerini birey seviyesindeki organizmalar olarak kabul eder. Bütün bireyler S genotip uzayında, L uzunluğundaki yani L tane genin bulunduğu genom zincirleri ile temsil edilir ve $\mathbf{S}^\alpha, \mathbf{S}^\beta, \dots$ gibi sembollerle ifade edilir. Burada $\mathbf{S}^\alpha = \{S^{\alpha_1}, S^{\alpha_2}, \dots, S^{\alpha_L}\}$ ve her bir $S^{\alpha_i} \in \{1, 0\}$ şeklindedir. Başka bir ifade ile, her bir $\mathbf{S}^\alpha$ bireyinin L uzunluklu genom zincirini oluşturan genleri, ikilik tabanda 1 ve 0 ile gösterilir. 1 ve 0 değerlerini alabilen S^{α_i} 'lerin her biri genetik olarak *iki alelli gen* olarak yorumlanabilir. Bireylerin ait olduğu her bir tür ise $\mathbf{S}^a, \mathbf{S}^b, \dots$ ile gösterilir ve L boyutlu hiperküp ile temsil edilen genotip uzayında 2^L tane olası konumdan birine yerleştirilir. Bu konumlar hiperküpün köşelerinden birisidir. Her bir tür birden fazla bireye sahip olabileceğinden dolayı genotip uzayındaki her bir $\mathbf{S}^a$ konumda, $\mathbf{S}^\alpha, \mathbf{S}^\beta, \dots$ gibi birden fazla birey bulunabilir.



Şekil 3.1. $L = 4$ boyutlu hiperküp ile temsil edilen genotip uzayı örneği.

Şekil 3.1’de $L = 4$ boyutlu hiperküp ile temsil edilen genotip uzayı örneği gösterilmiştir. Genotip uzayında $L = 4$ genom zinciri uzunluğuna sahip olası tür sayısı $2^4 = 16$ ’dır. Bu türler hiperküpün $2^4 = 16$ tane köşesine yerleştirilmiştir. Hiperküpün köşelerindeki 1 ve 0 dizileri, türlerin temsil edildiği genom zincirleridir. Genom zincirlerinin bulunduğu konumlar, renkli çizgilerle gösterilen hiperküpün kenarları ile birbirine bağlanmıştır. Genotip uzayında dolu olan konumların yani toplulukta varolan farklı türlerin sayısı *çeşitlilik* olarak adlandırılmaktadır. Hiperküpün köşelerine yerleştirilen her bir tür birden fazla bireye sahip olabileceğinden dolayı, her bir türün sahip olduğu bireylerin toplam sayısı ise *popülasyon* olarak adlandırılmaktadır. Hiperküpün köşelerine yerleştirilen türlerin 4 tane en yakın komşusu vardır. Türler tek bir mutasyonla yani genom zincirlerindeki 0 ve 1’lerden sadece bir tanesinin değişmesiyle bu komşu türlere ulaşabilirler.

3.2.2. Ağırlık Fonksiyonu ve Üreme İhtimaliyeti

Genotip uzayında dolu olan bir $\mathbf{S}^a$ konumunun $\mathbf{S}^a$ bireylerinin üreme kabiliyetleri $H_a = H(\mathbf{S}^a)$ denilen *ağırlık fonksiyonu* ile belirlenir ve t zamanında,

$$H_a(t) = \frac{1}{cN(t)} \sum_{b=1}^{2^L} J_{ab} n_b(t) - \mu N(t) \quad (3.1)$$

denklemleri ile ifade edilir. Burada sabit bir değer olan c , etkileşmelerin şiddetini tanımlayan kontrol parametresidir. $N(t)$, t zamanındaki toplulukta bulunan bireylerin toplam sayısı yani popülasyondur. Alınan toplam ifadesi ise, genotip uzayındaki 2^L tane olası konum üzerinden alınmıştır. J_{ab} genotip uzayındaki $\mathbf{S}^a$ ve $\mathbf{S}^b$ konumları arasındaki ilişkiyi gösteren $\mathbf{J}$ etkileşme matrisinin elemanıdır. $n_b(t)$, genotip uzayında dolu olan $\mathbf{S}^b$ konumunda bulunan bireylerin sayısıdır. $\mu N(t)$ terimi topluluğun içinde bulunduğu fiziksel çevrenin taşıma kapasitesidir. μ ise, sabit taşıma kapasitesi parametresidir.

Ağırlık fonksiyonu fiziksel olarak, Ising modeldeki Hamiltonyen ifadesi ile benzerlik göstermektedir. Bireyler arasındaki J_{ab} etkileşme terimi, Ising modelde spinler arasındaki karşılıklı ferromanyetik ve antiferromanyetik etkileşmeleri karakterize eden terim olarak düşünülebilir. Bir diğer benzerlik, Ising modelin Hamiltonyen ifadesindeki spinlerin dış manyetik alan ile etkileşimi, bireylerin fiziksel çevre ile etkileşimi olarak yorumlanabilir. Buna bağlı olarak μ terimi, dış manyetik alan olarak düşünülebilir.

Ağırlık fonksiyonun'daki ilk terim

$$I_a(t) = \frac{1}{cN(t)} \sum_{b=1}^{2^L} J_{ab} n_b(t) = \frac{1}{c} \sum_{b=1}^{2^L} J_{ab} \rho_b(t) \quad (3.2)$$

etkileşme terimidir. Burada $\rho_b(t) = n_b(t)/N(t)$, $\mathbf{S}^b$ konumunda bulunan bireylerin *yoğunluğudur*. Denklem (3.2), $\mathbf{S}^a$ konumunda bulunan bir bireyin, genotip uzayında yaşayan diğer bütün bireylerle nasıl bir etkileşim içerisinde olduğunu belirler. $\mathbf{J}$ etkileşme matrisinin her bir J_{ab} elamanı,

$$J_{ab} = \begin{cases} T_{ab}R_{ab} & a \neq b \quad \text{ise,} \\ 0 & a = b \quad \text{ise.} \end{cases} \quad (3.3a)$$

$$(3.3b)$$

ifadesi ile üretilir. Burada T_{ab} , genotip uzayındaki $\mathbf{S}^a$ ve $\mathbf{S}^b$ konumları arasında bir etkileşme olup olmadığını belirleyen 2^L uzunluklu *bağılantılık matris dizisidir*. Bu matris dizisinin elemanları T_{ab} ise,

$$T_{ab} = T_{ba} = \begin{cases} 1 & \theta \quad \text{ihtimaliyeti ile,} \\ 0 & 1 - \theta \quad \text{ihtimaliyeti ile.} \end{cases} \quad (3.4a)$$

$$(3.4b)$$

şeklinde bir θ ihtimaliyeti ile belirlenir. θ ihtimaliyeti, genotip uzayındaki herhangi bir konumun diğer konumlardan ne kadarı ile etkileşebileceğini belirler. $\mathbf{T}$ matris dizisi simetriktir. $T_{ab} = T_{ba} = 1$ olduğundan, $\mathbf{S}^a$ konumu $\mathbf{S}^b$ konumu ile ve $\mathbf{S}^b$ konumu da $\mathbf{S}^a$ konumu ile etkileşiyordur. Eğer $T_{ab} = T_{ba} = 0$ iken, konumlar arasında herhangi bir etkileşme bulunmamaktadır. Bu tez çalışmasında $\theta = 0.25$ olarak alınmıştır. Bunun anlamı, genotip uzayındaki herhangi bir konum bütün konumların $1/4$ 'ü ile etkileşim yapabilir. R_{ab} ise, $\mathbf{S}^a$ ve $\mathbf{S}^b$ konumları arasındaki etkileşme değerini belirler. Konumlar arasındaki etkileşme değerinin simetrik olmaması yani konumların karşılıklı olarak aynı etkileşme değerine sahip olmaması için R_{ab} ,

$$R_{ab} = R_1 R_2 \quad (3.5)$$

şeklinde her biri 2^L uzunluğuna sahip iki matris dizisinin çarpımı olarak tanımlanır. R_1 ve R_2 matris dizilerinin elemanları, $[-1, +1]$ aralığından rastgele belirlenir. Simülasyonun başlangıcında $\mathbf{J}$ etkileşme matrisi bir kez oluşturulur ve simülasyon boyunca sabit kalır. Bu bilgilerden sonra, J_{ab} etkileşme matrisi elemanlarının nasıl üretileceğini verelim. İlk önce, $\mathbf{S}^a$ ve $\mathbf{S}^b$ konumlarının sahip oldukları ikilik tabandaki genom zincirlerine XOR operatörü uygulanır. XOR operatörü, konumların genom zincirlerinde bulunan bütün

genleri tek tek karşılaştırır. Eğer genler farklıysa 1, aynıysa 0 sonucunu vererek ikilik tabanda bir sayı oluşturur. Oluşturulan bu sayının onluk tabandaki karşılığı elde edilir. Onluk tabandaki bu tam sayı 0 ile $2^L - 1$ arasında bir sayıdır. Daha sonra, XOR operasyonu sonucunda elde edilen onluk tabandaki tam sayı ile **T** matris dizisinin hangi elemanına bakılacağı belirlenir. Eğer **T** matris dizisinin elemanı 0 ise, S^a ve S^b konumları arasında etkileşme bulunmadığı ve $J_{ab} = 0$ olduğu sonucuna varılır. DDM'nin amacı eş-evrimleşmeyi incelemek olduğu için aynı genetik konumda olan bireyler arasındaki etkileşmeler hesaplamanın dışında tutulmuştur yani $J_{aa} = 0$ alınmıştır. Eğer **T** matris dizisinin elemanı 1 ise, S^a ve S^b konumları arasındaki etkileşme değeri R_{ab} belirlenir. R_{ab} 'nin belirlenmesi için, R_1 R_2 çarpımının belirlenmesi gerekmektedir. R_1 matris dizisinin hangi elemanının alınacağı, **T** matris dizisinde olduğu gibi XOR operasyonu neticesinde elde edilen onluk tabandaki tam sayı ile belirlenir. R_2 matris dizisinin hangi elemanının alınacağı ise, S^b konumunun sahip olduğu genom zincirinin ikilik tabandaki değerinin onluk tabandaki karşılığı alınarak elde edilen tam sayı ile belirlenir. R_1 ve R_2 matris dizilerinden elde edilen değerlerin çarpımı sonucu, S^a ve S^b konumları arasındaki etkileşme değeri olan R_{ab} değeri elde edilir. Yapılan bu işlemler sonucunda, $J_{ab} \in [-1, +1]$ olacak şekilde etkileşme matrisi elemanları üretilmiş olur.

J etkileşme matrisi, genotip uzayında bulunan herhangi iki konum arasında olabilecek bütün etkileşme çeşitlerini içerisinde barındırmaktadır. *Düşmanca etkileşme (Competition)* olarak adlandırılan $J_{ab} < 0$ ve $J_{ba} < 0$ etkileşme çeşidinde, S^a ve S^b konumlarının her ikisinde zarar gördüğü bir etkileşmedir. Bu etkileşmede aynı kaynak (yiyecek, mekan) için rekabet durumu söz konusudur. *İş birlikçi veya Karşılıklı etkileşme (Mutualism)*'yi ifade eden $J_{ab} > 0$ ve $J_{ba} > 0$ etkileşme çeşidinde, S^a ve S^b konumlarının karşılıklı fayda sağlaması durumu söz konusudur. *Av-avcı (Prey-Predator)* veya *Konak-Parazit (Host-Parasite) etkileşme* olarak adlandırılan $J_{ab} < 0$ ve $J_{ba} > 0$ veya $J_{ba} < 0$ ve $J_{ab} > 0$ etkileşme çeşidinde, konumlardan biri fayda sağlarken diğer konum zarar görmektedir. *Tek taraflı fayda (Commensalism) etkileşmesi* olarak adlandırılan $J_{ab} > 0$ ve $J_{ba} = 0$ veya $J_{ba} > 0$ ve $J_{ab} = 0$ etkileşme çeşidinde, S^a ve S^b konumlardan biri fayda sağlarken diğer konum ne fayda sağlamakta nede zarar görmektedir. *Tek taraflı zarar (Amensalism) etkileşmesi* olarak adlandırılan $J_{ab} < 0$ ve $J_{ba} = 0$ veya $J_{ba} < 0$ ve $J_{ab} = 0$ etkileşme çeşidinde, S^a ve S^b konumlardan biri zarar görürken diğer konum ne fayda

sağlamakta nede zarar görmektedir.

S^a konumu için bu etkileşmelerin yapacağı etkiler, S^b konumunun belirli bir t zamanındaki $\rho_b(t) = n_b(t)/N(t)$ yoğunluğuna bağlı olacaktır. Denklem (3.2)'deki etkileşme teriminin değeri, sabit c değerine bağlı olmakla birlikte $J_{ab} = \pm 1$ ve $\rho_b(t) = 1$ limit değerlerinde $[-1/c, +1/c]$ aralığında değişmektedir. c değeri küçük iken $I_a(t)$ etkileşme terimi c 'nin büyük değerlerine göre daha geniş değer aralığına sahip olacaktır. Bu durum, konumların etkileşmeler neticesinde birbirlerini çok daha farklı şekilde etkileyecekleri anlamına gelmektedir.

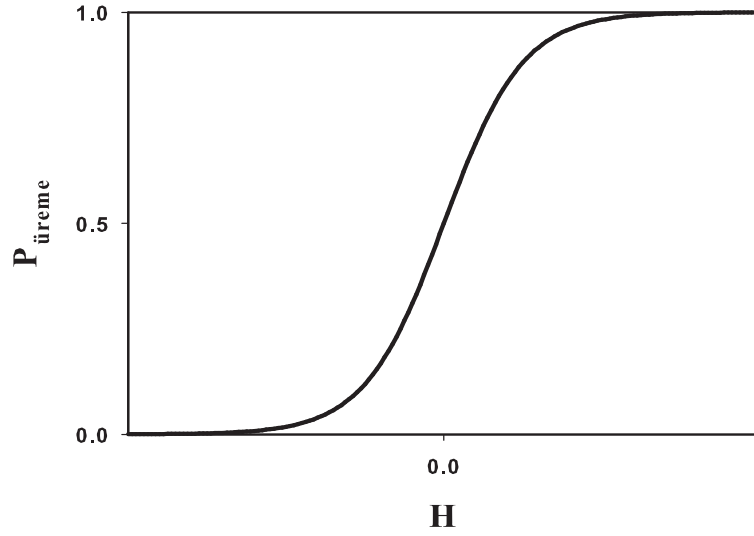
Denklem (3.1)'de verilen ağırlık fonksiyonun'daki ikinci terim $-\mu N(t)$, fiziksel çevrenin bütün koşullarını içerisinde barındıran taşıma kapasitesidir. Topluluğun taşıma kapasitesi, fiziksel çevrenin ne kadar büyük bir topluluğu barındıracağını belirler. Toplam birey sayısının küçük dalgalanmalar dışında hemen hemen sabit kalmasını sağlar. $\mu > 0$ alınarak, taşıma kapasitesinin konumlardan bağımsız olması yani genotip uzayında bulunan konumların hepsinin fiziksel çevreden aynı şekilde etkilenmesi sağlanır. μ taşıma kapasitesi parametresi, fiziksel çevrenin koşullarını belirler. Böylelikle μ değerindeki artış ya da azalış, sırasıyla daha çetin ya da daha ılımlı çevre koşullarının oluşmasına neden olur. Örneğin, hava sıcaklığının çok yüksek değerlere ulaşması sonucu kuraklık meydana gelir. Kuraklık neticesinde, daha çetin çevre koşulları oluşacağı ve buna bağlı olarak bireylerin yaşamlarını devam ettirebilmesi için gerekli olan kaynaklar azalacağı için topluluğun toplam birey sayısında azalma olacaktır.

Denklem (3.1)'de ağırlık fonksiyonu verilen ve S^α bireyleri ile dolu olan S^a konumunun üreme ihtimaliyeti,

$$P_{\text{üreme}}(H_a(t)) = \frac{e^{H_a(t)}}{1 + e^{H_a(t)}} \quad (3.6)$$

denklemleri ile verilir. Eşeyli ve eşeysiz üremenin her ikisi içinde $P_{\text{üreme}}$ ihtimaliyeti $[0,1]$ aralığında değerler alabilir. Şekil 3.2'de $P_{\text{üreme}}$ ihtimaliyetinin H ağırlık fonksiyonuna göre değişimi verilmiştir.

$P_{\text{üreme}}(H_a(t))$ ihtimaliyeti ile eşeysiz üreyen bir S^α bireyi, kendi genom zincirini kopyalayarak $S^{\beta 1}$ ve $S^{\beta 2}$ gibi iki yavru verir. Başka bir ifade ile, eşeysiz üremede bireylerin doğurganlık oranı 2'dir. Her iki yavrunun genom zinciri birbirinden bağımsız



Şekil 3.2. Üreme ihtimaliyetinin ağırlık fonksiyonuna göre değişimi. $H \rightarrow -\infty$ iken $P_{\text{üreme}} \rightarrow 0$ ve $H \rightarrow +\infty$ iken $P_{\text{üreme}} \rightarrow 1$ değerlerine eksponansiyel olarak değişmektedir.

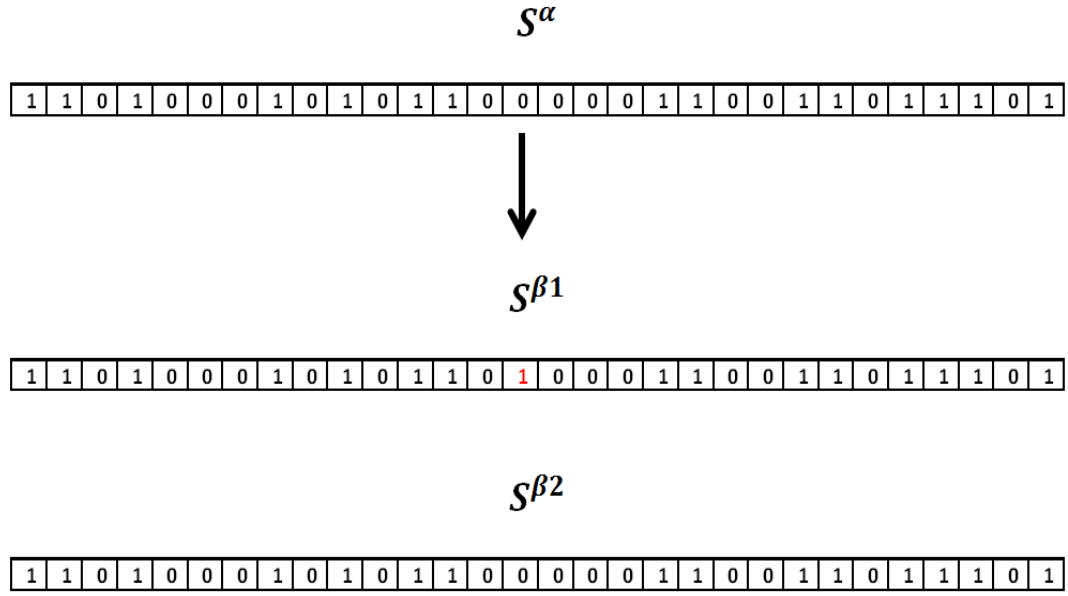
olarak,

$$S_i^{\beta r} = \begin{cases} S_i^\alpha & 1 - P_{mut} \text{ ihtimaliyeti ile} \\ \overline{S}_i^\alpha & P_{mut} \text{ ihtimaliyeti ile} \end{cases} \quad (3.7a)$$

$$(3.7b)$$

şeklinde belirlenir. Burada $i = 1, \dots, L$ ve $r = 1, 2$ değerlerini almaktadır. P_{mut} ise, gen başına gerçekleşen sabit bir mutasyon oranıdır. Denklem (3.7a)'da, yavruların genom zincirlerindeki i . genin $1 - P_{mut}$ ihtimaliyeti ile mutasyona uğramayarak S^α bireyinin i . genine eşit olabileceği görülmektedir. Ancak Denklem (3.7b)'de ise, yavruların genom zincirlerindeki i . genin P_{mut} ihtimaliyeti ile mutasyona uğrayarak S^α bireyinin i . geninden farklı olabileceği görülmektedir. Denklem (3.7b)'de, farklı olan i . gen " $\overline{S}_i^\alpha$ " şeklinde gösterilmiştir. Üreme süreci tamamlandığında S^α bireyi topluluktan kaldırılmaktadır. Şekil 3.3'de eşeysiz üreme için örnek bir gösterim verilmiştir.

Şekil 3.3'de eşeysiz bir üreme süreci gösterilmiştir. S^α bireyi, 1 ve 0'lardan oluşan $L = 30$ genom zinciri uzunluğuna sahiptir. Üreme sonucunda S^α bireyi, $S^{\beta 1}$ ve $S^{\beta 2}$ olmak üzere iki yavru vermektedir. $S^{\beta 1}$ yavrusunun kırmızı renkle gösterilen 15. geni mutasyona uğrayarak, S^α bireyinin 15. geninden farklı bir gene sahip olmuştur. $S^{\beta 2}$ yavrusu ise, mutasyona uğramayarak S^α bireyinin sahip olduğu genlere sahiptir yani S^α bireyinin genetik özelliğini aynen taşımaktadır.



Şekil 3.3. Eşeyli üremede yavruların oluşumu.

Eşeyli üreme durumunda ise, genotip uzayından rastgele seçilen S^a konumundaki S^α bireyi ile S^b konumundaki S^β bireyin birbirlerine olan genetik benzerliklerine bakılmalıdır. Bu nedenle Hamming mesafesi denilen bir genetik benzerlik hesabı yapılır. $d^{\alpha\beta}$ ile gösterilen Hamming mesafesi,

$$d^{\alpha\beta} = \frac{1}{2} \sum_{i=1}^L |S_i^\alpha - S_i^\beta| \quad (3.8)$$

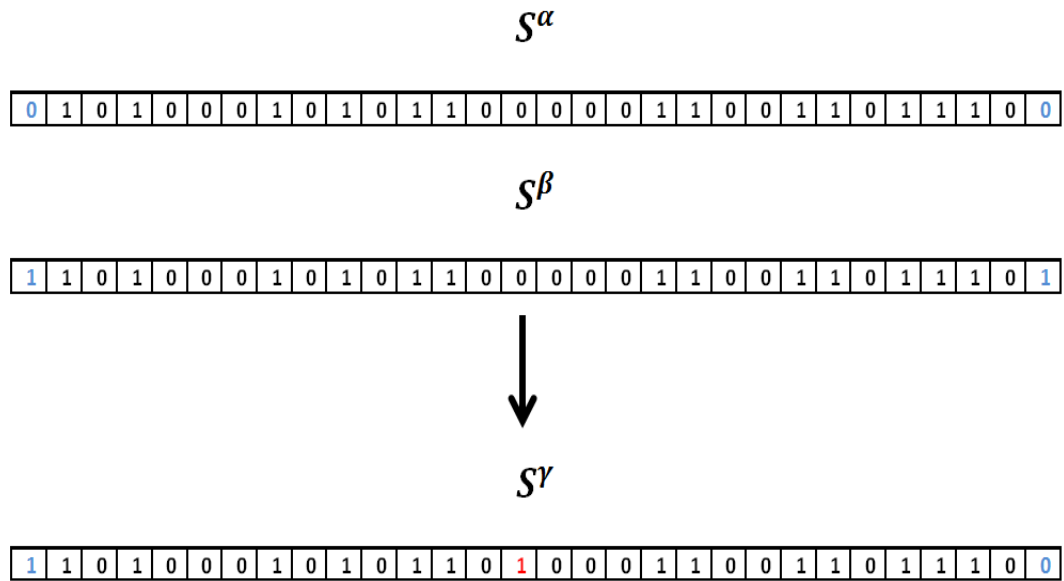
ifadesi ile elde edilir [140]. Burada S_i^α ve S_i^β , S^α ve S^β bireylerinin genom zincirlerindeki genlerdir. Denklem (3.8)'den anlaşılacağı üzere S^α ve S^β bireylerinin Hamming mesafesi hesaplanırken, bireylerin genom zincirlerindeki bütün genler birer birer karşılaştırılmaktadır. Bu karşılaştırma sonunda farklı genlerinin toplam sayısı, bireyler arasındaki Hamming mesafesini vermektedir. S^α ve S^β bireyleri arasında eşeyli üremenin gerçekleşebilmesi için, $d^{\alpha\beta}$ mesafesinin $d^{\alpha\beta} \leq d_{max}$ şeklinde sınırlanması gerekmektedir. Sabit bir değer olan d_{max} mesafesi, bireyler arasındaki farklı genlerin sayısının olabilecek en büyük değeridir. Bu değerin üzerinde bireylerin genetik benzerlikleri, üreme gerçekleştirmek için uygun olmayacaktır. Hamming mesafesi üreme gerçekleştirmek için uygun olan S^a konumundaki S^α bireyi ve S^b konumundaki S^β bireylerinin üreme ihtimaliyeti ise,

$$P_{üreme}(H_a(t), H_b(t)) = \sqrt{P_{üreme}(H_a(t))P_{üreme}(H_b(t))} \quad (3.9)$$

ifadesi ile verilir. S^α ve S^β bireyleri, üreme sonucunda bir tane S^γ yavrusu verir. Başka bir ifade ile, eşeyli üremede bireylerin doğurganlık oranı 1'dir. Bu yavrunun genom zinciri ise

$$S_i^\gamma = \begin{cases} S_i^\alpha & \% 50 \text{ ihtimaliyetle} \\ S_i^\beta & \% 50 \text{ ihtimaliyetle} \end{cases} \quad (3.10a)$$

şeklinde belirlenir. Görüldüğü gibi yavrunun genom zincirindeki genler, ebeveynlerinin genom zincirindeki genlerden rastgele seçilerek belirlenir. Ebeveynler, sahip oldukları genleri yavrusuna aktarmak için eşit şansa sahiptirler. Eğer $S_i^\alpha = S_i^\beta$ ve yavrunun i . geni mutasyona uğramıyorsa otomatik olarak yavrunun i . geni ebeveynleri ile aynı olacaktır. Yavrunun ebeveynlerinden farklı bir gene sahip olabilmesi için mutasyona uğraması gerekmektedir. Ancak bu mutasyon, sadece $S_i^\alpha = S_i^\beta$ durumunda gerçekleşirse S_i^γ 'nin yani yavrunun i . geni ebeveynlerinden farklı olabilir. Çünkü $S_i^\alpha \neq S_i^\beta$ durumunda, yavru mutasyona uğrasa bile mecburen $S_i^\gamma = S_i^\alpha$ veya $S_i^\gamma = S_i^\beta$ olacağından, yavru ebeveynlerinden farklı bir gene sahip olamaz.



Şekil 3.4. Eşeyli üremede yavrunun oluşumu.

Şekil 3.4'de eşeyli bir üreme süreci gösterilmiştir. S^α ve S^β bireyleri, 1 ve 0'lardan oluşan $L = 30$ genom zinciri uzunluğuna sahiptir. Bu bireylerin mavi ile gösterilen 1. ve 30. genleri birbirinden farklıdır. Bu durumda S^α ve S^β bireyleri arasındaki Hamming mesafesi $d^{\alpha\beta} = 2$ 'dir. Üreme sonucunda oluşan S^γ yavrusu 1. genini % 50 ihtimaliyetle S^β bireyinden, 30. genini ise % 50 ihtimaliyetle S^α bireyinden almıştır. Ayrıca S^γ

yavrusunun kırmızı ile gösterilen 15. geni, P_{mut} ihtimaliyeti ile mutasyona uğramıştır. S^γ yavrusunun 1., 15. ve 30. genleri dışında diğer bütün genleri, mutasyona uğramadığı için S^α ve S^β bireylerinin genleri ile aynıdır. Üreme süreci tamamlandığında ebeveynler topluluktan kaldırılmamaktadır.

Bu tez çalışmasında, sadece eşeysiz üreme üzerinde durulmuştur. Bu durumun sebebi, eşeyli üremede bireyler arasındaki Hamming mesafesinin belirlenmesinin çok uzun hesaplama süresi gerektirmesidir. Çünkü Hamming mesafesi belirlenirken, bireylerin L uzunluklu genom zincirlerindeki bütün genler karşılaştırılmaktadır. Bu karşılaştırma, üremenin gerçekleşebilmesi için uygun Hamming mesafesini bulana kadar devam edecek ve üreme gerçekleştirebilecek bireyler belirlenecektir. Bu süreç simülasyonun adımını yavaşlatacaktır. Christensen et al. [36] tarafından yapılan bir DDM çalışmasında, eşeyli ve eşeysiz üreme ile yapılan simülasyonların benzer sonuçlar verdiği elde edilmiştir. Her iki üreme durumunda da konumların yaşam süresi dağılımlarının, fosil kayıtları ile uyumlu olarak bir kuvvet kanunu dağılımı sergilediği bulunmuştur.

Ağırlık fonksiyonu genotip uzayındaki konumların bir fonksiyonu olduğu için, üreme ihtimaliyeti de genotip uzayındaki konumların bir fonksiyonu olacaktır. Eğer S^a konumundaki bireylerin üreme ihtimaliyeti $P_{üreme}(H_a(t))$, S^b konumundaki bireylerin üreme ihtimaliyeti $P_{üreme}(H_b(t))$ değerinden büyükse, S^a konumundaki bireylerin topluluğa S^b konumundaki bireylerden daha iyi uyum sağladığı yani uygunluk değerinin daha yüksek olduğu söylenebilir. Buradan da anlaşılabacağı üzere, konumların $P_{üreme}$ ihtimaliyeti aynı zamanda uygunluk değerine özdeşir.

Anlaşıldığı üzere, üreme olayında ebeveynlerin genom zincirlerinin kopyalanması söz konusudur. Doğacak olan yavrunun genom zinciri, ebeveynlerin genom zincirinin gen gen kopyalanmasıyla belirlenir. Ancak bu kopyalama sırasında L uzunluklu genom zincirindeki her bir gende sabit bir P_{mut} ihtimaliyeti ile mutasyon meydana gelebilmektedir. Bu sebeple genotip uzayındaki konumlarda bulunan bireylerin gerçekleştirecekleri bir üremenin başarılı olabilmesi için bireylerin, sahip oldukları genlerin olabildiğince büyük bölümünü yavrularına aktarması gerekmektedir. Başka bir ifadeyle, üreme sonucunda bireyle yavrusu arasındaki genetik benzerlik ne kadar fazla olursa, gerçekleştirilen üreme o kadar başarılı olur. Başarılı üreme sonucunda, bireylerin

ait oldukları konumlar topluluğa daha fazla uyum sağlayacaktır. $P_{üreme}$ ihtimaliyeti aynı zamanda uygunluk değerine özdeş olduğu için, başarılı üreme gerçekleştiren bireyin ait olduğu konumun $P_{üreme}$ ihtimaliyeti artacaktır.

Konumların başarılı bir üreme gerçekleştirecek, genetik kalıtımı sağlıklı bir şekilde devam ettirebilmesi için P_{mut} mutasyon ihtimaliyetinin belirli bir değerin altında olması gerekmektedir. Bu değeri belirleyebilmek için eşeysiz üreme durumunu inceleyelim:

Genotip uzayındaki $\mathbf{S}^a$ konumunda bulunan bireylerin toplam sayısındaki değişim ΔN_a , diğer konumlarda bulunan toplam birey sayısındaki değişim ise $\Delta N_{\bar{a}}$ ile gösterilsin. $\mathbf{S}^a$ konumundaki bir birey eşeysiz üreme gerçekleştirdiğinde, iki yavru vermekte ve kendisi topluluktan kaldırılmaktaydı. Bu durumda, yavruların genom zincirine bakılmadan $\mathbf{S}^a$ konumundaki birey sayısının 1 azaldığı yani $\Delta N_a = -1$ söylenebilir. Yavruların genom zincirlerindeki her bir gende meydana gelebilecek P_{mut} mutasyon ihtimaliyetine göre ΔN_a için 3 olasılık bulunmaktadır.

1) Yavruların ikisinin de mutasyona uğramama ihtimali

$$P_1 = \binom{2}{0} (1 - P_{mut})^L (1 - P_{mut})^L = (1 - P_{mut})^{2L} \quad (3.11)$$

şeklindedir. Bu durumda $\Delta N_a = -1 + 2 = +1$ ve $\Delta N_{\bar{a}} = 0$ olacaktır.

2) Yavruların birinin mutasyona uğrayıp diğerinin uğramama ihtimali

$$P_2 = \binom{2}{1} (1 - P_{mut})^L [1 - (1 - P_{mut})^L] = 2(1 - P_{mut})^L [1 - (1 - P_{mut})^L] \quad (3.12)$$

şeklindedir. Bu durumda $\Delta N_a = -1 + 1 = 0$ ve $\Delta N_{\bar{a}} = +1$ olacaktır.

3) Yavruların ikisinin de mutasyona uğrama ihtimali

$$P_3 = \binom{2}{2} [1 - (1 - P_{mut})^L][1 - (1 - P_{mut})^L] = [1 - (1 - P_{mut})^L]^2 \quad (3.13)$$

şeklindedir. Bu durumda $\Delta N_a = -1 + 0 = -1$ ve $\Delta N_{\bar{a}} = +2$ olacaktır.

Böylelikle bireyin eşeysiz üremesi sonucunda $\mathbf{S}^a$ konumundaki toplam birey sayısındaki değişim

$$\Delta N_a = \Delta N_a P_1 + \Delta N_a P_2 + \Delta N_a P_3 \quad (3.14)$$

ifadesi ile elde edilir. P_2 ihtimalinde $\Delta N_a = 0$ olduğu için

$$\begin{aligned}\Delta N_a &= \Delta N_a P_1 + \Delta N_a P_3 \\ &= (1 - P_{mut})^{2L} - [1 - (1 - P_{mut})^L]^2 \\ &= 2(1 - P_{mut})^L - 1\end{aligned}\quad (3.15)$$

olarak elde edilir.

S^a konumu dışında diğer bütün konumlardaki toplam birey sayısındaki değişim

$$\Delta N_{\bar{a}} = \Delta N_{\bar{a}} P_1 + \Delta N_{\bar{a}} P_2 + \Delta N_{\bar{a}} P_3 \quad (3.16)$$

ifadesi ile elde edilir. P_1 ihtimalinde $\Delta N_{\bar{a}} = 0$ olduğu için

$$\begin{aligned}\Delta N_{\bar{a}} &= \Delta N_{\bar{a}} P_2 + \Delta N_{\bar{a}} P_3 \\ &= 2(1 - P_{mut})^L [1 - (1 - P_{mut})^L] + 2[1 - (1 - P_{mut})^L]^2 \\ &= 2 - 2(1 - P_{mut})^L\end{aligned}\quad (3.17)$$

olarak elde edilir. Üreme sonucunda, S^a konumundaki bireyin genetik kalıtımının korunması için

$$\Delta N_a \geq \Delta N_{\bar{a}} \quad (3.18)$$

olması gereklidir. Denklem (3.15) ve Denklem (3.17), Denklem (3.18)'de yerine yazılırsa

$$P_{mut} \leq 1 - \left(\frac{3}{4}\right)^{1/L} \quad (3.19)$$

ifadesi elde edilir. Denklem (3.19), eşeysiz üremede bireylerin başarılı bir üreme gerçekleştirebilmesi için P_{mut} mutasyon ihtimaliyetinin alabileceği en büyük değeri belirler. Bu değer üzerinde, kalıtım özelliği kaybolacaktır. Bu nedenle, bireyler topluluğa uyum sağlayamayacak ve genotip uzayında rastgele dağılmış bir yapı ortaya çıkacaktır. $L = 30$ için, P_{mut} mutasyon ihtimaliyetinin alabileceği en büyük değer

$$P_{mut} \leq 0.00859 \quad (3.20)$$

şeklindedir.

Bireyler üzerinde gerçekleştirilecek bir diğer olay, öldürme işlemidir. Öldürme işleminde, bireyler sabit bir $P_{öldür}$ ihtimaliyeti ile topluluktan direkt olarak kaldırılırlar. Ortalama $P_{üreme}$ ihtimaliyeti $P_{öldür}$ ihtimaliyetine eşit olması durumunda, toplulukda

bulunan konumların sahip oldukları birey sayılarının toplamı yani popülasyon sayısı değişmeyecektir. Bu durumda Denklem (3.6)

$$P_{üreme}(H_a(t)) = \frac{e^{H_a(t)}}{1 + e^{H_a(t)}} = P_{öldür} \quad (3.21)$$

şeklinde olacaktır. Denklem (3.21) basit matematiksel işlemlerle

$$H_a(t) = \ln\left(\frac{P_{üreme}(H_a(t))}{1 - P_{üreme}(H_a(t))}\right) = \ln\left(\frac{P_{öldür}}{1 - P_{öldür}}\right) \quad (3.22)$$

şeklinde yazılabilir. $c \rightarrow \infty$ limit değerinde, etkileşme matrisi elamanları $J_{ab} = 0$ yani organizmalar arasında etkileşme olmayacaktır. Bu durumda Denklem (3.1)

$$H_a(t) = -\mu N(t) = \ln\left(\frac{P_{öldür}}{1 - P_{öldür}}\right) \quad (3.23)$$

olacak ve

$$N(t) = -\frac{1}{\mu} \ln\left(\frac{P_{öldür}}{1 - P_{öldür}}\right) \quad (3.24)$$

olarak elde edilecektir. $N(t)$ 'nin bu değeri, toplam popülasyonun etrafında dalgalanmalar gerçekleştirdiği değeri verir.

3.2.3. Zaman Adımı ve Nesilsel Zaman

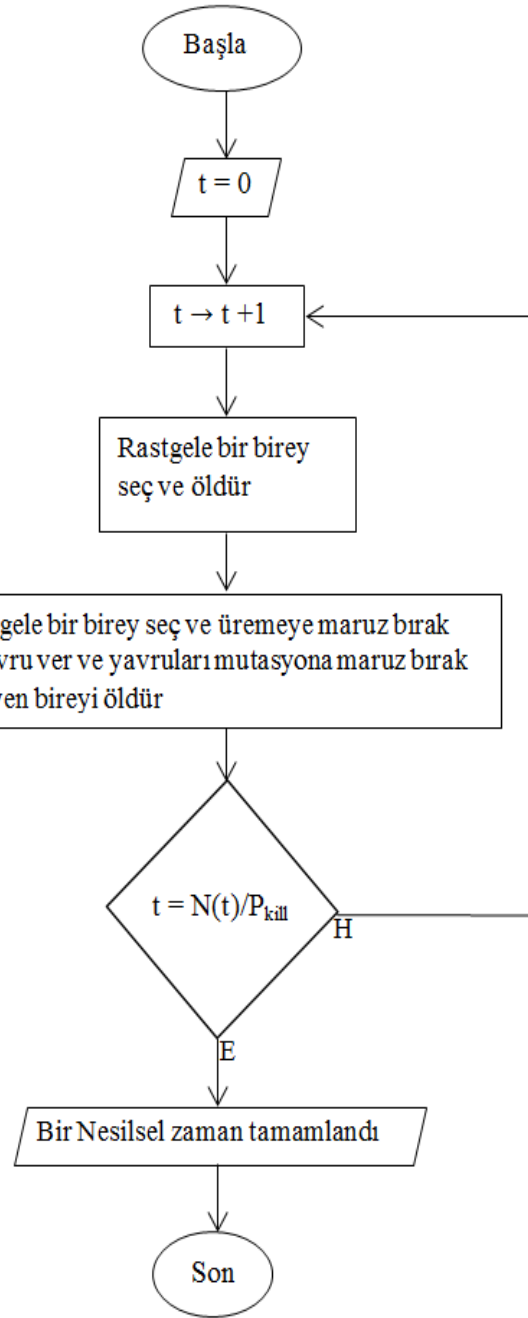
Simülasyonları oluşturmak için temel bir zaman adımının tekrarlanması gerekmektedir.

Bu temel zaman adımı, doğumu ve ölümü içermelidir. Bir zaman adımının süreçleri,

- 1) Topluluktan rastgele bir birey seç ve o bireyi sabit $P_{öldür}$ ihtimaliyeti ile öldür
- 2) Rastgele başka bir birey seç ve $P_{üreme}$ ihtimaliyeti ile eşeysiz üremeye maruz bırak
 - 2.1) Üreyen bireyin 2 yavrusunu P_{mut} ihtimaliyeti ile mutasyona maruz bırak
 - 2.2) Üreyen bireyi $P_{öldür}$ ihtimaliyeti ile öldür

şeklinde dir.

Biyolojik topluluklarda dinamik incelemeler için kullanılan zaman skalası *nesil* olarak adlandırılır. Yapılan simülasyonlar ise *nesilsel zaman* ile ölçülür. Bir nesilin tamamlanması için ise $t = N(t)/P_{öldür}$ adet zaman adımının tekrarlanması gerekmektedir. Bunun anlamı ise, belirli bir zamanda yaşayan bütün bireylerin öldürülmesi için gerekli ortalama zamandır. Daha net anlaşılması için Şekil 3.5'de bir nesilsel zaman için akış diyagramı verilmiştir.



Şekil 3.5. Eşeysiz üreme durumunda bir nesilsel zaman için akış diyagramı. $t = N(t)/P_{öldür}$ adet zaman adımı tekrarlanınca bir nesilsel zaman tamamlanmış olur.

Şekil 3.5'deki akış diyagramı, $t = 0$ zaman adımının sıfırlanması ile başlar. Zaman adımı bir artırılarak tekrarlama başlar. Rastgele seçilen bir bireyin öldürülmesi ve ardından rastgele seçilen başka bir bireyin üremeye maruz bırakılması ile bir zaman adımı yapılması gereken işlemler tamamlanmıştır. Bir nesil bitene kadar yani $t = N(t)/P_{öldür}$

olana kadar her bir zaman adımında bu işlemler tekrarlanır.

2002 yılında DDM ile biyolojik evrimdeki türlerin yok olma oranı ve zamanla tür bolluğunun değişimi araştırılmıştır [141]. Yok olma oranındaki büyük bir azalma neticesinde, üretilen ekosistemin daha karmaşık olması ile birlikte ortama daha iyi adapte olduğu görülmüştür. 2003 yılında DDM ile, mutasyonun eşik değeri için detaylı bir teorik tanımlama yapılmıştır. [142]. Modelin mikroskobik dinamiklerinin daha önce makroskobik seviyede tanımlanan olayları nasıl açıkladığı gösterilmiştir. 2005 yılında, evrimsel bir ekoloji modelindeki tür bolluğu incelenmiş ve DDM'nin ağ özelliklerine oldukça temel bir giriş yapılmıştır [143]. Tür bolluğunun sadece türler arasındaki bağlantının fazla olduğu durumlar için gerçekçi olduğu bulunmuştur. 2006 yılında, DDM ilk defa fenotip uzayında çalışılmış ve bireylerin ebeveynlerinin özelliklerini taşıdığı gözlemlenmiştir [144]. 2006 yılındaki bir diğer çalışmada, DDM ile biyolojik sistemlerde kuvvetli etkileşen organizmalar için çeşitlenmenin sadece belirli bir etkileşme değeri üzerinde olacağı ve bu değer altında rekabetçi dışlamanın olacağı görülmüştür [145]. 2007 yılında ise türlerin yaşam süreleri incelenmiş ve DDM'nin fosil kayıtları ile uyumlu olarak $P(s) \sim s^{-2}$ şeklinde bir kuvvet kanunu dağılımına sahip olduğu görülmüştür [146].

4. BÖLÜM

SİMÜLASYON SONUÇLARI

Bu bölümde, DDM ile incelenen biyolojik topluluklar için yapılan simülasyon sonuçları verilecektir. İlk olarak tek bir simülasyonda, topluluğun dinamik yapısı eş-evrimleşme modeli olan DDM ile incelenecektir. Eş-evrimleşme olayını anlamak için, türler arasındaki etkileşmelerle ağ sistemi kurulacaktır. Dinamik bir yapıya sahip olan ağ sistemini analiz edebilmek için, bir takım ağ parametrelerinin değerlerinin zamanla değişimi sunulacaktır. Tek bir simülasyonda son olarak, belirli bir nesil aralığında topluluğun soy ağacı kurulacak ve aynı topluluk yapısı için Molecular Evolutionary Genetics Analysis (MEGA) programı kullanılarak [147], ML ve NJ metotlarının tahmin ettiği soy ağaçları ile karşılaştırma yapılacaktır. Bu karşılaştırmanın, TOPD/FMTS (TOPological Distance/From Multiple To Single) yazılımını [148] kullanarak yapılan nümerik analiz sonuçları verilecektir. Daha sonra biyolojik toplulukların genel doğası hakkında anlamlı sonuçlar elde etmek için, çok sayıda simülasyon hesaplarının ortalamasını alarak elde edilen sonuçlar sunulacaktır.

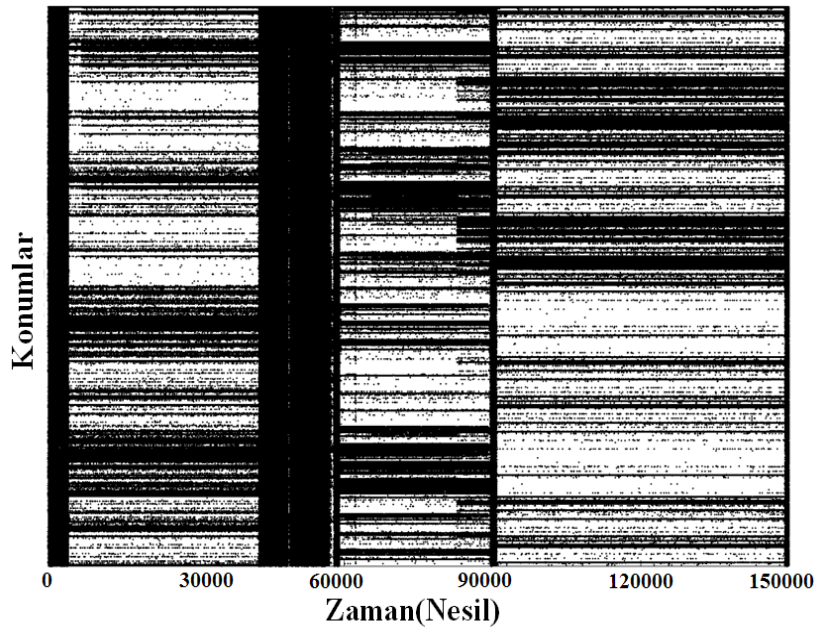
Bu tez çalışması kapsamında yapılmış olan bütün simülasyon çalışmaları için gerekli olan nümerik hesaplamalar, C programlama dili kullanılarak yapılmıştır. Gerekli kaynak kodun tamamı Prof. Dr. Osman CANKO tarafından üretilmiştir. Simülasyonda kullanılan rastgele sayıların üretilmesi için ise Mersenne-Twister rasgele sayı üreticine [149] başvurulmuştur. Bu rastgele sayı üreticinin kullanılmasının nedeni ise, üretilen sayılar arasında bir periyodiklik olmamasından yani üretilen sayıların rastgeleliğinin yüksek olmasından dolayıdır. Programın çalışması hızlıdır. Ancak modelin doğası gereği sistem oldukça büyük etkileşme matrisi barındırdığı için simülasyonun sonuçlanması vakit almaktadır.

4.1. Tek Bir Simülasyon Sonucu

Tek bir simülasyonun yapıldığı bu kesimde, eş-evrimleşme nedeniyle topluluk yapısının zamanla nasıl değiştiği ağ sistemleri ile incelenecektir. Yapılan simülasyon sonucunun kabul edilebilmesi için, gerçek dünyadaki olaylarla uyumlu olması gerekmektedir. Bu çerçevede ilk düşünülecek faktör, topluluğun büyüklüğüdür. Topluluk ne kadar büyük olursa yani toplulukda bulunabilecek farklı tür sayısı ne kadar fazla olursa simülasyon sonucu o kadar gerçekçi olacaktır. Ayrıca bireylerin sahip oldukları genom zincirlerinin uzunluğu yani gen sayısı da simülasyonun gerçekçiliğini etkileyen bir faktördür. Gerçek dünyada bulunan türlerden örneğin meyve sineğinde yaklaşık olarak 13600 [150] tane gen bulunmaktadır. Ancak bakterilerden türeyen bazı basit yaşam formlarındaki gen sayısı 20'den azdır. Elimizdeki bilgisayarların kapasiteleri nedeniyle bu çalışmada tek bir simülasyon için, toplulukdaki türlerin bireylerinin sahip oldukları genom zinciri uzunlukları yani gen sayıları $L = 30$ olarak alınmıştır. Bu değer meyve sineği ile karşılaştırıldığında oldukça düşük olmasına rağmen basit yaşam formları için kabul edilebilir bir değerdir. $L = 30$ için toplulukda 2^{30} tane olası tür bulunabilecektir. Gerçek dünyadaki toplam tür sayısı ise, en iyi tahminle 10-14 milyon aralığında olduğu düşünülmektedir [151]. Ancak biyologlar tarafından tanımlanan tür sayısı yaklaşık olarak 2 milyondur [151]. Fosil kayıtlarından ise şu ana kadar yaklaşık olarak 250 bin tür tanımlanmıştır [151].

DDM'deki c etkileşme şiddeti parametresi ve μ fiziksel çevrenin taşıma kapasitesi parametresi, topluluktaki toplam birey sayısının belirlenmesinde önemli role sahiptirler. Hall et al. [141] tarafından, bireylerin genom zinciri uzunluğunun $L = 20$ alınarak yapılan bir çalışmada, DDM'nin parametre uzayı taranmıştır. Toplulukdaki toplam birey sayısının bu parametrelerdeki değişime çok hassas olduğu görülmüştür. Hall et al., μ ve c değerlerindeki azalışın, toplulukdaki toplam birey sayısını artıracaklarını belirtmişlerdir. Bu tez çalışmasında kullanılacak c ve μ değerleri bu parametre uzayı tartışması çerçevesinde seçilmiştir. Tek bir simülasyon hesabında bu parametreler, $c = 0.5$ ve $\mu = 0.0002$ olarak alınmıştır. Önemli bir evrim mekanizması olan mutasyon ihtimaliyeti, gen başına $P_{mut} = 0.002$ ve toplulukdaki ölüm oranı $P_{kill} = 0.2$ olarak alınmıştır. Türlerin etkileşme oranı ise $\theta = 0.25$ olarak yani genotip uzayındaki bir tür varolan diğer bütün türlerin $1/4$ 'i ile etkileşecek şekilde alınmış ve simülasyon boyunca

sabit tutulmuştur. Ayrıca, bireyler arasındaki etkileşme değerlerini veren **J** etkileşme matrisi de simülasyonun başlangıcında rastgele oluşturulmuş ve simülasyon boyunca sabit tutulmuştur. Simülasyon, genom uzayından rasgele seçilen tek bir konuma Denklem (3.24)'e göre $N(t = 0) = -\frac{1}{\mu} \ln\left(\frac{P_{öldürme}}{1 - P_{öldürme}}\right) \cong 6931$ tane bireyin yerleştirilmesi ile başlatılmıştır. Bu konumun genotip uzayındaki komşu konumlara yaptığı mutasyonlar neticesinde, simülasyonun başlangıcında hızlı bir çeşitlenme gözlenmiştir.



Şekil 4.1. Genotip uzayının doluluğunun zaman ile değişimi.

Şekil 4.1'de genotip uzayındaki doluluğun zamanla değişimi gösterilmektedir. Her konumun, 1 ve 0 ile gösterilen genom zinciri bulunmaktadır. Genom zincirinin ikilik tabandaki değerinin onluk tabana çevrilmesiyle elde edilen bir tam sayı ile konumlara etiket verilmiştir. Konumların etiketleri $[0, 2^L - 1]$ aralığındaki sayılardır. Konumların etiketleri y eksenini, boyunca lineer olarak yerleştirilmiştir. Ancak y ekseninin fiziksel bir anlamı bulunmamaktadır. x eksenine ise, nesilsel olarak geçen zamanı göstermektedir. Genotip uzayındaki dolu olan yani en az bir bireye sahip olan konumlar nokta ile temsil edilmiştir. Bu kesikli yapı, topluluğun uzun zaman periyodu boyunca küçük değişimler dışında pek bir değişiklik göstermediği yani *q-ESS* (*quasi evolutionary stable strategies*) periyotları ve *q-ESS* periyoduna göre oldukça kısa zaman periyoduna sahip *hektik* (*hectic*) periyotlardan oluşmaktadır [36, 141]. *q-ESS* periyodu, genotip uzayında uzun süre boyunca evrimsel olarak sanki kararlıymış gibi davranan konumlardan

oluşmaktadır. Bu konumlar, mutasyonlara karşı dayanıklılık göstermektedir ve evrimsel aktivite düşüktür. Başka bir ifade ile, bu konumlardaki birey sayısı çok yüksektir. Bu nedenle mutasyon bu konumdaki bütün bireylerin yok olmasına neden olamamaktadır. Bu dayanıklılığı gösteremeyen ve topluluğa uyum sağlayamamış bazı konumlar gözükp kaybolmaktadırlar. Bu gözükp kaybolan konumlar, dayanıklı konumlardan yapılan mutantlardır. Bu mutantlar da mutasyon yaparak dayanıklı konumların etrafına toplanır ve onlardan fayda sağlamaya çalışır. Ancak q-ESS periyodunun sonunda, mutasyonlara karşı dayanıklı olan konumların yok olmasıyla toplulukta bulunan konumlar mutasyonlara karşı daha hassas hale gelir. Böylelikle topluluğun q-ESS yapısı, evrimsel aktivitenin çok yüksek olduğu hektik periyot ile son bulur. Hektik periyot, evrimsel olarak oldukça aktif konumlardan oluştuğundan, çok sayıda yeni konumlar meydana gelmektedir. Ancak genotip uzayındaki konumlar, avantajlı bir duruma sahip olamadıkları için topluluğa uyum sağlayamaz ve yok olur. Bu nedenle, topluluğun ortalama üreme ihtimaliyeti ölme ihtimaliyetine eşit yani $\langle P_{\text{üreme}} \rangle = P_{\text{öldürme}}$ 'dir. Bu sebeple hektik periyotdaki konumların yaşam süreleri oldukça kısadır. Konumların sürekli olarak gözükp kaybolmaları nedeniyle topluluğun sahip olduğu konumlarda dalgalanmalar meydana getirecektir. Bu dalgalanmalar neticesinde hektik periyodun sonunda, sistem başka bir q-ESS durumu bulacak ve topluluk mutasyonlara karşı dayanıklı konumlardan oluşacaktır. Sistemin bulduğu bu q-ESS durumunun, topluluğun evrimsel süreçte sahip olduğu son durum olduğu anlamına gelmez. Çünkü topluluğun evrimsel dinamiği farklı q-ESS durumlarında, uygunluk peyzajında belirli uygunluk değerlerine sahip farklı konumlarla karakterize edilmektedir.

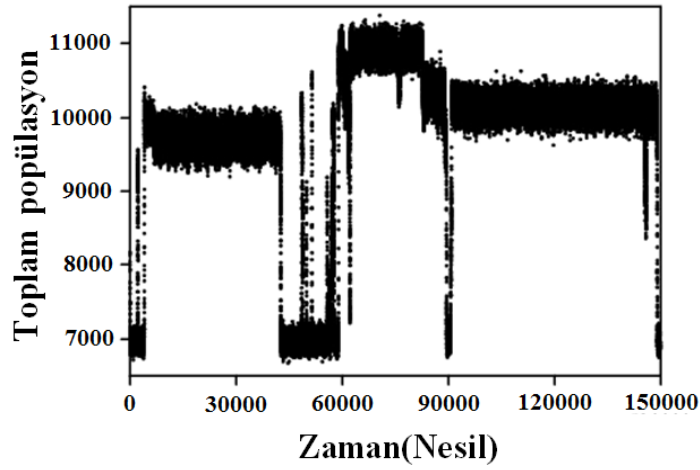
Denlem (3.19) ile verilen P_{mut} değeri için eşik değer aşıldığında, topluluk hektik periyottan çıkamayacak ve hiç bir zaman q-ESS durumu bulamayacaktır. Çünkü genotip uzayında bulunan bütün konumlar, mutasyon etkisinden kurtulamayacak ve sürekli evrimleşme gerçekleştirecektir. Büyük c değerlerinde etkileşme çok şiddetli olacağı için, konumlar topluluğa uyum sağlayamayacak ve q-ESS durumu gözlenmeyecektir. Ayrıca büyük μ değerlerinde de, topluluğun içinde bulunduğu fiziksel çevrenin koşulları çok sert olacağı için konumlar topluluğa uyum sağlayamazlar. Bunun sonucunda sistem q-ESS yapısı bulamaz.

4.1.1. Toplam Popülasyon ve Çeşitlilik

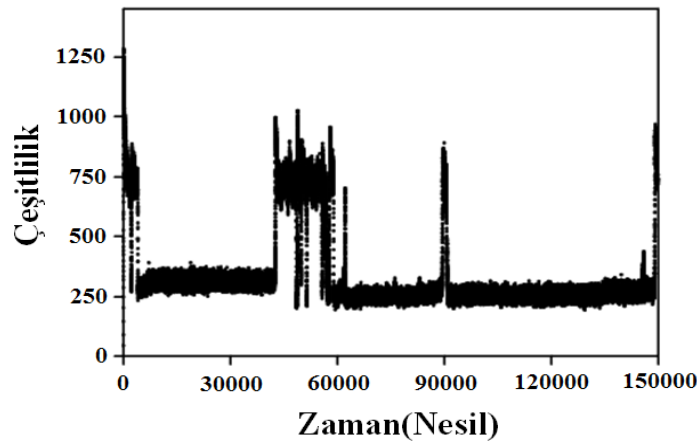
Genotip uzayında belirli bir t zamanında, dolu olan i konumunun sahip olduğu birey sayısı yani popülasyonu $n_i(t)$ olmak üzere, topluluğun sahip olduğu toplam popülasyon

$$N(t) = \sum_i n_i(t) \quad (4.1)$$

ifadesi ile elde edilmektedir. Topluluğun çeşitliliği ise, dolu olan bu konumların yani yaşayan farklı türlerin sayısıdır. Evrimsel süreçte bu ifadeler zamanla değişim göstermektedir.



(a)



(b)

Şekil 4.2. Şekil 4.1'deki simülasyon için, (a) Toplam popülasyonun ve (b) Çeşitliliğin zamanla değişimi.

Şekil 4.2'de, Şekil 4.1'deki simülasyon için topluluğun sahip olduğu toplam

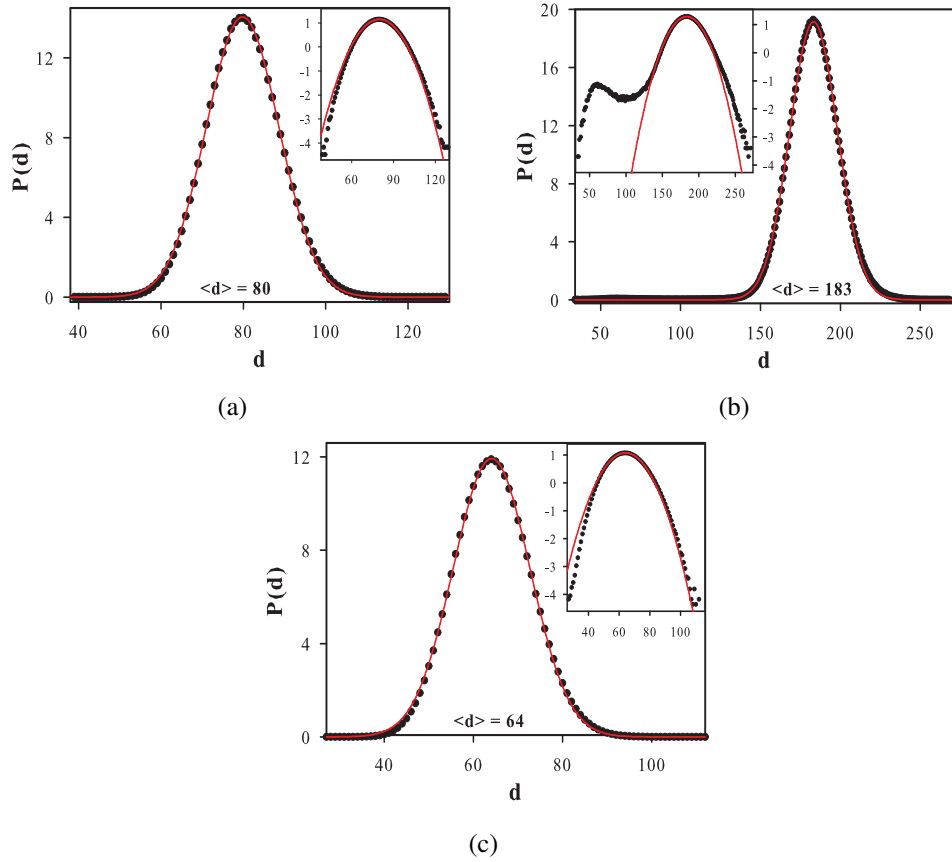
popülasyon sayısının ve çeşitliliğinin dinamik olarak değişimi verilmiştir. q-ESS periyodunda, topluluğa uyum sağlayan ve mutasyonlara karşı dayanıklı olan az sayıda tür bulunmaktadır. Türler uygunluk peyzajında, yaşamlarını uzun süre devam ettirebilecekleri uygun konumlarda oldukları için üreme olasılıkları yüksektir. Bu nedenle topluluğun toplam popülasyon sayısı yüksektir. q-ESS periyodunun sonunda toplam popülasyona bakıldığında, kitlesel bir yok olma olayının gerçekleştiği ve sistemin hektik periyoda girdiği görülmektedir. Bu kitlesel yok olmanın sebebi, toplulukta çok sayıda bireye sahip olan türlerin yok olmasıdır. Çünkü bu türler yok olduğunda, bu türün uygunluk peyzajında desteklediği türler de yok olacaktır. Evrimsel olarak aktif olan hektik periyotta toplam popülasyon sayısı azalırken, çeşitlilikte artma gözlenmektedir. Çeşitlilikteki bu artışın sebebi, hektik periyotta kısa yaşam sürelerine sahip çok sayıda türün ortaya çıkmasıdır. Ancak bu türlerin hiçbirinin, uygunluk peyzajında yaşamlarını uzun süre devam ettirebilecekleri uygun konumlar bulamamasından dolayı az sayıda bireylere sahip olacaklardır. Türlerin az sayıda bireylere sahip olmalarından dolayı, gerçekleşen mutasyon sonucu bu konumda hiç bir birey kalmayacaktır. Sistem uygunluk peyzajında gezinip duracaktır. Bu nedenle toplam popülasyon sayısında azalma meydana gelecektir. Hektik periyodun sonunda bu türlerin yok olması ile çeşitlilik azalacak ve toplam popülasyonda hızlı bir artış meydana gelecek şekilde sistem genotip uzayında başka bir q-ESS durumu bulacaktır. Şekil 4.2(b)'de toplulukdaki çeşitliliğin simülasyon ortalaması, 187.087477 ± 46.340815 dir. Ortalama çeşitlilik değerindeki standart sapmanın büyük olmasının nedeni, simülasyonun başlangıcında evrimsel aktivitenin çok yüksek olmasından dolayı, sistemin ilk kez kararlı bir yapıya ulaşmasına kadarki geçen bu sürede çeşitliliğin çok yüksek olmasıdır.

4.1.2. Dinamik Ağ Yapısı

DDM'nde genotip uzayındaki konumlar ister dolu ister boş olsun bütün konumlar arasındaki etkileşimler, simülasyonun başlangıcında $\theta = 0.25$ ihtimaliyeti ile **J** etkileşme matrisi rastgele oluşturulmuş ve simülasyon boyunca sabit tutulmuştur. Ancak topluluğun ağ yapısını oluşturmak için, genotip uzayında sadece dolu olan konumlar yani yaşayan türler arasındaki etkileşimlerin alınması gerekmektedir. Böylelikle, biyolojik toplulukdaki yaşayan türlerin köşeler ve yaşayan türler arasındaki doğrudan

etkileşmelerin kenarlar olarak düşünülmesi ile topluluğun ağ yapısı kurulabilir. Eğer verilen bir zamanda toplulukta yaşayan türler arasında bir etkileşme varsa, ağ sistemini kurmamızı sağlayan **A** etkileşme matrisinin elemanları 1 yoksa 0 olarak alınmaktadır. Bu şekilde kurulan ağ temel olarak yönelimsiz ve ağırlıksız bir ağ olacaktır. **A** etkileşme matrisinin boyutu ise, toplulukta yaşayan farklı türlerin toplam sayısı olarak tanımlanan çeşitliliğin değerine eşittir. Eş-evrimleşen türlerin ağ yapılarının kurulması sırasında önemli olan hususlar, doğal seçilim etkisi ile bazı türlerin nesli tükeneyeceği için dinamik olarak topluluktan kaldırılması ve mutasyon neticesinde bazı yeni türlerin topluluğa katılmasıdır. Bu hususlar, ağın etkileşme matrisinin boyutunun ve elemanlarının zamanla değişmesine neden olacaktır.

Sonuç olarak biyolojik toplulukların evrimsel süreci, zamana bağlı dinamik ağ yapısının oluşmasına neden olacaktır. Bu kesimde, DDM ile yapılan tek bir simülasyondan elde edilen ve Şekil 4.1’de gösterilen topluluğun dinamik ağ parametreleri sonuçları verilecektir.

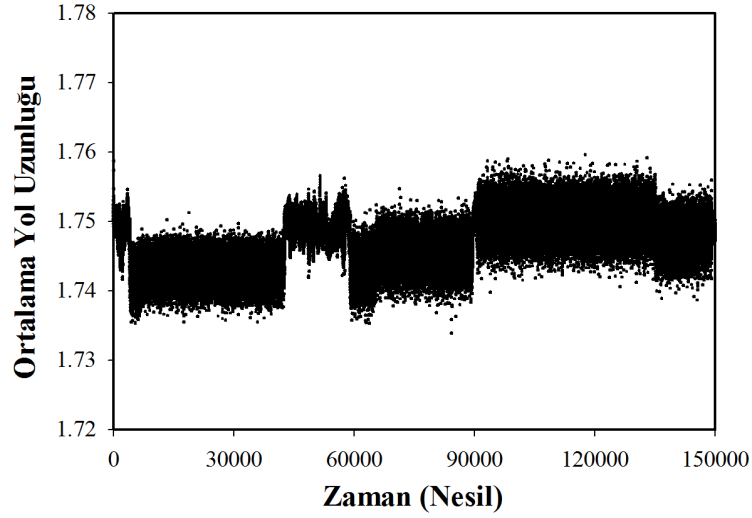


Şekil 4.3. Şekil 4.1’deki simülasyon için köşelerin derece dağılımları. Kırmızı çizgiler 3 parametrelili Gaussian fit eğrileridir.

Şekil 4.3’de, Şekil 4.1’deki simülasyonun 3 farklı zaman periyotlarındaki derece dağılımları verilmiştir. Şekil 4.3(a)’daki veriler simülasyonun ilk q-ESS periyodundan (4000 - 35000 nesilsel zaman aralığından), Şekil 4.3(b)’deki veriler hektik periyodundan (50000 - 55000 nesilsel zaman aralığından) ve Şekil 4.3(c)’deki veriler ise son q-ESS periyodundan (100000 - 145000 nesilsel zaman aralığından) elde edilmiştir. Her üç dağılımda da verilerin üzerine, kırmızı çizgi ile gösterilen 3 parametrelili Gaussian eğrileri fit edilmiştir. Şekillerin içerisine, her dağılımın yarı logaritmik şekilleri yerleştirilmiştir. Yerleştirilen bu şekillere bakıldığında, düşük **P(d)** değerlerinde Gaussian eğrilerinden sapmalar olduğu görülmektedir. Şekil 4.2’de ilk ve son q-ESS’ler deki çeşitlilik değerlerinin birbirine benzer olduğu görülmektedir. Bu sebeple bu periyotların Şekil 4.3(a) ve Şekil 4.3(c)’deki derece dağılımlarındaki **<d>** değerleri arasındaki fark, Şekil 4.2(b)’de son q-ESS’deki çeşitlilik değerinin ilk q-ESS’deki çeşitlilik değerinden küçük olmasından kaynaklanabilir. Ancak Şekil 4.2’de hektik periyot boyunca çeşitlilik değerinde gerçekleşen nispeten büyük artıştan dolayı Şekil 4.3(b)’deki derece dağılımındaki **<d>** ve **P(d)** değerleri, Şekil 4.3(a) ve Şekil 4.3(c)’ye göre daha büyük değerlere sahip olmaktadır. Şekil 4.3(b)’deki dağılımda ikili Gaussian gibi görünen yapının nedeni, hektik periyoda daha ayrıntılı çözünürlükte bakıldığında bu periyodun içinde bazı küçük q-ESS periyotlarının bulunmasıdır.

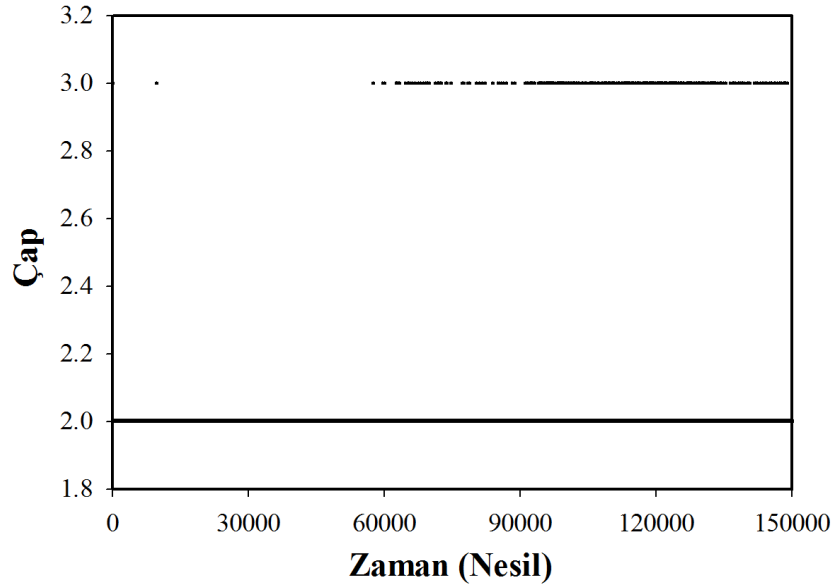
Biyolojik topluluklarda eş-evrimleşme sürecinde türler arasındaki ortalama yol uzunluklarının kısa olması beklenmektedir. Yapılan simülasyonda, ortalama yol uzunluğu ve çap hesapları yapılırken kendisine olan mesafelerin sonsuz olacağı izole türler hiç bir zaman gözlenmemiştir. Herbir nesilde ilk önce, toplulukta bulunan bütün canlı türlerin arasındaki en kısa yol uzunlukları bulunmuştur. Daha sonra bu yol uzunluklarının topluluk ortalaması alınarak ortalama yol uzunluğu hesaplanmıştır.

Şekil 4.4’de, Şekil 4.1’deki simülasyon için biyolojik toplulukların ortalama yol uzunluklarının zaman göre değişimi verilmiştir. Ortalama yol uzunluğu simülasyon boyunca 2’den küçük ve hemen hemen sabittir. Değerler q-ESS - hektik ve hektik - q-ESS geçişlerinden sonra artma eğilimi göstermektedir. Ortalama yol uzunluğunun simülasyon ortalaması değeri 1.746839 ± 0.003683 dür. Tablo 1.5.’e bakıldığında DDM ile elde edilen ortalama yol uzunluğu ifadesinin, rasgele ağlardaki değerler kadar düşük olduğu görülmektedir.



Şekil 4.4. Ortalama yol uzunluğunun zamanla değişimi.

Ortalama yol uzunluğu için bir sınır belirleyen ϕ ifadesi, herbir nesilde birbirine en uzak iki tür arasındaki en kısa mesafe alınarak hesaplanmış ve zamanla değişimi ise Şekil 4.5’de verilmiştir.

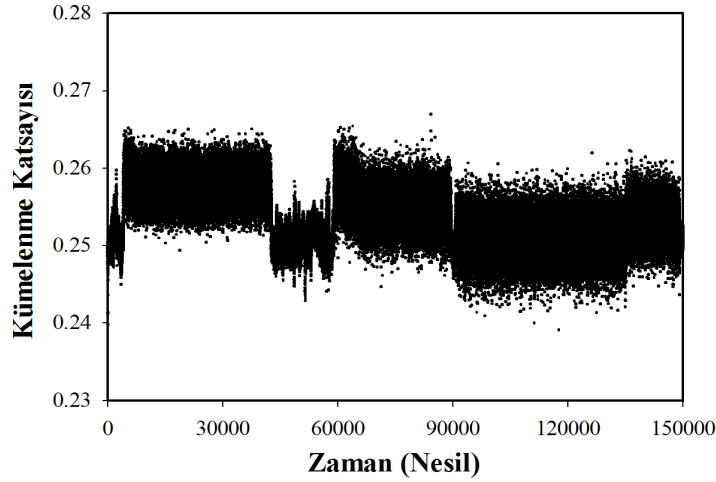


Şekil 4.5. Çapın zamanla değişimi.

ϕ ifadesi simülasyonun başlangıcında genellikle 2’ye eşittir. Ancak daha sonraki zamanlarda 2 yada 3 değerlerinden birine eşit olmaktadır. Çapın sahip olduğu bu değerler, ortalama yol uzunluğunun 2’den küçük olmasının nedenini açıklamaktadır.

Kümelenme katsayısı, yönelimli [109], yönelimsiz [108, 109] ve izole köşelerin çıkarılışı [108] ağlarda, yönelimli ve yönelimsiz ağları uygun şekilde birleştiren formüllerle

hesaplanmaktadır. Bu tez çalışmasında kümelenme katsayısı, DDM ile her üç ağ formülasyonu için de hesaplanmış ve aynı sonuçlar bulunmuştur. Sonuçların aynı çıkmasının sebebi, yapılan simülasyonda topluluğun yoğun bir şekilde etkileşmesidir. Bu nedenle, toplulukta izole yani kimseyle etkileşmeyen türler bulunmamaktadır.



Şekil 4.6. Ortalama kümelenme katsayısının zamanla değişimi.

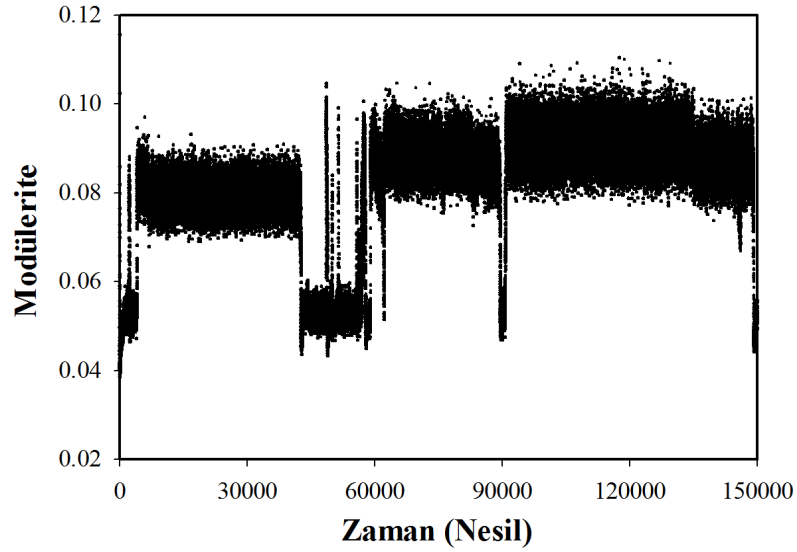
Şekil 4.6’da, Şekil 4.1’deki simülasyon için her nesilde biyolojik topluluklarda bulunan türlerin sahip oldukları kümelenme katsayılarının topluluk ortalamasının zaman göre değişimi verilmiştir. Kümelenme katsayısının simülasyon ortalaması değeri 0.253559 ± 0.003659 dur. Tablo 1.6.’ya bakıldığında DDM ile elde edilen bu değer, rasgele ağlardaki değerlerden oldukça büyük olduğu görülmektedir.

Yapılan simülasyonda, Tablo 1.7.’deki değerler ışığında DDM’nin davranışı küçük dünya ağı olarak kabul edilebilir. Çünkü küçük dünya ağlarında, ortalama yol uzunluğu rastgele ağlar kadar düşük iken kümelenme katsayısı düzenli ağlar kadar büyüktür.

Modülerite değerini hesaplamak için Blondel et al. [113] tarafından geliştirilen hızlı ve etkili optimizasyon metodu kullanılmıştır.

Şekil 4.7’de, Şekil 4.1’deki simülasyon için her nesilde hesaplanan modülerite değerinin zamana göre değişimi verilmiştir. Modüleritenin simülasyon ortalaması değeri 0.081274 ± 0.013016 dır. DDM ile elde edilen bu modülerite değerinin tartışılması için Tablo 4.1’de bazı gerçek dünya ağlarındaki değerler verilmiştir.

Tablo 4.1’e bakıldığında, gerçek dünya ağlarındaki modülerite değerlerinin DDM ile



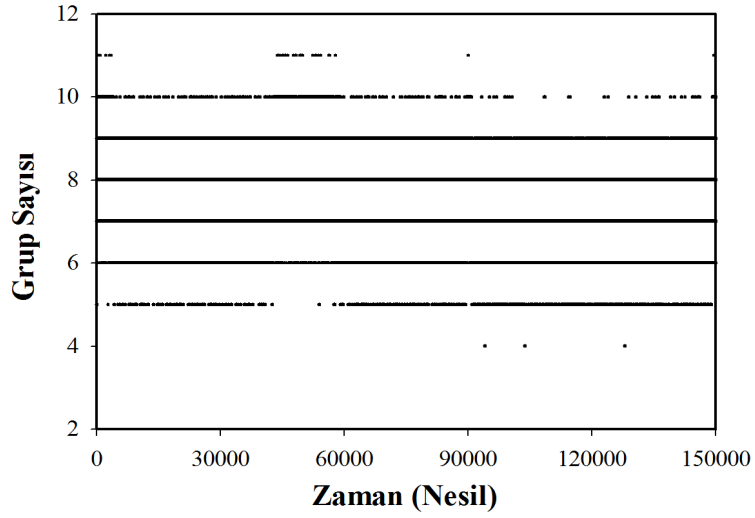
Şekil 4.7. Modüleritenin zamanla değişimi.

Tablo 4.1. Bazı gerçek dünya ağlarının modülerite değerleri.

Ağ	Modülerite (Q)		
	Ref. [113]	Ref. [114]	Ref. [117]
Karate	0.381	0.42	0.419
Arşiv	0.772	0.813	
Jaz müzisyenler	0.439		0.442
Metabolik	0.402		0.435
E-mail	0.494		0.572
İnternet	0.692	0.781	
Fizikçi	0.668		0.723

bulunan değerden oldukça yüksek olduğu görülmektedir. DDM'nin düşük modülerite değerine sahip olmasının nedenlerinden bir tanesi, toplulukta bulunan gruplar arasındaki etkileşmelerin grup içi etkileşmelere yakın büyüklükte olmasıdır. Bir başka neden ise, DDM'nde bitki gibi diğer türlerin avı olan başlangıç üreticilerinin düşünülmemesidir. Oysa topluluk esas olarak bu başlangıç üreticileri civarında gruplanabilir [152]. Ancak modülerite değeri düşük değere sahip olsa bile, DDM'nde topluluk hala az da olsa gruplara ayrılabilir.

Şekil 4.8'de, Şekil 4.1'deki simülasyon için her nesilde hesaplanan modülerite neticesinde toplulukta oluşan grup sayılarının zamana göre değişimi verilmiştir. Hektik periyottaki grup sayılarının q-ESS periyotlarındaki grup sayılarından daha çok olduğu görülmektedir.



Şekil 4.8. Modülerite neticesinde toplulukda oluşan grupların sayısının zamanla değişimi.

4.1.3. Soy Ağacı

Soy ağacı, evrimsel süreç neticesinde biyolojik türler arasındaki ilişkilerin anlaşılması için çok fazla bilgi barındırır. Evrimsel sürecin gerçekleşmesi için çok uzun zaman gerektiğinden, soy ağacı çıkarımında fosil kayıtlarından faydalanılmaktadır. Ancak fosil kayıtlarındaki eksiklikler doğru soy ağacı çıkarımında problem teşkil etmektedir. Bu yüzden üreme, mutasyon ve doğal seçim gibi temel unsurların bulunduğu evrimsel süreci bilgisayar ortamına taşıyarak, türler arasındaki evrimsel ilişkilerin elde edilmesi ve soy ağacı çıkarımının yapılabilmesi için simülasyon yapmak oldukça kullanışlıdır. Bu kesimde Şekil 4.1'deki simülasyonun başlangıcından ilk q-ESS periyodunun sonuna kadar (0 - 40000 nesil) geçen zamanda gerçekleşen mutasyonlar incelenerek gerçek soy ağacı analizi yapılacaktır. Elde edilen soy ağacı, Bölüm 2.1 ve 2.2'de anlatılan ML ve NJ metotları ile elde edilen soy ağaçları ile karşılaştırılarak bu metotların doğrulukları test edilecektir.

Genotip uzayındaki türlerin konumları, sahip oldukları ikili genom dizilerinin onluk tabandaki karşılığı olan sayılarla etiketlenmektedir. Soy ağacında, bütün türlerin ortak bir atası olması gerektiği düşüncesi nedeni ile Şekil 4.1'deki simülasyonun başlangıcında, genotip uzayından rasgele seçilen $N(t = 0) = 6931$ tane bireye sahip 712889241 numaralı tek bir tür bulunmaktadır. Bu türün genom dizilimi

1 0 1 0 1 0 0 1 1 1 1 1 0 1 1 1 0 1 0 0 1 1 1 0 0 1 1 0 0 1

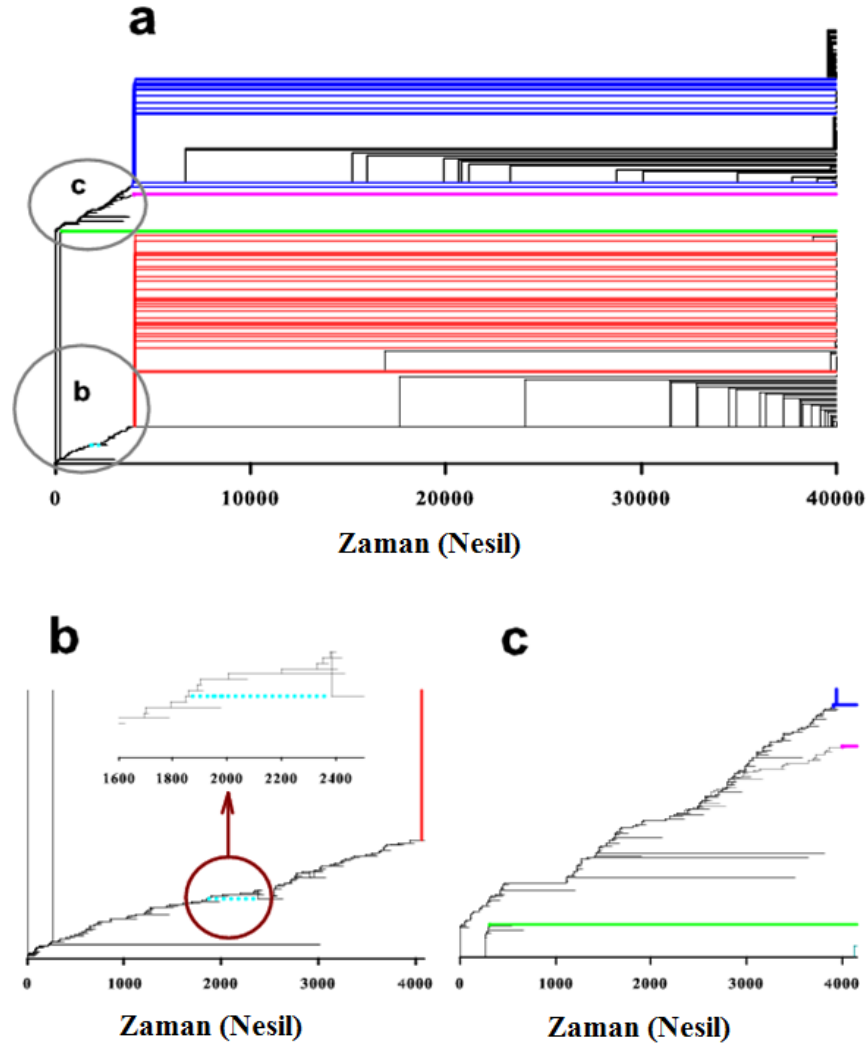
şeklindedir. 83 nesil yaşayan başlangıç türünün ilk neslin sonunda komşu konumları olan, 779998105 ve 712823705 numaralı konumlara mutasyon vermesi ile ilk çeşitlenme gerçekleşmekte ve toplulukta iki yeni tür gözlenmektedir. 779998105 ve 712823705 numaralı yeni türler sırasıyla 99 ve 103 nesil yaşamalarına rağmen soy ağacında, sonraki nesillerde özellikle ilk q-ESS boyunca uzun süre yaşayacak olan türlerin ataları olarak yerlerini alacaklardır. Bu türlerin genom dizileri ise

779998105	1 0 1 1 1 0 0 1 1 1 1 1 0 1 1 1 0 1 0 0 1 1 1 0 0 1 1 0 0 1
-----------	---

712823705	1 0 1 0 1 0 0 1 1 1 1 1 0 0 1 1 1 0 1 0 0 1 1 1 0 0 1 1 0 0 1
-----------	---

şeklindedir. Başlangıç türünden meydana gelen bu türlerde mutasyona uğrayan genler koyu renkle gösterilmiştir. Simülasyonun sonunda yani 40000. nesilde toplulukta 187 tür yaşamaktadır. Ancak bu türlerin sadece 45 tanesi ilk q-ESS periyodu boyunca (35000 nesilden fazla) yaşamaktadır. Diğer türler, genel olarak ilk q-ESS boyunca bu 45 türde gerçekleşen mutasyonlar neticesinde meydana gelmektedir.

Şekil 4.9 a'da ilk q-ESS'in sonunda yaşayan 187 türün soy ağacı verilmiştir. Ağacın dört ana dalı bulunmaktadır. Kırmızı ve yeşil renkle gösterilen çizgiler 779998105 numaralı türden meydana gelen türleri gösterirken, pembe ve mavi renkle gösterilen çizgiler ise 712823705 numaralı türden meydana gelen türleri göstermektedir. Bu renkli çizgilerle gösterilen ağaç dalları, ilk q-ESS periyodu boyunca yaşayan türlerin sahip olduğu dalları göstermektedir. q-ESS periyodunda bu türlerden meydana gelen türlerin sahip oldukları dallar siyah çizgilerle gösterilmiştir. Siyah renkli dallarla gösterilen bu türler q-ESS periyodundaki 45 ana türe göre kısa yaşamalarına rağmen 40000. nesilde varlıklarını sürdürmektedirler. Ancak yapılacak hesaplamalarda bu türler hesaba dahil edilmemiştir. Ayrıca q-ESS periyodu boyunca kısa süreli yaşayan yani q-ESS periyodunda ortaya çıkmış ve q-ESS periyodunda ölmüş olan türler, 40000. nesilde yaşayan türlerin ataları konumunda değilse ağaçta gösterilmemiştir. Çünkü bu türler genellikle q-ESS periyodu boyunca yaşayan 45 ana türden mutasyonla ortaya çıkmakta ve çevrelerine uyum sağlayıp yaşamlarını sürdürememektedir. Bu sebeple çalışmanın odak noktası, ilk q-ESS periyodu



Şekil 4.9. **a** İlk q-ESS periyodunun sonunda yaşayan 187 türün gerçek soy ağacı gösterimi. Düşey eksen 712889241 numaralı başlangıç türünden meydana gelen türlerin numaralarını göstermektedir. Ancak bu numaralar keyfi olarak yerleştirilmiş olup, artan yada azalan bir sıra konusu değildir. İlk q-ESS periyodu boyunca yaşayan 45 türün ağaçta sahip olduğu dallar dört farklı renk (kırmızı, yeşil, pembe ve mavi) ile gösterilmiştir. **b** ve **c** Soy ağacında ilk türleşmelerin gerçekleştiği dönemin büyütülmüş halleri.

boyunca yaşayan 45 ana tür olacaktır. Anlatım kolaylığı için, Şekil 4.9 **a**'daki bu türler, en alttaki kırmızı daldan en üstteki mavi dala doğru S1'den S45'e kadar numaralandırılmıştır. Bu türlerin genom dizileri ise Tablo 4.2'de verilmiştir.

Her ne kadar kırmızı ve yeşil türler aynı ortak atadan meydana gelmiş olsada Tablo 4.2'ye bakıldığında, yeşil türün genom dizisinin soy oluşsal olarak pembe ve mavi renkli türlere daha yakın olduğu görülmektedir. Türlerin birbirine olan yakınlıkları yada uzaklıkları,

Tablo 4.2. Şekil 4.9 a’da verilen soy ağacındaki 45 ana türün genom dizileri.

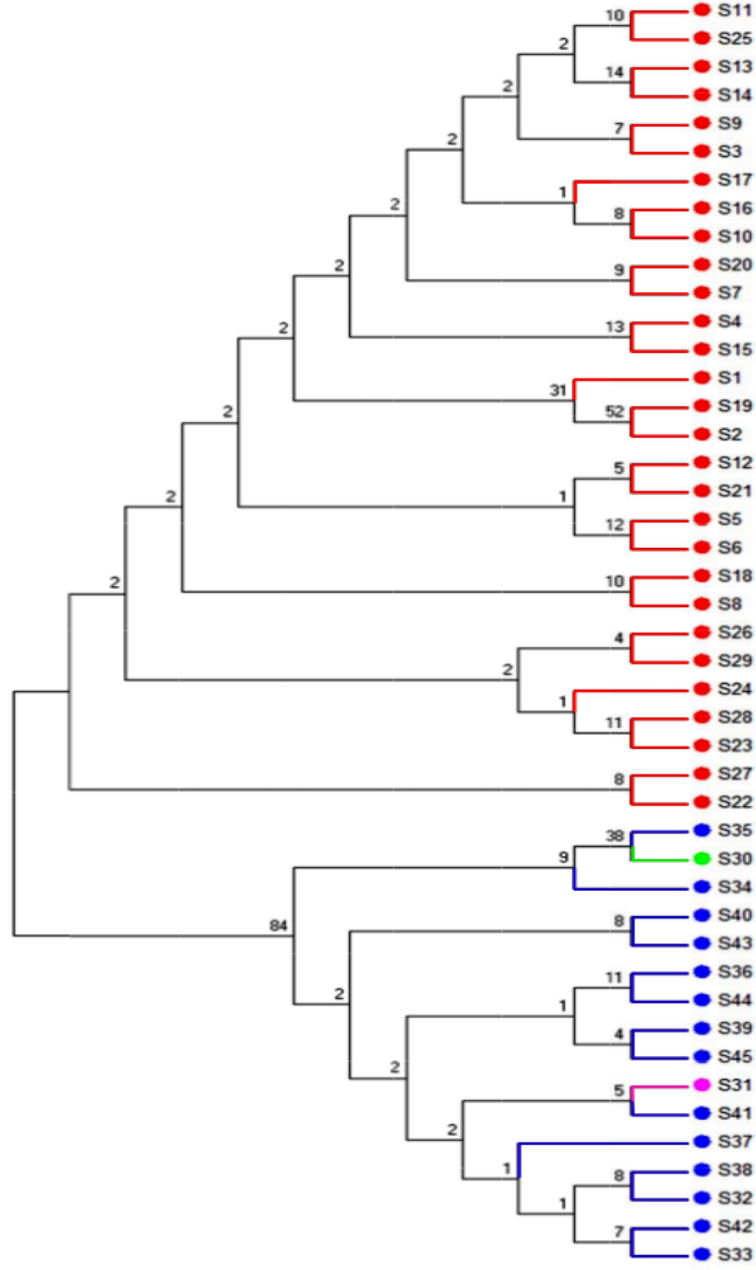
S1	203407359	001100000111111011111111111111
S2	203341823	001100000111101011111111111111
S3	203407351	001100000111111011111111111011
S4	203399159	001100000111111001111111111011
S5	203276279	001100000111011011111111111011
S6	203374583	001100000111110011111111111011
S7	203405303	001100000111111011011111111011
S8	203407223	00110000011111101111111011011
S9	203407335	001100000111111011111111100111
S10	203407319	001100000111111011111111101011
S11	203423735	001100000111111111111111110111
S12	203145207	001100000110111011111111111011
S13	211795959	001100100111111011111111111011
S14	740278263	101100000111111011111111111011
S15	203406327	001100000111111011101111111011
S16	203407287	00110000011111101111111011011
S17	471842807	011100000111111011111111111011
S18	202358775	001100000011111011111111111011
S19	203341815	001100000111101011111111111011
S20	203403255	001100000111111010111111111011
S21	203406839	001100000111111011111011111011
S22	220184567	001101000111111011111111111011
S23	236961783	001110000111111011111111111011
S24	205504503	001100001111111011111111111011
S25	207601655	001100010111111011111111111011
S26	69189623	000100000111111011111111111011
S27	202883063	001100000101111011111111111011
S28	136298487	001000000111111011111111111011
S29	203407095	001100000111111011111011111011
S30	1073741823	111111111111111111111111111111
S31	192413695	001011011101111111111111111111
S32	459800575	011011011001111111111111111111
S33	460849151	011011011101111111111111111111
S34	527958015	011111011101111111111111111111
S35	997720063	111011011101111111111111111111
S36	460783615	011011011101101111111111111111
S37	456654847	011011001101111111111111111111
S38	326631423	010011011101111111111111111111
S39	460832767	011011011101111011111111111111
S40	460845055	011011011101111110111111111111
S41	460840959	011011011101111110111111111111
S42	460848127	011011011101111111101111111111
S43	460847103	011011011101111111101111111111
S44	460849023	011011011101111111111110111111
S45	460849143	011011011101111111111111111011

genom dizileri arasındaki Hamming mesafesi ile belirlenmektedir. Hamming mesafesi küçük olan türler birbirine daha yakındır.

Simülasyonun başlangıcında ilk q-ESS yapısı kurulana kadar, meydana gelme ve yok

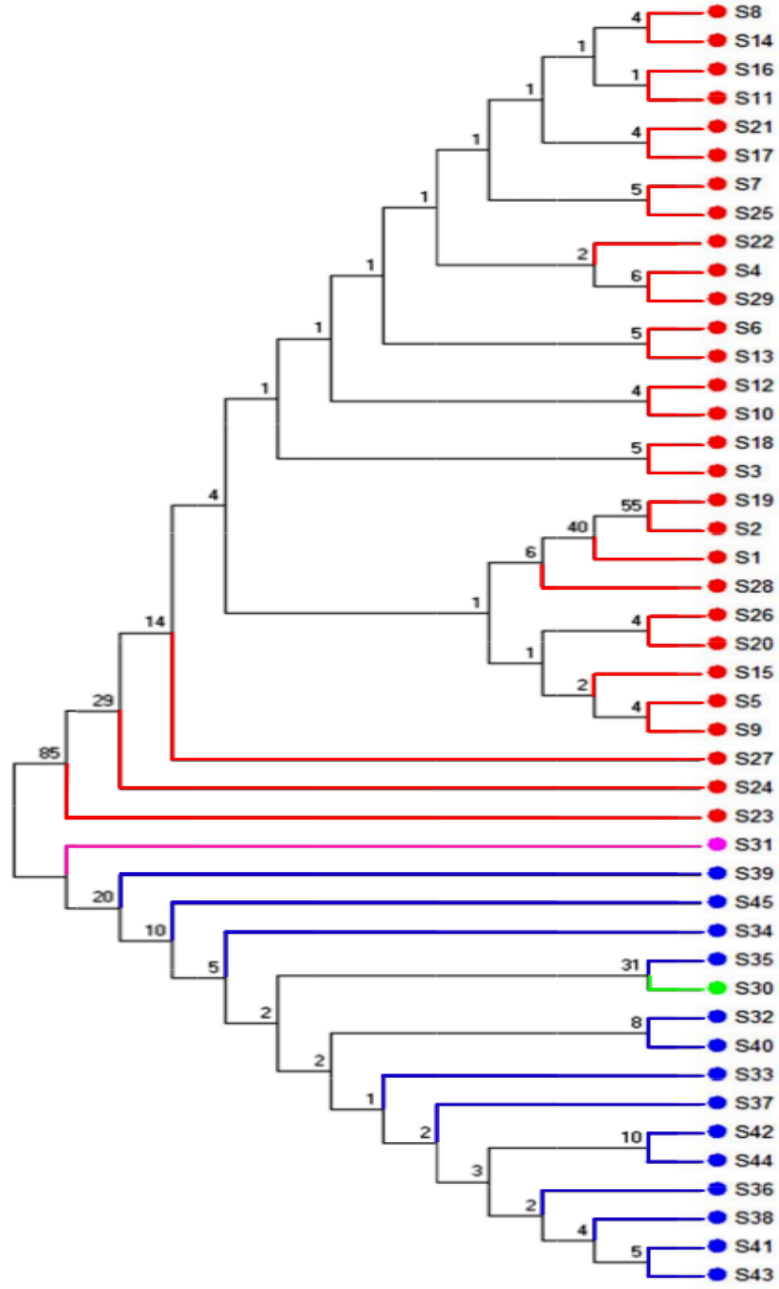
olma oranlarının yüksek olduğu yüksek seviyede evrimsel aktivite ile karşılaşmıştır. Şekil 4.9 **b** ve **c**'de görüldüğü gibi, eş-evrimleşmenin gerçekleştiği bu kısa hektik periyodunda, ortaya çıkıp kaybolan birçok geçiş türü bulunmaktadır. Bu geçiş türleri kısa siyah çizgilerle gösterilmiştir. Bu türlerin çoğu, sadece 40000. nesilde yaşayan türleri başlangıç türüne ulaştıracak olan ara türlerdir. DDM'nin bir dez avantajı Şekil 4.9 **b**'de turkuaz rekli kesikli çizgi ile gösterilen ters mutasyonlardır. Ters mutasyon, bir türden mutasyon neticesinde oluşan mutant türün tekrar orjinal türe mutasyon vermesidir. Bu durumun nedeni, DDM'nde çalışılan genom uzayının yeteri kadar büyük olmamasıdır. Tutarlı bir amaç için, DDM ile yapılan soy ağacı çalışmasında ters mutasyonlar ihmal edilmiş ve bütün türler her ne zaman meydana gelirlerse gelsinler ilk ataları hangisi ise ona bağlanmıştır.

Soy ağacı tahmininde geniş çapta kullanılan ML metodunun genel anlatımı Bölüm 2.1'de verilmiştir. ML metodunun en yüksek olabilirliğe sahip soy ağacını vermesi avantaj olarak görülebilir. Ancak çok sayıda DNA dizisi için soy ağacı tahmini yaparken çok uzun hesaplama süresi gerektirmektedir. Çünkü ML metodu soy ağacı tahmini yaparken, olası bütün ağaç topolojileri üzerinden arama yapmaktadır. Bir soy ağacı tahmini yapılacağı zaman, sıklıkla birden çok metot kullanılır ve sonuçlar karşılaştırılır. Soy ağacı tahmininde sıklıkla kullanılan bir diğer metot olan NJ metodunun genel anlatımı da Bölüm 2.2'de yapılmıştır. Bazı durumlarda NJ metodunun verdiği yerel topolojiler, ML metodundan daha iyi olmaktadır. Ancak alternatif topolojilerin test edilmemesi NJ metodunun önemli bir eksikliğidir. Bu çalışmada, Şekil 4.9 **a**'daki 45 ana türün, hem ML metodu hemde NJ metodu ile soy ağacı çıkarımını yapmak için MEGA 5 programı [147] kullanılmıştır. MEGA 5 programı bu metotlar için soy ağacı çıkarımı yaparken, Kimura 2-parametre modelini [153] kullanmaktadır. Bu model, DNA dizilimlerindeki pürin (A,G) ve pirimidin (C,T) türevleri olan bazlarda gerçekleşebilecek mutasyonların, kendi gruplarına (pürin - pürin yada pirimidin - pirimidin) mı yoksa başka gruba (pürin - pirimidin yada pirimidin - pürin) mı olacağını belirlemektedir. MEGA 5 programında hem ML hemde NJ metotları ile için ayrı ayrı 100 farklı ağaç topolojisi üretilmiş ve sonuçta bu ağaçlardan, dallanma noktalarında bulunan ara türlere göre herbir metot için ayrı ayrı ortak bir soy ağacı elde edilmiştir. Şekil 4.9 **a**'daki 45 ana tür için, Şekil 4.10 ve 4.11'de sırası ile ML ve NJ metotları ile elde edilen soy ağaçları verilmiştir.



Şekil 4.10. Şekil 4.9 a'daki 45 ana tür için ML metodunun tahmin ettiği soy ağacı. Dallanma noktalarında verilen sayılar, ilgili dallanma noktalarının üretilen 100 farklı ağaçtan kaç tanesinde aynı noktada olduğunu belirtmektedir. Türler, Şekil 4.9 a'daki gerçek soy ağacında sahip oldukları renklere göre dairelerle gösterilmiştir.

Şekil 4.10 ve 4.11'de yatay eksen genetik değişim miktarına karşılık gelmektedir. Başka bir ifade ile, herbir dallanma noktalarında yeni genetik karakterler meydana gelmektedir. Yatay eksen meydana gelen bu yeni karakterlerin miktarını belirtmektedir. Her iki şekilde de dikey eksenler herhangi bir fiziksel anlam içermez. Sadece türlerin düzenli bir sırada



Şekil 4.11. Şekil 4.9 a'daki 45 ana tür için, NJ metodunun tahmin ettiği soy ağacı. Dallanma noktalarında verilen sayılar, ilgili dallanma noktalarının üretilen 100 farklı ağaçtan kaç tanesinde aynı noktada olduğunu belirtmektedir. Türler, Şekil 4.9 a'daki gerçek soy ağacında sahip oldukları renklere göre dairelerle gösterilmiştir.

ayrılması için kullanılır. Her iki şekil de, dallanma noktalarında bulunan ve S1, S2, ..., S45 ile gösterilen türlerin ata türleri olan ara türler hakkında herhangi bir bilgi içermemektedir. Bu dallanma noktaları üzerindeki sayılar ise, ilgili dallanma noktalarının her iki metot için ayrı ayrı üretilen üretilen 100 farklı ağaçtan kaç tanesinde aynı noktada olduğunu

belirtmektedir. Her iki şekilde de bu sayılardan çoğunun küçük olmasının nedeni, türlerin genom dizilerinin birbirine çok yakın olmasıdır. Örneğin her iki şekilde de S1, S2 ve S19 türlerinin dallanma noktalarındaki sayılar büyük değere sahiptir. Bu türlerin Hamming mesafeleri S1 ile S2 arasında 1 ve S2 ile S19 arasında 1 dir. Diğer taraftan S1 türünün S2 ve S3 türü dışında diğer bütün kırmızı renkle gösterilen türlerle arasındaki Hamming mesafesi 2 dir. Bu sonuçlara rağmen S1 ve S3 türü aynı dallanma noktasına sahip değildir. Bunun nedeni, S3 türünün S2 türü dışında diğer bütün kırmızı renkle gösterilen türlerle arasındaki Hamming mesafesinin 1 olmasıdır. Sonuç olarak, her iki şekilde de S3 türünün sahip olduğu dallanma noktalarındaki değerin küçük olmasının nedeni, S3 türü ile diğer türlerin genetik olarak çok yakın olmasıdır. Benzer durum S30, S34 ve S35 türleri için de geçerlidir.

ML ve NJ metotları ile elde edilen soy ağaçları, Şekil 4.9 a'daki gerçek ağaç ile karşılaştırıldığında, S30 türünün her iki ağaçta da yanlış bağlandığı, ancak S31 türünün NJ metodu ile elde edilen ağaçta ML metodu ile elde edilen ağaca göre daha iyi bağlandığı görülmektedir. Yine de, ML ve NJ metodu ile elde edilen ağaçların hangisinin, gerçek ağaca daha yakın olduğunu belirlemek için nümerik analiz yapmak faydalı olacaktır. Bu sebeple Puigbo et al. [148] tarafından geliştirilen TOPD/FMTS (TOPological Distance/From Multiple To Single) yazılımı kullanılmıştır. TOPD/FMTS yazılımı, ağaçların farklı özelliklerini karşılaştıran farklı metotlar içermektedir. Bu yazılım sayesinde köksüz ağaçların sahip oldukları özelliklerin değerleri karşılaştırılmakta ve örtüşen türlerin yüzdesi elde edilmektedir. Bu sayede Şekil 4.9 a'daki DDM ile elde edilen gerçek ağaç ML ve NJ metotlarının tahmin ettikleri ağaçlar ile karşılaştırılmış ve sonuçlar Tablo 4.3'de verilmiştir.

Tablo 4.3. TOPD/FMTS yazılımı ile DDM ile elde edilen soy ağacının, ML ve NJ metotlarının tahmin ettikleri ağaçlar ile karşılaştırılması.

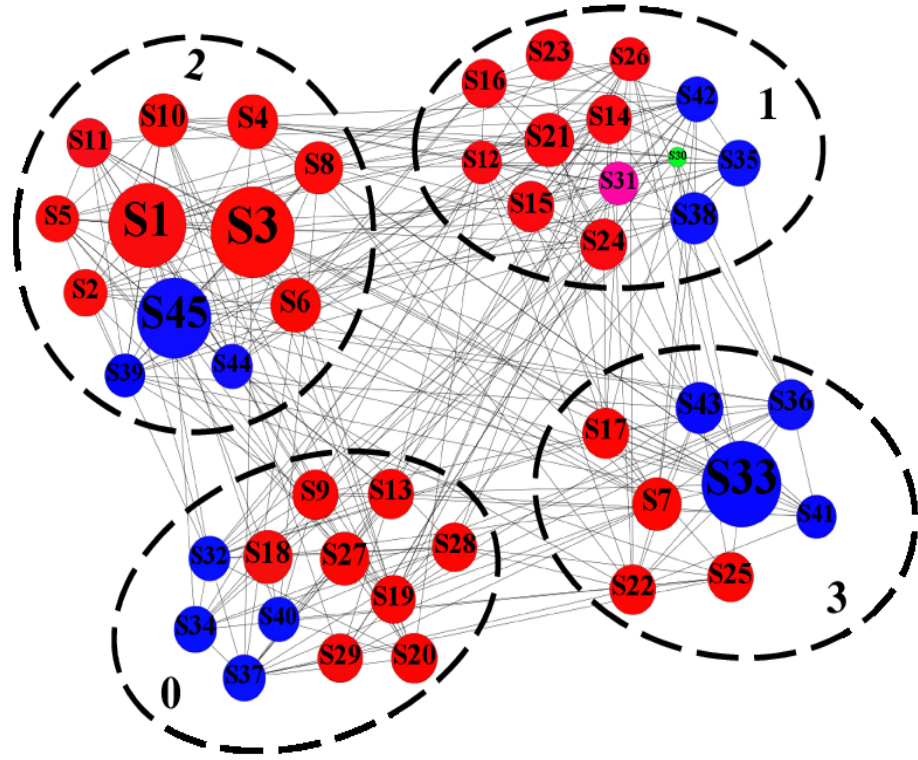
Özellik	DDM - ML	DDM - NJ
Türlerin örtüşme yüzdesi	% 95.6	% 100.0
Düğüm Mesafesi (budanmış/budanmamış)	(5.89115 $\pm$ 0.13434 / 6.15297 $\pm$ 0.14031)	(5.51975 / 5.51975)
Ayırma Farkı [farklar/olasılıklar]	0.99375 $\pm$ 0.011 [79.5 $\pm$ 0.866 / 80]	0.97619 [82 / 84]
Uyuşmazlık [uyuşmayan tür / bütün türler]	[38.0 $\pm$ 1.0 / 43]	[43 / 45]

Karşılaştırılan ağaçların özelliklerinden biri olan düğüm mesafesi, ağaçdaki türler arasında bulunan dallanma noktaları sayısına göre kurulan matris ile belirlenir. Karşılaştırılan ağaçların herbiri için ayrı ayrı kurulan matrislerin farklarının karesi alınarak bir mesafe matrisi elde edilir. Bu mesafe matrisinin değeri, karşılaştırılan ağaçlar arasındaki düğüm mesafesini belirler. Bu değer sıfıra yaklaştıkça, iki ağacın birbirine olan benzerlikleri artmaktadır. Ayrıca ağaçların örtüşen dallarında aynı tür bulunmazsa, her iki ağaçta da örtüşen bu dal budanmaktadır. DDM - NJ karşılaştırmasında bütün türler örtüşürken, DDM - ML karşılaştırmasında türlerin % 95.6'sı örtüşmektedir. Tablo 4.3'de DDM - ML ve DDM - NJ karşılaştırmaları için hem budanmış hemde budanmamış düğüm mesafesi değerleri verilmiştir. Bir diğer özellik ayırma farkıdır. Bu özellikte, ağaçların sahip oldukları dallar farklı noktalardan ayrılırlar. Olabilecek bütün ayrımlar yapıldıktan sonra, karşılaştırılan ağaçların bu ayrımlardan ne kadarında farklı olduğuna bakılır. Bu fark azaldıkça, karşılaştırılan ağaçların benzerlikleri artmaktadır. Son özellik uyumsuzluktur. Uyumsuzluk tekrarlamaların olduğu bir özelliktir. Karşılaştırılan ağaçlardan, her ikisinde de aynı olmak koşulu ile bir tür çıkarılarak ağaçların uyuşup uyuşmadığına bakılır. Bütün türler için bu süreç tekrar edilerek iki ağaç arasında aynı topoloji elde edilene kadar uygulanır. Ağaçların uyuşması ayırma farkındaki değişim ile belirlenir. Ayırma farkı sıfır olduğunda tekrarlama durur. Bu noktada ağaçların uyuşması için çıkarılan türlerin sayısına bakılır.

Sonuç olarak Tablo 4.3'e bakıldığında, DDM - NJ karşılaştırmasındaki türlerin uyuşma yüzdesi, düğüm mesafesi ve ayırma farkı değerlerinin, DDM - ML karşılaştırmasındaki değerlerden daha düşük bu yüzden daha iyi olduğu görülmektedir. Bu sonuç ile, NJ metodunun yaptığı soy ağacı tahminin, gerçek soy ağacına daha yakın bir sonuç olduğu görülmektedir.

Bu kesimde ayrıca, Şekil 4.9 a'daki 45 ana türün topluluk yapısının incelenmesi için, bu türlerden oluşan topluluğun yönelimsiz ve ağırlıksız bir ağ yapısı kurulmuş ve bu ağ sisteminin modüleritesi hesaplanarak topluluk gruplara ayrılmıştır.

Şekil 4.12'de, Şekil 4.9 a'daki 45 ana türün topluluk yapısı incelenmiş ve topluluğun dört gruptan oluştuğu görülmüştür. Şekil 4.12'de, ağın modülerite değeri 0.19 olarak bulunmuştur. Bu değer yüksek olmadığı için, topluluk yeterince açık olarak topluluklara



Şekil 4.12. Şekil 4.9 a'daki 45 ana türün topluluk yapısı. Toplulukta kesikli dairelerle gösterilen dört grup bulunmaktadır. Türler Şekil 4.9 a'da sahip oldukları renklere göre farklı büyüklükteki dairelerle gösterilmiştir. Dairelerin büyüklükleri, türlerin sahip oldukları $n_i(t)$ birey sayıları ile $50 \times \log(10 \times n_i(t))$ bağıntısına göre belirlenmiştir.

ayrılammaktadır. Çünkü gruplar arasında yoğun etkileşimler bulunmaktadır. Modülerite değerinin yüksek olması ve topluluğun net bir şekilde gruplara ayrılabilmesi için gruplar arasında grup içine göre daha az etkileşimler bulunmalıdır. Şekil 4.12'de türler arasındaki toplam bağlantı sayısı 275 dir. Fakat DDM gereği toplulukta bulunan türlerin 1/4'i etkileştiği için yapması gereken toplam bağlantı sayısı $495 (= \frac{45 \times 44}{4})$ olmalıdır. Bu sonuç, q-ESS periyodunun iskeletini oluşturan 45 ana türün birbirleri ile olan etkileşimlerinin düşük, q-ESS periyodundaki diğer türlerle olan etkileşimlerinin yüksek olduğunu göstermektedir. Şekil 4.12'den çıkarılacak bir diğer sonuç, grupların üye sayılarının yaklaşık olarak birbirine eşit olduğudur. Bütün bunların yanında bulunan grup sayısının, Tablo 4.4'de gösterildiği gibi gerçek dünyadaki yiyecek ağları için topluluklarda bulunan grup sayısı ile tutarlı olduğu görülmektedir.

Sonuç olarak soy ağacının yiyecek ağlarından çıkarılamayacağı sonucuna ulaşılmıştır. Bunun nedeni ise iki türün aynı ortak atadan meydana gelmesine rağmen, ekosistemde

Tablo 4.4. Gerçek dünyadaki yiyecek ağları için topluluklarda bulunan grup sayıları. Ref.[152]

Ağ	Tür sayısı	Grup sayısı
Little Rock gölü	181	4
Ythan nehri	92	3
St. Martin adası	44	5
Chesapeake koyu	33	3

varlıklarını sürdürebilmek için farklı fizyolojilere sahip olabilecek olmasıdır.

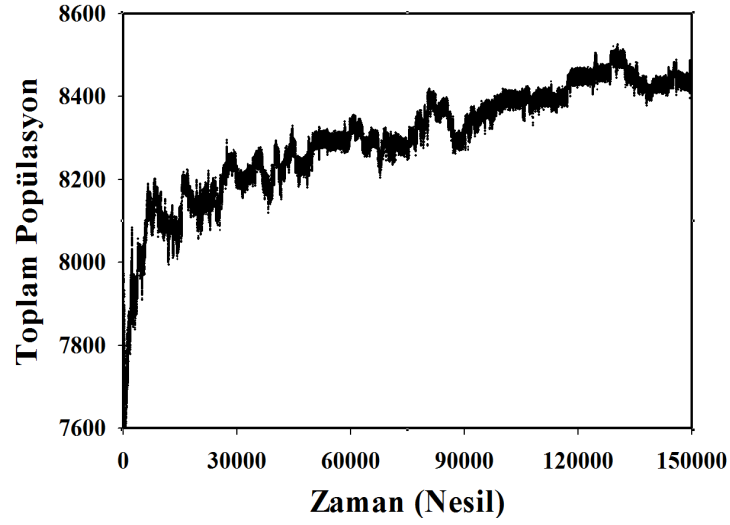
4.2. Çok Sayıda Simülasyonun Ortalama Sonuçları

İncelenen sistem için anlamlı sonuçlar almak ve sistemin genel davranışı hakkında bilgi sahibi olmak için çok sayıda simülasyon yapıp bu simülasyonların ortalamasını almamız gerekmektedir. Ancak artan simülasyon sayısı hesaplama sürelerinde artışa neden olacağı için bireylerin genom uzunlukları düşürülmüş yani $L = 20$ olarak alınmıştır. Herbir simülasyonda $\mathbf{J}$ etkileşme matrisi yeniden üretilmiştir. Bütün simülasyonlarda türlerin etkileşme oranı, tek bir simülasyonda olduğu gibi yine $\theta = 0.25$ olarak alınmıştır. Bu kesimde yapılacak herbir simülasyonun başlangıcında, $N(t = 0) = 500$ birey genotip uzayında rasgele seçilen konumlara yerleştirilmiştir.

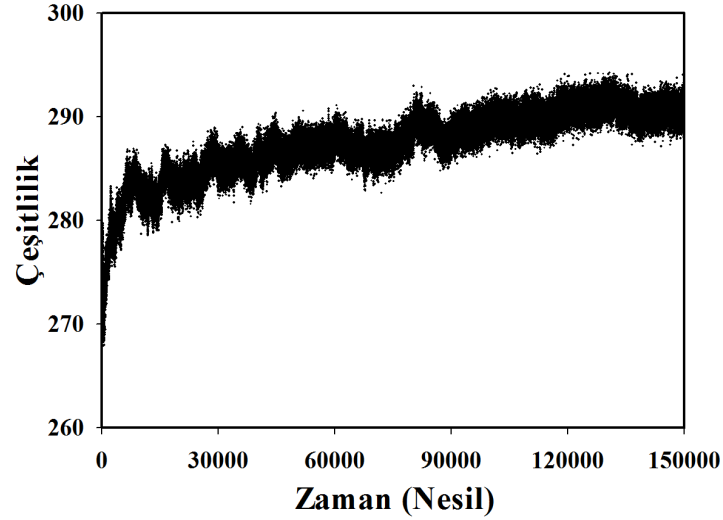
4.2.1. Toplam Popülasyon ve Çeşitlilik

DDM ile biyolojik toplulukların toplam popülasyon ve çeşitlilik değerlerinin hesaplanacağı bu kesimde, modelin ilgili parametreleri parametre uzayından [141] $c = 0.01$, $\mu = 0.0025$, $P_{mut} = 0.0095$ ve $P_{kill} = 0.2$ olarak belirlenerek 200 simülasyon üzerinden ortalama alınmıştır.

Şekil 4.13'da 200 simülasyonun ortalamasının alınması sonucunda Şekil 4.13(a)'da toplam popülasyonun ve Şekil 4.13(b)'de çeşitliliğin ortalama değerlerinin uzun zaman periyodundaki değişimi verilmiştir. Hem toplam popülasyon hem de çeşitlilik geçen zaman içerisinde artma eğilimi sergilemektedirler. Bu sonuç hem fosil kayıtları [76] ile hem de Hall et al. [141] tarafından yapılan DDM çalışması ile uyumludur. Toplam



(a)



(b)

Şekil 4.13. 200 simülasyonun ortalaması için, (a) Toplam popülasyonun ve (b) Çeşitliliğin zamanla değişimi.

popülasyondaki artışı anlamak için, Denklem (3.1) ve Denklem (3.2) ile

$$N(t) = \frac{1}{\mu}(I_a(t) - H_a(t)) \quad (4.2)$$

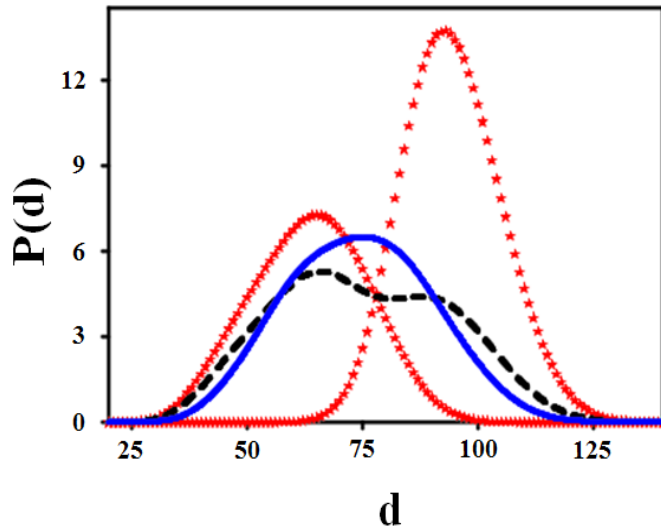
ifadesi yazılabilir. Denklem (4.2)'deki $H_a(t)$ ifadesi, simülasyon ortalamasında sabit kalmaktadır. Ancak $I_a(t)$ ifadesi artacaktır. Bunun nedeni evrimsel dinamiklerin, konumların karşılıklı etkileşimlerden daha iyi fayda sağlayabilmesi için genotip uzayında birbirleri ile etkileşen konumların sayısını artırma eğiliminde olmasıdır. $H_a(t)$ ve $I_a(t)$ ifadelerinin geçen zamanla sergiledikleri bu davranışlar sonucunda toplam popülasyonda artış gözlenmektedir. Evrimsel dinamiklerin karşılıklı etkileşen konumların sayısını

artırması, çeşitliliğin artmasına neden olacaktır. Çeşitliliğin artması ise ekosistemin kompleksliğinin artması olarak yorumlanabilir.

4.2.2. Dinamik Ağ Yapısı

DDM ile biyolojik toplulukların dinamik ağ yapılarının inceleneceği bu kesimde, L hariç modelin diğer parametreleri tek bir simülasyonda olduğu gibi $c = 0.5$, $\mu = 0.0002$, $P_{mut} = 0.002$ ve $P_{kill} = 0.2$ olarak alınmıştır.

Derece dağılımının doğru kararlı durum davranışını tahmin edebilmek için, sistemin sahip olduğu farklı q-ESS periyotlarının ortalaması alınmalıdır. Bu amaçla, herbir simülasyonun en az iki geniş q-ESS periyodu içerdiği 50 simülasyonun ortalaması alınmıştır. Hesaplama sonucunda, birinci ve ikinci q-ESS periyotlarının ortalama derece dağılımları Şekil 4.14’da aynı şekil üzerinde gösterilmiştir.

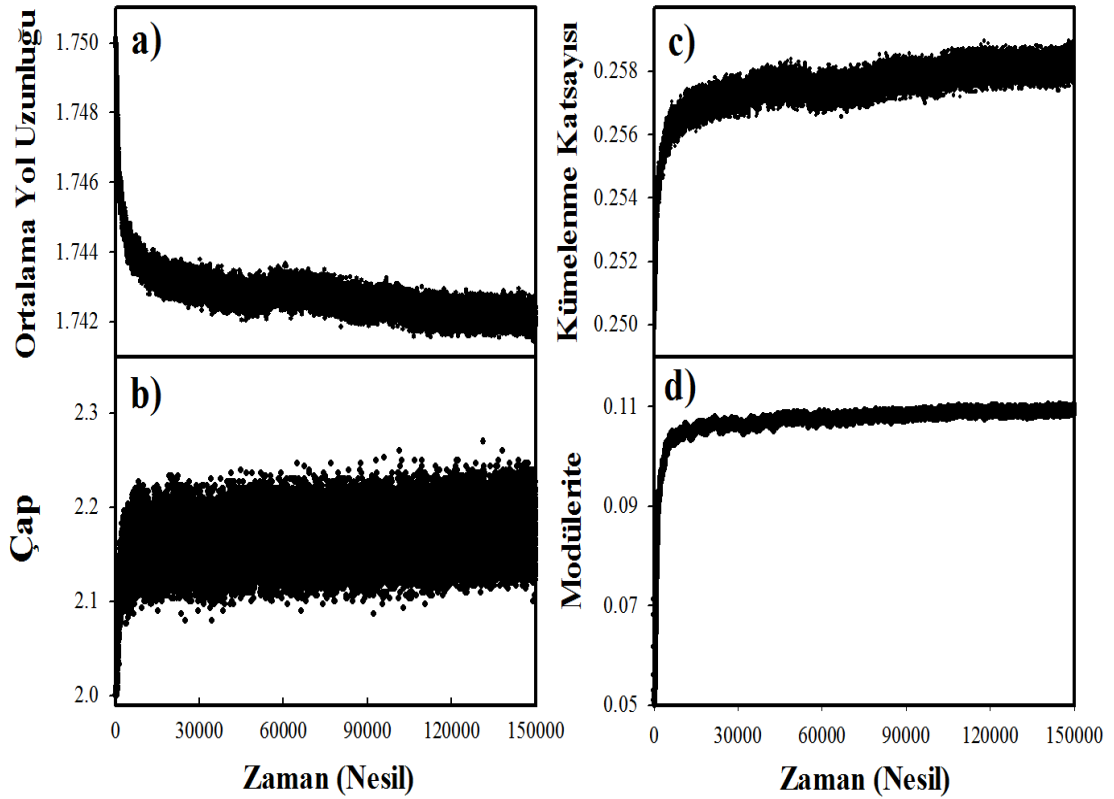


Şekil 4.14. 50 simülasyon üzerinden ortalaması alınan birinci ve ikinci q-ESS periyotlarının derece dağılımları. Siyah kesikli ve mavi sürekli çizgiler sırası ile birinci ve ikinci q-ESS periyotlarının ortalama derece dağılımlarıdır. Kırmızı yıldız semboller, birinci q-ESS periyodunun ortalama derece dağılımının iç detaylarını göstermektedir.

Şekil 4.14’da birinci q-ESS periyodunun ortalama derece dağılımına bakıldığında (siyah kesikli çizgi), dağılımın hörgüç yapısına sahip bir dağılım olduğu görülmektedir. Bu

hörgüç yapısı için birinci q-ESS periyodunun ortalama derece dağılımının kırmızı yıldız sembollerle gösterilen iç detaylarına bakıldığında, dağılımın iki farklı değer etrafında toplandığı görülmektedir. 50 simülasyon içerisinde, bu değerlerden 36'sının düşük $P(d)$ değerine 14'ünün yüksek $P(d)$ değerine sahip olduğu bulunmuştur. İkinci q-ESS periyodu birinci q-ESS periyodunu takip etmesine rağmen, ikinci q-ESS periyodunun ortalama derece dağılımın (mavi sürekli çizgi) düzgün normal dağılım sergilediği görülmektedir.

Ortalama yol uzunluğu, çap, kümelenme katsayısı ve modülerite için 300 simülasyon yapılmış ve bu simülasyonların ortalaması alınmıştır. Bu parametrelerin ortalama değerlerinin zamanla değişimi Şekil 4.15'de verilmiştir.



Şekil 4.15. 300 simülasyon üzerinden ortalamaları alınan ağ parametrelerinin zamanla değişimi. **a)** Ortalama Yol Uzunluğu **b)** Çap **c)** Kümelenme Katsayısı **d)** Modülerite.

Şekil 4.15'e bakıldığında, zaman geçtikçe ortalama yol uzunluğu azalma ve kümelenme katsayısı ise artma davranışı sergilemektedir. Bu durum DDM ile incelenen biyolojik toplulukların dinamik ağ yapısının git gide küçük dünya davranışı sergilediğini göstermektedir. Şekil 4.15 **a)**'daki ortalama yol uzunluğunun ortalama değeri 1.7428 ± 0.0004 ve Şekil 4.15 **c)**'deki kümelenme katsayısının ortalama değeri ise 0.2575 ± 0.0004

dür. Bu değerler tek bir simülasyonda bulunan değerlerden daha büyük ve standart sapmaları daha küçüktür. Şekil 4.15 b)'deki çap ifadesi çok küçük de olsa artış eğiliminde ve ortalama 2.1649 civarında dalgalanmaktadır. Şekil 4.15 d)'deki modülerite değeri de çok küçük bir eğimle artış sergilemekte ve ortalama olarak 0.1070 ± 0.00068 değerine sahiptir. Bu değer tek bir simülasyonda bulunan değerden daha büyük olsa da hala gerçek dünya ağlarına göre oldukça küçüktür. Bunun sebebi yine yoğun etkileşimlerden kaynaklanmaktadır.

5. BÖLÜM

TARTIŞMA, SONUÇ VE ÖNERİLER

Çok fazla sayıda cisim yada varlıkların birbirleri ile etkileşim içerisinde olduğu kompleks sistemlerin incelenmesi oldukça zordur. Bu nedenle kompleks sistemlerin incelenmesi için birçok matematiksel modeller geliştirilmiştir. Bu modeller neticesinde simülasyonlar yapılarak kompleks sistemlerin işleyişi hakkında bilgi edinmeye çalışılmıştır. Bu amaç doğrultusunda Dolaşık Doğa Modeli (DDM)'de, kompleks bir sistem olan ekosistemde bulunan türler arasındaki evrimsel ilişkiyi bazı varsayımlar ve kısıtlamalar altında incelemek için kurulmuştur [36]. Birey tabanlı eş-evrimleşme modeli olan DDM, ekosistemde bulunan türleri fiziksel çevrelerinden bağımsız olarak inceler. Bu durumun nedeni, türlerin kendi aralarındaki etkileşimler neticesinde birbirlerinin evrimleşmelerini nasıl etkilediklerini açıklamaya çalışmaktır. DDM hem matematiksel bir model hem de nümerik bir deneydir. Karşılıklı çıkar ilişkisi, av-avcı yada konak-parazit ilişkisi ve karşılıklı rekabet ilişkisi, ekosistemdeki türler arasında yaygın olarak görülen etkileşme türleridir. Biyolojik topluluklar, ekosistemde görülen etkileşmelerin alt bölümleri olan bu etkileşmelerin herbiri için çeşitli modellerle incelenmiştir. DDM, ekosistemdeki bütün etkileşme çeşitlerini içerisinde barındıran ve bu yüzden ekosisteme daha geniş perspektiften bakan bir modeldir. Bu nedenle DDM, biyolojik evrimin gözlemlenen bir çok özelliğini elde edebilir. Elde ettiği özelliklerden bir tanesi, evrimsel süreç içerisinde bireyler arasındaki etkileşmelerden dolayı, Gould ve Eldredge [154] tarafından tanımlanan topluluğun dinamik olarak kesikli denge (punctuated equilibria) davranışdır [36, 141]. Bu kesikli denge, evrimsel olarak sanki kararlıymış gibi davranan türlerin bulunduğu q-ESS (quasi Evolutionary Stable Strategies) periyotlarının, evrimsel olarak aktif türlerin bulunduğu hektik periyotlar tarafından ayrılmasıyla oluşur. Bu davranış, fosil kayıtlarında gözlemlenen davranışlarla uyum içerisinde [154]. q-ESS

periyodunda bulunan türlerin çoğunluğu topluluğa yüksek derecede uyum sağlamışlardır. Bu nedenle mutasyonlara karşı dayanıklılık göstermekte ve uzun süre boyunca varlıklarını sürdürebilmektedirler. q-ESS periyodunda, çeşitlilik sayısı az ancak toplam popülasyon sayısı fazladır. Küçük dalgalanmalar olmakla birlikte, toplam popülasyon ve çeşitlilik hafif artış sergiler. q-ESS periyodunda topluluğa yüksek derecede uyum sağlayan türlerin yok olmasıyla, topluluk hektik periyoda girmektedir. Hektik periyotta, türler mutasyonlara karşı oldukça hassastır. Çeşitlilik sayısı fazla ancak toplam popülasyon sayısı azdır. Birçok tür kısa zaman içinde ortaya çıkıp kaybolmaktadır. Bu nedenle, toplam popülasyon ve çeşitlilikte büyük dalgalanmalar oluşmaktadır. q-ESS periyotlarının süreleri zamanla artış göstermektedir [141]. Bu artışın bir kuvvet kanunu dağılımı sergilediği ve dağılımın üstelinin yaklaşık olarak -1,8 olduğu gözlenmiştir [36]. Fosil kayıtlarında ise, yok olan tür sayısının zamanla azaldığı görülmektedir. Bu azalmanın da, -2 üsteline sahip bir kuvvet kanunu dağılımı olduğu bilinmektedir [89]. Bu nedenle DDM’ndeki q-ESS sürelerindeki artış, fosil kayıtlarındaki yok olma oranının azalması ile uyumluluk göstermektedir. Çünkü yok olan tür sayısının azalması, türlerin topluluğa daha fazla uyum sağlayarak daha uzun yaşaması anlamına gelmektedir. Ayrıca bu sonuçların eşeyli ve eşeysiz üreme için benzer sonuçlar verdiği elde edilmiştir [36,141,146,155]. Sonuçların çok fazla değişmemesinden dolayı bu çalışmada, bireylerin üremesi eşeysiz olarak incelendi. Bu durumun sebebi, eşeyli üremede bireyler arasındaki Hamming mesafesinin belirlenmesinin çok uzun hesaplama süresi gerektirmesidir. Çünkü Hamming mesafesi belirlenirken, bireylerin L uzunluklu genom zincirlerindeki bütün genler karşılaştırılmaktadır. Bu karşılaştırma, üremenin gerçekleşebilmesi için uygun Hamming mesafesini bulana kadar topluluktaki bütün bireyler için devam edecektir. Bu süreç simülasyonun adımını yavaşlatacaktır.

Biyolojik topluluğun genel davranışı hakkında anlamlı bilgi elde edebilmek için, çok sayıda simülasyonun ortalaması alınmıştır. q-ESS periyotlarındaki ortama yüksek derecede uyum sağlayan türlerin ölmesi sonucunda, toplulukta kitlesel yok olmalar meydana gelmektedir. Çünkü topluluktaki diğer türler varlıklarını bu türler sayesinde devam ettirebilmektedirler. Ancak topluluğun evrimsel dinamiğine bakıldığında, toplam popülasyonun ve çeşitliliğin zamanla arttığı görülmektedir. Bu davranış, Sepkoski tarafından [76] deniz canlılarının fosil kayıtları üzerinde yapılan çalışmayla uygunluk

göstermektedir. Fosil kayıtlarına bakıldığında, çeşitli jeolojik zamanlarda kitlesel yok olmalar gerçekleşmekte ve bu zamanlarda çeşitlilik ciddi oranda azalmaktadır. Ancak evrimsel sürecin genel davranışına bakıldığında, çeşitliliğin zamanla artış sergilediği görülmektedir. DDM’nde q-ESS sürelerinin, çeşitliliğin ve toplam popülasyon sayısının artmasına bağlı olarak varolan türler arasındaki etkileşmelerin artması topluluğun kompleksliğini de artırmaktadır. Komplekslikteki bu artışın, toplulukta bulunan türlerin eş-evrimleşmesini nasıl etkilediği incelenmesi gereken bir konu olacaktır. DDM eş-evrimleşme sürecini anlamak için, ekosistemdeki biyolojik toplulukların dinamik kompleks bir ekolojik ağ sistemi ile incelenmesine olanak tanımaktadır.

Yapılan simülasyonlar sonucunda, topluluğun sahip olduğu ağ yapısı için hektik ve q-ESS periyotlarında etkileşmelerin derece dağılımları elde edildi. Hem q-ESS periyotlarındaki hemde hektik periyottaki derece dağılımlarının bir Gaussian dağılımına sahip olduğu görüldü. Hektik periyotta, çeşitlilik q-ESS periyoduna göre daha fazladır. Bu nedenle, hektik periyotta bulunan türlerin ortalama dereceleri, q-ESS periyodunda bulunan türlerin ortalama derecelerinden daha yüksektir. Genellikle evrimleşen kompleks gerçek dünya ağlarının derece dağılımları, kuvvet kanunu sergileyen üstel bir dağılım olarak kabul edilir [42, 127]. Ölçeksiz ağ olarak adlandırılan bu ağlarda, tercihli eklenme yani ortama iyi uyum sağlamış türlerle etkileşim kurma durumu söz konusudur. Ancak doğada bulunan yiyecek ağlarında tercihli eklenmenin gerekli olup olmadığı tartışma konusudur. Gerçek dünyada sık karşılaşılan ağlardan bir tanesi de küçük dünya ağlarıdır. Küçük dünya ağlarındaki derece dağılımlarının, çok sayıda küçük derecelere sahip bir Poisson dağılımına sahip olduğu görülmüştür [101, 126]. Ağda merkez yani derecesi yüksek türler bulunmaktadır. Topluluktaki diğer türler bu merkez türlerle etkileşim kurmaktadır.

Hektik ve q-ESS periyotlarında, toplulukta yaşayan türler arasındaki ortalama yol uzunluğu ve kümelenme katsayısı birbirine zıt davranışlar sergilemektedir. Hektik periyotta topluluğa uyum sağlayamayan türler kısa yaşam sürelerine sahip olmaktadır. Toplam popülasyon sayısı azaldığı için toplulukta varolan etkileşmelerin sayısı azalmaktadır. Etkileşme sayısının azalması, ağda bulunan bazı türler arasındaki bağlantıların kopmasına ve türler arasındaki mesafelerin artmasına neden olacaktır. Bu nedenle ortalama yol uzunluğu artmakta ve kümelenme katsayısı azalmaktadır. q-ESS periyodunda ise, topluluğa uyum sağlayan türlerin bulunması varolan etkileşmelerin

sayısını artırmaktadır. Topluluk genellikle bu türler etrafında toplanmaktadır. Bu nedenle, ağdaki türler arasındaki bağlantılar artmakta ve mesafeler kısalmaktadır. Bunun sonucunda ortalama yol uzunluğu azalmakta ve kümelenme katsayısı artmaktadır. Topluluğun evrimsel dinamiğinde, ortalama yol uzunluğu genel olarak azalma eğiliminde ve kümelenme katsayısı ise artma eğilimindedir. Çünkü topluluğun sergilediği genel davranışta, çeşitlilik ve varolan etkileşme sayısı artış göstermektedir. Ortalama yol uzunluğu için 1.7428 ve kümelenme katsayısı için 0.2575 değerleri elde edilmiştir. İskoçya'daki Ythan nehrindeki gerçek bir yiyecek ağı için inceleme yapılmıştır [156]. Ağdaki türler arasında yönelimli konak-parazit ilişkisi bulunmaktadır. Yapılan bu incelemede, ortalama yol uzunluğu için 2.43 ve kümelenme katsayısı için 0.22 değerleri bulunmuştur. Küçük dünya ağ modeli ile, ipliksi solucanın (*Caenorhabditis elegans*) sinir sistemi ağı için de incelemeler yapılmıştır [41]. Sinir sistemindeki nöron deney sinir hücreleri birbirlerine mesaj iletimini sağlayan sinapslarla bağlanarak, yönelimsiz ve ağırlıksız bir sinir ağı oluşturulmuştur. Yapılan incelemeler sonucunda, ortalama yol uzunluğu 2.65 ve kümelenme katsayısı 0.28 olarak bulunmuştur. Aynı ağ rastgele ağ modeli ile de incelenmiş ve ortalama yol uzunluğu 2.25, kümelenme katsayısı ise 0.05 olarak elde edilmiştir [41]. Görüldüğü üzere küçük dünya ağlarındaki, ortalama yol uzunluğu rastgele ağlardaki kadar küçük ancak kümelenme katsayısı rastgele ağlardan çok büyüktür. Bu sonuç küçük dünya ağlarının temel esasıdır. DDM ile elde edilen sonuçların gerçek dünya ağları ile uyumlu olduğu görülmektedir. Böylelikle DDM ile incelenen topluluğun ağ yapısının küçük dünya ağı davranışı sergilediği sonucuna ulaşılmaktadır.

Çap, ağdaki en uzak iki tür arasındaki en kısa mesafeyi verir ve bu sebeple ortalama yol uzunluğu için bir üst limit belirler. Ağdaki bağlantı sayısının yani türler arasındaki varolan etkileşmelerin sayısının artması ağın çapını azaltacaktır. Çap ifadesi için hektik periyotta 3, q-ESS periyodunda 2 değeri elde edilmiştir. Topluluğun evrimsel dinamiğinin genel davranışı incelendiğinde, çap değeri 2.16 civarında dalgalanma göstermektedir. Yapılan bir çalışmada 35 türe sahip Skipwith göletinde, 181 türe sahip Little Rock gölünde, 75 türe sahip Bridge Brook gölünde, 33 türe sahip Chesapeake koyunda, 92 türe sahip Ythan nehrinde, 30 türe sahip Coachella deresinde ve 44 türe sahip St. Martin adasında olmak üzere yedi farklı yönelimli gerçek kompleks yiyecek ağı incelenmiştir [157]. Bu ağlarda

ortalama olarak, türlerin % 26'sı tek bağlantı ile doğrudan birbirleri ile, % 80'i 2 bağlantı ve % 90'ı 3 bağlantı ile etkileşmektedir. Ağlardaki bağlantı sayısının artmasının, türler arasındaki mesafeyi azalttığı gözlenmiştir. Türler arasındaki en uzun mesafenin 2 ya da 3 olduğu görülmüştür. Topluluktaki türlerin % 25'inin etkileştiği DDM sonuçlarının, küçük dünya davranışı sergileyen bu gerçek yiyecek ağlarındaki sonuçlarla uyduğu görülmüştür.

Biyolojik toplulukların sergiledikleri ağ yapılarının incelenmesinde, toplulukta oluşan gruplanmanın anlaşılabilmesi için modülerite parametresi incelenmiştir. Hektik periyotta, türler topluluğa uyum sağlayamadıkları için aralarındaki etkileşimler rastgele olmaktadır. Bu nedenle, birbirleri ile daha çok etkileşen türlerin olduğu grup yapılarının oluşması zordur. Buna bağlı olarak ağın modülerite değeri azalmaktadır. q-ESS periyodunda, türler topluluğa uyum sağlayan türlerin etrafında toplanmaktadır. Bu nedenle q-ESS periyodunda toplulukta grup yapısının gözlenmesi hektik periyoda göre daha olasıdır. Buna bağlı olarak ağın modülerite değeri artmaktadır. Topluluğun evrimsel dinamiğinin genel davranışına bakıldığında, modülerite değeri 0.1070 olarak bulunmuştur. Toplulukta kabul edilebilir bir grup yapısının gözlenebilmesi için modülerite değerinin 0.3'ten büyük olması gerekmektedir [110]. Buna göre bu çalışmada elde edilen modülerite değeri düşüktür. Bu durumun bir sebebi, topluluktaki türler arasında yoğun etkileşmelerin bulunması olabilir. Başka bir sebebi, bu çalışmada besin zincirindeki bitki gibi sadece dış kaynaktan beslenen ilk üreticilerin düşünülmemesi olabilir. Çünkü topluluk temel olarak bu üretici türlerin etrafında toplanabilir [152]. Ancak bu düşük modülerite değeri, ekosistemde net bir gruplanmanın olmadığı [37] düşüncesi ile uyumludur.

DDM ile, biyolojik topluluklardaki türlerin sabit bir mutasyon altında üremesi ile soy ağacı çıkarımı yapıldı. Mutasyon sonucu oluşan mutant türler genellikle topluluğa uyum sağlayamayarak çok çabuk kaybolmaktadır. Ancak bazı durumlarda seçim baskısından dolayı topluluğa uyum sağlayarak varlığını sürdürebilmektedir. Bu şekilde elde edilen soy ağacında, mutant türlerin ataları ile birlikte belirli bir süre beraber yaşadıkları görülmüştür. Bu sonuç, bu türlerin aynı fosil kayıtlarında bulunmasına ve dolayısıyla ortaya çıkış zamanlarının ayırt edilememesi sorununa neden olmaktadır. Maximum Likelihood (ML) [137] ve Neighbor-Joining (NJ) [138] gibi soy ağacı tahmininde kullanılan metotlar ata türleri, ortama uyum sağlayamadığı için çok çabuk

ölen geçiş türlerini ve ortak atayı tanımlayamazlar. Sadece verilen bir ağacın olabilirliğini değerlendirirler. Ancak bilinen türlerden ürettikleri soy ağaçları, topolojik olarak gerçek soy ağacına benzerlik göstermektedir. Bu nedenle, DDM ile belirli bir nesil aralığında elde edilen soy ağacı, DDM ile aynı veriler kullanmak üzere ML ve NJ metotlarının tahmin ettikleri soy ağaçları ile karşılaştırıldı. Bu karşılaştırma TOPD/FMTS yazılımı [148] kullanılarak nümerik olarak yapıldı. Sonuç olarak NJ metodu ile tahmin edilen soy ağacı topolojisi ve bu ağacın dalları, DDM ile elde edilen gerçek ağaca daha çok benzediği görüldü. Ayrıca soy ağacında bulunan türlerin topluluk yapısı modülerite ile incelendi ve topluluğun 4 gruptan oluştuğu görüldü. Yapılan bir çalışmada, gerçek dünyada bulunan 181 türe sahip Little Rock gölündeki, 33 türe sahip Chesapeake koyundaki, 92 türe sahip Ythan nehrindeki, ve 44 türe sahip St. Martin adasındaki yiyecek ağlarında topluluk yapısı incelemesi yapılmıştır [152]. Ağda türler arasındaki etkileşmeler, grupların belirlenmesini etkilediği için ağırlıklı olarak alınmıştır. Yapılan incelemeler sonucunda, ağlarda bulunan grup sayılarının 1 ile 6 arasında olduğu bulunmuştur. Soy ağacı ile ilgili bu çalışma yayınlandı [158].

Bu tez çalışmasında, biyolojik topluluklardaki eş-evrimleşme mekanizmasını ağ yapısı ile inceledik. Topluluğun oluşturduğu ağ yapısının, küçük dünya ağı olduğunu bulduk. Bu nedenle, ağda birçok alternatif kısa yolların var olduğunu ve ağın rastgele saldırılara karşı dayanıklı olduğunu bulduk. Toplulukta bulunan bireyler arasındaki yoğun etkileşmelerin, toplulukta ayırt edilebilir bir grup yapısının oluşmasına müsaade etmediği sonucuna ulaştık.

KAYNAKLAR

1. Arthur, W. B., 1999. Complexity and the economy. **Science**, **284**(5411):107–109.
2. Weinberger, E. D., 1991. Local properties of kauffman's n-k model: a tunably rugged energy landscape. **Physical Review A**, **44**(10):6399–6413.
3. Wilde, G., Gorler, G. P., Willnecker, R., Dietz, G., 1994. Thermodynamic properties of pd40ni40p20 in the glassy, liquid, and crystalline states. **Applied Physics Letters**, **65**(4):397–399.
4. Nieuwenhuizen, T. M., 1998. Thermodynamic description of a dynamical glassy transition. **Journal of Physics A: Mathematical and General**, **31**(10):L201–L207.
5. Nieuwenhuizen, T. M., 1998. Thermodynamics of the glassy state: effective temperature as an additional system parameter. **Physical Review Letters**, **80**(25):5580–5583.
6. Nieuwenhuizen, T. M., 2000. Thermodynamic picture of the glassy state gained from exactly solvable models. **Physical Review E**, **61**(1):267–292.
7. Derrida, B., Peliti, L., 1991. Evolution in a flat fitness landscape. **Bulletin of Mathematical Biology**, **53**(3):355–382.
8. Kauffman, S. A., Johnsen, S., 1991. Coevolution to the edge of chaos: coupled fitness landscapes, poised states, and coevolutionary avalanches. **Journal of Theoretical Biology**, **149**(4):467–505.
9. Martijna, H., Paulien, H., 1994. Pattern generation in molecular evolution: exploitation of the variation in rna landscapes. **Journal of Molecular Evolution**, **39**(1):71–79.
10. Kim, J. H., Yum, B. J., 1993. A heuristic method for solving redundancy optimization problems in complex systems. **IEEE Transactions on Reliability**, **42**(4):572–578, 1993.
11. Kuo, W., Prasad, V. R., 2000. An annotated overview of system-reliability optimization. **IEEE Transactions on Reliability**, **49**(2):176–187.
12. Salazar, D., Rocco, C. M., Galvan, B. J., 2006. Optimization of constrained multiple-objective reliability problems using evolutionary algorithms. **Reliability Engineering and System Safety**, **91**(9):1057–1070.

13. Parisi, G., 1980. The order parameter for spin glasses: a function on the interval 0-1. **Journal of Physics A: Mathematical and General**, **13**(3):1101–1112.
14. Beck, C., Schlogl, F., 1993. Thermodynamics of chaotic systems. Cambridge University Press, Cambridge 306 s.
15. Edwards, S. F., Anderson, W., 1975. Theory of spin glasses. **Journal of Physics F**, **5**(5):965–974.
16. Parisi, G., 1980. Magnetic properties of spin glasses in a new mean field theory. **Journal of Physics A: Mathematical and General**, **13**(5):1887–1895.
17. Sherrington, D., Kirkpatrick, S., 1975. Solvable model of a spin-glass. **Physical Review Letters**, **35**(26):1792–1796.
18. Guerra, F., 2003. Broken replica symmetry bounds in the mean field spin glass model. **Communications in Mathematical Physics**, **233**(1):1–12, 2003.
19. Parisi, G., 1979. Infinite number of order parameters for spin-glasses. **Physical Review Letters**, **43**(23):1754–1756.
20. Proulx, S. R., Promislow, D. E. L., Phillips, P. C., 2005. Network thinking in ecology and evolution. **TRENDS in Ecology and Evolution**, **20**(6):345–353.
21. Moon, D. C., Moon, J., Keagy, A., 2010. Direct and indirect interactions. **Nature Education Knowledge**, **3**(10):50.
22. Kauffman, S. A., 1993. The Origins of Order. Oxford University Press, New York 715 s.
23. Drossel, B., 2001. Biological evolution and statistical physics. **Advance in Physics**, **50**(2):209–295.
24. Prugel, B. A., Shapiro, J. L., 1994. Alysis of genetic algorithms using statistical mechanics. **Physical Review Letters**, **72**(9):1305–1309.
25. McGhee, G. R., 2007. The Geometry of evolution: adaptive landscapes and theoretical morphospaces. Cambridge University Press, New York 216 s.
26. Derrida, B., 1981. Random-energy model: an exactly solvable model of disordered systems. **Physical Review B**, **24**(5):2613–2626.
27. Kauffman, S. A., Weinberger, E. D., 1989. The nk model of rugged fitness landscapes and its application to maturation of the immune-response. **Journal of Theoretical Biology**, **141**(2):211–245.

28. Flyvbjerg, H., Lautrup, B., 1992. Evolution in a rugged fitness landscape. **Physical Review A**, **46**(10):6714–6723.
29. Vermeij, G. J., 1994. The evolutionary interaction among species: Selection, escalation and coevolution. **Annual Review of Ecology and Systematics**, **25**:219–236.
30. Janzen, D. H., 1980. When is it coevolution ? **Evolution**, **34**(3):611–612.
31. Neumann, J. V., Morgenstern, O., 1953. Theory of games and economic behavior. Princeton University Press, Princeton 672 s.
32. Smith, J. M., 1982. Evolution and the theory of games. Cambridge University Press, Cambridge 224 s.
33. Kauffman, S. A., Levin, S., 1987. Towards a general theory of adaptive walks on rugged landscapes. **Journal of Theoretical Biology**, **128**(1):11–45.
34. Bak, P., Sneppen, K., 1993. Punctuated equilibrium and criticality in a simple model of evolution. **Physical Review Letters**, **71**(24):4083–4086.
35. Amaral, L. A. N., Meyer, M., 1999. Environmental changes, coextinction, and patterns in the fossil record. **Physical Review Letters**, **82**(3):652–655.
36. Christensen, K., Hall, M., Collobiano, S. A., Jensen, H. J., 2002. Tangled nature: a model of evolutionary ecology. **Journal of Theoretical Biology**, **216**:73–84.
37. Montoya, J. M., Pimm, S. L., Sole, R. V., 2006. Ecological networks and their fragility. **Nature**, **442**:259–264.
38. Erdős, P., Rényi, A., 1959. On random graphs, i. **Publicationes Mathematicae Debrecen**, **6**:290–297.
39. Erdős, P., Rényi, A., 1960. On the evolution of random graphs. **A Magyar Tudományos Akademia. Matematikai Kutató Intézet Közleményei**, **5**:17–61.
40. Milgram, S., 1967. The small-world problem. **Psychology Today**, **1**(1):61–67.
41. Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. **Nature**, **393**:440–442.
42. Barabási, A. L., Albert, R., 1999. Emergence of scaling in random networks. **Science**, **286**:509–512.
43. Lamarck, J. B., 1809. Philosophie zoologique: ou exposition des considérations relative á l'histoire naturelle des animaux. Dentu, Paris 411 s. (İngilizce çevirisi;

- Elliot, H. Zoological philosophy: an exposition with regard to the natural history of animals, 1914, MacMillan, London).
44. Koonin, E. V., Wolf, Y. I., 2009. Is evolution darwinian or/and lamarckian? **Biology Direct**, **4**(1):42–56.
 45. Darwin, C., 1859. On the origin of species by means of natural selection or the preservation of favoured Races in the struggle for life. John Murray, London 224 s.
 46. Lewontin, R. C., 1970. The units of selection. **Annual Review of Ecology and Systematics**, **1**:1–18.
 47. Wallace, A. R., 1889. Darwinism. An exposition of the theory of natural selection with some of its applications. MacMillan, London 494 s.
 48. Weismann, A., 1892. Das Keimplasma: Eine Theorie der Vererbung. Fischer, Jena, Berlin 524 s. (İngilizce çevirisi; Parker, W. N., Rönfeldt, H. The germ-plasm: a theory of heredity, 1893, Walter Scott, London).
 49. Kutschera, U., Niklas, K. J., 2004. The modern theory of biological evolution: an expanded synthesis. **Naturwissenschaften**, **91**:255–276.
 50. Mendel, G., 1865. Versuche über pflanzen-hybriden. **Verhandlungen des naturforschenden Ver-eines in Brünn**, **4**:3–47 (İngilizce çevirisi; 1901. Experiments in plant hybridization, **Journal of the Royal Horticultural Society of London**, **26**:1–32).
 51. Tamm, C., Herman, T., Shapiro, S., Lipschitz, R., Chargaff, E., 1953. Distribution density of nucleotides within a desoxyribonucleic acid chain. **Journal of Biological Chemistry**, **203**(2):673–688.
 52. Watson, J. D., Crick, F. H. C., 1953. Molecular structure of nucleic acids. **Nature**, **171**(4356):737–738.
 53. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. Molecular Biology of the Cell. Garland Science, New York 1616 s.
 54. Knippers, R., 1977. Molekulare Genetik. Thieme Press, Stuttgart 567 s.
 55. Huxley, J. S., 1942. Evolution: the modern synthesis. George Allen ve Unwin Ltd., London 645 s.
 56. Drake, J. W., Charlesworth, B., Charlesworth, D., Crow, J. F., 1998. Rates of spontaneous mutation. **Genetics**, **148**(4):1667–1686.

57. Nachman, M. W., Crowell, S. L., 2000. Estimate of the mutation rate per nucleotide in humans. **Genetics**, **156**(1):297–304.
58. Sawyer, S. A., Parsch, J., Zhang, Z., Harti, D. L., 2007. Prevalence of positive selection among nearly neutral amino acid replacements in drosophila. **Proceedings of the National Academy of Science of the United States of America**, **104**:6504–6510.
59. Ellstrand, N. C., 2014. Is gene flow the most important evolutionary force in plants? **American Journal of Botany**, **101**(5):737–753.
60. Mayr, E., 1963. Animal species and evolution. Harvard University Press, Boston, Massachusetts, USA 797 s.
61. Ehrlich, P. R., Raven, P. H., 1969. Differentiation of populations. **Science**, **165**:1228–1232.
62. Masel, J., 2011. Genetic drift. **Current Biology**, **21**(20):R837–R838.
63. Excoffier, L., Foll, M., Petit, R. J., 2009. Genetic consequences of range expansions. **Annual Review of Ecology, Evolution, and Systematics**, **40**:481–501.
64. Hallatschek, O., Hersen, P., Ramanathan, S., Nelson, D. R. Genetic drift at expanding frontiers promotes gene segregation. **Proceedings of the National Academy of Science of the United States of America**, **104**(50):19926–19930.
65. Zimmer, C., Emlen, D. J., 2010. Evolution: making sense of life. Roberts and Company Publishers, Greenwood Village 720 s.
66. Sober, E., 1984. The Nature of selection. MIT Press, Cambridge 383 s.
67. Thompson, J. N., 1994. The coevolutionary process. University of Chicago Press, Chicago 383 s.
68. Yip, K. Y., Patel, P., Kim, P. M., Engelman, D. M., McDermott, D., Gerstein, M., 2008. An integrated system for studying residue coevolution in proteins. **Bioinformatics**, **24**(2):290–292.
69. Van, V. L., 1973. A new evolutionary law. **Evolutionary Theory**, **1**:1–30.
70. Dawkins, R., Krebs, J. R., 1979. Arms races between and within species. **Proceedings of The Royal Society B**, **205**:489–511.
71. Dawkins, R., Krebs, J. R., 2010. Links between global taxonomic diversity, ecological diversity and the expansion of vertebrates on land. **Biology Letters**, **6**:544–547.

72. Cappellari, S. C., Schaefer, H., Davis, C. C., 2013. Evolution: pollen or pollinators. which came first? **Current Biology**, **23**(8):R316–R318.
73. Larsson, T. B., 2001. Ecological bulletins, biodiversity evaluation tools for european forests. Wiley, 240 s.
74. Phillips, J., 1860. Life on the Earth: its origin and succession. Macmillan, London 255 s.
75. Sepkoski, J. J., 1993. Ten years in the library: new data confirm paleontological patterns. **Paleobiology**, **19**(1):43–51.
76. Sepkoski, J. J., 1997. Biodiversity: past, present and future. **Journal of Paleontology**, **71**(4):533–539.
77. Norell, M. A., Novacek M. J., 1992. The fossil record and evolution: comparing cladistic and paleontologic evidence for vertebrate history. **Science**, **255**:1690–1693.
78. Benton, M. J., Storrs G. W., 1994. Testing the quality of the fossil record: paleontological knowledge is improving. **Geology**, **22**:111–114.
79. Benton, M. J., 1993. The Fossil Record 2. Chapman and Hall, London 846 s.
80. Futuyma, D. J., 1997. Evolutionary Biology. Sinauer Associates, Massachusetts 763 s.
81. Sepkoski, J. J., 1992. A compendium of fossil marine animal families. Milwaukee Public Museum, 156 s.
82. Newman, M. E. J., Sibani, P., 1999. Extinction, diversity and survivorship of taxa in the fossil record. **Proceeding of The Royal Society of London B**, **266**(1428):1593–1599.
83. Ayala, F. J., Rzhetsky, A., Ayala, F. J., 1998. Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. **Proceedings of the National Academy of Science of The United States of America**, **95**:606–611.
84. Valentine, J. W., Awramik, S. M., Signor, P. W., Sadler, P. M., 1991. The biological explosion at the precambrian-cambrian boundary. **Evolutionary Biology**, **25**:279–356.
85. Runnegar, B., 2000. Loophole for snowball earth. **Nature**, **405**:403–404.
86. Alvarez, L. W., Alvareza, W., Asaro, F., Michel, H. V., 1980. Extraterrestrial causes for the cretaceous-tertiary extinction. **Science**, **208**:1095–1108.

87. Hut, P., Alvarez, W., Elder, W. P., Handsen, T., Kauffman, E. G., Keller, G., Shoemaker, E. M., Weissman, P. R., 1987. Comet showers as a cause of mass extinctions. **Nature**, **329**:118–126.
88. Raup, D. M., 1986. Biological extinction in earth history. **Science**, **231**(4745):1528–1533.
89. Sole, R., Bascompte, J., 1996. Are critical phenomena relevant to large-scale evolution? **Proceeding of the Royal Society of London B**, **263**:161–168.
90. Simon, H. M., 1962. The architecture of complexity. **Proceeding of the American Philosophical Society**, **106**(6):467–482.
91. Linnaeus, C. C. L., 1735. *Systema Naturæ*. Lugduni Batavorum: Apud Theodorum Haak: Ex Typographia Joannis Wilhelmi de Groot, 13 s.
92. Smith, T. C., 2010. Deep phylogeny, ancestral groups and the four ages of life. **Philosophical Transactions of the Royal Society B**, **365**:111–132.
93. Newman, M. E. J., 2010. *Networks. An Introduction*. Oxford University Press, New York 740 s.
94. Newman, M. E. J., Strogatz, S. H., Watts, D. J., 2001. Random graphs with arbitrary degree distributions and their applications. **Physical Review E**, **64**:026118–026135.
95. Yook, S. H., Jeong, H., Barabási, A. L., 2001. Weighted evolving networks. **Physical Review Letters**, **86**(25):5835–5838.
96. Dorogovtsev, S. N., Mendes, J. F. F., 2002. Evolution of networks. **Advances in Physics**, **51**(4):1079–1187.
97. Boccaletti, S., Latora, V., Moreno, Y., Chaves, M., Hwang, D. U., 2006. Complex networks: structure and dynamics. **Physics Reports**, **424**:175–308.
98. Price, D. J. S., 1965. Networks of scientific papers. **Science**, **149**(3683):510–515.
99. Albert, R., Jeong, H., Barabási, A. L., 1999. Diameter of the world wide web. **Nature**, **401**:130–131.
100. Amaral, L. A. N., Scala, A., Barhélémy, M., Stanley, H. E., 2000. Classes of small-world networks. **Proceedings of the National Academy of Science of the United States of America**, **97**(21):11149–11152.
101. Albert, R., Barabási, A. L., 2002. Statistical mechanics of complex networks. **Reviews of Modern Physics**, **74**:47–97.

102. Freeman, L. C., 1977. A set of measures of centrality based on betweenness. **Sociometry**, **40**(1):35–41.
103. Fronczak, A., Fronczak, P., Holyst, J. A., 2004. Average path length in random networks. **Physical Review E**, **70**:056110–056117.
104. Bogñá, M., Satorras, R. P., 2003. Class of correlated random networks with hidden variables. **Physical Review E**, **68**:036112–036125.
105. Girvan, M., Newman, M. E. J., 2002. Community structure in social and biological networks. **Proceedings of the National Academy of Science of the United States of America**, **99**(12).
106. Girvan, M., Newman, M. E. J., 2009. Clustering in weighted networks. **Social Networks**, **31**:155–163.
107. Serrano, M. A., Boguñá, M., 2006. Clustering in complex networks. i. general formalism. **Physical Review E**, **74**:056114–056122.
108. Kaiser, M., 2008. Mean clustering coefficients: the role of isolated nodes and leafs on clustering measures for small-world networks. **New Journal of Physics**, **10**:083042–083052.
109. Fagiolo, G., 2007. Clustering in complex directed networks. **Physical Review E**, **76**:026107–026114.
110. Newman, M. E. J., Girvan, M., 2004. Finding and evaluating community structure in networks. **Physical Review E**, **69**:026113–026127.
111. Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., Parisi, D., 2004. Defining and identifying communities in networks. **Proceedings of the National Academy of Science of the United States of America**, **101**(9):2658–2663.
112. Pons, P., Latapy, M., 2006. Computing communities in large networks using random walks. **Journal of Graph Algorithms and Applications**, **10**(2):191–218.
113. Blondel, V. D., Guillaume, J. L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. **Journal of Statistical Mechanics: Theory and Experiment**, **P10008**:P10008–P10020.
114. Clauset, A., Newman, M. E. J., Moore, C., 2004. Finding community structure in very large networks. **Physical Review E**, **70**:066111–066116.
115. Wu, F., Huberman, B. A., 2004. Finding communities in linear time: a physics approach. **The European Physical Journal B**, **38**:331–338.

116. Newman, M. E. J., 2006. Finding community structure in networks using the eigenvectors of matrices. **Physical Review E**, **74**:036104–036122.
117. Newman, M. E. J., 2006. Modularity and community structure in networks. **Proceedings of the National Academy of Science of the United States of America**, **103**(23):8577–8582.
118. Newman, M. E. J., 2004. Analysis of weighted networks. **Physical Review E**, **70**:056131–056139.
119. Newman, M. E. J., 2004. Fast algorithm for detecting community structure in networks. **Physical Review E**, **69**:066133–066138.
120. Bollobás, B., 1981. Degree sequence of random graphs. **Discrete Mathematics**, **33**:1–19.
121. Chung, F., Lu, L., 2001. The diameter of sparse random graphs. **Advances in Applied Mathematics**, **26**:257–279.
122. Martinez, N. D., 1991. Artifacts or attributes? effects of resolution on the little rock lake food web. **Ecological Monographs**, **61**(4):367–392.
123. Davidsen, J., Ebel, H., Bornholdt, S., 2002. Emergence of a small world from local interactions: modeling acquaintance networks. **Physical Review Letters**, **88**(12):128701–128704.
124. Pastor, S. R., Vázquez, A., Vespignani, A., 2001. Dynamical and correlation properties of the internet. **Physical Review Letters**, **87**:258701.
125. Fell, D. A., Wagner, A., 2000. The small world of metabolism. **Nature Biotechnology**, **18**:1121–1122.
126. Barrat, A., Weigt, M., 2000. On the properties of small-world network models. **The European Physical Journal B**, **13**:547–560.
127. Barabási, A. L., Albert, R., Jeong, H., 1999. Mean-field theory for scale-free random networks. **Physica A**, **272**:173–187.
128. Bollobás, B., Riordan, O., 2004. The diameter of a scale-free random graph. **Combinatorica**, **24**(1):5–34.
129. Huxley, J. S., 1940. The new systematics. Oxford University Press, Oxford 583 s.
130. Mayr, E., 1942. Systematics and the origin of species. Columbia University Press, New York 334 s.

131. Simpson, G. G., 1961. Principles of animal taxonomy. Columbia University Press, New York 247 s.
132. Hennig, W., 1966. Phylogenetic systematics. University of Illinois Press, Urbana 280 s.
133. De Queiroz, K., Donoghue, M. J., 1988. Phylogenetic systematic and the species problem. **Cladistics**, **4**(4):317–338.
134. Mayr, E., 1969. The biological meaning of species. **Biological Journal of the Linnean Society**, **1**(3):311–320.
135. Saitou, N., Imanishi, T., 1989. Relative efficiencies of the fitch-margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. **Molecular Biology and Evolution**, **6**(5):514–525.
136. Eck, R. V., Dayhoff, M. O., 1966. Atlas of protein sequence and structure. National Biomedical Research Foundation, Silver Springs, Maryland 215 s.
137. Felsenstein, J., 1981. Evolutionary trees from dna sequences: A maximum likelihood approach. **Journal of Molecular Evolution**, **17**(6):368–376.
138. Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, **4**(4):406–425.
139. Felsenstein, J., 1978. The number of evolutionary trees. **Systematic Biology**, **27**(1):27–33.
140. Higgs, P. G., Derrida, B., 1992. Genetic distance and species formation in evolving populations. **Journal of Molecular Evolution**, **35**:454–465.
141. Hall, M., Christensen, K., Collobiano, S. A., Jensen, H. J., 2002. Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution. **Physical Review E**, **66**:011904–011914.
142. Collobiano, S. A., Christensen, K., Jensen, H. J., 2003. The tangled nature model as an evolving quasi-species model. **Journal of Physics A: Mathematical and General**, **36**:883.
143. Anderson, P. E., Jensen, H. J., 2005. Network properties, species abundance and avolution in a model of evolutionary ecology. **Journal of Theoretical Biology**, **232**:551–558.

144. Laird, S., Jensen, H. J., 2006. The tangled nature model with inheritance and constraint: evolutionary ecology restricted by a conserved resource. **Ecological Complexity**, **3**:253–262.
145. Lawson, D., Jensen, H. J., Kaneko, K., 2006. Diversity as a product of inter-specific interactions. **Journal of Theoretical Biology**, **243**:299–307.
146. Rikvold, P. A., 2007. Self-optimization, community stability and fluctuations in two individual-based models of biological coevolution. **Journal of Mathematical Biology**, **55**:653–677.
147. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. Mega5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Molecular Biology and Evolution**, **28**(10):2731–2739.
148. Puigbo, P., Garcia, V. S., McInerney, J. O., 2007. Topd/fmts: a new software to compare phylogenetic trees. **Bioinformatics**, **23**:1556–1558.
149. Matsumoto M., Nishimura, T., 1998. A 623-dimensionally equidistributed uniform psuedo-random number generator. **ACM Transactions on Modeling and Computer Simulation**, **8**:3–30.
150. Adams, M. D. et.all, 2000. The genome sequence of drosophila melanogaster. **Science**, **287**(5461):2185–2195.
151. Miller, G. T., Spoolman, S., 2012. Environmental Science. Cengage Learning Press, USA 576 s.
152. Ann, E. K., Kenneth, A. F., Doran, M. M., Robert, E. U., William, W. T., 2003. Compartments revealed in food-web structure. **Nature**, **426**:282–285.
153. Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. **Journal of Molecular Evolution**, **16**:111–120.
154. Gould, S. J., Eldredge, N., 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. **Paleobiology**, **3**:115–151.
155. Rikvold, P. A., Zia, R. K. P., 2003. Punctuated equilibria and 1/f noise in a biological co-evolution model with individual based dynamics. **Physical Review E**, **68**:031913–031924.

156. Huxham, M., Beaney, S., Raffaelli, D., 1996. Do parasites reduce the chances of triangulation in a real food web? **Oikos**, **76**(2):284–300.
157. Williams, R. J., Berlow, E. R., Dunne, J. A., Barabási, A. L., Martinez, N. D., 2002. Two degrees of separation in complex food webs. **Proceedings of the National Academy of Science of the United States of America**, **99**(20):12913–12916.
158. Canko, O., Taşkın F., Arın, K., 2015. Phylogenetic tree and community structure from a tangled nature model. **Journal of Theoretical Biology**, **382**:216–222.

ÖZGEÇMİŞ

Kamil ARGİN

kmlargn@gmail.com

Kişisel Bilgiler

Uyruğu	TÜRKİYE CUMHURİYETİ (T.C)
Doğum tarihi	08 / 06 / 1987
Doğum yeri	GAZİANTEP

Eğitim Bilgileri

2005 - 2009	Lisans, Erciyes Üniversitesi, Fen Edebiyat Fakültesi, Fizik Bölümü.
2009 - 2011	Yüksek Lisans, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Fizik Bölümü.
2011 - 2015	Doktora, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Fizik Bölümü.

Projeler

Araştırmacı	Yüksek Lisans Tez Projesi, Proje No: FBY-10-3371, "Dört Alt Örgülü Antiferromanyetik Ising Modelinin, İncelenmesi" Erciyes Üniversitesi.
Bursiyer	TÜBİTAK Projesi, Proje No: 111T735 "Biyolojik Toplulukların ve Gelişen Ağ Yapılarının Denge Dışı İstatistik Fiziksel Metodlar Simülasyonu ile İncelenmesi" .

Yayınlar

1. Kamil Argın, Osman Cankö, **"Blume-Capel model on the antiferromagnetic fcc lattice"**, Physica A **391**, 2556-2563, 2012
2. Osman Cankö, Ferhat Taşkın, Kamil Argın, Ahmet Erdinç, **"Hysteresis behavior of Blume-Capel model on a cylindrical Ising nanotube"**, Solid State Communications **183**, 35-40, 2014
3. Osman Cankö, Ferhat Taşkın, Kamil Argın, **"Phylogenetic Tree and Community Structure from a Tangled Nature Model"**, Journal of Theoretical Biology **382**, 216-222, 2015

Bildiriler ve Katılımlar

30 Haziran - 2 Temmuz 2011	18. İstatistiksel Fizik Günleri , İstanbul
5 - 10 Eylül 2011	Feza Gürsey Institute - Imperial College International Summer School and Research Workshop on Complexity , İstanbul
5 - 10 Eylül 2013	Matematiksel Evrim Çalıştayı , İzmir-Şirince