



**T.C.
YALOVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**PROTEİN PROTEİN ETKİLEŞİM AĞLARINDA GÜVENİLİRLİK
ANALİZLERİ**

**DOKTORA TEZİ
Volkan ALTUNTAŞ
165305001**

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Murat GÖK

TEMMUZ 2020



T.C.
YALOVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**PROTEİN PROTEİN ETKİLEŞİM AĞLARINDA GÜVENİLİRLİK
ANALİZLERİ**

DOKTORA TEZİ
Volkan ALTUNTAŞ
165305001

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Murat GÖK

TEMMUZ 2020

YALOVA Üniversitesi Fen Bilimleri Enstitüsü'nün 165305001 numaralı Doktora Öğrencisi **Volkan ALTUNTAŞ**, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “**PROTEİN PROTEİN ETKİLEŞİM AĞLARINDA GÜVENİLİRLİK ANALİZLERİ**” başlıklı tezini aşağıda imzaları olan jüri önünde sunmuş ve oy birliği/oy çokluğu ile başarılı bulunmuştur.

İmza / **Kanaati**
(Kabul/Red)

Tez Danışmanı :

..... /

Jüri Üyeleri :

..... /

..... /

..... /

..... /

Teslim Tarihi : 24 Haziran 2020
Savunma Tarihi : 27 Temmuz 2020

ÖNSÖZ

Değerli bilgi ve deneyimlerinden yararlandığım tez danışmanım Sayın Doç. Dr. Murat Gök hocama, Sayın Prof. Dr.Tamer Kahveci hocama ve Dr. Öğr. Üyesi Sayın Osman Hilmi Koçal hocama teşekkür ederim. Maddi ve manevi desteklerini hiçbir zaman esirgemeyen eşim Seda Altuntaş ve Aileme şükranlarımı sunarım.

Temmuz 2020

Volkan Altuntaş
(Bilgisayar Yüksek Mühendisi)





İÇİNDEKİLER

	Sayfa No
KISALTMALAR	xi
ÇİZELGE LİSTESİ.....	xv
ŞEKİL LİSTESİ	xvii
ÖZET	xix
ABSTRACT	xxi
1. GİRİŞ	1
1.1 Tezin Amacı	2
1.2 Tezin Önemi	2
2. GENEL BİLGİLER.....	5
2.1 Protein Nedir?	5
2.2 Protein-Protein Etkileşimi ve Önemi	7
2.3 PPI Tespit Yöntemleri	8
2.3.1 In Vivo Yöntemler	8
2.3.2 In Vitro Yöntemler	10
2.3.3 In Silico Yöntemler	14
2.4 PPI Veri Tabanları ve Erişim Kaynakları	17
2.5 PPI Veri Tabanları Arasındaki Farklar	21
2.6 Literatür Özeti	23
3. MATERYAL ve METOT	27
3.1 Ağ Güvenilirliği	27
3.1.1 Ortalama Kümelenme Katsayısı (CC).....	28
3.1.2 Rastsal Yürüyüş (RW)	28
3.1.3 Difüzyon Durum Mesafesi (DSD)	29



3.1.4 Komşu Benzerliği (NS).....	29
3.2 Geliştirilen Mutasyon Algoritması.....	29
3.3 Uygunluk Ölçümleri.....	35
3.3.1 Bağlı Hata (RE).....	35
3.3.2 K-L Iraksaması (DKL)	37
3.3.3 Pearson Korelasyon Katsayısı (PCC).....	38
3.4 Geliştirilen Lyapunov Üs Tabanlı Ağ Katsayı Çıkarım Algoritması....	39
3.5 Düğüm Merkezliği	42
4. ARAŞTIRMA BULGULARI	45
4.1 Veri Seti.....	45
4.1.1 Sentetik Veri Seti	45
4.1.2 Gerçek Veri Seti	46
4.2 Deneysel Ortam.....	46
4.3 Deneysel Çalışma 1	47
4.3.1 Sentetik Veriler Üzerine Sonuçlar.....	47
4.3.1.1 Yöntem karşılaştırması.....	47
4.3.1.2 Ağ parametrelerinin değerlendirilmesi.....	48
4.3.1.3 Mutasyon oranının değerlendirilmesi.....	52
4.3.1.4 Ağ boyutunun değerlendirilmesi	57
4.3.2 Biyolojik Veriler Üzerine Sonuçlar.....	59
4.3.2.1 Organizmaların değerlendirilmesi	59
4.3.2.2 Mutasyon oranının değerlendirilmesi.....	61
4.4 Deneysel Çalışma 2	63
4.4.1 Sentetik Veriler Üzerine Sonuçlar.....	63
4.4.1.1 Ağ parametrelerinin değerlendirilmesi.....	63
4.4.1.2 Mutasyonların değerlendirilmesi.....	65
4.4.1.3 Sıralamanın değerlendirilmesi.....	66



4.4.2 Biyolojik Veriler Üzerine Sonuçlar.....	69
4.4.2.1 Organizmaların değerlendirilmesi.....	69
4.4.2.2 Sıralamanın değerlendirilmesi.....	70
5. TARTIŞMA ve SONUÇ	73
KAYNAKLAR.....	75
EKLER	
ÖZGEÇMİŞ	



KISALTMALAR

AD	: Activation Domain
BA	: Barabasi-Albert
BBSRC	: Biotechnology and Biological Sciences Research Council
BIND	: Biomolecular Interaction Network Database
CC	: Clustering Coefficient
co-IP	: Co-Immunoprecipitation
CIHR	: Canadian Institutes of Health Research
CPR	: Novo Nordisk Foundation Center for Protein Research
DSD	: Diffusion State Distance
DBD	: DNA-Binding Domain
DIP	: Database of Interacting Proteins
EMBL	: European Molecular Biology Laboratory
ER	: Erdos Renyi
GA	: Genetic Algorithm
GO	: Gene Ontology
HC	: Hill Climbing
HINT	: High-Quality Interactomes
HPID	: Human Protein Interaction Database
IBI	: Institute of Bioinformatics
IMEx	: International Molecular Exchange
K-L	: Kullback-Leibler
KU	: University of Copenhagen
LE	: Lyapunov Exponent
LEC	: Lyapunov Exponent Based Network Coefficient
MIMD	: Multiple Instruction Multiple Data
MIPS	: Munich Information Center for Protein Sequences
NCBI	: National Center for Biotechnology Information
NIH	: National Institutes of Health
NLM	: National Library of Medicine
NMR	: Nükleer Manyetik Rezonans
NMV	: Neighbor Majority Vote
NP	: NondeterministicPolynomial time
NS	: Neighbourhood Similarity
OBC	: Order Betweenness Centrality
OCC	: Order Closeness Centrality
OCLC	: Order Clustering Coefficient
ODC	: Order Closeness Centrality
OLC	: Order Load Centrality
ORA	: Order Random Centrality
PCC	: Pearson Correlation Coefficient
PPI	: Protein Protein Interaction
PSI	: Proteomics Standards Initiative
RAM	: Random Access Memory
RE	: Relative Error
RW	: Randow Walk
Y2H	: Yeast Two-Hybrid
TAP-MS	: Tandem Affinity Purification-Mass Spectroscopy
TUD	: Technische Universität Dresden



SIB : Swiss Institute of Bioinformatics
UZH : University of Zurich
WS : Watts Strogats





ÇİZELGE LİSTESİ

Sayfa No

Çizelge 4.1 Biyolojik Ağ Parametreleri.	46
Çizelge 4.2 Sentetik Ağlarda Algoritma Karşılaştırması	48





ŞEKİL LİSTESİ	Sayfa No
Şekil 1.1 Moleküler Biyolojinin Santral Dogması	6
Şekil 3.1 En İyi Mutasyon Kenar Kümesi Bulma Algoritması	34
Şekil 3.2 LEC Hesaplama Algoritması	41
Şekil 4.1 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi (RE^{RW})	50
Şekil 4.2 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi (RE^{CC})	51
Şekil 4.3 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi (DKL^{RW})	52
Şekil 4.4 Sentetik Ağlarda Mutasyon Oranının Değerlendirilmesi (RE^{RW})	54
Şekil 4.5 Sentetik Ağlarda Mutasyon Oranının Değerlendirilmesi (RE^{CC})	55
Şekil 4.6 Sentetik Ağlarda Mutasyon Oranının Değerlendirilmesi (DKL^{RW})	56
Şekil 4.7 Sentetik Ağlarda Ağ Boyutunun Değerlendirilmesi (RE^{RW})	57
Şekil 4.8 Sentetik Ağlarda Ağ Boyutunun Değerlendirilmesi (RE^{CC})	58
Şekil 4.9 Sentetik Ağlarda Ağ Boyutunun Değerlendirilmesi (DKL^{RW})	58
Şekil 4.10 Organizmaların Değerlendirilmesi	60
Şekil 4.11 Mutasyon Oranının Gerçek Ağlarda Değerlendirilmesi	62
Şekil 4.12 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi	64
Şekil 4.13 Sentetik Ağlardaki Mutasyonların Değerlendirilmesi	65
Şekil 4.14 Sentetik Ağlarda Düğüm Sırasının Değerlendirilmesi	67
Şekil 4.15 Organizmaların Değerlendirilmesi	69
Şekil 4.16 Biyolojik Ağlarda Düğüm Sırasının Değerlendirilmesi	71



PROTEİN PROTEİN ETKİLEŞİM AĞLARINDA GÜVENİLİRLİK ANALİZLERİ

ÖZET

Çoklu moleküllerin etkileşimleri biyolojik aktiviteleri ortaya çıkarmaktadır. Biyolojik ağlar, temel işlemlerin hücrel aktiviteyi nasıl kontrol ettiğini anlamak için, eşsiz bilgi birikimiyle büyük bir potansiyel sağlar. Gürültülü ağlardan elde edilen hesaplama bilgisi güvenilir değildir ve bilginin türüne göre ağ topolojisinin etkisi olabilmektedir. Protein-protein etkileşim (PPI) ağları, yüksek oranlarda yanlış pozitif ve yanlış negatif kenarlar içeren belirsiz topolojiler ve gürültüye sahiptir. Bu tez çalışmasında bir ağ topolojisindeki olası mutasyonların bu ağın içerdiği bilgiyi ve ağ üzerinde çalışan algoritmaları etkileyebileceği hipotezi önerilmektedir. Tezin amacı ağda bulunabilecek mutasyonların ağ güvenilirliğini nasıl etkilediğinin araştırılmasıdır. Tez çalışmasında, mutasyonların hem gerçek hem de sentetik ağlar üzerindeki etkisini göstermek amacıyla, belirli bir ağda etkisi yüksek yapay mutasyonlar oluşturmak için yöntem geliştirilmiştir. Hedef ağın mutasyonlar karşısındaki duyarlılığını değerlendirmek için uygunluk ölçümleri geliştirilmiştir. Ağların kırılabilirliğinin veya kararlılığının matematiksel tanımlamasını yaparak sentetik ağların ve biyolojik PPI ağlarının mutasyonlar karşısındaki kıyaslamaları gerçekleştirilmiştir. Lypunov üstelleri ile ağ güvenilirliği arasındaki ilişkiler incelenerek, aralarındaki korelasyon varlığı araştırılmıştır. Yüksek işlem yükü gereksinimi olan ağda mutasyon oluşturma yöntemi ile elde edilen güvenilirlik sonuçlarının daha hızlı ve kolay tespit edilmesini sağlamak için Lypunov üstel temelli öz nitelik geliştirilmiştir.

Anahtar Kelimeler: Protein Etkileşim Ağları; Ağ Analizleri; Güvenilirlik; Lyapunov üsteli.



PROTEIN INTERACTION NETWORKS AND NETWORK STABILITY ANALYSIS

ABSTRACT

Interactions of multiple molecules reveal biological activities. Biological networks provide great potential with unmatched knowledge to understand how basic processes control cellular activities. Computing information obtained from noisy networks is unreliable and depending on the type of information, network topology can have an effect. Protein-protein interaction (PPI) networks have indeterminate topologies and noise with high rates of false positive and false negative edges. In this thesis, it is suggested that possible mutations in a network topology may affect the information contained in this network and the algorithms running on the network. The aim of the thesis is to investigate how mutations in the network affect network reliability. In the thesis, in order to show the effect of mutations on both real and synthetic networks, a method has been developed to create highly effective artificial mutations in a particular network. Conformity measurements have been developed to evaluate the sensitivity of the target network to mutations. By comparing the fragility or stability of the networks, comparisons of synthetic networks and biological PPI networks against mutations have been performed. Relationships between Lyapunov's exponentials and network reliability were investigated to see if there was a correlation between them. Lyapunov's exponential-based attribute has been developed to ensure faster and easier detection of reliability results obtained by the method of generating a mutation in the network, which requires high processing load.

Keywords: Protein Interaction Networks; Network Analysis; Reliability; Lyapunov exponent.



1. GİRİŞ

Gürültülü ağlardan elde edilen hesaplama bilgisi güvenilir değildir ve ağ topolojisi güvenilirliği belirler. Protein-protein etkileşim ağları, yüksek oranda yanlış pozitif ve yanlış negatif kenarlar içeren belirsiz topolojilere ve gürültüye sahiptir. Bu çalışmada, bir ağ topolojisindeki mevcut mutasyonların o ağın güvenilirliği üzerine etkileri analiz edilmiştir. Ağ güvenilirliğinin değerlendirilmesi için, ağın kararlılığını matematiksel olarak tanımlayıp, ağ kararlılığına dayanarak uygunluk ölçüleri türetilmiştir. Ağda etkili kenar kümesinin aranması NP-Tam zor bir problemdir. Hesaplama problemi yeni bir meta-sezgisel optimizasyon yöntemi geliştirilerek çözülmüş ve ağda etkili mutasyonlar zaman verimli hesaplanmıştır. Açık veri tabanlarından hem sentetik hem de gerçek ağlar üzerinde yapılan deneylerimiz, yöntemimizin her tür ağ topolojisi için rakiplerinden daha iyi sonuçlar elde ettiğini göstermiştir. Yapılan çalışma protein-protein etkileşim ağlarının ve sentetik ağların güvenilirliğinin değerlendirildiği ilk çalışmadır. Çalışmamız, biyolojik - sentetik ağların kararlılığı ile ilgili önemli biyolojik sonuçları belirlemektedir. Biyolojik ve sentetik ağların topolojilerindeki mutasyonlar, ağın kararlılığı üzerinde önemli bir etkiye sahiptir. Ağ kararlılığı ağ modelinden etkilenmektedir. Ağ topolojisi, ağın deterministik ve deterministik olmayan karaktere sahip olmasını sağlayabilir ve bu durum ağın kararlılığı ile ilgili olabilir. Çalışmanın ikinci kısmında ağ topolojisinin ağın Lyapunov üsleri üzerindeki etkisini ve ağ kararlılığı ile ilişkisini araştırdık. Ağ topolojisinin etkisini değerlendirmek için biyolojik – sentetik ağlardan Lyapunov üslerini elde etme için yöntem geliştirdik. Lyapunov üsleri ile ağın güvenilirliği arasındaki ilişkiyi değerlendirmek için her bir ağın kararlılığı ölçülmüştür. Açık veri tabanlarından elde edilen hem sentetik hem de gerçek ağlarda uygulandığını gösterdiğimiz tekniğimiz, her tür ağ topolojisine uygulanabilir. Bu çalışma Lyapunov üslerinin topolojik mutasyonlar altında değerlendirildiği ve ağ analizi için kullanıldığı ilk çalışmadır. Analizlerimiz ağların Lyapunov üs tepkisi ve ağların kararlılığı ile ilgili önemli biyolojik sonuçlar sunmaktadır. Ağ topolojilerinin ağların Lyapunov üsleri üzerinde önemli bir etkisi vardır. Ağ kararlılığı ağ modelinden etkilenir ve bir ağın kararlılığı Lyapunov üsleri tarafından ölçülebilir. Lyapunov üsleri ağ kararlılığı ile güçlü bir korelasyona sahiptir ve her ikisi de ağ modelinden etkilenir. Ağların kararlılığı, ağ analizi için önemli olan zamandan ve işlem gücünden tasarruf edilerek daha hızlı bir şekilde belirlenebilir.

1.1 Tezin Amacı

Çoklu moleküllerin etkileşimleri biyolojik aktiviteleri ortaya çıkarmaktadır [1, 2]. Biyolojik ağlar, temel işlemlerin hücrel aktiviteleri nasıl kontrol ettiğini anlamak için eşsiz bilgi birikimiyle büyük bir potansiyel sağlar[3, 4]. Gürültülü ağlardan elde edilen hesaplama bilgisi güvenilir değildir ve bilginin türüne göre ağ topolojisinin etkisi olabilmektedir [5, 6, 7, 8]. Protein-protein etkileşimi (PPI) ağları, yüksek oranlarda yanlış pozitif ve yanlış negatif kenarlar içeren belirsiz topolojiler ve gürültüye sahiptir [5, 9]. Bu tez çalışmasında ilk olarak bir ağ topolojisindeki olası mutasyonların bu ağın içerdiği bilgiyi ve ağ üzerinde çalışan algoritmaları etkileyebileceği hipotezi önerilmektedir. Tezin amacı ağda bulunabilecek mutasyonların ağ güvenilirliğini nasıl etkilediğinin araştırılmasıdır. Tez çalışmasında, mutasyonların hem gerçek hem de sentetik ağlar üzerindeki etkisini göstermek amacıyla, belirli bir ağda etkisi yüksek yapay mutasyonlar oluşturmak için yöntem geliştirilmiştir. Hedef ağın mutasyonlar karşısındaki duyarlılığını değerlendirmek için uygunluk ölçümleri geliştirilmiştir. Ağların kırılabilirliğinin veya kararlılığının matematiksel tanımlamasını yaparak sentetik ağların ve biyolojik PPI ağlarının mutasyonlar karşısındaki kıyaslamaları gerçekleştirilmiştir. İkinci hipotez olarak ağların Lyapunov üstel tepkisi ile ağ güvenilirliği arasında ilişki olabileceği önerilmiştir. Ağ güvenilirliği ve Lyapunov üstel arasındaki ilişkinin araştırılması için ağın Lyapunov üstel cevabını hesaplayacak yöntem geliştirilmiştir. Geliştirilen yöntem sayesinde yüksek işlem yükü gereksinimi olan ağda mutasyon oluşturma yöntemi ile elde edilen güvenilirlik sonuçlarının, hızlı ağ analizleri ile daha kolay tespit edilebilmesi sağlanmıştır.

1.2 Tezin Önemi

Literatürde az sayıda çalışma, biyolojik sistemlerin karmaşıklığı ve sağlamlığını, belirsizlikler ya da belirsizlik koşulları altında modellemeyi ve test etmeyi amaçlamaktadır [10,11,12]. Hem matematiksel modelleme hem de deneysel gözlemler yoluyla yapılan değerlendirmeler, biyolojik sistemlerin çeşitli bozulmalara karşı oldukça dirençli olduğunu göstermektedir. Mevcut çalışmaların çoğu topolojik bozulmaları göz önünde bulundurmaktadır. Rastgele veya seçici düğümlerin ağdan çıkarılmasına odaklanılmıştır [6,7,8]. Yöntemler, çoğu karmaşık sistemin, genetik mutasyonlar, proteinlerin yanlış katlanması, genlerin veya proteinlerin yıkılmasına

neden olan çevresel deęişiklikler gibi çeşitli faktörlere baęlı olarak gürültü içerdęi şeklindeki varsayımı test etmektedir. Düzümlemlerin aędan çıkarılmasının, ortalama jeodezik uzunluk, kümelenme katsayısı ve aę çapı gibi çeşitli aę özellikleri üzerindeki etkileri deęerlendirilmiştir. Rastgele model, derece merkezlilik ve arada olma merkezlilik temelli modeller gibi çeşitli seçici düğüm silme stratejileri geliştirilmiştir.

Yapılan çalışmalar, ölçülen aę özelliklerinin saęlamlığı ile aę topolojisi arasında güçlü bir korelasyon olduğunu göstermektedir. Özellikle, ölçek bağımsız aęlar rasgele düğümlerin aędan çıkarılmasına karşı saęlamdır. Bununla birlikte, seçici düğümlerin uzaklaştırılmasına karşı kırılğandır.

Çalışmalar yoğun işlem yükü sebebi ile aęlardaki kenar mutasyonlarını dikkate almadıkları için oluşabilecek bozulmaları sınırlı olarak ele almaktadır. Çalışmalarda, bir düğüm ile ilişkili tüm etkileşimlerin, bu düğüm mevcut olduğu sürece var olduğu varsayılmaktadır. Li ve arkadaşları, protein etkileşim aęları üzerinde konsantrasyon bozulmasının etkisini incelemiştir [13]. Denge durumlarını, bozulma öncesi ve sonrasında, geliştirdikleri hesaplamalı model ile karşılaştırmışlardır. Yakın ve uzak komşuların denge durumlarını etkilemek için, protein konsantrasyonlarının nasıl deęiştiğini ölçmüşlerdir.

Az sayıda çalışmada kenar mutasyonlarının etkisi araştırılmıştır. Home ve arkadaşları aędan kenar silinmesinin, jeodezik uzunluk ve aęın büyüklüğü üzerindeki etkisini araştırmışlardır [8]. Diğer çalışmalarda, hastalık genlerinin ve ilaç hedeflerinin tanımlanması ve önceliklendirilmesi için tahmin algoritmalarının performansını test etmek amacıyla aędan rastgele kenar silme modeli kullanılmıştır [14], [15]. Zhang ve arkadaşları, [16] insan PPI aęındaki yapısal deęişikliklere göre vazgeçilmez düğümler kavramını tartışmışlardır. Etkileşimlerdeki güven puanları ile aęda seçici kenar silme bozulmaları oluşturmuşlardır. Lyapunov üsteli birçok alanda kullanım alanı bularak başarılı sonuçlar elde edebilmesine karşı biyolojik aęlarda kullanılmamıştır.

Bu tez çalışmasında, aę mutasyon modeli olarak kenar silme bozulması kullanılacaktır. Bu modelde, düğümün sahip olduğu tüm kenarların silinmesinin etkisi, ilgili düğümü silmek ile aynı etkiye sahip olduğundan, düğüm silme modelini de kapsamaktadır.

Mevcut literatürden farklı olarak, topolojik bozulmaların, global ağ özellikleri üzerine olan etkisi yerine, ağda bulunan düğümlerin komşuluk ilişkileri ile oluşan benzerlik ölçümlerine ve bu ölçümler ile yapılan tahmin sonuçları üzerine etkileri araştırılacak ve ağların güvenilirlikleri belirlenecektir. Çalışma, ağ topolojisindeki mutasyonların ağın içerdiği bilginin güvenilirliği üzerine olan etkilerin incelendiği ilk çalışmadır.

Etkileşimleri yöneten biyolojik sürecin doğal olarak belirsiz olaylardan kaynaklanması nedeniyle, biyolojik ağlar belirsiz topolojilere sahiptir [5]. Özellikle, protein etkileşim ağı verilerinde yanlış pozitif ve yanlış negatif oranlarının ağın yarısını sıklıkla aştığı bildirilmiştir [9]. Bu gibi çeşitli faktörlerden dolayı ağlarda bulunan gürültü ağın içerdiği bilgiyi ve bu bilgiyi kullanarak çıkarım yapan algoritmaları direk etkilemektedir. Bu etki sebebi ile protein etkileşim ağları ile yapılmış literatür çalışmaları hedef ağ üzerinde etkili olmakta ve tüm ağlar için kullanılamamaktadır. Yapılan bu tez çalışması ile protein etkileşim ağlarının güvenilirliklerinin belirlenmesi için yöntemler ve ölçümler geliştirilmiştir.

Protein etkileşim ağları için çeşitli veri tabanları bulunmaktadır. Farklı veri tabanlarından alınan aynı organizmaya ait PPI ağları arasında farklılıklar oluşabilmektedir. Farklılıklar, veri olarak kabul edilen literatür, deneysel çalışmayı yürüten kurum ve deney şartları gibi faktörlerden kaynaklanmaktadır [17,18].

Aynı veri tabanı içerisinde de hedef organizmaya ait PPI ağının oluşturulmasında alternatif yöntemler bulunmaktadır. İşbirliği yapılan veri tabanlarından veri transferi, yapısal - dizilimsel tabanlı veya komşuluk tabanlı gibi *in silico* ağ etkileşim tahminlerinin eklenmesi, deneysel sonuçların elde edildiği tekniklerin güvenilirliğine göre ağın filtrelenmesi gibi yöntemlerin kullanımı ile bir organizma için aynı veri tabanından farklı boyutlarda PPI ağları tedarik edilebilmektedir.

PPI ağlarında oluşan bu farklılıkların içerdiği bilgi ve üzerinden çalışan algoritmaların doğrulukları üzerine olan etkileri bilinmemektedir. Yapılan çalışma ile oluşturulan PPI ağının güvenilirliği test edilerek, geliştirilen benzerlik ölçümleri ve tahmin algoritmaları için hedef ağın durumu hakkında bilgi sağlanacaktır.

2. GENEL BİLGİLER

Önemli biyolojik aktiviteler tek bir molekülün sonucu değil, birbirleriyle etkileşime giren çoklu moleküllerin etkilerinin bir ürünü olarak ortaya çıkmaktadır. Protein-protein etkileşimlerinin belirlenmesi, ilgili proteinlere ait fonksiyonların tespit edilmesi için önemli bilgi sağlamaktadır. Genlerin ve proteinlerin büyük bir çoğunluğu işlevlerini birbirleriyle etkileşimleri sonucunda oluşturmaktadırlar.

Protein-protein etkileşimlerini incelemek için çok sayıda yöntem geliştirilmiştir. Etkileşimlerin tespitinde *in vitro*, *in vivo* ve *in silico* olarak adlandırılan 3 temel yaklaşım bulunmaktadır. *In vitro* ve *in vivo* yöntemlerin maliyet, zaman gibi sınırlamaları bulunur. *In silico* yöntemler deneysel yönlendirme ile maliyet ve zaman kazancı için geliştirilmiştir.

Yöntemler sonucunda oluşan veri setleri gürültülü, çok sayıda yanlış pozitif ve yanlış negatif değerler içermektedirler. Protein etkileşim tespit yöntemlerindeki gelişmeler hastalıkların tespit edilmesi, model organizmalara ait yolların ve protein komplekslerinin belirlenmesi gibi birçok alana doğrudan etki etmektedir. Yapılan çalışmalar sonucunda tespit edilen etkileşimler veri tabanlarında saklanmakta ve ücretsiz olarak erişilebilmektedir. Metotların hızlanması ile tespit edilen etkileşim sayısındaki artış, elde edilen bu verilerin analiz edilmesini, bir veya birden fazla metot ile sağlamasının yapılarak doğruluğunun belirlenmesini önemli hale getirmektedir.

Bu bölümde protein-protein etkileşim tespitinde kullanılan *in vitro*, *in vivo* ve *in silico* yöntemler, protein-protein etkileşim veri tabanları, içerdikleri bilgilerin sahip olduğu farklılıklar ve yapılan çalışmaya ait mevcut literatür hakkında bilgi paylaşılmaktadır.

2.1 Protein Nedir?

Bütün canlı organizmalar yaşayan hücrelerden oluşur ve temel hücresel mekanizmaları paylaşırlar. Bu hücre yapıları, bakteri veya insana ait olsa da, karmaşıklıkları bakımından farklılık içerseler de aynı yapı bloklarını (DNA, RNA ve protein) içermektedirler [19]. Protein sekansları DNA'da kodlanır ve moleküler biyolojinin santral dogması olarak anılan bir yolla sentezlenir. Her hücreye ait plan, DNA diziliminde kodlanmaktadır. DNA kopyalanır ve bir kısmı haberci RNA'ya (mRNA) aktarılır ve mRNA da proteinlere dönüştürülür (Şekil 1.1).

Bir mRNA'nın proteine tercümesinin ardından, proteinin işlevsel hale gelmeden önce geçmesi gereken birçok süreç bulunur. Protein önce katlanarak doğru üç boyutlu yapıyı oluşturmalı, daha sonra belirli bir hücresel bölgeye taşınmalıdır ve çoğu zaman ilgili proteine özgü değişiklikler yapılmalıdır [20].

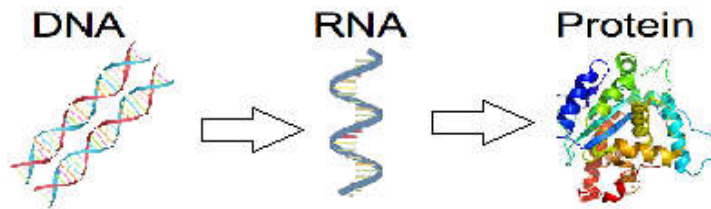
Bir hücrede birçok farklı tipte protein bulunmaktadır. Örneğin, model mikroorganizmalardan *Saccharomyces cerevisiae*'de 6000 farklı protein türü bulunmaktadır [21]. Her bir protein bir kopya veya binlerce kopya halinde sentezlenebilir.

Protein ifade seviyeleri sıkı bir şekilde düzenlenmektedir. Bu yüzden benzer hücrelerden, farklı koşullar altında sentezlenen proteinler belirgin farklılıklar içerebilir.

Proteinler, sinyal yollarının oluşturulması, hücresel süreçlerin düzenlenmesi, hücre zarı üzerinde seçici taşıyıcılık, kimyasal reaksiyonların hızlandırılması gibi hücresel süreçlerde önemli rol oynarlar [20]. Bu görevlerin çoğunda, proteinler, farklı boyutlarda kompleksler oluşturarak, birbirlerini değiştirerek ve diğer proteinleri transfer ederek birlikte çalışırlar [22].

Bu protein kompleksleri, hücrede çeşitli görevleri yerine getiren küçük makineler gibi davranmaktadırlar. Hücrenin metabolizmasına, sinyal iletimine, DNA transkripsiyon ve kopyalanmasına, DNA hasar onarımına ve daha birçok mekanizmaya katılırlar.

Bu hücresel makinelerin tam fonksiyonlarını tanımlamak, karakterize etmek ve aralarındaki etkileşimleri belirlemek, canlı bir hücrenin işlevselliğini anlamada önemlidir.



Şekil 1.1 Moleküler Biyolojinin Santral Dogması

2.2 Protein –protein etkileşimi ve önemi

Protein-protein etkileşimi (PPI) hücrenin metaboliksel gelişimi, kontrolü, hücreler arası iletişim gibi birçok biyolojik süreci temsil etmektedir [23]. Bu önemli süreçler sebebi ile PPI'lar sistem biyolojisinin temel konularından biridir. PPI'lar zincirler arasındaki kovalent olmayan temaslar ile oluşmaktadır [24]. Bu temasların sonucu olarak proteinler arası etkileşim ve ilişkiler meydana gelir. Yapısal ve işlevsel özelliklerine göre PPI'lar çeşitli şekillerde sınıflandırılabilir [25]. Etkileşim yüzeyleri temelinde homooligomerik veya heterooligomerik olarak, istikrarına göre mecburi veya mecburi olmayan ve kalıcılıklarına göre geçici veya kalıcı olarak sınıflandırılabilirler [26]. Belirli bir PPI, bu sınıfların bir kombinasyonu olabilir. Proteinler arasındaki geçici etkileşimler sinyallerin iletimini sağlayan sinyal yollarını oluşturulmakta ve kalıcı etkileşimler kararlı bir protein kompleksine dönüşmektedir [27].

In vivo ortamda yapılan deneysel çalışmalarda, doku ekstraktı veya ölü organizma kullanımı yerine canlı organizmalar veya hücreler kullanılır [28]. Proteinler *in vivo* ortamlarda fonksiyonlarını yerine getirirken izole edilmiş türler gibi davranmazlar [29]. Proteinlerin büyük bir kısmının yalnız çalışmadığı, yalnızca protein komplekslerinde faaliyet gösterebildiği ortaya çıkmıştır [30]. Doğrulanmış proteinlerin analizi, aynı hücresel süreçlerde yer alan proteinlerin art arda birbirleri ile etkileşime girdiğini ortaya koymaktadır [31]. PPI çalışmaları, proteinlerin hücre içindeki fonksiyonlarını anlamak için önemlidir. Tanımlanamayan bir proteinin fonksiyonları, fonksiyonları daha önce tespit edilmiş bir protein ile olan etkileşimleri tespit edilerek tahmin edilebilir.

PPI'ların ayrıntılı bir şekilde incelenmesi, hücresel süreçlerin anlaşılması için model geliştirme sürecini hızlandırmıştır [20]. Belli bir proteomdaki proteinlerin etkileşimlerinin modellenmesi, hücrenin biyokimyasını anlamak için fayda sağlayacaktır [20]. İki veya daha fazla proteinin belli bir fonksiyonu yerine getirmek amacı ile etkileşmesi çeşitli yollardan gerçekleşebilir. PPI'lar enzimlerin kinetik özelliklerini değiştirebilir, efektör molekülleri için yeni bir bağlanma yeri oluşturabilir, bir proteini inaktive edebilir, hücresel düzenleyici olarak rol alabilirler.

PPI'lar heterojen süreçler içerir ve oluşmaları için gerekli koşulların kapsamı geniştir. Hücredeki önemlerinin anlaşılabilmesi için, yapmış oldukları etkileşimler tanımlanmalı ve etkileşimlerin sonuçları belirlemelidir [20]. Farklı deneysel kaynaklar kullanılarak kapsamlı PPI ağları inşa edilebilmesi mümkündür. PPI deneylerinin yüksek çıktılı verileri laboratuvar doğrulamalarını zorlaştırmaktadır. Keşfedilmemiş proteinlerin işlevlerini anlamak için PPI ağlarının hesaplamalı analizi ile verilerin hacminin azaltılması ve deneysel yönlendirme artık zorunlu hale gelmiştir. Günümüzde, protein-protein etkileşimi modern sistem biyolojisinin gelişimi ve ilerlemesi için en önemli konulardan biridir.

2.3 PPI Tespit Yöntemleri

Protein-protein etkileşim tespit yöntemleri kategorik olarak *in vivo*, *in vitro* ve *in silico* yöntemler olarak sınıflandırılabilir. *In vivo* tekniklerinde belirli bir prosedür canlı organizma üzerinde [28], *in vitro* tekniklerde ise canlı organizmanın dışında kontrollü bir ortamda gerçekleştirilir [32]. *In silico* teknikler, bilgisayar simülasyonları aracılığı ile bilgisayar ortamında gerçekleştirilir [33]. Her yöntemin farklı avantaj ve dezavantajları bulunmaktadır. *In vivo* ve *in vitro* yöntemler maliyet, zaman gibi sınırlamalara sahip iken *in silico* yöntemlerde de deneysel doğrulama gereksinimi bulunmaktadır.

2.3.1 *In Vivo* Yöntemler

In vivo çalışmalarda, çeşitli biyolojik varlıkların etkileri, bir doku ekstraktı veya ölü organizma kullanımı yerine, canlı organizmalar veya hücreler (insanlar, hayvanlar ve bitkiler) üzerinde test edilir [28]. Hayvanlarda yapılan testler ve klinik çalışmalar *in vivo* araştırmanın temel unsurlarıdır. *In vivo* testler genellikle *in vitro* testlerden sonra uygulanır, çünkü bir deneyin canlı organizma üzerindeki genel etkilerini gözlemlemek için daha uygundur. *In vitro* testler bazen yanıltıcı sonuçlar ortaya çıkarabileceği için yapılan deneylerin *in vivo* etkinliğinin doğrulanması önemlidir. Literatürde *in vitro* ortamda etkin olan ve *in vivo* ortamda etkinlik kazanamayan moleküller [34,35], *in vitro* ortamda etkisiz iken *in vivo* ortamda ölümcül organizmaların olduğu [16,36] raporlanmıştır. *In vivo* çalışmaların kesin görüş sunma potansiyeline sahip olduğu düşünülse de, bu sonuçların yanıltıcı olduğu çeşitli durumlarda olabilir. Bir tedavinin kısa vadede fayda sağlarken uzun vadede zarar vermesi buna örnek olabilir.

Y2H (Yeast two-hybrid) :

Y2H yöntemi PPI'ların saptanmasında uygulanan *in vivo* bir yöntemdir. Y2H yönteminde, spesifik fonksiyonlara sahip olan iki protein bölgesi gereklidir:

- DNA'ya bağlanmaya yardımcı olan bir DNA bağlama alanı (DBD).
- DNA'nın transkripsiyonunu aktive etmekten sorumlu aktivasyon bölgesi (AD).

Her iki alan bir haberci genin transkripsiyonu için gereklidir. Y2H analizi, protein çiftleri arasında PPI'nın doğrudan tanınmasına olanak tanır. Ancak, yöntem çok sayıda yanlış pozitif etkileşimli sonuca neden olabilir. İlave olarak, Y2H yönteminde yanlış negatif sonuç olarak elde edilen pek çok gerçek etkileşim tespit edilemeyebilir. Y2H analizinde, etkileşimli proteinler çekirdeğe yerleştirilir. Proteinler işlevlerini yerine getirmek için post-translasyonel modifikasyonlara ihtiyaç duyar. Bu yüzden Y2H deneyinde proteinlerin normal davranması ve etkileşimde olma olasılığı düşer. Proteinler doğal fizyolojik ortamlarında olmadığında etkileşime uygun şekilde katlanamazlar [37] – [39].

Sentetik letalite (Synthetic lethality) :

Sentetik letalite fiziksel etkileşimden ziyade fonksiyonel etkileşimlere dayanır. Sentetik letalite; varyasyon, çevresel değişimler ve mutasyonlar gibi rastgele olaylara rağmen fenotipik kararlılığa izin veren mekanizmaları anlamaya çalışan bir *in vivo* genetik tarama türüdür. Bu yöntemde, tek başına uygulandığında zarar vermeyen fakat belirli koşullar altında birleştirildiğinde öldürücü olan iki veya daha fazla gende mutasyonlar üretilir. Sentetik letalite ile tespit edilen sonuçlar proteinler arasında fiziksel etkileşim gerektirmez. Bu tür ilişkiler fonksiyonel etkileşim olarak adlandırılır [40] – [44].

Afinite etiketleri (Affinity tags) :

Rekombinant proteinlerin ekspresyonu ve saflaştırılması, proteinlerin yapısını ve işlevini karakterize etmek için kullanılmaktadır. Yüksek saflıkta yeterli konsantrasyon elde etmek için ilgili proteini saflaştırmaya ihtiyaç vardır. Afinite bazlı protein saflaştırması, genetik olarak kaynaştırılmış bir afinite etiketinin seçici bağlanmasından yararlanır.

İlk olarak; hücreler, etikete kaynaşmış bir yem proteinini kodlayan bir plazmid ile transfekte edilir. Uygun bir ekspresyon döneminden sonra, hücreler parçalanır ve bağlı proteinler ile birlikte etiketli yem, katı bir desteğe bağlı spesifik bir kimyasal veya biyolojik ligand kullanılarak izole edilir. Ayırıştırılan proteinler daha sonra jel-elektroforez ile ayrılır ve spesifik olarak bağlı proteinler kütle spektrometresi ile tanımlanır. Kullanılacak etiketin nihai seçiminden önce dikkate alınması gereken kendine özgü avantajları ve dezavantajları vardır. Bu uygulama; özgüllük, çözünürlük, bağlama ve elüsyon koşulları için şartlara bağlıdır. Bu yöntemler, yüksek afinite ve yavaş kinetikler ile etkileşime giren proteinlere karşı yanlıdır. Afinite kromatografisine dayalı yöntemler, geçici protein etkileşimlerinin saptanması için uygun olmayabilir. Protein aktivitesinin translasyon sonrası kontrolünde yer alan geçici kompleksler saptanamazlar. Hücre içi ortam ile seyreltme tüpü arasındaki konsantrasyon farkları hatalı sonuçların sebebi olarak gösterilmektedir [45] – [51].

Kimyasal çapraz bağlama (Chemical crosslinking) :

Protein-protein etkileşimlerini yapısal olarak çözmek için alternatif bir strateji olan kimyasal çapraz bağlama, enzimatik sindirim ve reaksiyon ürünlerinin kütle spektrometresi ile analizidir. Çapraz bağlama, bir protein içindeki fonksiyonel grup çiftlerini kovalent olarak bağlayarak, bir dizi yapısal olarak tanımlanmış etkileşimin oluşturulmasını sağlar. Kimyasal çapraz bağların sağladığı kısıtlamalardan yararlanarak, düşük çözünürlüklü 3 boyutlu yapıların çıkarılmasında temel oluşturacak bir protein veya protein kompleksi içinde mesafe haritaları oluşturulabilir. Bu deneylerde kimyasal sabitleştirme gereklidir. Doğal durumunda hücrelerden kromatin yakalanamaz. Çapraz bağlı kromatin, bazı antikör epitoplarının antikör-kromatin bağlanmasını etkileyerek maskeleyebilir [52] – [55].

2.3.2 In Vitro Yöntemler

In vitro çalışmalar, normal biyolojik ortam dışındaki (canlı organizma dışı) mikroorganizmalar, hücreler veya biyolojik moleküller ile yapılır [28]. Test tüpü deneyleri olarak adlandırılan bu çalışmalar; test tüpleri, şişeler, petri tabakları ve mikrotiter plakalar gibi ekipmanlar ile laboratuvar ortamında yapılır [56].

Biyolojik çevreden izole edilmiş bir organizma kullanılarak yapılan çalışmalar, tüm organizmalarla yapılabilecek analizlerden daha detaylı ve daha uygun bir analiz yapmaya izin verir. Bununla birlikte, *in vitro* deneylerden elde edilen sonuçlar, deneyin bütün bir organizma üzerindeki etkilerini tam olarak ve doğru bir şekilde ortaya koymayabilir [34,35,36,57].

In vitro deneylerin zayıf yönlerinden biri, organizmanın hassas hücresel koşullarının oluşturulamamasıdır. *In vitro* deneylerin aksine *in vivo* çalışmalar biyolojik ortamda yapılan araştırmalardır. *In vitro* çalışmaların doğrulanması için *in vivo* çalışmalara başvurulur [32, 58].

TAP-MS (Tandem affinity purification-mass spectroscopy) :

Tandem afinite arıtma-kütle spektroskopisi TAP-MS, kromozomal lokusu üzerindeki ilgilenilen proteinin çift etiketlenmesine dayanır. Bu işlemi iki aşamalı bir saflaştırma ve kütle spektroskopik analizi takip eder. Kütle spektrometrik protein tanımlaması ile birleştirilmiş tandem-afinite saflaştırması, biyolojik olarak ilgili protein komplekslerinin bileşimini belirlemek için güçlü bir yaklaşımdır. TAP etiketi kaynaştırılan bir protein, hücrelerde ifade edilir ve TAP-etiketli proteini birleştiren protein komplekslerini saflaştırmak için bir yem olarak kullanılır. TAP-etiketlemenin önemi, çok çeşitli protein komplekslerini tanımlama ve *in vivo* olarak var olan monomerik veya multimerik protein komplekslerinin aktivitesini test etme yeteneğidir. TAP'ın kütle spektroskopisi ile kullanımı, protein etkileşimlerini ve protein komplekslerini tanımlamada güç kazandırmaktadır [59] – [63].

Afinite kromatografi (Affinity chromatography) :

Afinite kromatografi yüksek oranda duyarlıdır ve proteinlerdeki en zayıf etkileşimleri bile tespit edebilir. Etkileşim genellikle tersinirdir. Saflaştırma, bir yüzeyine sabitlenmiş olan moleküllerden (ligand) biri ile bifazik bir etkileşim yoluyla elde edilirken; eşi (hedef), karmaşık bir karışımın parçası olarak hareketli bir fazdadır. Yakalama adımı genellikle yıkama ve elüsyon ile takip edilir. Sonuçta yüksek oranda saflaştırılmış protein elde edilir. Hücresel sisteme dâhil olmasalar bile, proteinler arasındaki yüksek belirlilik nedeniyle yanlış pozitif sonuçlar ortaya çıkar. Bu nedenle elde edilen sonuçları çapraz kontrol etmek ve doğrulamak için başka yöntemler kullanılması gerekir [64].

Eş-immüno-çökeltme (Co-immunoprecipitation) :

Eş-immüno-çökeltme (co-IP), spesifik bir hedef proteine bağlı olan proteinleri dolaylı olarak yakalamak için kullanılır. Hedef protein-spesifik antikorların kullanımı ile fizyolojik olarak ilgili protein-protein etkileşimlerini tanımlayan bir tekniktir. Bilinen protein (antijen) yem proteini olarak adlandırılır ve etkileştiği protein av protein olarak adlandırılır. Hücreler denatüre edici olmayan koşullar altında tamamen parçalandıktan sonra, birbirine bağlanan proteinler korunur. Proteinlerin, doğal etkileşimlerinde gerekli olabilecek karmaşık hücresel bileşenler karışımında mevcut olan bütün bir hücre ekstresi kullanılır. Ökaryotik hücrelerin kullanımı post translasyonel modifikasyonu mümkün kılar. [65] – [67].

Mikrodizi (Microarrays) :

Bir protein mikrodizisi; üzerinde çeşitli şekillerde, proteinlerin ayrı yerlere düzenli şekilde bağlandığı bir cam parçasıdır. Protein mikrodizi gelişiminin ardındaki amaç, otomatik işlemlerle paralel olarak çok sayıda belirleme yaparak verimli protein analizleri gerçekleştirmektir. Protein mikrodizileri; proteinleri tespit etmek, ekspresyon seviyelerini izlemek ve protein etkileşimlerini ve fonksiyonlarını hızlı bir şekilde araştırmak için yaygın olarak kullanılmaktadır. İlk olarak hedef protein dizilimin fragmanlarına floresan işaretleyiciler eklenir. Bunların çipin problemleri ile reaksiyona girmelerine izin verilir. Etkileşen dizilimlere ait fragmanlar çipin problemlerine bağlanır. Çip yıkanarak temizlenir. Etkileşim kuran dizilimler bir lazer ışını ile tespit edilir. Bu teknik, çok sayıda dizilimin hızlı ve eş zamanlı olarak tanımlanmasını sağlar [68] – [72].

Parça tamamlama (Fragment complementation) :

Protein-parça tamamlama deneyleri protein-protein etkileşimlerini saptamak için oluşturulmuş analiz ailesidir. Basit ve doğrudan yollar sağlar. Herhangi bir moleküler ağırlığa sahip proteinler arasındaki PPI'yı tespit etmek ve endojen seviyelerinde ifade etmek için kullanılabilir. Kütle spektroskopisi kullanarak protein tanımlaması için peptid parmak izi ve geniş yelpaze (shotgun) proteomik yöntemleri kullanılır.

Peptit parmak izi ile az protein içeren saflaştırılmış örneklerde çalışılabilir. Geniş yelpaze proteomikde ise birçok protein içeren örnekler kullanılabilir. Karmaşık karışımları analiz etmek için güçlü bir yöntemdir. İncelenen fragmanların, bağlayıcı proteinlerin yokluğunda kendiliğinden birleşmemesi metodun doğruluğu için önemlidir. Bu durum, yanlış pozitiflere ve yöntemin yararsız hale gelmesine neden olabilir [73] – [76].

Faj ekranı (Phase display) :

Faj ekranı eski ve güçlü yöntemlerden biridir. Teknoloji; bakteriyofaj kaplama proteinlerine kaynaşmış polipeptidlerin, kodlama geni de içeren faj partikülleri üzerinde ifade edilebilmesine dayanmaktadır. Bu şekilde, genotip ve fenotip arasında fiziksel bir bağlantı kurulur. Basit moleküler biyoloji yöntemleri ile DNA kodlu peptidlerin veya proteinlerin çeşitli kütüphaneleri üretilebilir. Faj gösterimli kütüphaneler, bir bakteri konakçı kullanılarak çoğaltılabilir. Herhangi bir amino asit dizisi, DNA'nın faj partiküllerinin içinde dizilmesiyle kolayca tespit edilebilir [77] – [83].

X-ışını (X-ray) :

X-ışını kristalografisi protein yapılarının atomik düzeyde görselleştirilmesini ve bu şekilde protein fonksiyonunun anlaşılmasını sağlayan, yüksek çözünürlüklü bir mikroskopidir. Proteinlerin diğer moleküller ile nasıl etkileştiği ve enzimlerin durumlarındaki değişimler izlenebilir. Bir kristal yapısından, protein-protein etkileşimini ayırt etmek genellikle kolay değildir. Bunları ayırt etmek için bir dizi hesaplama gereklidir ve hata oranı yüksektir [84] – [87].

NMR (Nükleer manyetik rezonans) :

NMR spektroskopisi ile protein-protein etkileşiminin analizinde, bağlanma ara yüzünün yeri, protein etkileşim tayini için önemli bir özelliktir. NMR spektroskopisinin temeli; güçlü bir manyetik kaynağın yönlendirdiği manyetik olarak aktif çekirdeklerin, kimyasal çevreleri tarafından yönetilen karakteristik frekanslarda elektromanyetik radyasyonu absorbe etmesine dayanmaktadır. Yüksek alanlı spektrometre ve kriyojenik probalar ile tekniğin duyarlılığı artmıştır. Büyük ve biyolojik olarak ilişkili biyomoleküler kompleksleri yüksek doğrulukla karakterize etmek mümkündür. [88] – [90].

2.3.3 *In Silico* Yöntemler

In silico çalışmalar, bilgisayar veya simülasyonlar ile gerçekleştirilen çalışmaları ifade etmektedir [33]. Adını bilgisayar teknolojilerinde yoğun kullanılan silikona atfen almıştır. *In silico* çalışmaları, pahalı laboratuvar deneyleri ve klinik araştırmalara olan ihtiyacın azaltılması için avantaj sağlamaktadır.

In silico teknikler; protein tasarımı, hücre davranışlarının modellenmesi, biyoproses optimizasyonu ve benzeri alanlarda kullanılabilmektedir [91, 92]. Deneysel yaklaşımla etkileşim tespitine destek sağlama ve alternatif oluşturmak için çeşitli *in silico* yöntemler geliştirilmiştir.

In silico tespit; yapı, dizilim, gen komşuluğu, gen birleşimi, filogenetik ağaç ve gen ifadesine dayalı yaklaşımları içerir.

Yapı tabanlı yaklaşımlar (Structure based approaches) :

Yapıya dayalı yaklaşımda iki protein benzer bir yapıya sahipse, etkileşim kurulan proteinler arasında da aynı benzerliğin olacağı kabul edilmektedir. Etkileşim tahmini için ilk olarak proteinlerin yapıları çeşitli yöntemler ile oluşturulmalıdır. Yapıları ve aralarındaki etkileşimleri bilinen proteinler ile bir havuz oluşturulur. Etkileşimi tahmin edilmek istenen hedef proteinin yapısal olarak benzediği proteinler belirlenir. Belirlenen proteinlerin sahip olduğu protein-protein etkileşimleri, hedef protein içinde geçerli olduğu kabul edilir.

Protein yapısı ve yapılar arası benzerliklerin skorlanması için çeşitli algoritmalar geliştirilmiştir. Yapılan işlem protein yapı benzerliği ile protein – protein etkileşim tahminidir [93] – [97].

Dizilim tabanlı yaklaşımlar (Sequence based approaches) :

Protein sekansı verilerindeki hızlı artış, sekans bilgisi kullanılarak tahmin yapılması için gerekli olan verilerin oluşmasını sağlamıştır. Bilinen protein etkileşimleri protein homoloji bilgileri ile birleştirilerek bilinmeyen etkileşimler tahmin edilir. Bu yaklaşım, bir türün içinde bulunan bir etkileşimin diğer türlerdeki etkileşimi anlamak için kullanılabileceği kavramına dayanmaktadır. Protein birincil yapıları, ikincil yapıları, homoloji, gen ontolojileri gibi bilgilerin farklı teknikler ile birleştirilmesi ve daha güçlü tahminler yapılması için çeşitli algoritmalar bulunmaktadır [98] – [102].

Gen komşuluğu (Gene neighbourhood) :

Operonlar gibi işlevsel olarak ilişkili proteinler, genomlardaki yakın bölgelerde organize olma eğilimindedirler. Komşuluk ilişkisi birden fazla genom boyunca korunduğunda ilgili genler tarafından kodlanan proteinler arasındaki fonksiyonel bağlantının olma olasılığı daha yüksek olacaktır.

Bu yöntemde genlerin komşuluk ilişkileri bilinen protein-protein etkileşim ilişkileri ile birleştirilerek bilinmeyen etkileşimler tahmin edilmeye çalışılır. Komşuluk ilişkilerinin varlığı ve göreceli gen oryantasyonundan bağımsız olduğu deneysel sonuçlarla doğrulanmıştır. Komşuluk ilişkileri, transkripsiyon düzenleyicileri için hedef süreçleri ve düzenleyici özellikleri tahmin etmek için kullanılmaktadır. Gen komşulukları; bakteriyel genomlarda korunduğu için doğrudan tespit edilebilmesine karşın, diğer genomlarda zor bir problem olma özelliğini korumaktadır [103, 104].

Gen birleşimi (Gene fusion) :

Füzyon gen, iki ayrı gen veya parçalarının çeşitli sebeplerle birleşiminden oluşan hibrid bir gendir. Bu değişim sonucu üretilen füzyon proteinleri, bazı kanser türlerinin gelişmesine yol açmaktadır. Füzyon genleri ve proteinleri, kanserin tanı ve tedavisinde önemli rol oynamaktadır.

Yöntem füzyon proteinlerinin temel özelliği olan etkileşime dayanmaktadır. Bir organizmada tekil olan proteinler başka bir organizmada etkileşim ile protein kompleksleri oluşturabilir. Bu durumun tespit edilmesi, ilgili proteinler için fonksiyonel birlikteliklerin belirlenmesini sağlayacaktır. Metabolik yollara katılan proteinlerde birleşim daha sık rastlanmaktadır. Bu yöntemde farklı genomlara ait olan düzenleme bilgileri kullanılarak, protein-protein etkileşimleri tahmin edilir. Yöntem sadece alan düzenlemesinin mevcut olduğu füzyon proteinlerine uygulanabilir [105] – [107].

Filogenetik ağaçlar (Phylogenetic trees) :

Proteinin tahmini evrim süreci hakkında bilgi verir. Bu yöntemde temel olarak bu bilgi kullanılarak proteinler arası etkileşimler tahmin edilir. Yöntemdeki temel kabul, etkileşimli proteinlerin birlikte evrimleşmesi sebebi ile benzer filogenetik ağaçlar oluşturacağıdır.

Benzer filogenetik ağaçlar protein-protein etkileşiminin göstergesidir. Araştırılan proteinleri içeren organizmalar için filogenetik ağaçlar ve benzerlik matrisleri oluşturulur. Benzerlikler arasındaki ilişkiler ile proteinler arasındaki etkileşim tahmin edilir. Yüksek benzerlik, oluşturulan filogenetik ağaç benzerliğini dolayısıyla proteinlerin etkileşimli olduğunu göstermektedir [108] – [110].

Gen ifadesi (Gene expression) :

Bu yöntemde genlerin ifade modelleri ile kodladıkları proteinlerin etkileşimleri arasındaki ilişki kullanılarak protein-protein etkileşim tahmini yapılır. Gen ifadesi; belirli bir genin, farklı deney koşulları altında ve zaman aralıklarında bir hücre, doku veya organizma içinde ifade edildiği seviyenin belirlenmesidir.

Büyük ölçekli gen ifadesi ve protein-protein etkileşim verilerinin istatistiksel analizi ile birlikte, ifade edilen genler tarafından kodlanan protein çiftlerinin, birbirleriyle rastgele proteinlerden daha sık etkileştiği gösterilmiştir. Bu kabul ile gen ifade ve protein-protein etkileşim bilgileri birleştirilerek yapılan analizler ile bilinmeyen etkileşimler tespit edilmeye çalışılır.

Bu yöntem protein etkileşimini ortaya çıkarmanın dolaylı bir yolu olduğu için tespit edilen protein etkileşimlerinin doğruluğu hakkında çeşitli görüşler bulunmaktadır. Diğer deneysel yöntemlerle birlikte doğrulama amacı ile de kullanılabilir [111] – [113].

Metin madenciliği (Text mining) :

Metin madenciliğine dayalı hesaplama metodolojileri, proteinler ve protein-protein etkileşimlerinin tahmini için literatür ve biyolojik veri tabanlarını metin madenciliği metotları ile analiz etmektedir [114].

Literatürdeki üstel artış göz önüne alındığında, metin madenciliği tekniklerinin daha fazla önem kazanacağı görülmektedir. Tahmin edilen protein ve etkileşimlerin hatalarının azaltılması ile doğruluğun artırılması üzerine olan çalışmalar yöntemin uygulanabilirliği açısından önem arz etmektedir [115].

2.4 PPI Veri Tabanları ve Eriřim Kaynakları

PUBMED (Public Medicine) :

Yařam bilimleri ve biyomedikal konularla ilgili MEDLINE veri tabanına ve özetlere erişim saęlayan ücretsiz bir arama motorudur. Ulusal Saęlık Enstitüleri'ndeki Amerika Birleşik Devletleri Ulusal Tıp Kütüphanesi (NLM), Ulusal Biyoteknoloji Bilgi Merkezi (NCBI) tarafından saęlanan bir hizmettir. Bu veri tabanı, Entrez sisteminin bilgi alma parçası olarak kullanılmaktadır. PubMed, tüketiciler için kolay okunan özetler ve tam teknik raporlar ile klinik etkinlik araştırması incelemelerinde uzmanlařmıştır [116,117].

DIP (Database of Interacting Proteins) :

Etkileřen proteinlerin veri tabanı (DIP), deneysel olarak belirlenen protein-protein ikili etkileřimlerini içeren bir veri tabanıdır. Etkileřimler hakemli literatürden incelenir ve uzman sorumlular tarafından veri tabanına manuel olarak girilir. Manuel kayıtlara ek olarak, DIP'deki etkileřimler hesaplama deęerlendirmesine tabi tutulur. Deęerlendirme algoritmaları, bildirilen etkileřimin özelliklerini, DIP etkileřimlerinin en güvenilir özellikleriyle karşılařtırarak biyolojik olarak anlamlı olup olmadıklarını denetler. Deęerlendirme sonuçları etkileřimin bilgi sayfasında görüntülenir. Protein-protein etkileřim detaylarının kataloglanmasının ötesinde; DIP protein fonksiyonu ve protein-protein ilişkilerini anlamak, etkileřen proteinlerin aęlarının özelliklerini incelemek, protein-protein etkileřimlerinin kestirimlerini ölçmek ve protein-protein etkileřimlerinin evrimini incelemek için yararlıdır. Etkileřimli verilerin ana kamu saęlayıcılarından oluşan bir grup olan Uluslararası Moleküler Etkileřim Deęişim Konsorsiyum'unun (IMEx) [118] bir üyesidir [119,120].

IntAct (An Open Source Molecular Interaction Database) :

IntAct protein etkileřimlerinin depolanması, sunulması ve analizi için açık kaynak veri tabanı ve analiz araçları sunmaktadır. Web ara yüzü, protein etkileřimlerinin hem metinsel hem de grafiksel gösterimlerini sunar ve proteinlerinin GO(Gene Ontology) [121] ek açıklamalarına ulařılabilir. Tüm etkileřimler literatür madencilięi veya doğrudan kullanıcı gönderimlerinden türetilmiştir [17, 122].

MINT (The Molecular Interaction Database) :

Moleküler etkileşim veri tabanı (MINT), deneysel olarak elde edilen literatürden oluşturulan PPI verilerinin veri tabanıdır. Etkileşimlere ilave olarak, her etkileşim için kanıt ağırlığı bilgisi de sunulur. Web ara yüzü; kullanıcılara etkileşimli verileri arama, görselleştirme ve indirme olanağı sağlar. MINT, Uluslararası Moleküler Değişim Konsorsiyumu'nun (IMEx) [118] üyelerinden biridir. İyileştirme ve veri alışverişi için Proteomik Standart Girişimi (PSI-MI) standartlarının Moleküler Etkileşim Ontolojisi'ni benimsemektedir [17, 123, 124].

MIPS (Munich Information Center for Protein Sequences) :

Münih Bilgi Merkezi Protein Sekansları (MIPS), Neuherberg, Almanya'daki Biyoenformatik Enstitüsü'nde (IBI) bulunan bir araştırma merkezidir. Genom yönelimli biyoinformatik üzerine odaklanılmıştır. Bakteri, mantar ve bitki genomlarının sistematik karşılaştırmalı analizleri de sunulmaktadır. Yüksek ilişki düzeyine sahip seçilmiş veri kümeleri manuel kontrollere tabi tutulur, yüksek verimli veri kümeleri ise otomatik araçlarla etiketlenir. Alternatif olarak hesaplama araçlarının kullanılması ve ikincil bilgiler veri tabanı oluşturulması, bilginin daha verimli bir şekilde kullanılmasını sağlamaktadır [125, 126].

BIOGRID (The Biological General Repository for Interaction Datasets) :

Yüksek çıktılı veri kümelerinden ve birincil literatürden alınan, kapsamlı iyileştirme çalışmaları ile derlenen etkileşimleri içerir. Tüm veriler arama dizini aracılığıyla ücretsiz olarak temin edilmekte ve standart formatlarda indirilebilmektedir. Bünyesinde protein-protein etkileşimleri, genetik etkileşimler, kimyasal etkileşimler ve post-translasyonel modifikasyon bilgileri bulunur. Kanada ve Amerika Birleşik Devletleri'nde barındırılmaktadır. BBSRC, NIH ve CIHR tarafından finanse edilmektedir. Uluslararası Moleküler Değişim Konsorsiyumu'nun (IMEx) [118] üyesidir. İyi bir ara birime sahiptir. Kullanıcılar ilgi alanlarına giren proteinleri veya yayınları araştırıp, ek açıklamaların yanı sıra, rapor edilen, literatürde bulunan ve büyük ölçekli kontrol çalışmaları ile derlenen verilere kolaylıkla ulaşabilirler. Dinamik etkileşim ağ görüntüleyicisi, tüm genetik ve protein etkileşim verilerinin yanı sıra biyoaktif bileşikler ve bunların yerleşik hedefleri için kolay gezinme ve filtreleme sağlar [17], [127,128].

STRING (Search Tool for the Retrieval of Interacting Genes) :

Protein etkileşimlerinin hem fiziksel hem de fonksiyonel ilişkilerini içerir. Deneysel veriler, hesaplamalı tahmin yöntemleri ve genel metin koleksiyonları dâhil olmak üzere çok sayıda kaynaktan bilgi bulunmaktadır.

Veriler ağırlıklandırılmış ve tüm protein etkileşimleri için bir güven skoru ile bütünleştirilmiştir. GO [121], Pfam [129] ve KEGG [130] gibi fonksiyonel sınıflandırma sistemleri kullanarak protein listelerine fonksiyon bilgileri eklenmektedir.

Novo Nordisk Foundation Center for Protein Research (CPR), European Molecular Biology Laboratory (EMBL), University of Copenhagen (KU), Swiss Institute of Bioinformatics (SIB), Technische Universität Dresden (TUD) ve University of Zurich (UZH) gibi akademik kurumların bir konsorsiyumu tarafından geliştirilmiştir. Protein etkileşimlerinin hızlı ve kolay bir şekilde sunulduğu web ara birimi bulunmaktadır [131,132].

HPID (Human Protein Interaction Database) :

HPID; insan proteomundaki her bir protein için domain mimarisi, translasyon sonrası değişiklikler, etkileşim ağları ve hastalık birliği ile ilgili bilgileri görsel olarak göstermek ve entegre etmek için merkezi bir platformdur. Mevcut yapısal ve deneysel verilerden hesaplanan insan protein etkileşimlerine, proteinler arasındaki tahmin edilmiş etkileşimlere ve kullanıcılardan gelen yeni insan protein etkileşimi verilerine bir depo sağlamak üzere tasarlanmıştır. Tüm bilgiler; yayınlanmış verileri okuyan, yorumlayan ve analiz eden uzman biyologlar tarafından incelenmektedir. Web tabanlı ara yüzü sayesinde kullanıcılar protein etkileşimi ağlarını görselleştirip analiz edebilmektedir [133,134].

BIND (The Biomolecular Interaction Network Database) :

Proteomik kaynak oluşturma, diğer kaynaklardan veri madenciliğini sağlayacak bir platform oluşturma ve karmaşık moleküler etkileşimlerin görselleştirmelerini sunabilen bir platform oluşturma amaçları ile tasarlanmıştır. Biyomoleküler etkileşim, kompleks ve yolak bilgilerini içerir. Kayıtları sorgulamak ve görüntülemek için web tabanlı bir ara yüz bulunmaktadır. İlgilenilen bölgelere odaklanmaya yardımcı olacak bir etkileşim ağı kümeleme aracına sahiptir [135, 136].

HINT (High-Quality Interactomes) :

Farklı organizmalar için yüksek kaliteli protein-protein etkileşimlerinden oluşan bir veri tabanıdır. Veriler farklı kaynaklardan derlenir ve hatalı veya düşük kaliteli etkileşimleri gidermek için hem sistematik hem de manuel olarak filtrelenir.

Projenin amacı, hücresel ağların global özelliklerini analiz etmenin yanı sıra, belirli proteinler veya yollar hakkında spesifik hipotezler üretmek için kullanılabilir veri kümesi oluşturmaktır [137].

HitPredict (A Database of Quality Assessed Protein-Protein Interactions) :

DeneySEL olarak tanımlanmış fiziksel protein-protein etkileşimlerinin güven puanlarıyla birlikte yayınlandığı bir kaynaktır. Veri entegrasyonu sırasında karşılaşılabilecek sorunları çözmek için manuel doğrulama kullanılmıştır.

Etkileşime giren proteinlerin sekansı, yapısı ve işlevsel ek açıklamalarından elde edilen kanıtlar kullanılarak, hesaplanan güvenilirlik puanına dayalı olarak etkileşimlere bir güven düzeyi atanır. Protein aramak ve etkileşimlerini görselleştirmek için web ara yüzüne sahiptir, veriler çevrimdışı kullanım için indirilebilir [138, 139].

GeneMANIA :

GeneMANIA genleri analiz etmek ve gen fonksiyonlarını önceliklendirmek için geliştirilmiş web tabanlı bir kaynaktır. İşlevsel olarak benzer genler tespit edilebilir ve gen fonksiyonları tahmin edilebilir. Veriler BioGRID gibi farklı kaynaklardan derlenir. Cytoscape [140] yazılımı ile entegre çalışabilmektedir [141].

CORUM (The Comprehensive Resource of Mammalian Protein Complexes) :

CORUM memeli organizmalarından deneySEL olarak karakterize edilmiş protein komplekslerine ait veriler içerir. Veriler elle küratörlüğe tabi tutulmaktadır. Birimler arasında bilinen etkileşimleri gösteren Cytoscape [140] tabanlı grafiksel bir araç içermektedir. İleri çalışmalar için tüm veri setlerini farklı formatlarda indirilebilir [142].

iRefIndex :

Proteinlere ait etkileşim verileri farklı veri tabanlarında bulunabilmektedir. iRefIndex bu tür verileri aramayı kolaylaştırmak için tasarlanmıştır. iRefIndex veri kaynağı olarak BIND, BioGRID, CORUM, DIP, HPRD, IntAct, MINT gibi birçok birincil etkileşim veri tabanında mevcut olan protein etkileşimlerini kullanır.

Her bir etkileşim ve etkileşime katılan proteinler için benzersiz bir anahtar üretilmektedir. Anahtar üretim esnasında hatalı biçimlendirilmiş, kullanımdan kaldırılmış, belirsiz veya sorunlu referansların belirlenebilmesi için bir puanlama sunulmaktadır. İleri çalışmalar için tüm veri setlerini indirilebilir [143].

2.5 PPI Veri Tabanları Arasındaki Farklar

Veri tabanları arasında kullanım kolaylığı, veri görselleştirme ve kapsayıcılık gibi konularda farklılıklar bulunmaktadır.

Deneyisel ve hesaplamalı etkileşimleri sunan STRING veri tabanı mükemmel grafikler sağlar, kullanımı kolaydır, başarılı grafiksel etkileşim haritasına sahiptir, kanıtlar ayrıntılı olarak açıklanır.

Literatür küratörlüğünden veya doğrudan kullanıcı sunumlarından türetilmiş tablolardan oluşan IntAct, sonuçların sunumunda anlaşılması zor kodlamalar kullanmaktadır. Kolay tür seçimi yoktur ve grafiksel arabirime sahip değildir.

Literatürden ve yüksek verimli veri setlerinden elde edilen verileri paylaşan BIOGRID veri tabanı, mayalar için zengin veri kaynağına sahiptir. İyi bir ara yüze sahiptir, genellikle memeli genleri için diğer veri tabanlarına kıyasla daha az veri içerir. Referans ve açıklamalar iyi bir şekilde sunulmaktadır.

Kullanıcı arabirimi ve sonuçların kullanım kolaylığı konularında yaşanan farklılaşmaya veri tabanlarının içeriklerinde de karşılaşılmaktadır. Her veri tabanında yer alan canlı tür sayısı, protein sayısı, etkileşim sayısı farklı olmakla beraber aynı kaynaktan türetilen bilgi sayısında farklılıklar da gözlenmektedir. En az iki veri tabanı tarafından paylaşılan 14.899 yayının, 5.782'si (%39) farklı veri tabanlarında farklı sayıda etkileşimle bildirilmiştir [17].

Veri tabanlarında bulunan düşük oranlı farklılıklar, bu veriyi kullanan algoritmaların sonuçlarını verideki farklılıktan daha fazla oranda etkileyebilir [144]. Veri tabanlarının aynı yayından aldığı bilgilerde de önemli ölçüde farklılıklar olabilmektedir. Bu durumun tanımlayıcı eşlemesi kaynaklı olduğu düşünülmektedir. Farklı veri tabanlarının aynı yayından küratörlük ile elde ettiği bilgide, etkileşimlerin %42 ve proteinlerin %62 oranında aynı olduğu görülmüştür [145].

Yapılan bir diğer deneysel çalışmada Y2H kullanılarak insan proteinleri arasında 2.754 etkileşim tespit edilip veri tabanları ile kıyaslanmıştır. Yapılan kıyaslamaya göre, HPRD 2.371 etkileşim, IntAct 2.671 etkileşim ve MINT 2.463 etkileşim bildirildiği tespit edilmiştir [146].

Elde edilen sonuçların aynı olmaması, farklı güven kümesi veya eşik kullanımı gibi deneysel verinin farklı yorumlanmasını sağlayan parametrelerin aynı şekilde kullanılmamasından kaynaklı olabilir.

Veri tabanları arasındaki farklılıklar ve tespit yöntemleri incelendiğinde, gelişmelere bağlı olarak, protein-protein etkileşimlerinin saptanması için birçok yöntem geliştirildiği ve geliştirilmeye devam edeceği görülmektedir. Yöntemlerde kullanılan prensipler ve performansları çeşitlilik göstermektedir. Bazı durumlarda yöntemler birbirlerini tamamlamakta veya birlikte kullanılmaktadır. Eski yöntemlerin performanslarının artırılmasının yanında yeni yöntemler de geliştirilmiştir. Yeni yöntemler yüksek miktarda veri üretmekte ve bu verilerin istatistiksel olarak işlenmesi gerekmektedir. Deneylerin yürütülme şekli ve yorumlanması da yeni yöntemler için önemli kriterler arasında yer almaktadır.

Yeni veya eski kullanılabilir tüm yöntemler yüzde yüz doğrulukla etkileşimleri tahmin edememektedir. Birçoğu yüksek hata oranlarına sahiptir. Düşük hata oranlı yöntemler genellikle uzun zaman gereksinimi ve düşük verimli yöntemlerdir. Hesaplamalı yöntemler olası etkileşim kümelerini tahmin etmektedir. Tahmin edilen bu etkileşimler yapılması planlanan laboratuvar deneyleri için yol gösterici olarak kullanılmaktadır. Farklı kaynaklardan elde edilen bilgiler birlikte kullanılarak protein-protein etkileşimlerinin ve PPI ağlarının doğruluğu artırılmaya çalışılmaktadır.

Teknolojik gelişmeler deneysel süreçleri hızlandırmakla beraber deneysel doğruluğun sorgulanmasını zorunlu hale getirmiştir. Hesaplamalı yöntemler olası deneylerin sayısını azaltarak hedef deneylerin belirlenmesinde önemli rol üstlenir duruma gelmiştir.

Her grup tespit yöntemi içinde doğruluk analizi, tespiti ve iyileştirilmesi üzerine çalışmalar oldukça azdır. Eksikliği fark edilen ve iyileştirmeye açık olan bu alanlarda yapılacak çalışmaların sayısında yaşanacak artış ve hesaplamalı yöntemlerin tahmin doğruluklarının iyileştirilmesi ile yöntemlerin sonuç üretme doğruluğu ve sonuç üretme hızlarında önemli gelişmeler yaşanacaktır.

2.6 Literatür Özeti

Ağ kararlılığı ile ilgili çalışmalar: Bazı çalışmalar, biyolojik sistemlerin bozulma veya belirsizlik koşulları altında karmaşıklığını ve sağlamlığını modellemeyi ve test etmeyi amaçlamaktadır [10,11,12]. Hem matematiksel modelleme hem de deneysel gözlemler yoluyla yapılan değerlendirmeler, biyolojik sistemlerin çok çeşitli bozulmalara karşı oldukça dirençli olduğunu göstermektedir.

Mevcut çalışmaların çoğu topolojik bozulmaları ele almaktadır. Genellikle rastgele ve seçici düğüm çıkarılmasına odaklanılmıştır [6,7,8]. Bu yöntemler; çoğu karmaşık sistemin mutasyonlar, protein yanlış katlanması, genlerin veya proteinlerin parçalanmasına yol açan çevresel değişiklikler gibi çeşitli faktörler nedeniyle gürültü içerdiği varsayımını test eder.

Düğüm çıkarma işlemlerinin ortalama jeodezik uzunluk, kümeleme katsayısı ve ağ çapı gibi çeşitli ağ özellikleri üzerindeki etkileri değerlendirilmiştir. Rastgele model, derece merkezliği ve merkezlilik tabanlı modeller arasındaki çeşitli seçici düğüm çıkarma stratejileri test edilmiştir.

Sonuçlar, sağlamlık ve ağ topolojisi arasında güçlü bir korelasyon olduğunu göstermektedir. Özellikle, ölçeksiz ağlar rastgele düğüm çıkarma işlemlerine karşı dayanıklıdır. Bununla birlikte seçici düğüm çıkarılmasına karşı kırılabilirler. Bu çalışmaların sadece düğüm çıkarmayı içerdiği ve kenar çıkarmayı dikkate almadığı için büyük bir sınırlaması vardır. Bir düğümle ilişkili tüm etkileşimlerin, düğüm mevcut olduğu sürece var olduğunu varsayarlar.

Li ve arkadaşları [13] konsantrasyon pertürbasyonunun protein etkileşim ağları üzerindeki etkisini araştırmıştır. Hesaplama modelleri, pertürbasyon öncesi ve sonrası denge durumlarını karşılaştırır. Yakın ve uzak komşuların denge durumlarını etkilemek için bireysel proteinlerdeki konsantrasyon büyüklüğünün ağ üzerinde nasıl arttığı ölçülmüştür.

Gürültünün kenarlar üzerindeki etkisi çeşitli çalışmalarda araştırılmıştır. Home ve arkadaşları kenar silme işleminin jeodezik uzunluk ve altta yatan ağın büyüklüğü üzerindeki etkisini araştırmıştır [8]. Diğer çalışmalar, hastalık genlerinin ve ilaç hedeflerinin tanımlanması ve önceliklendirilmesi için tahmin algoritmalarının performansını test etmek amacıyla rastgele bir kenar çıkarma modelleri kullanmıştır [14], [15]. Zhang ve arkadaşları [16] insan PPI ağındaki yapısal değişikliklerle ilgili vazgeçilmez düğümler kavramını araştırmışlardır. Ağda seçici kenar çıkarma için etkileşimlerin güven puanlarını kullanmışlardır.

Bu çalışmada kenar çıkarma işlemi ağ mutasyon modeli olarak kabul edilmiştir. Düğüme bitişik tüm kenarları çıkarmanın etkisi, bu düğümü kaldırmakla aynı etkiye sahip olduğundan modelimiz düğüm çıkarma işlemini de kapsamaktadır. Çalışmanın birinci kısmında, mevcut literatürden farklı olarak, topolojik pertürbasyonların verilen ağı difüzyon durumu üzerindeki etkisi araştırılmıştır. Bu çalışma, topolojik mutasyonlar altında difüzyon durumunu değerlendiren ilk çalışmadır.

Ağların difüzyon durumu: Etkileşim ağlarındaki difüzyon analizi, gen önceliklendirme [147], [14], protein fonksiyon tahmini [148], [149], korunmuş ağ modüllerinin tespiti [150] ve protein-protein etkileşim (PPI) ağlarında alt ağ tespiti gibi birçok problemin çözümünde yaygın olarak kullanılmaktadır [151].

Kohler ve arkadaşları [147] kalıtsal bozukluklarla ilişkili aday genleri önceliklendirmek için yöntem geliştirdiler. Protein protein etkileşim ağlarındaki benzerlikleri tanımlamak için rastsal yürüyüşü kullandılar. Erten ve arkadaşları [14], aday hastalık genlerini önceliklendirmek amacıyla rastsal yürüyüş ile ağların topolojik benzerliklerini ölçmüşlerdir.

Protein fonksiyonunun tahmini için Cao ve arkadaşları [148], protein çiftleri arasındaki mesafenin bir ölçüsü olarak rastsal yürüyüşü kullanmışlardır. Cho ve arkadaşları [149], difüzyon durumları için boyut azaltma tekniği geliştirmişlerdir.

Jeong ve arkadaşları [150], rastgele yürüyüş modeline dayanarak PPI ağlarını karşılaştırmak için bir sistem tasarlamışlardır. Bu sistemi farklı türlerde korunmuş fonksiyonel ağ modüllerini tespit etmek için kullandılar. He ve arkadaşları [151] PPI ağlarındaki alt ağları tespit etmek için sınırlı K-yürüyüş algoritması kullanmışlardır.

Çalışmanın ikinci kısmında, bir ağda bulunabilecek en etkili mutasyonlar dikkate alınmıştır. Mevcut literatürden farklı olarak, topolojik pertürbasyonların verilen ağın Lyapunov Üstel yanıtı üzerindeki etkisi araştırılmıştır. Bu çalışma Lyapunov üslerinin ağ analizi için kullanıldığı ve topolojik mutasyonlar altında değerlendirildiği ilk çalışmadır.

Lyapunov üsteli: Bir sistemin Lyapunov Üsteli (LE), sonsuz derecede yakın yörüngelerin ayrılma oranını karakterize eden bir ölçüttür. Lyapunov üstelleri yakındaki yörüngelerin sapma oranını gösterir [152]. Lyapunov üsteli birçok çalışmada öznitelik çıkarma yöntemi olarak kullanılmıştır. Liu ve arkadaşları [153], en büyük LE'yi, insanın sessiz durması sırasındaki denge yeteneğinin bir ölçüsü olarak kullandılar. Çok değişkenli, en büyük LE ölçümüne dayanan insan vücudunun denge kabiliyetinin bir metriğini önerdiler. Biyolojik sistemlerde kritik geçişlerin tahmini için Nazarimehr ve arkadaşları [154], LE ölçümünü “kritik yavaşlama” göstergesi olarak önerdi. Gao ve arkadaşları [155], biyolojik verilerin çok ölçekli analizi için LE ölçümünü kullanmıştır. Konjestif kalp yetmezliğini ve nöbetlerini tespit etmek için kalp hızı değişkenliğini ve elektroensefalografi verilerini analiz ettiler. Valenza ve arkadaşları [156], baskın LE ölçümünü kullanarak duygusal görsel çıkarım sırasında kalp hızı değişkenliğinin karmaşıklığını karakterize etmişlerdir. Qingpeng ve arkadaşları [157], sağlıklı insanların kalp atış hızı değişkenliğini ve aritmi olanların kalp atış hızı değişkenliğini karşılaştırmalı olarak analiz etmişlerdir. Proteinlerdeki düzensiz bölgelerin tespiti için Gök ve arkadaşları [158] fizikokimyasal özellikleri LE kullanarak birbirine bağlayan yeni bir özellik kodlama tekniği geliştirmişlerdir.



3. MATERYAL ve METOD

3.1 Ağ Güvenilirliği

Bu bölümde, topolojik bozulmalar açısından ağların kararlılığı kavramı tanımlanmaktadır. Burada V ve E sırasıyla düğümler kümesini ve bu düğümler arasındaki etkileşim kümesini göstermek üzere hedef ağ $G = (V, E)$ ile gösterilmektedir.

Bu çalışmada PPI ağlarına odaklanılmıştır. Bu nedenle G , yönsüz bir ağdır. Verilen bir topolojik mutasyon oranı $\varepsilon[0:1]$ göz önüne alındığında, eğer $\frac{|E \cup E^\varepsilon - E \cap E^\varepsilon|}{|E|} \leq \varepsilon$ ise bu ağın G ağının pertürbasyonu olduğu söylenir ve $G^\varepsilon = (V, E^\varepsilon)$ ile temsil edilmektedir.

$Q: G \rightarrow R$ ile gösterilen kalite fonksiyonu için (bu fonksiyon bölümün ilerleyen kısımlarında ayrıntılı olarak açıklanmaktadır); G ağının topolojisini ε faktör oranı ile değiştirdikten sonra kalite fonksiyonundaki en büyük değişiklik miktarı olarak q fonksiyonu altındaki belirli bir pertürbasyon miktarı için G ağının stabilitesi hesaplanmaktadır. Böylece, mutant ağın sonuçtaki kararlılığı matematiksel olarak aşağıdaki şekilde (**Denk.1**) ifade edilmektedir.

$$\operatorname{argmax}_{G^\varepsilon} = \{|q(G) - q(G^\varepsilon)|\} \quad (1)$$

Yukarıdaki formülasyon, ağ durumunda en yüksek değişikliğe sahip G 'nin tüm olası ε sapmaları arasında ağ topolojisini arar. Belirli bir ε için daha küçük $|q(G) - q(G^\varepsilon)|$ değerine sahip ağ daha kararlıdır.

Belirli bir G ağının bozulmaları olan olası topolojilerin sayısı, G ağının büyüklüğü ve ε değeri ile katlanarak büyüdüğünü belirtmek gerekir. Düğüm sayısı n ve mutasyon kenar küme sayısı r ile temsil edilmek üzere problem karmaşıklığı C_r^n ile formüle edilir ve yaklaşık olarak $O(n^r)$ olarak hesaplanır. Bu nedenle, ağ durumu farkını en üst düzeye çıkaran topolojiyi bulmak, hesaplama açısından zor bir görevdir ve çözümü ilerleyen bölümlerde detaylı olarak anlatılacaktır.

Çalışma kapsamında dört kalite fonksiyonu tanımlanmıştır ve bu fonksiyonlar detaylı olarak açıklanacaktır.

3.1.1 Ortalama Kümelenme Katsayısı (CC)

Kümelenme katsayısı bir ağın yerel uyumluluğunu tanımlar [159]. Daha spesifik olarak, ortak bir komşuyla iki düğümün birbirine bağlanma olasılığını ölçer.

k_i komşusu olan belirli bir düğüm n_i için, komşuları arasında en fazla $e_{\max} = k_i * (k_i - 1) / 2$ olası kenar vardır. Bir n_i düğümünün komşuları arasında gözlenen kenar sayısını e_i ile belirtelim. n_i düğümüne ait kümeleme katsayısı c_i , gerçek kenar sayısı e_i nin olası maksimum kenar sayısı $e_{\max}$ değerine oranıdır (yani, $c_i = (2e_i / k_i * (k_i - 1))$). CC_G ile belirtilen belirli bir G ağının ortalama kümeleme katsayısı, o ağdaki tüm düğümlerin kümeleme katsayılarının aritmetik ortalamasıdır (**Denk.2**):

$$CC_G = \sum_{i=1}^N \frac{c_i}{N} \quad (2)$$

3.1.2 Rastsal Yürüyüş (RW)

Belirli bir ağ $G = (V, E)$ için, rastgele yürüme yöntemi belirli bir n_i düğümünden başlayan rastgele bir yürüyüşü simüle eder. Her adımda rastgele bir yön seçilir (n_i düğümünün bir kenarı) ve bu kenardan n_i düğümüne bağlı komşu n_j düğümüne geçilir. PPI ağları için, tüm yönler $1/\text{derece}(n_i)$ ile formüle edilmiş eşit olasılığa sahiptir.

Kaynak n_i düğümünden yürünebilecek adım mesafesini k kısıtlama parametresi ile gösterelim. G ağına ait komşuluk matrisini A_G , A_G komşuluk matrisi ile aynı boyutta ve köşegen elemanlarında $1/\text{degree}(n_i)$ değeri olan köşegen matrisi D ve her ikisinin çarpımını temsil eden matrisi de M ile gösterelim ($M = D_G * A_G$). k adım yürüme ile oluşan rastsal yürüyüş matrisini RW_k ile belirtiyoruz. $k=1$ için $RW_1 = D_G$ ve $k>1$ için $RW_k = M^T \cdot RW_{k-1}$ olarak hesaplanır [160].

Sonuçta elde edilen RW_k matrisinde $RW_k [i, j]$ elemanı, k adımda n_j düğümünden n_i düğümüne rastgele yürüyüş ile geliş olasılığını temsil eder. Bu nedenle, bu matrisin her bir sütunu, belirli bir n_j düğümünden gelen varış olasılıklarının tüm düğümlere olan dağılımını gösterir.

3.1.3 Difüzyon Durum Mesafesi (DSD)

Belirli bir düğüm çifti için (n_x, n_y) için DSD'yi, belirlenmiş k değeri için RW_k tarafından hesaplanan matris ile hesaplanır. Daha spesifik olarak, RW_k sütunlarının n_x ve n_y düğümlerine karşılık gelen değerlerin L1 normu hesaplanır.

Belirli bir n_i düğümüne karşılık gelen RW_k sonucunu $RW_k(n_i)$ ile gösterelim. Matematiksel olarak DSD aşağıdaki şekilde (**Denk.3**) tanımlanır [148].

$$DSD_k(n_x, n_y) = \|RW_k(n_x) - RW_k(n_y)\|_1 \quad (3)$$

3.1.4 Komşu Benzerliği (NS)

Protein fonksiyon tahmini biyoinformatik için önemli bir sorundur. Komşuluk ilişkileri fonksiyon tahmin yöntemlerinin temel dayanağıdır. Komşu çoğunluk oylama algoritması (NMV) [161] fonksiyon tahmini için yaygın olarak kullanılmaktadır.

Ağ topolojisindeki değişikliklerin komşu ilişkisi üzerindeki etkisini ölçmek için Eşleştirme Endeksi [162] kullanılır. Verilen iki düğüm arasındaki eşleşen indeksi, ortak komşularının sayısının tüm komşularının sayısına oranı olarak formüle edilir [163].

G ağındaki n_i düğümüne ait komşu kümesini nb_i ile gösterelim ve topolojik mutasyonlar sonrasında G ağından elde edilen G^e ağına ait komşu kümesini nb_i^e ile gösterelim. Bu değişiklikten sonra n_i düğümüne ait komşuluk değişimi aşağıdaki formülasyonla (**Denk.4**) hesaplanır:

$$NS_i = \frac{|nb_i \cap nb_i^e|}{|nb_i \cup nb_i^e|} \quad (4)$$

3.2 Geliştirilen Mutasyon Algoritması

Bu bölümde geliştirdiğimiz yöntem ayrıntılı olarak açıklanmaktadır. Yöntemimiz iki aşamada çalışır. İlk aşama, bir dizi başlangıç çözümü oluşturmak için genetik algoritma uygular. İkinci aşama, bu çözümleri daha fazla optimize etmek için tepe tırmanma stratejisini kullanır.

Şekil 3.1 de yer alan algoritmada, yöntemimizin sözde kodu sunulmuştur. Algoritmamız sekiz giriş parametresi alır; graf $G=(V, E)$, izin verilen kenar silme oranı ε , ağ kalitesi fonksiyonu $f()$, genetik algoritma popülasyon büyüklüğü ps , maksimum yineleme sayısı mi , minimum artış hızı mri , geçiş olasılığı pc ve mutasyon olasılığı pm dir. Bu parametreleri ve yöntemimizin bu parametreleri nasıl kullandığı bölümün devamında detaylı olarak açıklanacaktır. Algoritmamız, ağ kalitesi fonksiyonunda maksimum değişikliğe yol açan kenar silme kümesini sonuç olarak döndürür.

Aşama 1 Genetik Algoritma :

Bu aşamada, çözümler kümesi üretilmektedir. Her çözüm G de yer alan kenarlardan oluşup $E' \subseteq E$ ve $\varepsilon|E|$ eleman sayısındadır. E' çözümünün uygunluğunu hesaplamak için E' kümesinde yer alan kenarları G ağından çıkardıktan sonra ağ kalite fonksiyonundaki değişiklik hesaplanmaktadır.

Spesifik olarak, E' kümesinin uygunluğu $f(E'|G) = |q(G = (V, E)) - q(G' = (V, E - E'))|$ olarak tanımlanır. Ardından, sonraki iterasyonlarda yinelemeli olarak yeni ve potansiyel olarak daha iyi çözümler üretilmektedir. Algoritmada kullanılan ilk değer atama, seçim, çaprazlama, mutasyon ve elitleştirme adımları detaylı olarak anlatılacaktır.

İlk çözüm popülasyonu: İzin verilen kenar silme sayısını $cs = \varepsilon|E|$ olarak tanımlanmıştır. Algoritmamız ilk çözüm popülasyonu olan $X^0 = \{X^0_1, X^0_2, \dots, X^0_{ps}\}$ kümesini rastsal olarak oluşturur. Her $X^0_i \in X^0$ çözüm kümesi, E kenar kümesinden rastgele seçilmiş cs adet kenardan oluşan bir kümedir. Bu adımda üretilen tüm çözümlerin uygunluğu da hesaplanmaktadır.

Aşağıdaki adımlar, bu adımda üretilen çözümleri yinelemeli olarak iyileştirir. Her iterasyonda, bir önceki iterasyonda üretilen popülasyonu kullanarak yeni bir çözüm popülasyonu üretilir. X^t ile t inci iterasyonda üretilen çözüm popülasyonu temsil edilmektedir.

Aşağıda, X^{t-1} popülasyonunun zaten üretildiği varsayılmakta ve X^{t-1} 'den X^t popülasyonunun nasıl elde edildiği açıklanmaktadır. Bu adımda, X^{t-1} 'deki çözümlerin bir kombinasyonu olarak ps adet yeni çözüm üretilmektedir.

Seçim: İlk olarak mevcut popülasyondan nasıl iki ana çözümün seçildiği açıklanacaktır. Genetik algoritma, yüksek uyum değerine sahip ebeveynlerin yüksek uyum değerlerine sahip çocuklar oluşturma olasılıklarının yüksek olduğunu varsayar. Bu varsayım gereğince, ebeveyn çözümler uygunluklarıyla orantılı bir olasılıkla rastsal olarak seçilmektedir. Algoritmamızda iki aday ebeveyn çözüm F ve M ($F, M \in X^{t-1}$) olarak tanımlanmıştır (Satır 7).

Çaprazlama: Bu adımda, iki ana çözüm F ve M 'den bir çocuk çözümünün oluşturulmasında kullanılan çaprazlama yöntemi anlatılacaktır. Çaprazlama pc olasılığı ile uygulanır. Eğer hiçbir çaprazlama gerçekleşmezse, ebeveynleriyle aynı olan iki çocuk oluşturulur. Eğer çaprazlama gerçekleşirse, rasgele bir fs tamsayısı $[1; cs-1]$ üretilir. Bu sayı F ebeveyninden çocuğa ne kadar bilgi aktarılacağını gösterir. Daha sonra, çocuk çözümü oluşturmak için fs adet kenar F ebeveyninden ve $(cs - fs)$ adet kenarda M ebeveyninden rastgele seçilir.

F ve M arasından seçilen bazı kenarlar ortak olabilir. Bu durumda, çocuğun kenar sayısı cs adetten az olacaktır. Böyle bir durum gerçekleştiğinde E kenar kümesinden rastgele kenarlar seçilerek çocuk çözüme eklenir ve kenar sayısı cs parametresine tamamlanır. Bu şekilde iki çocuk çözüm üretilir ve sırasıyla B ve S ile isimlendirilir ($B, S \in X^t$) (Satır 8). Bir sonraki X^t çözüm popülasyonunda X^{t-1} de bulunan çözüm sayısına eşit yeni çözüm oluşana kadar adım tekrarlanır.

Mutasyon: Seçim ve çaprazlama işlemleri, her iterasyonda çözümleri tekrar tekrar iyileştirir. Bununla birlikte, bu operatörlerle üretilen çözümlerin yerel optimuma ulaşma olasılığı vardır. Bu adımda, böyle yerel optimumlardan kaçınmak için bir fırsat oluşturulmaktadır. Bunu yapmak için, çaprazlama tarafından üretilen her çocuk, düşük bir p_m olasılığı ile mutasyona uğratarak değiştirilir.

Spesifik olarak, X^t içinde bir çocuk çözüm göz önüne alındığında, p_m olasılığı ile bu çözümde yer alan tüm kenarlar rastgele seçilmiş bir $e_r \in E \setminus (B \text{ veya } S)$ kenarı ile değiştirilir (Satır 9).

Elitizm: X^t çözüm kümesinin X^{t-1} çözüm kümesine eşit çözüme sahip olana kadar yeni nesil X^t çözümlerini oluşturmak için yukarıda açıklanan seçim, çaprazlama ve mutasyon adımlarını tekrarlar. Çaprazlama adımının X^{t-1} den daha iyi X^t çözümleri üretme eğiliminde olacağı varsayılsa da, bu adımlardaki rastsallık nedeniyle, yeni nesildeki en iyi çözümün bir önceki nesilde yer alan en iyi çözüm kadar iyi olacağına dair bir garanti yoktur.

Elitizm, her neslin en iyi çözümünün gelecek nesilde korunmasını sağlar. Bunu yapmak için, X^t de yer alan uyum değeri en düşük çözüm, X^{t-1} de yer alan uyum değeri en yüksek çözüm ile değiştirilir (Satır 10 ve 11).

Her iterasyondan sonra, popülasyonda genellikle çok daha uygun kromozomlar vardır. İterasyonun sonunda mevcut artış oranı cri hesaplanır, $cri = |\max(F_i) - \max(F_i^{mut})| / \max(F_i)$ (Satır 12). İterasyon sayacı artırılır (Satır 13). Algoritma, iki sonlandırma kriterinden birine ulaşıncaya kadar yeni nesil çözümler üretmeyi tekrarlar; maksimum yineleme sayısı $ci < mi$ veya minimum artış oranı $cri > mri$ dir (Satır 5).

Aşama 2 Tepe Tırmanma Algoritması :

İlk aşamada, bir dizi aday çözüm üretilmektedir. İkinci aşama, her bir çözüme birer birer odaklanarak ilk adımda bulunan çözümleri ayrı ayrı optimize eder. Bunu yapmak için, tepe tırmanma algoritmasını kullanarak yerel bir optimizasyon stratejisi oluşturulur.

Daha spesifik olarak, bir aday çözüm verildiğinde, bu çözümdeki bir kenarı yeni bir kenar ile değiştirerek çözüm iyileştirilir. Bu işleme, çözümü iyileştirecek kenar kalmayana kadar devam edilerek çözüm optimize edilir. Bir $E' \subseteq E$ kenar kümesinin etkisi aşağıdaki şekilde hesaplanır.

E' kenar kümesini orijinal $G = (V, E)$ ağından çıkartılır ve uygunluk fonksiyonu $f(\{E'\} | G)$ hesaplanır. E' kenar kümesinin etkisi $err(E')$ ile gösterilmiş olup $err(E') = |f(\{e_i\} | G) - f(\emptyset | G)|$ ile hesaplanmaktadır (Satır 17). Bu tanım kullanılarak, her $e_i \in E$ kenarın etkisini $err(\{e_i\})$ olarak hesaplayıp, bu bilgiler orijinal ağdaki tüm kenarlar için saklanır. Artık mevcut çözümlerin optimize edilmesi gerekir.

İlk aşamada bulunan ES ile gösterilen bir çözümü ele alalım (Satır 19). Bu çözümdeki kenar kümesini $ES = \{e_1, e_2, \dots, e_{cs}\} \subseteq E$ olarak gösterelim. Yinelemeli olarak $E - ES$ kenar kümesinden bir kenar seçilir ve aşağıdaki gibi ES kenar kümesinden bir kenarla değiştirilir.

İlk olarak kenar başına ortalama hata $err(ES)/cs$ formülü ile hesaplanır. Ardından, her kenar $e_i \in (E - ES)$ için, $err(ES \cup \{e_i\}) / (cs + 1)$ formülü ile geçerli çözüme e_i kenarı eklendikten sonra ortalama hata hesaplanır. Yeni kenar eklendikten sonra ortalama hata artmazsa (yani, $err(ES \cup \{e_i\}) / (cs + 1) \leq err(ES) / cs$) e_i kenarı eklenmez ve listeden çıkarılır. Aksi durumda, e_i kenarı ES çözüm kümesinde tutulur (Satır 22).

Ardından, çözüm kümesi boyutunun değişmeden kalması için ES 'den silinecek bir kenar olan e' kenarı aranır. Bunu yapmak için, ES de yer alan kenarlar $err()$ etki değerlerine göre artan sırada sıralanır. En düşük etki değerine sahip $err(\{e'\})$ $e' \in ES$ kenarından başlayarak arama döngüsü başlatılır (Satır 24). ES kenar kümesinden e' kenarı silinerek yeni bir kenar kümesi oluşturulur (Satır 26). Ardından kenar başına ortalama hata $err(ES - \{e'\}) / (cs - 1)$ ile hesaplanır. Bu değer geçerli değerden ($err(ES)/cs$) büyükse, e' kenarının çözüme dâhil edilen yeni kenardan daha iyi olmadığı sonucuna varılır. Bu sebeple, e' kenarı mevcut çözümden kalıcı olarak silinir (Satır 27). Aksi takdirde, ES de yer alan tüm kenarlar için sıralanmış listedeki bir sonraki kenarı silinmeye çalışılır.

Her iterasyonun sonunda, sonuç olarak en yüksek uyum değerine sahip çözüm seçilmiş olur (Satır 29). Aşama 1'de üretilen her aday çözüm için aşama 2 tekrar edilir. Nihai sonuç olarak bulunan en iyi çözüm sonuç olarak sunulur.

Algoritma 1: En optimal mutasyonun bulunması

Girişler:

- graf $G = (V; E)$, kenar silme oranı edr
- objektif fonksiyon $f =$ (kümeleme katsayısı, k-l ıraksama, RW)
- maksimum iterasyon mi , popülasyon büyüklüğü ps , minimum artış hızı mri , geçiş olasılığı pc , mutasyon olasılığı pm

Çıktılar:

- hedef ölçümüne ait maksimum hata oranı

Aşama 1: Genetik Algoritma

2: kromozom boyutu $cs \leftarrow |E| \times edr / 100$

3: $X^0 \leftarrow ps$ boyutunda ilk rastgele popülasyonu oluşturur

4: $F^0 \leftarrow f(X^0)$ uygunluğu değerlendirir

5: while $t < mi$ and $cri > mri$ do

6: for $i = 0$ to $ps=2$ do

7: $(F, M) \leftarrow$ ranking selection ile ebeveyn kromozomları seç

8: $(B, S) \leftarrow p_c$ olasılığı ile (F, M) 'ye one point crossover işlemi uygula

9: $(X_t^i, X_{t+1}^i) \leftarrow p_m$ olasılığı ile (B, S) 'ye integer range mutation işlemini uygula

10: $F^t \leftarrow f(X^t)$

11: X_t en kötü çözüm ile X_{t-1} en iyi çözümü değiştirerek elitizm uygula

12: $cri \leftarrow \frac{|\max(F^{t-1}) - \max(F^t)|}{\max(F)}$ mevcut artış oranını hesapla

13: $t = t + 1$

14: Aşama 2: Tepe Tırmanma

15: En iyi 10 sonucu seç, $BR \leftarrow \max(X^t, 10)$

16: result $\leftarrow 0$, en iyi sonuca ilk değeri ata

17: EdgeScores değerlerini yükle

18: for each item in BR do

19: $ES \leftarrow$ item, başlangıç kenar kümesini ata

20: while select random e in E do

21: ES 'ye e kenarını ekle ve ES' oluştur

22: if $[err_{ES'}/(cs + 1)] > [err_{ES}/cs]$ then e yi tut

23: else continue

24: ES yi artan düzende EdgeScores a göre sırala

25: for e' in ES do

26: $\overline{ES}$ kümesini oluşturmak için e' kenarını ES' kümesinden çıkart

27: if $[err_{\overline{ES}}/(cs - 1)] > [err_{ES}/cs]$ then e' kenarını kalıcı olarak sil

28: else continue

29: if $f(ES) > result$ then result $\leftarrow f(ES)$

30: Girdi olan 10 aday içinden en iyi çözümü döndür

Şekil 3.1 En İyi Mutasyon Kenar Kümesi Bulma Algoritması

3.3 Uygunluk Ölçümleri

Ağ topolojisinde meydana gelen her mutasyon, ağın özelliklerini ve dolayısıyla ağdan hesaplanan matematiksel sonuçları değiştirir.

Bu çalışmada, matematiksel sonuçlarda meydana gelen değişikliklerin yoğunluğunu belirlemek için spesifik değişikliklere göre ölçümler tasarlanmıştır. Bu bölümde, uygunluk ölçümleri ayrıntılı olarak açıklanacaktır.

3.3.1 Bağlı Hata (RE)

Mutlak hata, başlangıç değeri ile son değer arasındaki farktır. Nesnel fonksiyonların mutlak hataları değişkendir, bu nedenle göreceli hata mutlak hatanın başlangıç değerine yüzde oranı olarak hesaplanır (**Denk.5**).

$$AE = x_i - x_f \text{ ve } RE = \frac{|AE|}{x_i} \times 100 \quad (5)$$

Bu hesap bize ilk değerden yüzde sapmayı verir. RE ölçümü, ‘CC’, ‘RW’, ‘DSD’ ve ‘NS’ gibi başlangıç ve mutasyon ile değişmiş objektif fonksiyon sonuçları arasındaki yüzde sapmayı ölçmek için kullanılır.

Ortalama cc hatası :

Topolojik mutasyonlardan sonra hedef ağdaki kümelerin değişimini değerlendirmek için ortalama CC hatası kullanılır. CC objektif fonksiyonunun sonucu, kümeleme yüzdesini temsil eden bir sayıdır (bkz. kalite fonksiyonları). Dolayısıyla, ortalama CC hatası daha önce tarif edilen RE formülü uygulanarak hesaplanır.

Verilen ağın başlangıç CC değerini CC_i ile ve mutasyona uğramış ağın CC değerini CC_f ile gösterelim. Ortalama CC hatası aşağıdaki formülasyon ile (**Denk.6**) hesaplanır.

$$RE_{ES}^{CC} = \frac{|CC_i - CC_f|}{CC_i} \quad (6)$$

Ortalama rw hatası :

Ortalama RW hatasını hesaplamak için önce RW_i ile belirtilen, verilen bir $G = (V, E)$ ağının başlangıç durumu ve $ES = (e_1, e_2, \dots e_\varepsilon)$ kenar kümesi ile mutasyona uğramış $G^\varepsilon = (V, E^\varepsilon)$ ağının nihai durumu RW_f için RW (bkz. Kalite fonksiyonları) değeri hesaplanır. Topolojik mutasyonlar, RW ile temsil edilen her bir düğüm $(n_x, n_y) \in V$ için bir düğümden diğerine ulaşma olasılığını değiştirir. Bu nedenle, RW_f matrisindeki her hücre, G ağının hassasiyeti nedeniyle RW_i matrisindeki hücrelerden farklı olabilir. Hücre tabanlı RW değişikliklerini ölçmek için, oransal değişikliği hesaplayan aşağıdaki formülasyon (**Denk.7**) kullanılır.

$$RE_{(x,y)}^{RW} = \frac{|(RW_i(n_x, n_y) - RW_f(n_x, n_y))|}{RW_i(n_x, n_y)} \times 100 \quad (7)$$

Son olarak, matris temelli değişiklikleri (global değişiklik) hesaplamak için, aşağıdaki formülasyon ile tüm hücre temelli değişikliklerin aritmetik ortalaması hesaplanır (**Denk.8**).

$$RE_{ES}^{RW} = \sum_{x=1}^N \sum_{y=x}^N \frac{RE_{(x,y)}^{RW}}{N}, N = \text{boyut}(V) \quad (8)$$

Ortalama dsd hatası :

DSD, RW değerinin L1 normudur (bkz. kalite fonksiyonları). L1 norm operatörünün doğası gereği, DSD değerindeki değişiklikler RW değerindeki değişikliklerden farklıdır. Ortalama DSD hatası ile L1 norm operatörünün RW üzerindeki etkisi değerlendirilir. Sırasıyla DSD_i ve DSD_f ile belirttiğimiz ilk ağ G ve mutasyona uğramış nihai ağ G^ε üzerinde DSD hesaplanır. DSD değerindeki hücre bazlı değişiklikleri değerlendirmek için aşağıdaki formülasyon (**Denk.9**) kullanılır.

$$RE_{(x,y)}^{DSD} = \frac{|(DSD_i(n_x, n_y) - DSD_f(n_x, n_y))|}{DSD_i(n_x, n_y)} \quad (9)$$

Bu formül, her DSD hücresi için değişiklik oranını verir. Orandaki toplam değişiklikleri hesaplamak için aşağıdaki formülasyon (**Denk.10**) tanımlanır.

$$RE_{ES}^{DSD} = \sum_{x=1}^N \sum_{y=x}^N \frac{RE_{(x,y)}^{DSD}}{N}, N = boyut(V) \quad (10)$$

Ortalama komşu hatası :

DSD rastgele yürüyüş tabanlı mesafe ölçümüdür. RW, DSD ve temel olarak tüm tahmin algoritmaları benzerliklerin belirlenmesi için komşu düğümler arasındaki mesafe ölçümlerini kullanırlar. Bu çalışmada, ortalama komşu hata ölçümü ile belirlenen komşulardaki değişiklikler değerlendirilmektedir. Ortalama komşu hatasını hesaplamak için, başlangıç ağı G ve mutasyona uğramış ağ G^e üzerinde her bir düğüm için en yakın 8 komşu belirlenir. Her n_i düğümü için, ilk ağ G ve mutasyona uğramış G^e ağını kullanarak komşu benzerliği NS_i hesaplanır ve n_i düğümünün komşu hata oranı (1-NS_i)×100 olarak tanımlanır. Son olarak, komşuluk açısından ağ genelindeki değişiklikler aşağıdaki formülasyon ile (**Denk.11**) değerlendirilir.

$$RE_{ES}^{NS} = \sum_{i=1}^N \frac{(1 - NS_i)}{N} \times 100, N = boyut(V) \quad (12)$$

3.3.2 K-L ıraksaması (DKL)

Kullback-Leibler (k-l) ıraksaması, istatistikte, p = p_i olasılık dağılımının bir model (aday) dağılımı q = q_i dağılımına ne kadar yakın olduğunu nicel olarak ifade eden bir ölçüdür (**Denk.12**) [164].

$$DKL(p||q) = E_i(p_i) \log \left(\frac{p_i}{q_i} \right) \quad (32)$$

DKL ölçümü, O ile temsil edilen, G ağının objektif fonksiyon sonuç matrisi f(G) ve O^e ile temsil edilen, mutasyona uğramış G^e ağının etkilenmiş objektif fonksiyon sonuç matrisi f(G^e) arasındaki dağılımın nasıl değiştiğinin ölçümünde kullanılır (bkz. kalite fonksiyonları). Histogramları oluşturmak için ilk olarak giriş verileri 1000 bölmeye ayrılır. Ardından, Hist(O) ve Hist(O^e) ile gösterilen, her bir bölmeye dâhil olan verilerin oluşma sayıları hesaplanır. Son olarak, daha önce tanımlanan DKL formülü ile k-l ıraksaması hesaplanır (**Denk.13**).

$$DKL(Hist(O)||Hist(O^e)) \quad (43)$$

DKL değerini hesaplamak için, hedef fonksiyon sonucu bir dizi veya matris olmalıdır. Bu nedenle, sadece RW ve DSD hesaplamalarının k-1 ıraksaması ardışık olarak DKL^{RW} ve DKL^{DSD} olarak tanımlanmıştır.

3.3.3 Pearson korelasyon katsayısı (PCC)

Korelasyon analizi, iki değişken arasındaki doğrusal ilişkiyi test etmek ve bu ilişkinin derecesini ölçmek için kullanılan istatistiksel bir yöntemdir. Analizin sonucu, doğrusal bir ilişkinin olup olmadığı ve varsa bu ilişkinin derece değeridir. Sonuç, r ile gösterilen korelasyon katsayısıdır ve -1 ile $+1$ arasındaki değer alabilir. $r=+1$ tam pozitif doğrusal bir ilişki, $r = -1$ tam negatif doğrusal bir ilişki ve $r = 0$ hiçbir ilişki olmadığı anlamına gelir. $r < 0$ olan tüm r değerleri için, hiçbir ilişki yok anlamına gelen $r = 0$ kullanılmıştır.

RW ve DSD nesnel işlev matrislerindeki satırlar, kaynak düğüm ile tüm düğümler arasındaki ilişkiyi gösterir (bkz. kalite fonksiyonları). Başlangıç objektif fonksiyon sonucu O ile nihai objektif fonksiyon sonucu O^e arasındaki yüzde değişimi daha önce belirtildiği gibi RE ile ölçülür. Ancak bu bilgi etkinin oranını değerlendirmek için yeterli değildir. RE değişikliğin büyüklüğünü verir, ancak düğümler arasındaki değişikliklerin benzerliği hakkında bilgi içermez.

Bu tezde, kaynak düğüm ile diğer tüm düğümler arasındaki ilişkinin etki ile nasıl değiştiğini değerlendirmek için PCC tabanlı uygunluk ölçümleri tanımlanmıştır. Bu hesaplamayı yapmak için, ilk önce, PCC hesaplaması başlangıç objektif fonksiyon O ile son objektif fonksiyon O^e arasında satır satır uygulanır. Değerlendirilen düğüme ait O_i ve mutasyon sonrası sonuç dizisi O_i^e için r_i aşağıdaki gibi formüle (Denk.14) edilmiştir.

$$r_i = \frac{\sum_{x=1}^N O_i[x] \cdot O_i^e[x] - \frac{\sum_{x=1}^N O_i[x] \cdot \sum_{x=1}^N O_i^e[x]}{N}}{\sqrt{(\sum_{x=1}^N O_i[x]^2 - \frac{\sum_{x=1}^N O_i[x]^2}{N}) \cdot (\sum_{x=1}^N O_i^e[x]^2 - \frac{\sum_{x=1}^N O_i^e[x]^2}{N})}} \quad (54)$$

Son olarak, toplam değişiklikleri ölçmek için aşağıdaki formülasyon (Denk.15) tanımlanıp, RW ve DSD ölçümlerinin PCC hesaplaması sırasıyla PCC^{RW} ve PCC^{DSD} olarak adlandırılmıştır.

$$PCC = r = \sum_{i=1}^N \frac{r_i}{N} \quad (65)$$

3.4 Geliştirilen Lyapunov Üs Tabanlı Ağ Katsayı Çıkarım Algoritması

Lyapunov üssü (LE) sistemin başlangıç koşullarına duyarlılığını ölçmektedir. Aslında, yakın yörüngelerin ayrılma oranını karakterize eden bir miktardır. Yakındaki yörüngeler, bir sinyal vektörünün faz uzayı boyunca izlediği yoldur.

Faz uzayının farklı yönelimleri için, ıraksama oranı farklı olabilir. Bu nedenle, faz uzayında bütün bir LE spektrumu vardır ve bunların sayısı faz uzay boyutlarının sayısına eşittir. Pozitif bir üs, ıraksama anlamına gelir. Yörüngeler başlangıçta birbirine yakındır ve zamanla birbirlerinden uzaklaşırlar. Ne kadar hızlı ayrıldıklarının oranı, pozitif bir üs büyüklüğü ile gösterilir. Aynı şekilde, negatif bir üs yörüngelerin birbirine yaklaşması yani yakınsama anlamına gelir [165].

Referans noktasını $s(n)$ ve yakın yörüngesindeki en yakın komşusunu $s(m)$ ile gösterelim. Her boyut için LE, faz uzayında aşağıdaki şekilde hesaplanır (**Denk.16**):

$$\tau = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \ln \frac{d(s(n+1), s(m+1))}{d(s(n), s(m))} \quad (76)$$

$d(s(n), s(m))$, en yakın komşular arasındaki başlangıç mesafesidir. $d(s(n+1), s(m+1))$, yörüngedeki bir sonraki komşu çifti olan $s(n+1)$ ve $s(m+1)$ arasındaki mesafedir [166].

LEC (Lyapunov exponent based network coefficient) olarak adlandırılan yöntemimiz bir parametre olarak yönsüz ağ olan $G = (V, E)$ grafını girdi olarak alır ve G için Lyapunov üst frekansını sonuç olarak döndürür. Sentetik ve biyolojik ağlar üzerindeki deneylerimiz, gömme boyutu olarak 6 kullanmanın ağların farklı yanıtlarını ayırt etmek için kullanılabileceğini göstermektedir. Hesaplamaları hızlandırmak için gömme boyutu 4'e düşürülebilir. Karşılaşılabilecek farklı özelliklere sahip ağların kapsamını azaltmaktan kaçınmak için önerdiğimiz yöntemde gömme boyutu azaltılmamıştır. Şekil 3.2 de yer alan algoritmada, yöntemimizin sözde kodu sunulmuştur.

G giriş ağında n adet düğüm $V = (v_1, v_2, \dots, v_n)$ bulunur. Belirli bir v_i düğümü için, v_i düğümünün diğer tüm düğümlere mesafesini temsil eden n adet kenar $E(v_i) = (e_1, e_2, \dots, e_n)$ vardır. Algoritmamız önce G giriş ağını her bir düğüm için olmak üzere n satıra böler. En büyük 6 Lyapunov üssünü $Le(v_i) = (ex_1, ex_2, ex_3, ex_4, ex_5, ex_6)$ ile gösterelim. Birinci v_1 düğümünden başlayarak algoritma, G giriş ağının her bir düğümü için Lyapunov üslerini $Le(G) = \{Le(v_1), Le(v_2), \dots, Le(v_n)\}$ olarak hesaplar.

Lyapunov üssü hesaplamasından sonra $Le(G)$ 'nin pozitif üs frekansı bulunur. G ağımızın pozitif üs frekanslarını $pf(G) = (f_0, f_1, f_2, f_3, f_4, f_5, f_6)$ olarak gösterelim. $pf(G)$ 7 değere sahiptir. Bunlar sırasıyla pozitif olmayan üs, ve 1, 2, 3, 4, 5, 6 pozitif üs frekanslarıdır.

Frekans değeri, ağdaki düğüm sayısı ile orantılıdır. Sonucu ağdaki düğüm sayısına bölerek frekans değerini 0-1 aralığına normalleştirilmiştir. Son olarak, verilen G ağının yüzde Lyapunov üs değerlerini hesaplamak için sonuç 100 ile çarpılır. Elde edilen vektör, herhangi bir ağ için Lyapunov üslerinin yüzdesini gösteren 7 elemanlı bir öz nitelik vektörü olarak kullanılabilir.

7 elemanlı öz nitelik vektöründe LE değerlerini temsil eden her Lyapunov üssü niteliği kümülatif bir değerdir. Bu nedenle $f_1 \geq f_2 \geq f_3 \geq f_4 \geq f_5 \geq f_6$ olur. Bu nitelikler için minimum değer 0, maksimum değer 100'dür. Ağların Lyapunov üs tepkisini temsil eden tek bir ölçüm üretmek için, sonuçta ortaya çıkan özellikler toplandıktan sonra normalleştirilir. Her ağ için Lyapunov üs tepkisi 0 ile 100 arasında bir ölçümdür. Lyapunov üssünün hesaplanması için TISEAN paketi kullanılmıştır [167]. Sano ve Sawada'nın çalışmalarına dayanan bu programla Lyapunov üslerinin tüm spektrumunu tahmin edilmektedir.

Sistemlere ait olmayan yanlış veri veya gürültü, sistemlerin dinamiklerini önemli ölçüde değiştirebilir, böylece sistemlerin Lyapunov üs tepkisini etkileyebilir [168]. Bu tezin ikinci kısmında sunduğumuz hipoteze göre, protein-protein etkileşim ağlarındaki yanlış tespitler, ağların topolojik dinamiklerini değiştirerek ağın Lyapunov üssünü etkileyebilir ve bu etki ağın stabilitesi ile ilişkili olabilir. $d(s(n+1), s(m+1))$ hesaplaması için gürültünün aşağıdaki etkisi vardır (**Denk.17**) [169].

$$d(s(n+1), s(m+1)) = \sqrt{\sum_{k=0}^{D-1} (x(n+k+1) + r(n+k+1) - x(m+k+1) - r(m+k+1))^2} \quad (87)$$

$$E[d(s(n+1), s(m+1))^2] = d(s(n+1), s(m+1))^2 + 2\sigma^2 D \quad (98)$$

Ağıdaki yanlış veriler, faz uzayındaki Lyapunov üssünün hesaplanmasında gürültü olarak kabul edilebilir. Faz uzayındaki gürültü etkisi bir önceki denklemde görülmektedir (**Denk.18**). Sonuç olarak, yanlış kenarlar denklemde gösterildiği gibi Lyapunov üssü hesaplaması üzerinde bir etkiye sahiptir.

Algoritma 2: LEC Hesaplama

Girişler:

- graf $G = (V; E)$

Çıktılar:

- Hedef ağa ait LEC ölçümü ve pozitif exponent frekansları

1: for each node in G do

2: $E(v_i) \leftarrow (e_1, e_2, \dots, e_n)$, mesafe sinyali vektörünü ata

3: $Le(v_i) \leftarrow (ex_1, ex_2, ex_3, ex_4, ex_5, ex_6)$, Lyapunov üst spektrumunu hesapla

4: end for

5: $Le(G) \leftarrow \{Le(v_1), Le(v_2), \dots, Le(v_n)\}$, G ağına ait Lyapunov üst matrisini oluştur

6: for each exponent in $Le(G)$ do

7: if exponent > zero then

8: increase $f_1, f_2, f_3, f_4, f_5, f_6$, ilgili frekans değerini güncelle

9: else

10: increase f_0 , pozitif olmayan frekans değerini güncelle

11: end if

12: end for

13: $pf(G) \leftarrow \{f_0, f_1, f_2, f_3, f_4, f_5, f_6\}$, G ağın ait frekans matrisini oluştur

14: $pfn(G) \leftarrow \text{normalize}_{0-1}(pf(G)) \times 100$, sonuçları 0-100 aralığına taşı

15: $LEC \leftarrow \text{mean}(pfn(G))$

16: return $LEC, pfn(G)$

Şekil 3.2 LEC Hesaplama Algoritması

3.5 D ğ m Merkezliđi

Graf teorisi ve ađ analizinde, bir ađdaki  nemli d ğ mleri tanımlamak i in d ğ m merkezi g stergeleri kullanılır. Ađ merkezlik g stergeleri mesafe  l  tleriyle yakından ili kilidir ve her bir d ğ m n ađ i indeki konumlarına g re  nemini karakterize etmek i in kullanılır. Merkeziyet kavramları ilk olarak sosyal ađ analizinde geli tirilmi tir ve merkezile meyi  l mek i in kullanılan terimlerin  ođu sosyolojik k kenlerini yansıtmaktadır [170]. D ğ mlerin merkezi olması ya da hangi d ğ mlerin diđerlerinden daha merkezi olduđunu belirlemesi ađ analizinde  nemli bir konudur [171].

Ger ek ađlardaki d ğ mlerin belirli bir sırası yoktur. D ğ m sırası veri tabanından veri tabanına veya veri tabanı i inde deđi iklik g sterebilir. D ğ m sırası, Lyapunov  st hesaplamasında kullanılacak sinyali deđi tireceđi i in sonu ları etkileyebilir. Bu  alı mada, d ğ m merkezlik  l  mleri ile ađlar i in bir sıralama standardı belirlenmi  ve sıralamanın sonu lar  zerindeki etkisi incelenmi tir.

$G = (V, E)$ ađındaki i d ğ m ne ait merkezliđi $c(v_i)$ ile g sterelim. Sıralama algoritmasının giri i, bir ađ olan G ve  ıktısı d ğ mleri hedef  l  me g re sıralanmı  yeni G_s ađıdır. Algoritma, hedef  l  m metriđini kullanarak belirli bir G ađındaki t m d ğ mlerin $\{c(v_1), \dots, c(v_n)\}$ merkezliyet deđerini hesaplar. İlk d ğ m sıra numarasını ve son d ğ m sıra numarasını sırasıyla o_1, o_n ile g sterelim. Algoritma, G ađındaki her d ğ m i in d ğ m merkezlik deđerlerine g re b y kl kten k   đe d ğ m sırasını belirler ve d ğ m sıra vekt r  $no = (o_1, o_2, \dots, o_n)$ olu turur. Son olarak algoritma, no sıralama vekt r n  kullanarak verilen G ađının d ğ m sırasını deđi tirir ve Lyapunov  s hesaplaması i in yeni sıralanmı  G_s ađını olu turur.

Derece Merkezliđi :

Derece merkezliđi, d ğ m merkezliyetini yalnızca d ğ mlerin etrafındaki yerel yapıyı kullanarak  l er. Y ns z ađdaki bir d ğ m n derecesi, dođrudan d ğ me bađlı olan biti ik d ğ mlerin sayısı ya da sadece o d ğ me bađlanan kenarların sayısı kullanılarak hesaplanabilir [172]. Ađdan akan bilgilere alternatif bir yol  retip  retmemeye a ısından bir avantaj veya risk sađladıđını temsil ettiđi i in derece merkezliđi  nemlidir [173]. Daha b y k dereceli bir d ğ m n, daha k   k dereceli bir d ğ mden daha y ksek etkisi olması muhtemeldir [174].

Arasındalık Merkezliđi :

Arasındalık merkezliđi, bir düđümün diđer iki düđüm arasındaki en kısa yol boyunca bir köprü görevi gördüğünün göstergesidir. Rastgele seçilen iki düđüm arasındaki en kısa rastgele seçilen yolda bulunma olasılığı yüksek olan düđümler yüksek bir ağırlığa sahiptir. Arasındalık merkezliđi, bir düđümün ağıdaki en kısa yollarda ne kadar önemli olduğunu ölçer [175]. Ölçülen düđümün bulunduğu ağıdaki en kısa tüm mesafelerin, ağıdaki en kısa tüm mesafelere oranı alınarak hesaplanır.

Belirlenen bir n düđümüne ait arasındalık merkezliđini hesaplamak için, bir çift düđüm seçilir ve bu düđümler arasındaki en kısa yollar bulunur. Daha sonra n düđümünü içeren en kısa yolların oranı hesaplanır [176]. Belirli bir ağı için $G = (V, E)$, σ_{st} ile s düđümünden t düđümüne toplam en kısa yol sayısını ve $\sigma_{st}(n)$ ile n düđümünden geçen s düđümünden t düđümüne toplam en kısa yol sayısını tanımlayalım. Hedef n düđümü için arasındalık merkezliđi aşağıdaki formülle (**Denk.19**) hesaplanır.

$$B(n) = \sum_{s \neq t \neq n \in V} \sigma_{st}(n) \quad (109)$$

Yakınlık Merkezliđi :

Bağılı bir ağıda, bir düđümün yakınlık merkezliđi, düđüm ile ağıdaki diđer tüm düđümler arasındaki en kısa yolun ortalama uzunluğudur. Böylece, yakınlık merkezliđi ne kadar fazlaysa, ilgili düđüm diđer tüm düđümlere o kadar yakın olacaktır [177]. Yakınlık merkezliđinin faydası, bir düđüm ağıdaki düđümlerin çoğuna daha yakın olması durumunda ilgili düđümleri daha merkezi olarak göstermesidir.

Bu ölçüm görsel merkez ile ilişkilidir. Bir ağı çizildiğinde ağıın merkezinde görünecek bir düđüm genellikle yüksek yakınlık merkezliđine sahiptir [175]. Belirli bir x ve y düđümleri arasındaki mesafeyi $d(y,x)$ ile gösterelim, x düđümü için yakınlık merkezi aşağıdaki formülle (**Denk.20**) hesaplanır.

$$C(x) = \frac{N - 1}{\sum_y d(x,y)} \quad (20)$$

Yük Merkezliği :

Yük merkezliği, bu düğümden geçen en kısa yolların bir oranıdır. Arasındalık merkezliği ölçümüne yakındır [178]. Yük merkezliği, varsayımsal bir akış süreci ile tanımlanan bir arasındalık benzeri ölçüdür [179]. Belirli bir ağ $G = (V, E)$ için, n düğümünden geçen ve düğüm s 'den düğüm t 'ye toplam en kısa yol sayısını $\sigma_{st}(n)$ ile tanımlayalım. Hedef n düğümü için yük merkezliği aşağıdaki formülle **(Denk.21)** hesaplanır.

$$L(n) = \sum \frac{1}{\sigma_{st}(n)} \quad (211)$$

Kümelenme Katsayısı Merkezliği :

Kümelenme katsayısı, bir ağdaki düğümlerin kümelenme derecesinin bir ölçüsüdür. Birçok gerçek ağda düğümler nispeten yüksek bir bağlantı yoğunluğu ile karakterize edilen sıkıca bağlı gruplar oluşturma eğilimindedir [159]. Bu merkezliğin sezgisel anlamları vardır. Kümelenme katsayısı, hedef düğümün komşularının da birbirlerinin komşuları olma derecesini yansıtır ve bu şekilde oluşan bir bağlantı çemberinin yoğunluğunu ölçer [180]. Belirlenen v düğümünün k adet komşuları ile içerdiği kenar sayısını k_v ve v düğümünün k adet komşusu ile içerebileceği kenar sayısını $p_v = k * (k - 1)$ ile tanımlayalım. Belirlenen v düğümü için kümeleme katsayısı merkezliği aşağıdaki formülle **(Denk.22)** hesaplanır.

$$C(v) = \frac{k_v}{p_v} \quad (22)$$

Rastsal Merkezlik :

Bu merkezlik ölçütünde, düğüm öncelikleri rastgele sayılarla belirlenir, böylece düğümler veri tabanlarında olduğu gibi rastgele sıradadır. Rastgele düğüm sırası oluşturmak için rastgele sayı üretici kullanılır. Her sayının bir kez kullanılmasına izin verilir, böylece sayılar tekrarlanmaz. Belirlenen n düğümü için üretilen rastgele sayı vektörünü $ro_n(o_1, o_2, \dots, o_n)$ ile tanımlayalım. Hedef ağ $G = (V, E)$ için, G_s sıralı ağını oluşturmak üzere rastgele sıra vektör ro kullanılarak G ağındaki düğümlerin sırası değiştirilir. G ve G_s aynı ağlar olmasına rağmen, G_s düğüm sırası G 'den farklıdır.

4. ARAŞTIRMA BULGULARI

4.1 Veri Seti

Çalışma kapsamındaki deneyler sentetik olarak üretilen ve gerçek biyolojik ağlar üzerinde kapsamlı olarak gerçekleştirilmiştir. Kullanılan ağlar ağ türü, düğüm sayısı, kenar sayısı ve ortalama düğüm derecesi özellikleri kullanılarak ifade edilir.

4.1.1 Sentetik Veri Seti

Mutasyonların farklı yapay ağlar üzerindeki etkisini gözlemlemek için farklı topolojik özelliklere sahip sentetik ağlar üzerinde deneyler yapılmıştır. Çalışmalar kapsamında Erdos Renyi (ER), Watts Strogats (WS), Barabasi-Albert (BA) adında üç tanınmış ağ modeli ve değişken parametreler (düğüm sayısı: 500, 1000, 1500, 2000, ortalama düğüm derecesi: 2; 3; 4, 5) kullanılmıştır. Her bir ağ için 5 tekrar oluşturularak, deneyler toplam 240 sentetik ağ üzerinde gerçekleştirilmiştir. Ek olarak, büyük ağların değerlendirilmesi için 60 büyük sentetik ağ üzerinde deneyler yapılmıştır. Bu kısımda, deneylerde kullanılan ağ modelleri kısaca açıklanmaktadır.

ER modeli rasgele bir ağ $G = (n, p)$ oluşturur; burada n kenar sayısıdır ve p olası kenarların olasılığıdır. Bu model düğüm çiftlerini p olasılığı ile bir kenarla bağlar. Ortaya çıkan ağdaki kenar sayısı $C_2^n \times p$ 'dir. ER modeli binom derece dağılımına sahiptir, ortalama yol uzunluğu çok küçüktür (küçük dünya) ve küçük kümelenme katsayısına sahiptir. Küçük dünya (SW) ağı olarak adlandırılan WS modeli, önce her bir düğümü tek boyutlu bir halka üzerinde düzenler ve daha sonra her düğümü en yakın $k/2$ komşu ile bağlar. Ortaya çıkan ağda, n düğüm için $n \times k / 2$ kenar vardır. WS ağı yüksek bir kümelenme katsayısına sahiptir ve ortalama yol uzunluğu sistem boyutuyla doğrusal olarak ölçeklenir. BA modeli kenarsız m düğüm ile başlar. Mevcut düğümlere kenar bağlantısı ile yeni düğümler eklenerek ağ büyütülür. Kenarlar tercihen yüksek dereceli düğümlere eklenir. Sonuçta ortaya çıkan ağda n düğüm için $(n - m) \times m$ kenar vardır. BA ağı, güç yasasını izleyen ölçeksiz (scale-free, SF) derece dağılımına sahiptir. Her ağ modeli için, hedef düğüm boyutu ve ortalama düğüm derecesi kullanılarak ağ oluşturulur ve bu parametreler kenar sayısını belirler. Ortaya çıkan ağın beklenenden daha az kenarı varsa işlem tekrarlanır. Eğer kenar sayısı daha fazla ise kenar sayısı kıstasını sağlamak için rastgele seçilen kenarlar silinir. Karmaşık sentetik ağların oluşturulması ve manipülasyonu için NetworkX Python paketi kullanılmıştır [178].

4.1.2 Gerçek Veri Seti

Çalışmada BioGRID [36] ve STRING [37] veri tabanlarından indirilen sekiz farklı organizmanın protein-protein etkileşim ağı kullanılmıştır. Ağlar çeşitli düğüm sayısına, kenar sayısına ve ortalama düğüm derecesine sahiptir.

Tekrarlı kenarlar kaldırdıktan ve tüm organizmalar için deneysel olarak türetilmiş fiziksel etkileşim PPI ağının en büyük bağlı bileşeni seçtikten sonra veri setleri üzerinde çalışılmıştır. Ortaya çıkan biyolojik ağların topolojik özellikleri aşağıdaki Çizelge 4.1'de özetlenmiştir.

Çizelge 4.1 : Biyolojik Ağ Parametreleri.

Organizma	Düğüm Sayısı	Kenar Sayısı	Ortalama Düğüm Derecesi
<i>S. cerevisiae</i> (7)	5.936	65.139	10,9
<i>Homo sapiens</i> (8)	6.122	14.426	2,3
<i>Arabidopsis thaliana</i> (6)	5.726	13.409	2,3
<i>Caenorhabditis elegans</i> (5)	2.937	5.333	1,8
<i>Plasmodium falciparum</i> (4)	1.262	2.598	2,0
<i>Helicobacter pylori</i> (1)	733	1.480	2,0
<i>Mus musculus</i> (2)	1.340	1.416	1,0
<i>Rattus norvegicus</i> (3)	631	704	1,1

4.2 Deneysel Ortam

Tüm deneyler Linux ve Windows işletim sistemleri üzerinde gerçekleştirilmiştir. Zaman karmaşıklığı sorununun üstesinden gelmek için tek komut çoklu veri (MIMD) dağıtık mimari tasarlanarak [181] algoritmalarımız 212 GB RAM ile 421 işlemci üzerinde çalıştırılmıştır.

4.3 Deneysel Çalışma 1

4.3.1 Sentetik Veriler Üzerine Sonuçlar

Farklı topolojik özellikleri temsil eden çeşitli parametreler altında mutasyonların sentetik ağlar üzerindeki etkisini gözlemlemek için sentetik ağlar üzerinden deneyler yapılmıştır. İlgili performanslar, önceden tanımlanmış uygunluk ölçümleri kullanılarak belirlenmiştir.

4.3.1.1 Yöntem karşılaştırması

Çalışmanın bu kısmındaki temel amacımız, seçilen kenar mutasyonlarının düşük bir yüzdesinin ağ topolojisi ve difüzyon durumu üzerindeki etkisini değerlendirmektir. Bu nedenle, çözülmesi gereken ilk problem farklı ağ topolojileri için uygun algoritmanın bulunmasıdır. Yapmış olduğumuz bu çalışma ile algoritmamızın performansının farklı ağ topolojileri için uygun olduğunu ve çeşitli ağ parametreleri altında en iyiye yakın bir sonuç üretmeyi garanti ettiğini göstermek amaçlanmıştır. Kıyaslamanın yapılması için metodumuzun ağda oluşturduğu yüzde CC değişimini ($\Delta CC(\%)$) Tepe Tırmanma (HC) ve Genetik Algoritma (GA) ile karşılaştırılmıştır. Mutasyon oranı tüm kenarların %1'ine sabitlenmiştir. Ağ modelleri olarak ER, WS, BA, düğüm sayısı olarak 500, 1000, 1500, 2000, ortalama düğüm derecesi sayısı olarak 2, 3, 4, 5 kullanılmış ve deneyler 5 tekrarlı olarak gerçekleştirilmiştir. Çizelge 4.2 karşılaştırma sonuçlarını özetlemektedir. Diğer parametrelerin ortalaması alınarak sonuçlar ağ modeline ve düğüm sayısına göre gruplanmıştır. Sonuçlar, HC algoritmasının daha fazla kenara ve yüksek CC'ye sahip ağlarda daha iyi performansa sahip olduğunu göstermektedir. GA algoritması, daha az kenarı olan ağlarda daha başarılı sonuçlara sahiptir. Geliştirdiğimiz algoritmamız, tüm ağ parametreleri için rakip algoritmalarından daha iyi sonuçlara sahiptir. Çalışmanın devamındaki tüm deneyler, tüm koşullara uygunluğu gösterilen algoritmamız ile yapılmıştır. Her üç ağ modeli içinde, ağ boyutları büyüdükçe mutasyonların etkisi ve CC değişimi artar. Özellikle, düğüm sayısının artırılması kararlılığı azaltır ve ortalama düğüm derecesinin arttırmak kararlılığı artırır. Mutasyonların WS ağlarının kararlılığı üzerinde ihmal edilebilir bir etkisi ve ER ağları üzerinde önemli bir etkisi vardır. Ayrıca, sonuçlar kararlılığın ağ boyutuna kıyasla ağ modelinden daha fazla etkilendiğini göstermektedir. Bunun altında yatan neden, kümelenme katsayısı değerinin, ağ boyutu yerine ağ modeli tarafından büyük ölçüde belirlenmesidir.

Sonuç olarak, deneysel sonuçlarımız ağ kararlılığının ağın modeline ve topolojik özelliklerine bağlı olduğunu göstermektedir. Çeşitli topolojik özellikler altında mutasyonlar, farklı büyüklükteki hedef ağların kararlılığı üzerinde bir etkiye sahiptir. CC ölçümüne göre, topolojik mutasyonlara karşı en dirençli ağ modeli WS ağ modelidir.

Çizelge 4.2 : Sentetik Ağlarda Algoritma Karşılaştırması.

Model	Düğüm Sayısı	Başlangıç CC	Tepe Tırmanma $\Delta CC(\%)$	Genetik Algoritma $\Delta CC(\%)$	Genetik+Tepe $\Delta CC(\%)$
ER	500	0,01245	49,20	55,55	55,55
	1000	0,00619	70,27	76,07	76,07
	1500	0,00438	80,56	84,37	84,64
	2000	0,00302	89,42	89,98	90,33
WS	500	0,43369	2,79	2,93	2,94
	1000	0,44013	3,18	3,27	3,28
	1500	0,43507	3,33	3,38	3,44
	2000	0,43851	3,37	3,32	3,44
BA	500	0,044897	28,98	32,14	32,14
	1000	0,03204	40,53	44,44	44,44
	1500	0,02301	49,20	53,99	54,15
	2000	0,01845	54,35	58,76	58,84

4.3.1.2 Ağ parametrelerinin değerlendirilmesi

Cevaplamak istediğimiz sorulardan bir diğeri, mutasyon etkisinin sentetik ağların topolojik özelliklerine ve mutasyon yanlılığına bağlı olarak değişip değişmediğidir. Bu deneyde, bu soruya cevap vermeyi ve ağ topolojisini yöneten farklı parametrelerin ağın kararlılığını nasıl etkilediğini değerlendirmek amaçlanmıştır.

Topolojik deęişikliklerin etkisini gözlemlemek için mutasyon oranı kenarların %1'ine sabitlenmiştir. Ağ modelleri olarak ER, WS, BA, düğüm sayısı olarak 500, 1000, 1500, 2000, ortalama düğüm derecesi sayısı olarak 2, 3, 4, 5 kullanılmış ve deneyler 5 tekrarlı olarak gerçekleştirilmiştir.

Mutasyon yanlılığını değerlendirmek için, mutasyon algoritmamız, başlangıç ve mutasyona uğramış ağlar arasındaki RE^{CC} , RE^{RW} ve DKL^{RW} objektif fonksiyonlarını maksimize ederek çalıştırılmıştır. Hedef ağların kararlılığını değerlendirmek için RE^{RW} , RE^{DSD} , DKL^{RW} , DKL^{DSD} , PCC^{RW} , PCC^{DSD} , NS^{RW} ve NS^{DSD} ölçümlerini kullanılmıştır. Şekil 4.1, Şekil 4.2 ve Şekil 4.3 sonuçları göstermektedir.

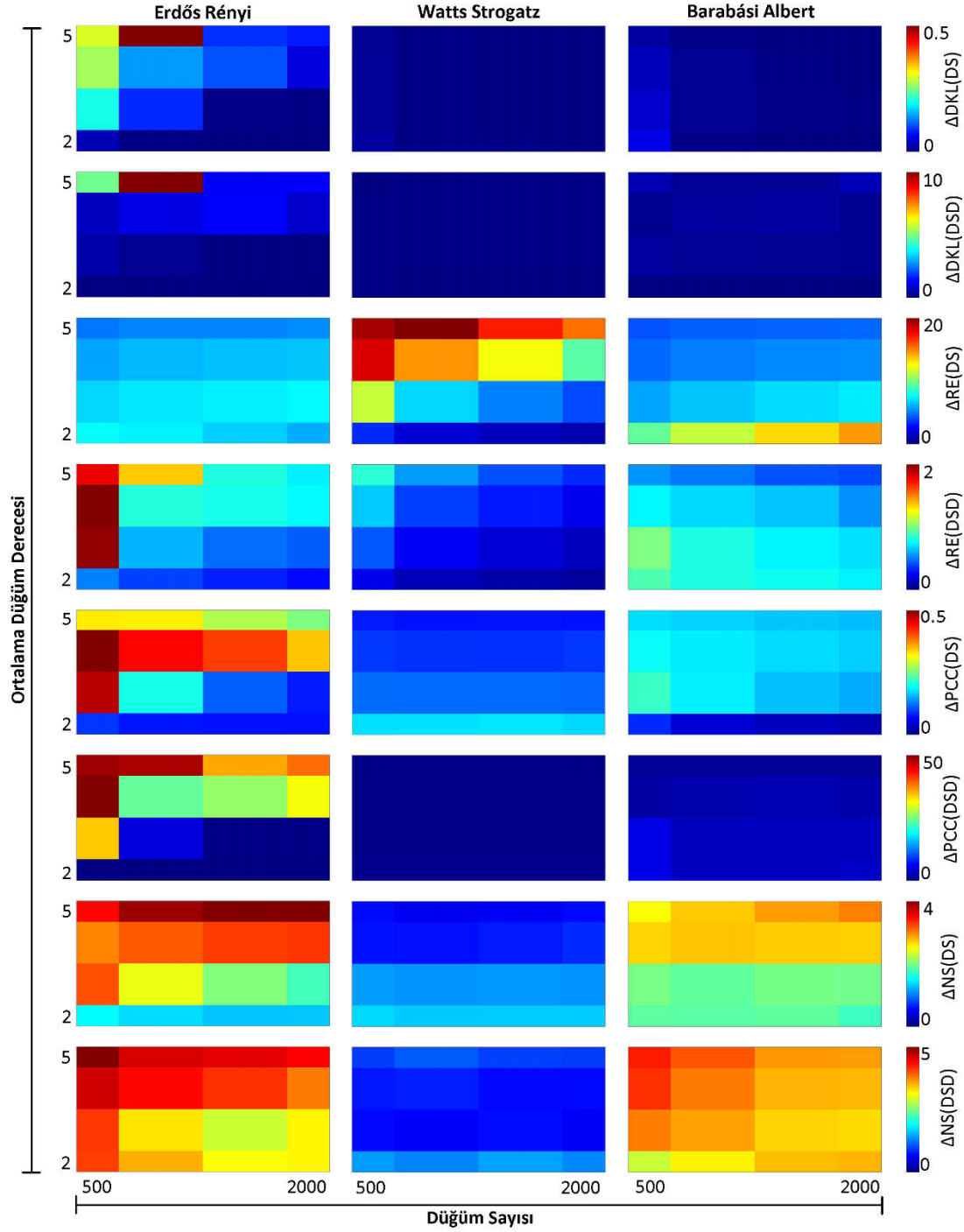
ER, WS ve BA sentetik ağlarının RW ve DSD dağılımlarının DKL ölçümleri açısından mutasyonlara karşı dayanıklı olduğu görülmektedir. DKL^{RW} maksimizasyon deneyinde DKL^{DSD} ölçümleri dışında tüm sonuçlar sıfıra yakındır. RW dağılımını hedef alan deneylerde, ER ve BA ağlarının DSD dağılımları bozulma eğilimindedir. WS ağı en sağlam ve BA ağı en kırılgandır.

Tüm deneylerdeki ölçümlerde, RE^{RW} deęişimleri anlamlı ve RE^{DSD} deęişimleri göz ardı edilebilir düzeydedir. En büyük deęişimler RE^{RW} maksimizasyon deneylerinde gözlenmiştir. Bu deneylerde, WS ağları en kırılgandır ve kırılganlık ağın ortalama düğüm derecesi ile artar. RE^{DSD} 'nin kararlılığı, RW'nin L1 normu hesaplamasında pozitif ve negatif deęişikliklerin nötralizasyonundan kaynaklanır.

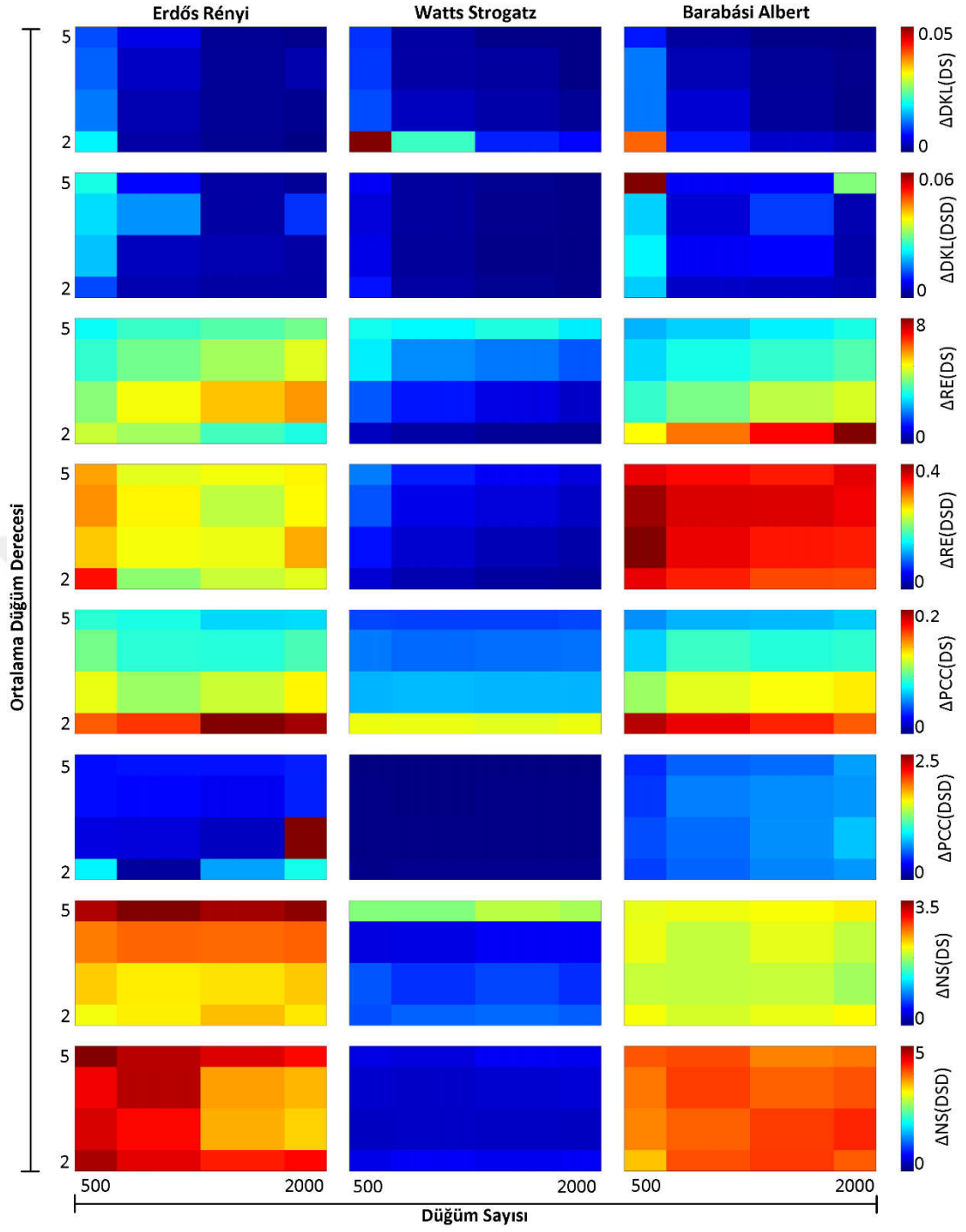
RE^{RW} , RE^{DSD} ölçümlerinden farklı olarak, PCC^{RW} ölçümlerindeki deęişiklikler önemsizdir ve PCC^{DSD} ölçümleri anlamlıdır. Gözlenen bu ölçümler, RW'deki deęişikliklerin, düğümler arasındaki benzer ilişkili deęişimlerden ve DSD'deki deęişimlerin düğümden düğüme farklılık gösteren deęişikliklerden kaynaklandığını göstermektedir. PCC^{DSD} 'nin kararsızlığı, düğümler arasında uygulanan L1 normunun aritmetik ortalama işleminden kaynaklanır. Bu ölçümlere göre, WS ağları en sağlam ve ER ağları en kırılgandır.

Tüm deneylerde düğümlerin komşuları deęişmiştir. DSD, NS ölçümü açısından RW'den daha kırılgandır. Yapılan ölçümlere göre, WS ağları en sağlam ve ER ağları en kırılgandır.

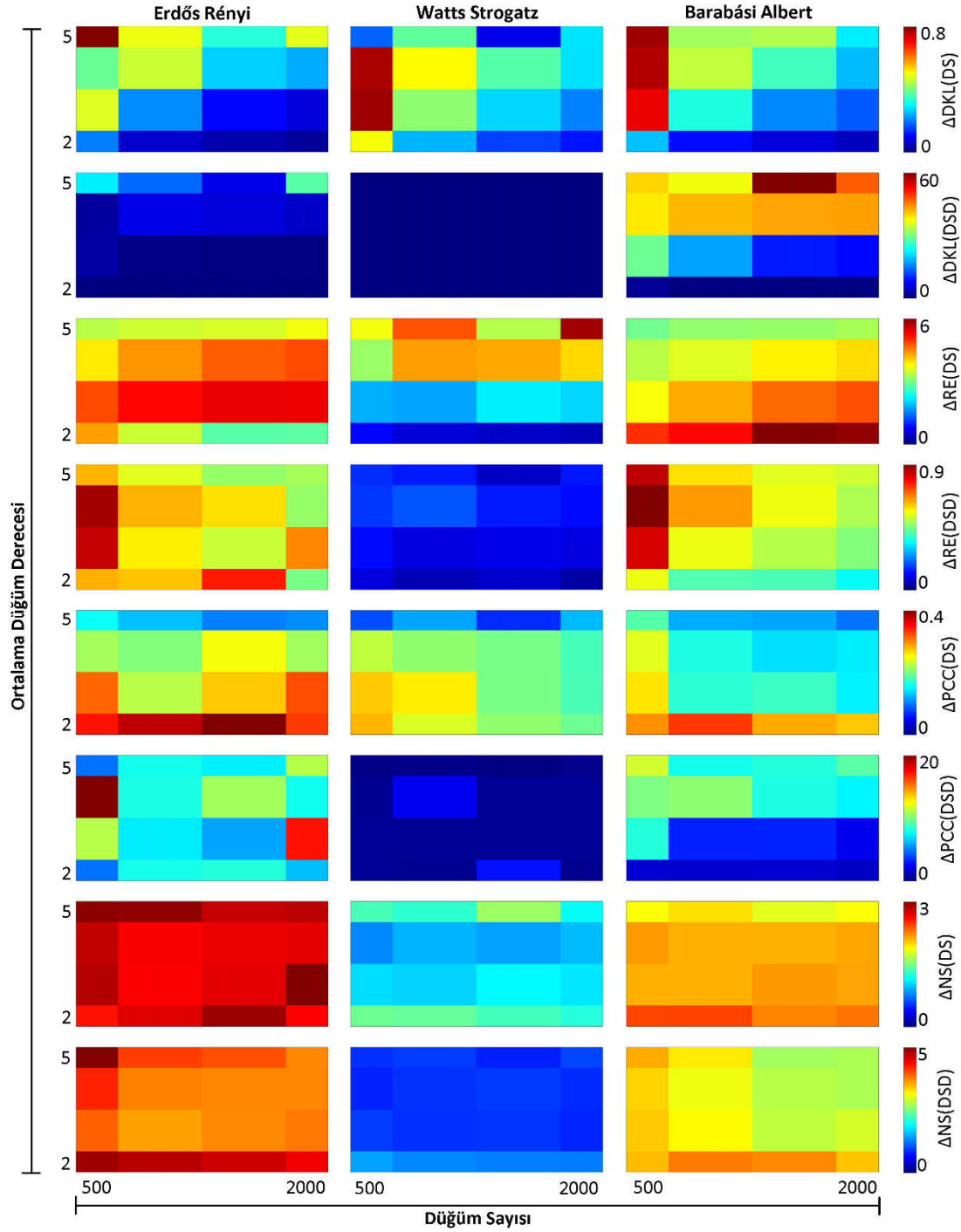
Sonuç olarak, deneysel sonuçlarımız, RE^{RW} maksimizasyon deneylerinin hedef ağlarda kenar mutasyonları için en etkili kenar silme kümelerini ürettiğini göstermektedir. Tüm deneylerde yer alan tüm ölçüm türleri için, sonuçlar ağ türlerine göre benzerlik göstermektedir. Bu sonuçlar, kenar mutasyonlarının etkisinin, sentetik ağların topolojik özelliklerine, özellikle ağ modeline ve mutasyon yanlılığına bağlı olarak değiştiğini göstermektedir.



Şekil 4.1 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi (RE^{RW}).



Şekil 4.2 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi (RE^{CC}).



Şekil 4.3 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi (DKL^{RW}).

4.3.1.3 Mutasyon oranının değerlendirilmesi

Bu çalışma için motivasyonlarımızdan biri protein etkileşim ağlarının gürültülü olması ve değişken oranlarda yanlış pozitif ve yanlış negatif bilgiye sahip olmasıdır. Sıradaki sorun ağ kararlılığının mutasyon oranına göre değişip değişmediğidir.

Bu deneyde, mutasyon oranının ağ kararlılığı üzerindeki etkisi ölçülmüştür. Düğüm sayısı ve ortalama düğüm derecesi sırasıyla 1000 ve 3'e sabitlenmiştir. Etkiyi incelemek için mutasyon oranı 0,5'ten 4'e kadar çıkarılmıştır. Her ağ modeli için 5 tekrarlı olarak %0,5, %1, %2, %4 mutasyon oranları kullanılmıştır. Mutasyon algoritmamız, başlangıç ve mutasyona uğramış ağlar arasında RE^{CC} , DKL^{RW} ve RE^{RW} objektif fonksiyonlarındaki değişimleri maksimize edecek şekilde çalıştırılmıştır. Hedef ağların kararlılığını değerlendirmek için RE^{RW} , RE^{DSD} , DKL^{RW} , DKL^{DSD} , PCC^{DSD} , PCC^{RW} , NS^{RW} ve NS^{DSD} ölçümleri kullanılmıştır. Şekil 4.4, Şekil 4.5 ve Şekil 4.6 sonuçları göstermektedir.

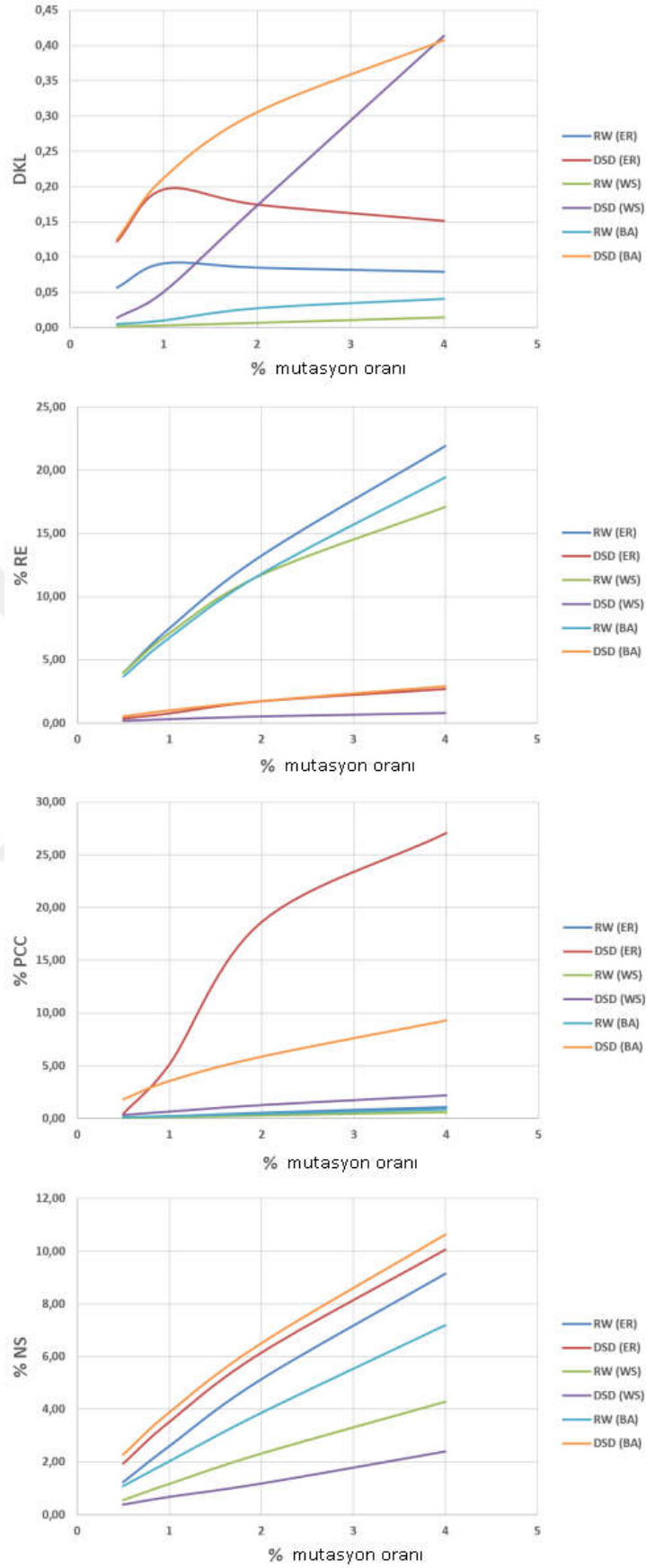
Gürültü oranının arttırılması, tüm sentetik ağlar ve tüm deneyler için DKL ölçümleri üzerinde ihmal edilebilir bir etkiye sahiptir. BA ağ modelinin DKL^{DSD} ölçümü hariç, tüm sonuçlar sıfıra yakındır.

Tüm RE ölçüm sonuçları artan mutasyon seviyesi ile artmaktadır. RW hata oranları, DSD hata oranlarından daha fazla artmaktadır. Ağ modelleri arasında WS en kararlı ve ER en az kararlı olan ağlardır.

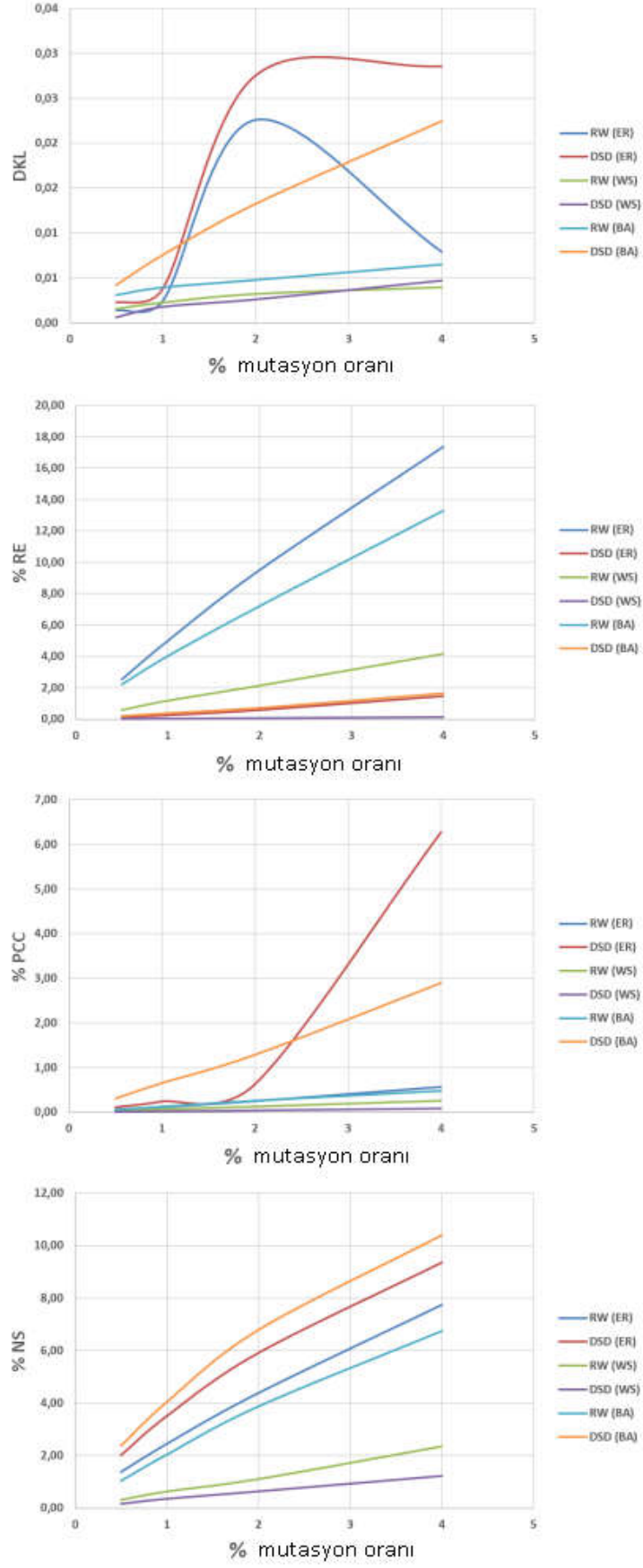
PCC ölçümlerine göre, mutasyon oranının arttırılması ER ve BA ağ modellerinin PCC^{DSD} ölçümlerinde çarpıcı bir artışa neden olur. Tüm ağların PCC^{RW} ölçümü sabit kaldığı görülmektedir.

Tüm deneylerde komşu benzerlik hataları artan mutasyon oranı ile artmaktadır. NS ölçümleri açısından DSD, RW'den daha fazla etkilenir. WS ağ modeli en sağlam, BA ve ER ağ modelleri kırılgandır. Özellikle BA ağ modeline ait NS^{DSD} ölçümü önemli ölçüde artmaktadır.

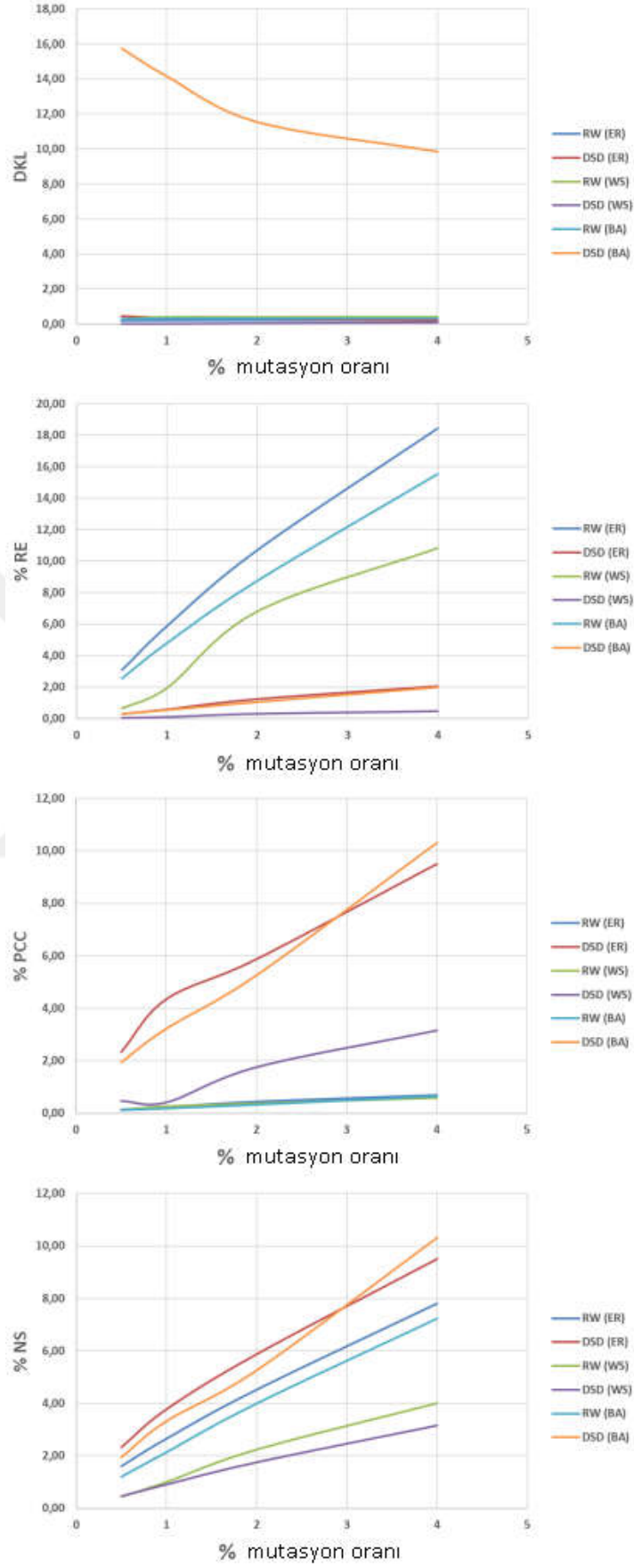
Sonuç olarak, deneysel çalışmalarımız tüm ağ modellerinin hata oranlarının, artan mutasyon oranlarıyla birlikte arttığını göstermektedir. Mutasyonların etkililik oranları ağ modellerine ve bilgi çıkarma yöntemlerine bağlı olarak değişmektedir.



Şekil 4.4 Sentetik Ağlarda Mutasyon Oranının Değerlendirilmesi (RE^{RW}).



Şekil 4.5 Sentetik Ağlarda Mutasyon Oranının Değerlendirilmesi (RE^{CC}).



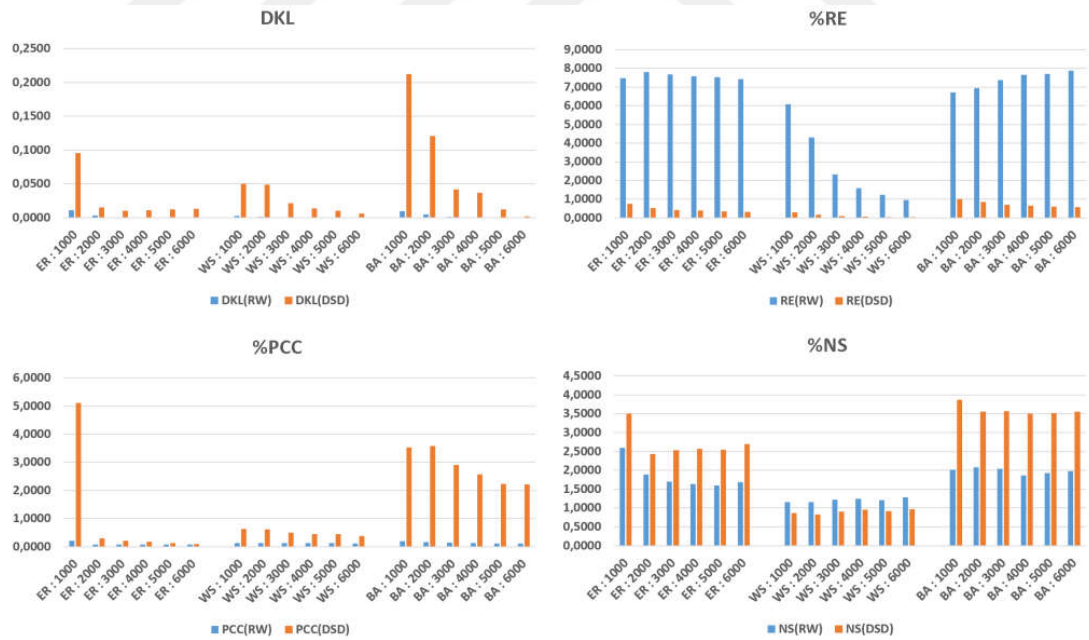
Şekil 4.6 Sentetik Ağlarda Mutasyon Oranının Değerlendirilmesi (DKL^{RW}).

4.3.1.4 Ağ boyutunun değerlendirilmesi

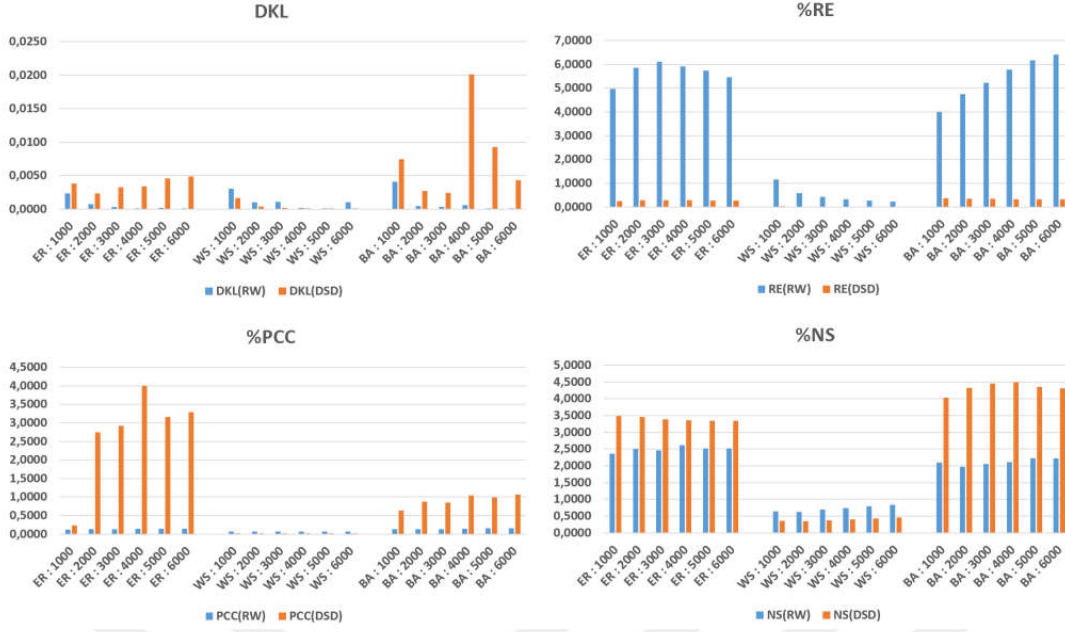
Biyolojik ağlardaki düğüm sayısı değişir ve yüksek sayılara ulaşabilir. Problemimiz sentetik ağ kararlılığının yüksek oranda artan düğüm sayısına göre değişip değişmediğidir. Bu deneyde, düğüm boyutunun ağ kararlılığı üzerindeki etkisi ölçülmektedir. Ortalama düğüm derecesi ve mutasyon oranı sırasıyla 3 ve 1'e sabitlenmiştir. Etkiyi incelemek için düğüm sayısı 1000'den 6000'e kadar yükseltilmiştir. Mutasyon algoritmamız başlangıç ve mutasyona uğramış ağlar arasında RE^{CC} , DKL^{RW} ve RE^{RW} objektif fonksiyonlarını maksimize edecek şekilde çalıştırılmıştır. Her ağ modeli için deneyler 5 kez tekrarlanmıştır.

Şekil 4.7, Şekil 4.8 ve Şekil 4.9 sonuçları göstermektedir. Daha büyük ağlar için, WS ağları daha güçlü olma eğiliminde ve BA ağları RE ölçümü açısından daha kırılgan olma eğilimindedir. NS ölçümü ağ boyutundan etkilenmez.

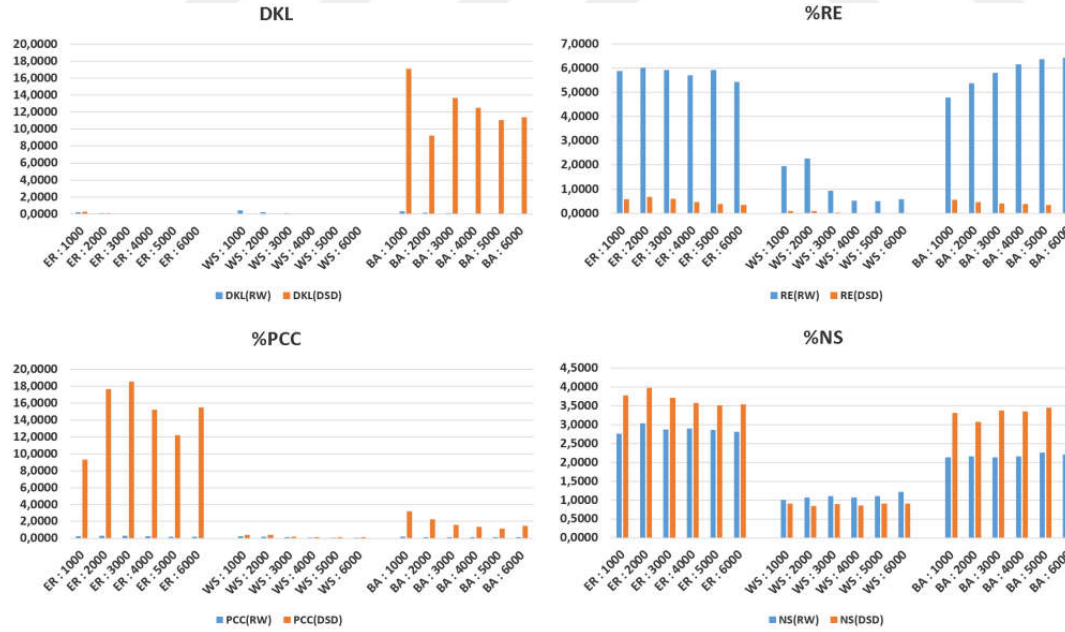
Sonuç olarak, bu deneyler düğüm sayısının artırılmasının tüm deney ve ölçüm türleri üzerinde ihmal edilebilir bir etkiye sahip olduğunu, ağ model ve topolojisinin ağ kararlılığı için düğüm sayısından daha önemli olduğunu göstermektedir.



Şekil 4.7 Sentetik Ağlarda Ağ Boyutunun Değerlendirilmesi (RE^{RW}).



Şekil 4.8 Sentetik Ağlarda Ağ Boyutunun Değerlendirilmesi (RE^{CC}).



Şekil 4.9 Sentetik Ağlarda Ağ Boyutunun Değerlendirilmesi (DKL^{RW}).

4.3.2 Biyolojik Veriler Üzerine Sonuçlar

Şimdiye kadar, topolojik mutasyonlar altında çeşitli parametrelerle sentetik ağların kararlılığı incelenmişti. Bununla birlikte, topolojik mutasyonlar altında biyolojik ve sentetik ağların kararlılığı arasında farklılıklar olabilir.

Bu bölümde topolojik mutasyonların biyolojik ağlar üzerindeki etkisini gözlemlemek için çeşitli düğüm sayısı, kenar sayısı ve ortalama düğüm derecelerine sahip farklı organizmalara ait sekiz protein etkileşim ağı kullanılmıştır. Sentetik ağlar üzerinde gerçekleştirilen RE^{RW} maksimizasyon deneyleri, hedef ağlardaki kenar mutasyonları için en etkili kenar silme kümesinin üretildiğini göstermektedir. Bu nedenle, bu bölümde mutasyon algoritmamız, biyolojik veri kümelerinde RE^{RW} objektif fonksiyonunu maksimize edecek şekilde kullanılmıştır. Ağların performansları, önceden tanımlanmış olan uygunluk ölçümleri kullanılarak incelenmiştir.

4.3.2.1 Organizmaların değerlendirilmesi

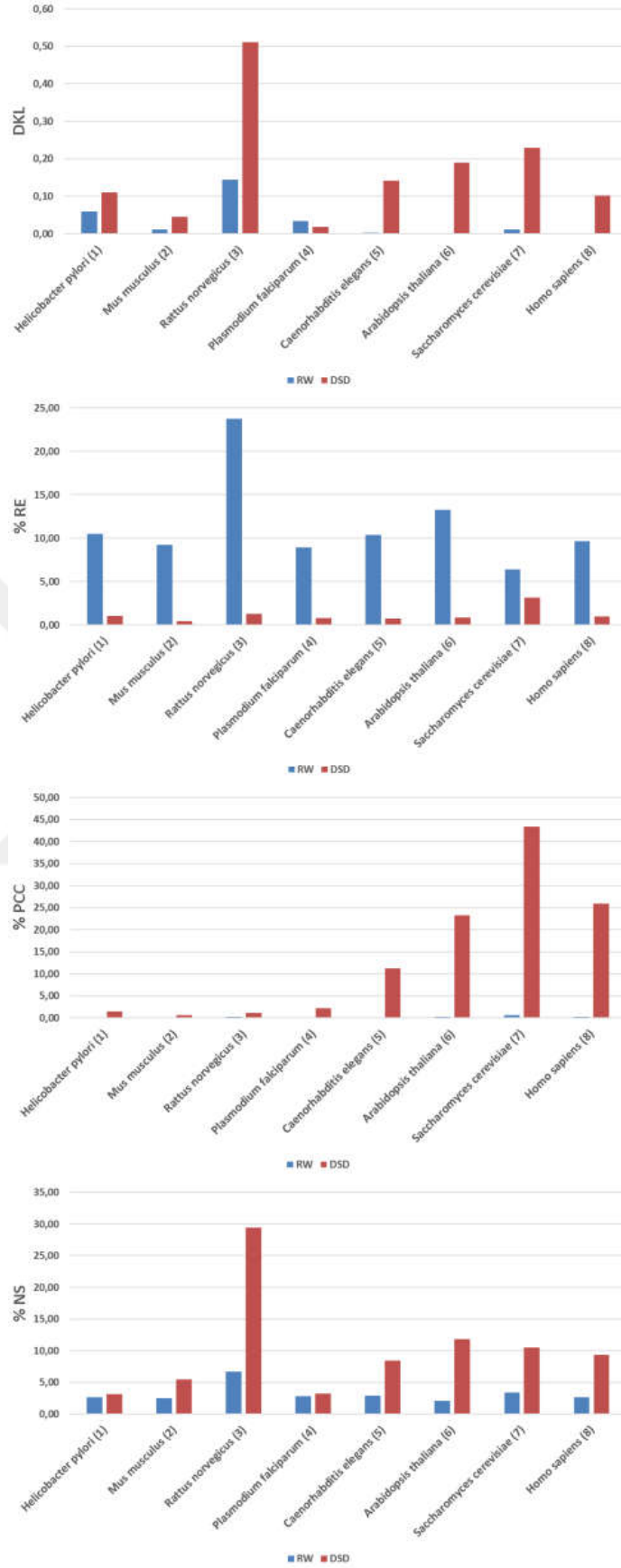
Çalışmamızdaki bir sonraki soru, sentetik ve biyolojik ağların kararlılık açısından farklı olup olmadığıdır. Bu deneyde, farklı biyolojik ağların kararlılığı incelenmiştir.

Topolojik değişikliklerin etkisini gözlemlemek için mutasyon oranı kenarların % 1'ine sabitlenmiştir. Bir mutasyonun biyolojik ağlar üzerindeki etkisini değerlendirmek için, mutasyon algoritmamız ilk ve mutasyona uğramış ağlar arasındaki RE^{RW} objektif fonksiyonunu maksimize edecek şekilde çalıştırılmıştır. Hedef ağların kararlılığını değerlendirmek için RE^{RW} , RE^{DSD} , DKL^{RW} , DKL^{DSD} , PCC^{RW} , PCC^{DSD} , NS^{RW} ve NS^{DSD} ölçümleri kullanılmıştır. Şekil 4.10 sonuçları göstermektedir.

Tüm organizmalara ait PPI ağlarının RW ve DSD dağılımlarının, mutasyonlara karşı değişen etki oranlarında kırılan olduğu görülmektedir.

Sentetik ağların deneysel sonuçlarına benzer şekilde, RE^{RW} ölçümleri anlamlıdır ve RE^{DSD} ihmal edilebilir. En kırılğan ağ "*Rattus norvegicus* (3)" ve en sağlam ağ "*Saccharomyces cerevisiae* (7)" dir.

RE ölçümlerinin aksine, PCC^{RW} , NS^{RW} ve DKL^{RW} değişiklikleri ihmal edilebilir ve PCC^{DSD} , NS^{DSD} ve DKL^{DSD} ölçümleri anlamlıdır. Biyolojik ağlar sentetik ağlarla aynı özellikleri göstermektedir.



Şekil 4.10 Organizmaların Değerlendirilmesi.

NS ölçümü açısından DSD, RW'den daha kararsızdır. Tüm düğümlerin komşuları mutasyonlar altında yüksek değişim oranlara sahiptir. PCC, NS ve DKL ölçümleri bu sonuçların kanıtıdır. Özellikle DKL ölçümü komşu değişime (NS) oldukça benzemektedir. En kararsız ağ "*Rattus norvegicus* (3)" ve en kararlı ağ "*Helicobacter pylori* (1)" dir.

Sonuç olarak, deneysel çalışmalarımız organizmaların RW ve DSD dağılımlarının kırılma olduğunu göstermektedir. Doğası gereği, RW dağılımları yüksek miktarda ve homojen RE değişiklikleri gösterme eğilimindedir ve DSD dağılımları mutasyonlar altında düşük miktarda ve heterojen RE değişikliklerine eğilimlidir. RW dağılımlarının homojen değişiminin doğası, onları DSD dağılımlarından daha kararlı hale getirir. RE ve DKL ölçümleri, olası protein mutasyonlarına karşı hedef protein etkileşim ağındaki komşu değişikliklerini tahmin etmek için kullanılabilir. Mutasyonlar altında, organizmaların PPI ağlarının kırılma eğilimi değişir.

4.3.2.2 Mutasyon oranının değerlendirilmesi

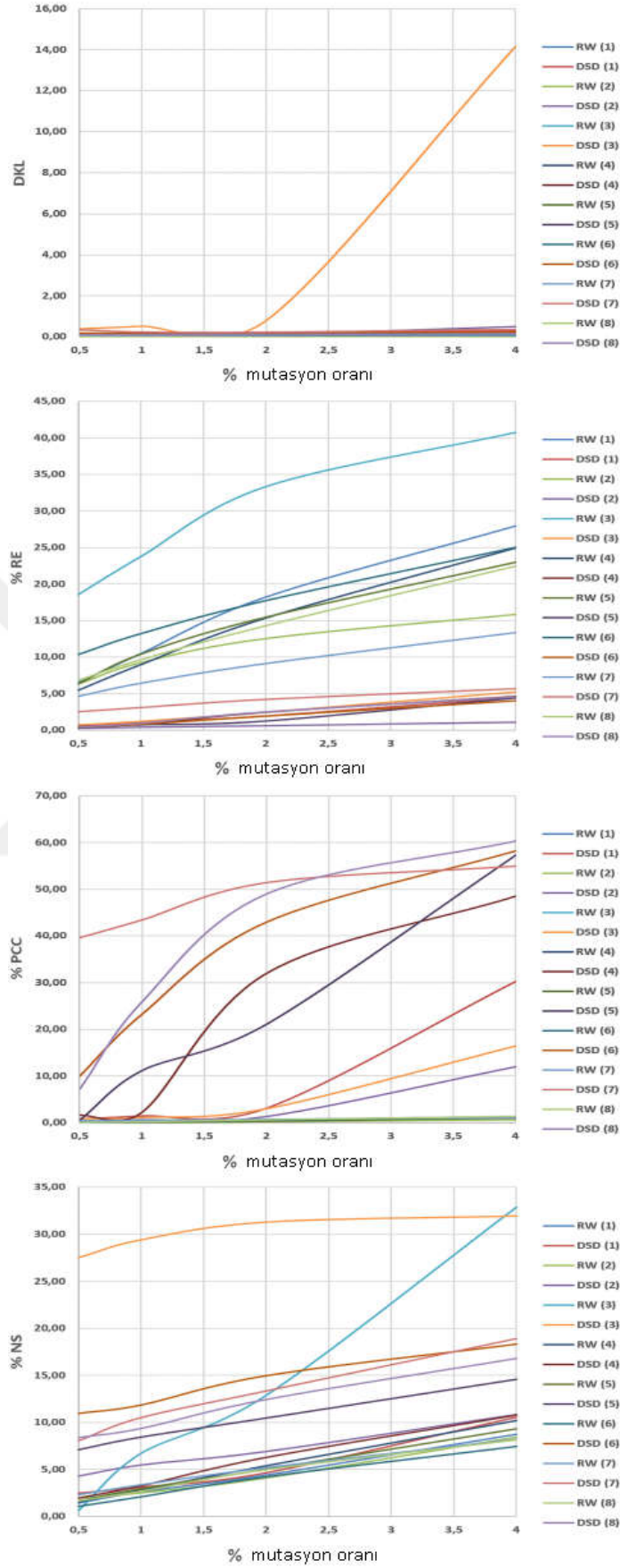
Protein etkileşim ağları değişen mutasyon oranlarına sahiptir. Çalışmamızdaki bir sonraki soru, gerçek ağ kararlılığının mutasyon oranına göre değişip değişmediğidir. Bu deneyde mutasyon oranının biyolojik ağ kararlılığı üzerindeki etkisi ölçülmüştür.

Etkinin araştırılması amacıyla her bir biyolojik ağ için %0,5, %1, %2, %4 mutasyon oranları kullanılmıştır. Mutasyon algoritmamız başlangıç ve mutasyon geçirmiş ağlar arasındaki RE^{RW} objektif fonksiyonunu maksimize etmek için çalıştırılmıştır. Hedef ağların kararlılığını değerlendirmek için RE^{RW} , RE^{DSD} , DKL^{RW} , DKL^{DSD} , PCC^{DSD} , PCC^{RW} , NS^{RW} ve NS^{DSD} ölçümleri kullanılmıştır. Şekil 4.11 sonuçları göstermektedir.

Mutasyon oranının artırılmasının DKL ölçümleri üzerinde önemsiz etkiye sahiptir ve mutasyon oranını arttırmak DKL sonuçlarını da artırır.

Tüm RE sonuçları artan mutasyon oranı ile artar. RW hata oranları, DSD hata oranlarından daha fazla artar. RE ölçümüne göre "*Rattus norvegicus* (3)" en kararsız ve "*Mus musculus* (2)" en kararlı ağlardır.

PCC ölçümlerine göre, mutasyon oranının artırılması PCC hatasının önemli ölçüde artmasına neden olur. RW dağılımı DSD dağılımından daha kararlıdır.



Şekil 4.11 Mutasyon Oranının Gerçek Ağlarda Değerlendirilmesi.

Tüm organizmaların komşu hatası mutasyon oranı ile artar. NS ölçümü açısından DSD dağılımı mutasyonlardan daha fazla etkilenir. “*Rattus norvegicus* (3)” en kararsız ve “*Arabidopsis thaliana* (6)” en kararlı organizmalardır.

Sonuç olarak, deneysel sonuçlarımız tüm organizmaların hata oranının artan mutasyon oranlarıyla birlikte arttığını göstermektedir. Mutasyonlar altında hata oranları organizmadan organizmaya değişir. Biyolojik ağ sonuçları sentetik ağ sonuçlarına benzerdir. Dolayısıyla, ağların özellikleri topolojik mutasyonlar altında ağların kararlılığı üzerinde bir etkiye sahiptir.

4.4 Deneysel Çalışma 2

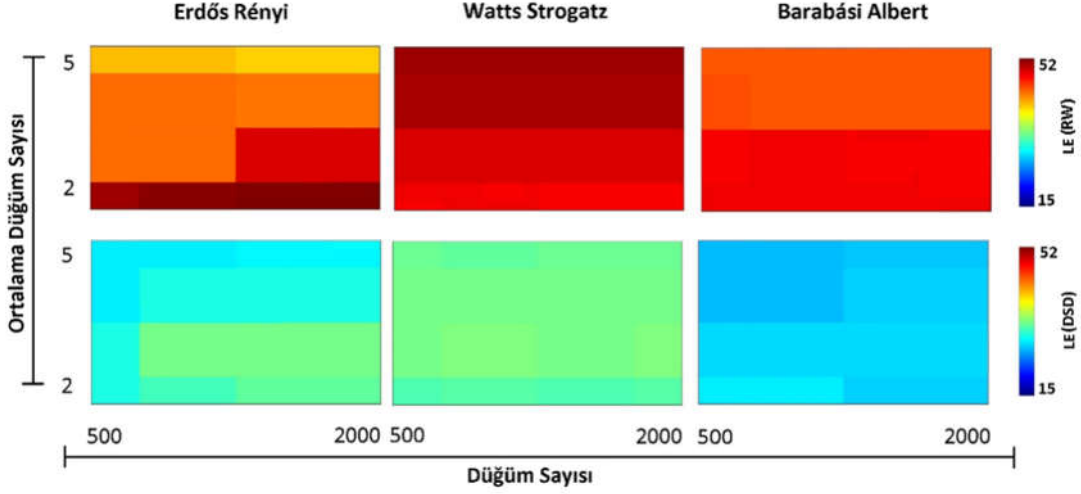
4.4.1 Sentetik Veriler Üzerine Sonuçlar

Lyapunov üssünün sentetik ağlar üzerindeki yanıtını araştırmak için sentetik ağlar kullanılmıştır. Çalışma, farklı topolojik özellikleri temsil eden çeşitli ağ parametreleri ve bu ağlar üzerinde çeşitli ağ mutasyonlarını içermektedir. Lyapunov üssüne ait sonuçlar, orijinal ve mutant ağlarda önceden tanımlamış olduğumuz Lyapunov üs ölçüsü kullanılarak ölçülmüştür.

4.4.1.1 Ağ parametrelerinin değerlendirilmesi

Bu çalışma, ağ topolojisi üzerinde Lyapunov üst tepkisini araştırmaya odaklanmaktadır. Bu nedenle, önce Lyapunov üst yanıtının sentetik ağların topolojik özelliklerine bağlı olup olmadığı araştırılmıştır. Bu deneyin amacı, ağ topolojisini düzenleyen farklı parametrelerin hedef ağın Lyapunov üst tepkisini nasıl etkilediğini değerlendirmektir.

Topolojik değişikliklerin etkisini gözlemlemek için, ağ modeli olarak ER, WS ve BA, düğüm sayısı olarak 500, 1000, 1500 ve 2000, ortalama düğüm derecesi sayısı olarak 2, 3, 4 ve 5 parametreleri kullanılarak deneyler beş tekrarlı olarak gerçekleştirilmiştir. Lyapunov üssünün hedef ağlardaki yanıtını değerlendirmek için Le(G) ölçümü kullanılmıştır. Şekil 4.12 sonuçları göstermektedir.



Şekil 4.12 Sentetik Ağlarda Ağ Parametrelerinin Değerlendirilmesi.

Şekil 4.12'de gösterildiği gibi, ER, WS ve BA sentetik ağlarının RW ve DSD dağılımları farklı Lyapunov üs tepkisine sahiptir. Tüm deneylerde RW dağılımı, DSD dağılımından daha yüksek bir Lyapunov üs ölçümüne sahiptir. Her ağ modeli ve dağılımı için, ortalama düğüm derecesi ve düğüm sayısı topolojik özelliklerinin Lyapunov üs tepkisi üzerinde neredeyse hiçbir etkisi yoktur ve tüm sonuçlar hemen hemen aynıdır.

Tüm deneylerde, WS sentetik ağları ER ve BA sentetik ağlarından daha yüksek Lyapunov üs ölçümüne sahiptir. ER ve BA sentetik ağlarının sonuçları çok benzerdir.

Sonuç olarak, bu çalışma büyük ölçüde ağ topolojisine bağımlı olan, ağların Lyapunov üs tepkisini anlamamıza katkıda bulunmaktadır. Bu gözlemler, Lyapunov üs tepkisinin sentetik ağların özellikle ağ model ve dağılım topolojik özelliklerine bağlı olarak değiştiğini göstermektedir.

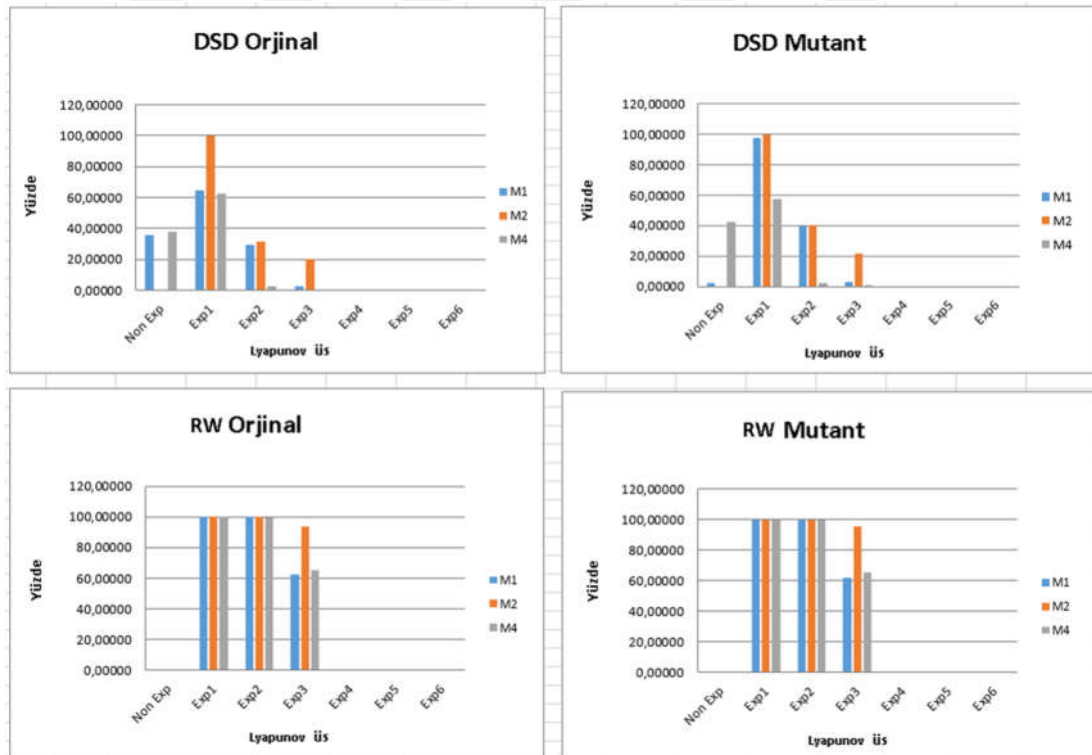
Sonuçlar ağ kararlılığının sonuçlarıyla karşılaştırıldığında, aralarında benzerlikler vardır. Karşılaştırmalar, Lyapunov üssünün, yüksek kararlılığa sahip ağlarda daha yüksek olduğunu göstermektedir.

4.4.1.2 Mutasyonların değerlendirilmesi

Bu çalışma için motivasyonlarımızdan biri protein etkileşim ağlarının gürültülü olmasıdır. Değişken oranlarda yanlış pozitif ve yanlış negatif kenarlar içerirler. Bu kapsamda bir sonraki soru ağların Lyapunov üslerinin topolojik mutasyonlara göre değişip değişmediğidir.

Bu deneyde, mutasyonların hedef ağın Lyapunov üsleri üzerindeki etkisi ölçülmüştür. Topolojik mutasyonların etkisini gözlemlemek için ER, WA, BA ağ modelleri kullanılmıştır. Toplam düğüm sayısı olarak 500, 1000, 1500, 2000 ve ortalama düğüm derecesi olarak 2, 3, 4, 5 parametreleri kullanılmıştır. Tüm deneyler 5 tekrarlı olarak gerçekleştirilmiştir. Daha sonra, mutant ağ G' oluşturmak için önceden tanımlanmış mutasyon algoritması kullanılarak maksimum etkili mutasyonlar her ağ G için oluşturulmuştur.

Lyapunov üssünün hedef mutant ağlar üzerindeki yanıtını değerlendirmek için $Le(G')$ ölçümü kullanılmıştır. Şekil 4.13, orijinal ve mutant ağların karşılaştırmalı sonuçlarını göstermektedir.



Şekil 4.13 Sentetik Ağlardaki Mutasyonların Değerlendirilmesi.

Mutasyon operatörü, ağdaki en etkili değişikliği sağlayan minimum sayıdaki kenarın silinmesini sağlar. Süreç, değişiklik önemsiz etkiye sahip olana kadar devam eder. Ortaya çıkan ağ mutasyonlara karşı kararlıdır.

Tüm deneylerde, kenar mutasyonları sonucu oluşan ağlar daha yüksek oranda Lyapunov üslerine sahiptir. Her tür ağ için RW dağılımının topolojik mutasyonlara DSD dağılımından daha kararlı olduğu görülmektedir. Kararlılık sonuçlarına benzer şekilde tüm ağ türleri için, daha az kararlı DSD dağılımı mutasyondan sonra daha yüksek oranda Lyapunov değişimi ve mutasyonlara karşı daha dirençli olan RW dağılımı daha düşük bir Lyapunov değişikliği sergilemiştir.

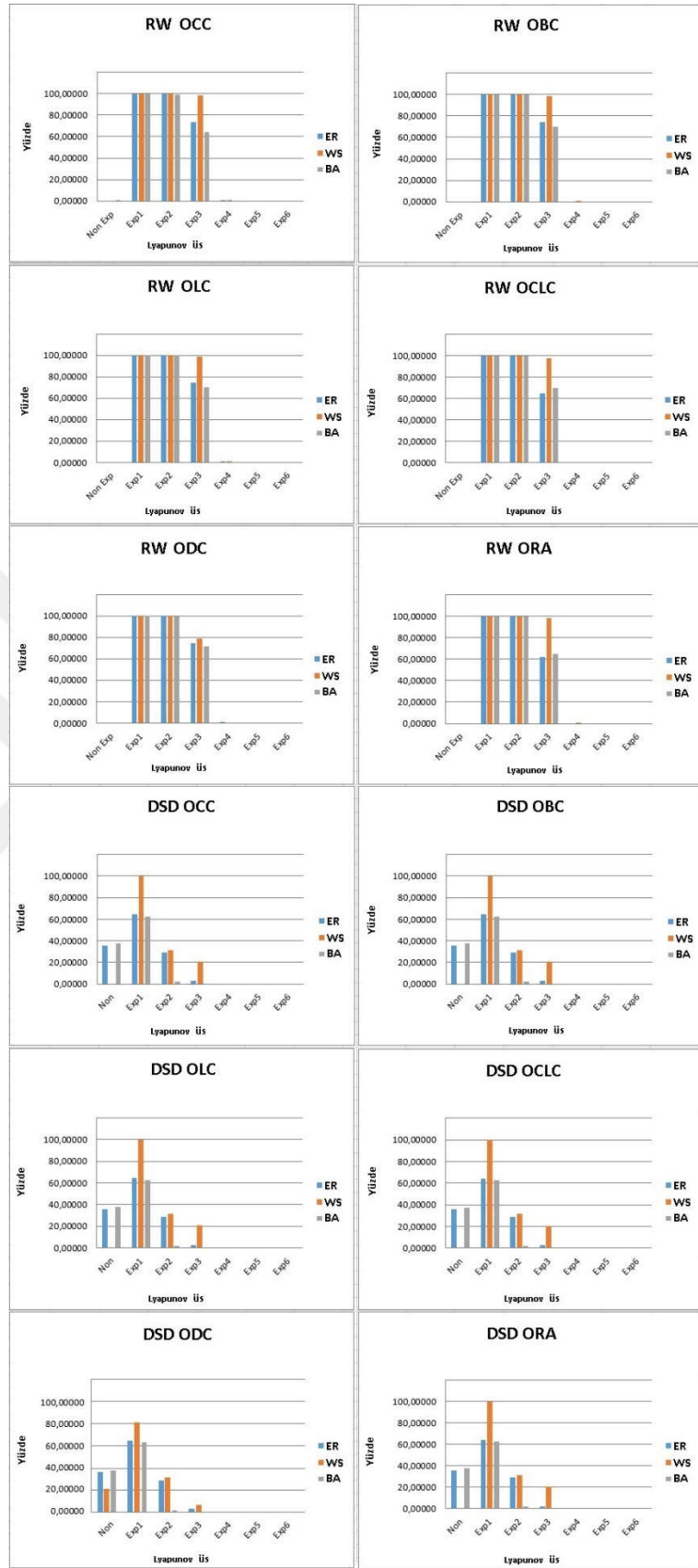
Tüm deneylerde WS ağları mutasyonlara karşı daha dayanıklı, daha kararlı ve daha düşük Lyapunov değişikliğine sahiptir.

Bulgularımız, topolojik mutasyonlarla daha kararlı hale gelen ağların daha yüksek Lyapunov üslerine sahip olduğunu kanıtlamaktadır. RW dağılımı mutasyonlara karşı daha dirençlidir, daha kararlıdır ve mutasyonlardan sonra daha az Lyapunov üs değişimine sahiptir. DSD dağılımı mutasyonlara karşı kırılgandır, daha az kararlıdır ve mutasyonlardan sonra daha fazla Lyapunov üssü değişikliğine sahiptir. Lyapunov üslerinin ağ kararlılığı ile bir korelasyonu vardır ve her ikisi de ağ modeline ilişkili olarak etkilenir.

4.4.1.3 Sıralamanın değerlendirilmesi

Sorunlardan biri, sentetik ağların düğüm sırasına bağlı olarak Lyapunov üs yanıtının değişip değişmediğidir. Bu sebeple, farklı düğüm sıralamalarının Lyapunov üs tepkisini nasıl etkilediği incelenmiştir.

Düğüm sırasının etkisini gözlemlemek için ağ oluşturma modeli olarak ER, WS, BA modelleri kullanılmıştır. Toplam düğüm sayısı olarak 500, 1000, 1500, 2000 ve ortalama düğüm derecesi olarak 2, 3, 4, 5 kullanılmıştır. Tüm deneyler 5 tekrarlı olarak yapılmıştır. Daha sonra, her bir ağ G için, düğüm sırası yeniden düzenlenmiş ağ G^r ağını oluşturmak üzere önceden tanımlanmış düğüm merkezi algoritmaları kullanılarak düğüm sırası değiştirilmiştir. Lyapunov üssünün yeniden sıralanmış ağ üzerindeki yanıtını değerlendirmek için $Le(G^r)$ ölçümü kullanılmıştır. Şekil 4.14, orijinal ve yeniden sıralanmış ağların karşılaştırmalı sonuçlarını göstermektedir.



Şekil 4.14 Sentetik Ağlarda Düğüm Sırasının Değerlendirilmesi.

Elde edilen sonuçlara göre, ER, WS ve BA sentetik ağlarının RW ve DSD dağılımlarına ait Lyapunov üs tepkisinin farklı düğüm sırasından etkilenmediği açıktır. Tüm deneylerde, düğüm sıra parametresinin Lyapunov üs tepkisi üzerinde etkisi vardır, ancak etki ihmal edilebilir düzeydedir.

Her ağ modeli için, hem RW hem de DSD dağıtımları düğüm sırasından etkilenir. OCC, OBC, OLC, OCLC ve ORA sıralamaları kullanılarak elde edilen Lyapunov üs sonuçları, önceden tanımlanmış Lyapunov üs sırası sonuçlarıyla hemen hemen aynıdır. Bu sıralamalardaki değişiklikler sıfıra yakındır. ODC sıralaması Lyapunov üssü üzerinde önemsiz bir etkiye sahiptir. Tüm ağ modelleri ODC sırası ile benzer sonuçlar üretmiştir. Rastgele düğüm sırası ORA on tekrarlı üretilmiştir ve sonuçlar arasında fark yoktur.

Sonuç olarak, deneysel sonuçlarımız düğüm sırasının hedef ağın Lyapunov üsleri üzerinde önemsiz bir etkiye sahip olduğunu göstermektedir. Her zaman farklı bir sıralama türünün Lyapunov üs sonuçlarının önemini bozma olasılığı vardır. Hedef ağın Lyapunov üs hesaplamasını etkileyecek bir sıraya sahip olma olasılığını önlemek için Lyapunov üs hesaplanmadan önce düğümler rastgele sırada düzenlenebilir.

Ağdan üretilen sinyalin sıralamaya olan duyarsızlığı, protein-protein etkileşim ağı veri yapısından kaynaklanmaktadır. PPI ağ matrisi tekrarlı bir matristir ve karşılıklı değerler eşittir. Bu özellik sebebi ile ağ benzer serilerden oluşur.

Örneğin, 6 etkileşim değerinden oluşan 3 protein düğümü içeren örnek etkileşim ağı aşağıdaki gibi tanımlanır $[(n_{1,2}, n_{1,3})(n_{2,1}, n_{2,3})(n_{3,1}, n_{3,2})]$. Bu 3 satır, 2 sütun ve 6 bağımsız etkileşim değerinden oluşan bir matristir. $n_{1,2}$ ve $n_{2,1}$ gibi etkileşim sonuçlarının aynı olduğu düşünüldüğünde, ağ matrisi benzer seriler haline gelir $[(n_{1,2}, n_{1,3})(n_{1,2}, n_{2,3})(n_{1,3}, n_{2,3})]$. Diğer bir faktör düğüm benzerliğidir. PPI ağları topolojik ve biyolojik olarak benzer düğümler içerir (örneğin $n_1 = n_2$). Bu dizi benzerliğini ve tekrarları artırır $[(n_{1,2}, n_{1,3})(n_{1,2}, n_{1,3})(n_{1,3}, n_{1,3})]$.

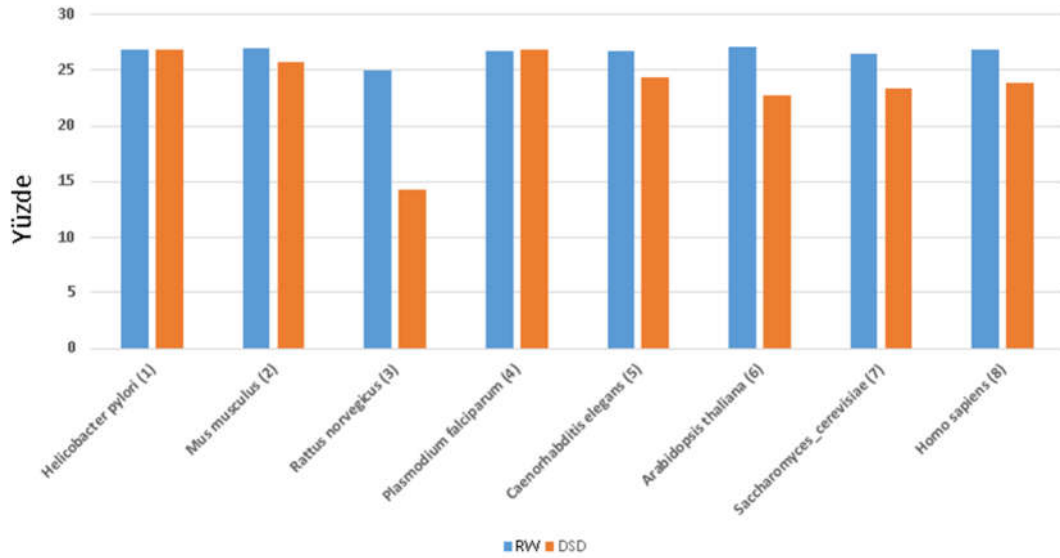
Ağ matrislerinin seri tekrarları ve benzerliği nedeniyle, düğüm sırası değiştiğinde, birçok seri için yalnızca ağdaki konum değişir. Bu özellik, ağdan üretilen sinyal kümesinin sıralamadan bağımsız olmasını sağlar.

4.4.2 Biyolojik Veriler Üzerine Sonuçlar

Şimdiye kadar, farklı düğüm sırası ve topolojik mutasyonlar ile çeşitli parametrelere sahip sentetik ağların Lyapunov üs tepkisi gösterilmiştir. Bununla birlikte, biyolojik ve sentetik ağların Lyapunov üs tepkileri arasında farklılıklar olabilir. Bu amaçla, bu bölümde farklı organizmalara ait sekiz protein-protein etkileşim ağını kullanarak biyolojik ağlar üzerindeki Lyapunov üs tepkisi incelenmiştir.

4.4.2.1 Organizmaların değerlendirilmesi

Bir sonraki başlığımız sentetik ve biyolojik ağların Lyapunov üs tepkisi açısından benzer olup olmadığını tespit etmektir. Bu deneyle, farklı biyolojik ağların Lyapunov üs tepkisi ortaya çıkarılmıştır. Lyapunov üssünün biyolojik ağlar üzerindeki yanıtını değerlendirmek için $Le(G)$ ölçümü kullanılmıştır. Şekil 4.15 sonuçları göstermektedir.



Şekil 4.15 Organizmaların Değerlendirilmesi.

Biyolojik ağların RW ve DSD dağılımının farklı Lyapunov üs tepkisine sahip olduğu görülmektedir. Sentetik ağların deneysel sonuçlarına benzer şekilde, tüm deneylerde, RW dağılımı, DSD dağılımından daha yüksek bir Lyapunov üs yüzdesine sahiptir.

Ağların boyutları ve Lyapunov üs tepkisi ile korelasyon yoktur. “*Rattus norvegicus* (3)” en az Lyapunov üs oranına ve “*Helicobacter pylori* (1)” en çok Lyapunov üs oranına sahip ağlardır.

Deneyisel sonuçlarımıza dayanarak, RW ve DSD dağılımlarının sentetik ağlara benzer bir Lyapunov üs tepkisine sahip olduğu tespit edilmiştir. Deneyisel sonuçlar, Lyapunov üs tepkisinin biyolojik ağların topolojik özelliklerine bağlı olarak değiştiğini göstermektedir.

Sonuçlar ağ kararlılığının sonuçlarıyla karşılaştırıldığında, aralarında benzerlikler vardır. Karşılaştırmalar, Lyapunov ölçümünün, yüksek kararlılığa sahip ağlarda daha yüksek olduğunu göstermektedir.

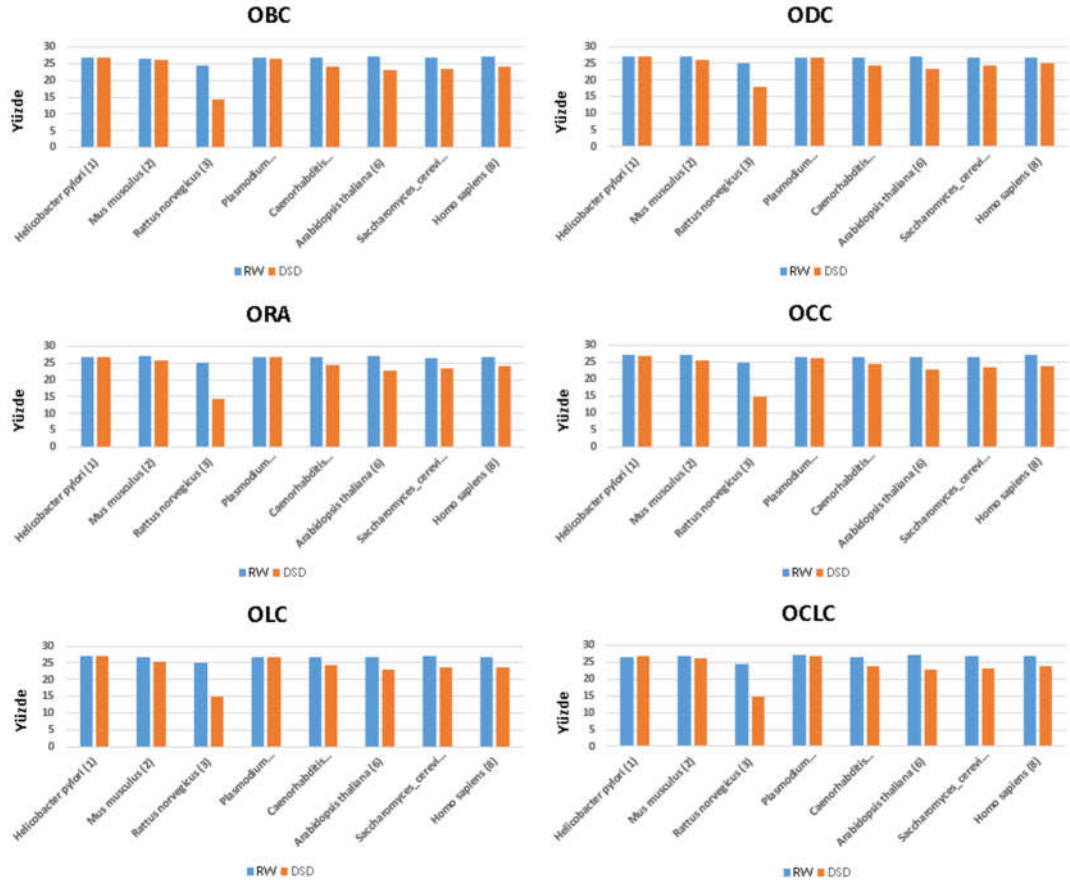
4.4.2.2 Sıralamanın değerlendirilmesi

Proteinlerin etkileşim verileri, her bir veri tabanında farklı sıralarda sunulur. Bir sonraki soru, biyolojik ağların düğüm sırasına bağlı olarak Lyapunov üs yanıtlarının değişip değişmediğidir. Bu soruya cevap verilmesi ve farklı düğüm sıralamasının biyolojik ağların Lyapunov üs tepkisine verdiği yanıtı nasıl etkilediğinin gözlemlenmesi amaçlanmıştır.

Farklı organizmalara ait sekiz protein etkileşim ağı kullanılmıştır. Daha sonra, her biyolojik ağ G için, düğüm sırası yeniden düzenlenmiş G^r ağını oluşturmak için önceden tanımlanmış düğüm merkezliği algoritmaları kullanılarak düğüm sırası değiştirilmiştir. Yeniden sıralanmış biyolojik ağ üzerindeki Lyapunov üs yanıtını değerlendirmek için $Le(G^r)$ ölçümü kullanılmıştır. Şekil 4.16, orijinal ve yeniden sıralanan biyolojik ağların karşılaştırma sonuçlarını göstermektedir.

Biyolojik ağların RW ve DSD dağılımlarının, farklı düğüm sırası altında benzer Lyapunov üs tepkisine sahip olduğu görülmektedir. Tüm deneylerde, düğüm sırası parametresinin Lyapunov üs tepkisi üzerinde etkisi vardır, ancak etki ihmal edilebilir düzeydedir.

Her biyolojik ağ için, hem RW hem de DSD dağıtımları düğüm sırasından etkilenir. OCC, OBC, OLC, OCLC ve ORA sıralamaları kullanılarak elde edilen Lyapunov üs sonuçları, önceden tanımlanmış Lyapunov üs sırası sonuçlarıyla hemen hemen aynıdır. Bu sıralamalardaki değişiklikler sıfıra yakındır. ODC sıralaması Lyapunov üs üzerinde önemsiz bir etkiye sahiptir. Sentetik ağların aksine, ODC sıralamasının etkisi daha azdır. Rastgele düğüm sırası ORA 10 tekrar ile üretilmiştir ve sonuçlar arasında fark bulunmamaktadır.



Şekil 4.16 Biyolojik Ağlarda Düğüm Sırasının Değerlendirilmesi.

Sonuç olarak, deneysel sonuçlarımız düğüm sırasının hedef biyolojik ağların Lyapunov üsleri üzerinde önemsiz bir etkiye sahip olduğunu göstermektedir. Her zaman farklı bir sıralama tipinin Lyapunov üs sonuçlarının önemini bozma olasılığı vardır. Hedef ağın Lyapunov üs hesaplamasını etkileyecek bir sıraya sahip olma olasılığını önlemek için düğümler hesaplamadan önce rastgele olarak sıralanabilir.



5. TARTIŞMA ve SONUÇ

Bu çalışmanın ilk kısmında ağ topolojisindeki mevcut mutasyonların ağın kararlılığını etkilediği gösterilmiştir. Bu, topolojik mutasyonlar tarafından bozulan bir ağın kararlılığını değerlendiren ilk çalışmadır.

Topolojik mutasyonların hem biyolojik hem de sentetik ağlar üzerindeki etkisini göstermek amacıyla, önce ağın kararlılığı tanımlanmış ve ağın hassasiyetini değerlendirmek için uygunluk ölçümleri geliştirilmiştir. Mevcut çalışmaların çoğunun aksine, ağ mutasyon modeli olarak kenar silme mutasyonu kullanılmıştır. Bu model ayrıca düğüm silme modelini de kapsamaktadır.

Bir ağ üzerinde etkili bir kenar kümesi aramak, hesaplama açısından zor bir problemdir. Bu sorunun üstesinden gelmek için, belirli bir ağ $G = (V, E)$ üzerinde etkili yapay mutasyonlar yapan yeni bir meta-sezgisel optimizasyon yöntemi geliştirilmiş ve en etkili kenar kombinasyonları zaman-verimli olarak hesaplanmıştır. Yöntemimiz, hem genetik algoritma hem de tepe tırmanma stratejisi kullanan iki aşamadan oluşur. İlk aşama, bir dizi aday çözüm kümesi üretmek için genetik algoritmayı kullanır ve ikinci aşama, bu çözümleri iyileştirmek için bir tepe tırmanma algoritmasını kullanır. Deney sonuçları, yöntemimizin tüm ağ topolojileri için en etkili kenarları bulduğunu doğrulamaktadır.

Ağ kararlılığı ağ boyutuna kıyasla ağ modelinden daha fazla etkilenir, bu nedenle kararlılık ağ modeli ve topolojisine bağlıdır. Tüm organizmaların RW ve DSD dağılımları kırılgandır ve kırılganlık oranı organizmadan organizmaya farklılık gösterir. Mutasyon ne kadar çok olursa, o kadar fazla hata oluşmaktadır.

Deney sonuçlarına göre, protein-protein etkileşim ağı topolojisindeki mutasyonların, ağın kararlılığı üzerinde önemli bir etkisi vardır. Bu nedenle, hedef ağdan edinilen bilgileri kullanarak tahmin yapan teknikler, o ağın kararlılığını ve tahmin yöntemlerinin kararlılığını göz önünde bulundurmalıdır.

Çalışmanın ikinci kısmında doğa verilerinde Lyapunov üssü başarısının varlığını gösteren teorik ve deneysel kanıtlar sunan çalışmalardan esinlenerek ağ analizi için Lyapunov üsleri araştırılmıştır.

Bu kapsamda, ağ topolojisi ve mutasyonların Lyapunov üsleri üzerindeki etkisi ve ağ kararlılığı ile ilişkisi değerlendirilmiştir. Bu, Lyapunov üslerinin ağ analizi için kullanıldığı ve topolojik mutasyonlar altında değerlendirildiği ilk çalışmadır.

Topolojik mutasyonların hem biyolojik hem de sentetik ağlar üzerindeki etkisini göstermek amacıyla, ilk olarak çeşitli ağlar için Lyapunov üs tabanlı ölçümleri elde etmek için yeni bir yöntem tanımlanmıştır. Ardından, her bir ağın kararlılığını ve Lyapunov üs tepkisini ölçerek Lyapunov üsleri ile ağın kararlılığı arasındaki ilişki değerlendirilmiştir. Yapılan deneyler tekniğimizin her tür ağ topolojisine uygulanabileceğini göstermiştir.

Ağ modeli ve mutasyonlar ağın Lyapunov üsleri üzerinde önemli bir etkiye sahiptir. Lyapunov üslerinin ağ kararlılığı ile korelasyonu vardır ve her ikisi de ağ modelinden etkilenir. Bir ağın kararlılığı Lyapunov üsleri ile ölçülebilir.

DeneySEL sonuçlar, Lyapunov üslerinin ağ kararlılığının tespiti için kullanılma potansiyelini göstermiştir. Geliştirdiğimiz yeni yöntem sayesinde ağların kararlılığı zamandan ve işlem gücünden tasarruf edilerek daha hızlı bir şekilde belirlenebilir. Biyolojik ağların kaynak çeşitliliği ve büyüklüğü göz önüne alındığında, ağ analizinin hızı büyük önem taşımaktadır. Ek olarak, sonuçlar, Lyapunov üssünün ağ analizinde kullanılabilecek potansiyel bir aday ölçüm olduğunun göstergesidir. Elde edilen sonuçlar, ağ analizi üzerine gelecekteki araştırmalar için müteakip yeni çalışmalar başlatabilir.

KAYNAKLAR

- [1] Y. Kavurucu, Network motifs and indexing techniques in biological networks, *Journal of Naval Science and Engineering*. 8 (2012) 87–102.
- [2] X. Ma, L. Gao, Biological network analysis: insights into structure and functions, *Briefings in Functional Genomics*. 11 (2012) 434–442.
- [3] C. Bachmaier, U. Brandes, F. Schreiber, Biological networks, in: R. Tamassia, E. B. Raton, *Handbook of graph drawing and visualization*, Taylor & Francis, 2014, pp. 621–651.
- [4] X. Zhu, M. Gerstein, M. Snyder, Getting connected: analysis and principles of biological networks, *Genes & Development*. 21 (2007) 1010–1024.
- [5] H. Gabr, J.C. Rivera-Mulia, D.M. Gilbert, T. Kahveci, Computing interaction probabilities in signaling networks, *EURASIP Journal on Bioinformatics and Systems Biology*. 1 (2015) 10.
- [6] R. Albert, A.L. Barabasi, Statistical mechanics of complex networks, *Reviews of Modern Physics*. 74 (2002) 47.
- [7] R. Albert, H. Jeong, A.L. Barabasi, Error and attack tolerance of complex networks, *Nature*. 406 (2000) 378–382.
- [8] P. Holme, B.J.Kim, C.N. Yoon, S.K. Han, Attack vulnerability of complex networks, *Physical Review E*. 65(2002) 056109.
- [9] M.P. Stumpf, C. Wiuf, Incomplete and noisy network data as a percolation process, *Journal of the Royal Society Interface*. 7 (2010) 1411–1419.
- [10] U. Alon, Biological networks: the tinkerer as an engineer, *Science*. 301 (2003) 1866–1867.
- [11] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.L. Barabasi, The large-scale organization of metabolic networks, *Nature*. 407 (2000) 651–654.
- [12] J. Stelling, U. Sauer, Z. Szallasi, F.J. Doyle, J. Doyle, Robustness of cellular functions, *Cell*. 118 (2004) 675–685.

- [13] F. Li, P. Li, W. Xu, Y. Peng, X. Bo, S. Wang, Perturbationanalyzer: a tool for investigating the effects of concentration perturbation on protein interaction networks, *Bioinformatics*. 26 (2010) 275–277.
- [14] S. Erten, G. Bebek, M. Koyuturk, Vavien: an algorithm for prioritizing candidate disease genes based on topological similarity of proteins in interaction networks, *Journal of Computational Biology*. 18 (2011) 1561–1574.
- [15] A. Vinayagam, T.E. Gibson, H.J. Lee, B. Yilmazel, C. Roesel, Y. Hu, Y. Kwon, A. Sharma, Y.Y. Liu, N. Perrimon, Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets, *Proceedings of the National Academy of Sciences*. 1 (2016) 201603992.
- [16] X. Zhang, H. Wang, Y. Yang, Robustness of indispensable nodes in controlling protein-protein interaction network, *arXiv Preprint*. 1609 (2016) 02637.
- [17] B. Lehne, T. Schlitt, Protein-protein interaction databases: keeping up with growing interactomes, *Human genomics*. 3 (2009) 291.
- [18] L. Sam, Discovery of protein interaction networks shared by diseases, *Biocomputing*. 1 (2007) 6-87.
- [19] D.E Ingber, The origin of cellular life, *Bioessays*. 22(2000) 1160-1170.
- [20] H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell , *Molecular cell biology* 4th edition, National Center for Biotechnology Information, Bookshelf, 2000.
- [21] L. Lu, A.K. Arakaki, H. Lu, J. Skolnick, Multimeric threading-based prediction of protein–protein interactions on a genomic scale: application to the *Saccharomyces cerevisiae* proteome, *Genome Research*. 13 (2003) 1146-1154.
- [22] S. Ghaemmaghami, W.K. Huh, K. Bower, R.W. Howson, A. Belle, N. Dephoure, J.S. Weissman, Global analysis of protein expression in yeast, *Nature*. 425(2003), 737.
- [23] P. Braun, A.C. Gingras, History of protein–protein interactions: From egg-white to complex networks, *Proteomics*. 12(2012), 1478-1498.

- [24] C. Yan, F. Wu, R.L. Jernigan, D. Dobbs, V. Honavar, Characterization of protein–protein interfaces, *The protein journal*. 27 (2008) 59-70.
- [25] I.M.A. Nooren, J.M. Thornton, Diversity of protein–protein interactions, *The EMBO journal*. 22(2003), 3486-3492.
- [26] A. Zhang , *Protein interaction networks: computational analysis*, Cambridge University Press, USA, 2009.
- [27] M. Iqbal, *Introductory Chapter: Protein-Protein Interactions and Assays, Protein-Protein Interaction Assays 1, BoD--Books on Demand, Deutschland*, 2018.
- [28] S. Klein, The use of biorelevant dissolution media to forecast the in vivo performance of a drug, *The AAPS journal*. 12 (2010) 397-406.
- [29] M. Yanagida, Functional proteomics; current achievements, *Journal of Chromatography B*. 771 (2002) 89-106.
- [30] T. Berggård, S. Linse, P. James, Methods for the detection and analysis of protein–protein interactions, *Proteomics*. 7 (2007) 2833-2842.
- [31] M.C. Von, R. Krause, B. Snel, M. Cornell, S.G. Oliver, S. Fields, P. Bork, Comparative assessment of large-scale data sets of protein–protein interactions, *Nature*. 417 (2002) 399.
- [32] G.M. Rishton, Reactive compounds and in vitro false positives in HTS, *Drug discovery today*. 2 (1997) 382-384.
- [33] S. Vivona, J.L. Gardy, S. Ramachandran, F.S. Brinkman, G.P.S Raghava, D.R. Flower, F. Filippini, Computer-aided biotechnology: from immuno-informatics to reverse vaccinology, *Trends in biotechnology*. 26 (2008) 190-200.
- [34] A.G. Atanasov, B. Waltenberger, E.M. Pferschy-Wenzig, T. Linder, C. Wawrosch, P. Uhrin, J.M. Rollinger, Discovery and resupply of pharmacologically active plant-derived natural products: A review, *Biotechnology advances*. 33 (2015) 1582-1614.
- [35] C. Lipinski, A. Hopkins, Navigating chemical space for biology and medicine, *Nature*. 432 (2004) 855.

- [36] D.A. Relman, Detection and identification of previously unrecognized microbial pathogens, *Emerging infectious diseases*. 4 (1998) 382.
- [37] P. Uetz, L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, E.A. Qureshi, A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*, *Nature*. 403 (2000) 623.
- [38] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, Y. Sakaki, A comprehensive two-hybrid analysis to explore the yeast protein interactome, *Proceedings of the National Academy of Sciences*. 98 (2001) 4569-4574.
- [39] S. Oliver, Proteomics: guilt-by-association goes global, *Nature*. 403 (2000) 601.
- [40] S.L. Rutherford, From genotype to phenotype: buffering mechanisms and the storage of genetic information, *Bioessays*. 22 (2000) 1095-1105.
- [41] J.L. Hartman, B. Garvik, L. Hartwell, Principles for the buffering of genetic variation, *Science*. 291 (2001) 1001-1004.
- [42] A. Bender, J.R. Pringle, Use of a screen for synthetic lethal and multicopy suppressor mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*, *Molecular and cellular biology*. 11 (1991) 1295-1305.
- [43] S.L. Ooi, X. Pan, B.D. Peyser, P. Ye, P.B. Meluh, D.S. Yuan, J.D. Boeke, Global synthetic-lethality analysis and yeast functional profiling, *TRENDS in Genetics*. 22 (2006) 56-63.
- [44] J.A. Brown, G. Sherlock, C.L. Myers, N.M. Burrows, C. Deng, H.I. Wu, J.M. Brown, Global analysis of gene function in yeast by quantitative phenotypic profiling, *Molecular systems biology*. 2(2006).
- [45] J. Arnau, C. Lauritzen, G.E. Petersen, J. Pedersen, Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins, *Protein expression and purification*. 48(2006) 1-13.
- [46] D.S. Waugh, Making the most of affinity tags, *Trends in biotechnology*. 23 (2005) 316-320.

- [47] S. Fields, O. Song, A novel genetic system to detect protein–protein interactions, *Nature*. 340 (1989) 245.
- [48] D. Bray, S. Lay, Computer-based analysis of the binding steps in protein complex formation, *Proceedings of the National Academy of Sciences*. 94 (1997) 13493-13498.
- [49] G. Drewes, T. Bouwmeester, Global approaches to protein–protein interactions, *Current opinion in cell biology*. 15 (2003) 199-205.
- [50] T. Bouwmeester, A. Bauch, H. Ruffner, P.O. Angrand, G. Bergamini, K. Croughton, C. Hopf, A physical and functional map of the human TNF- α /NF- κ B signal transduction pathway, *Nature cell biology*. 6 (2004) 97.
- [51] A.C. Gingras, R. Aebersold, B. Raught, Advances in protein complex analysis using mass spectrometry, *The Journal of physiology*. 563 (2005) 11-21.
- [52] A. Sinz, The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks, *Expert review of proteomics*. 11 (2014) 733-743.
- [53] J. Rappsilber, The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes, *Journal of structural biology*. 173 (2011) 530-540.
- [54] J.E. Bruce, In vivo protein complex topologies: Sights through a cross-linking lens, *Proteomics*. 12 (2012) 1565-1575.
- [55] J.J. Serpa, C.E. Parker, E.V. Petrotchenko, J. Han, J. Pan, C.H. Borchers, Mass spectrometry-based structural proteomics, *European Journal of Mass Spectrometry*. 18 (2012) 251-267.
- [56] J. Kuramochi, Y. Sakakibara, Intensive in vitro experiments of implementing and executing finite automata in test tube, *International Workshop on DNA-Based Computers*, Springer, Berlin, 2005.
- [57] H. Smith, Discovery of the anthrax toxin: the beginning of studies of virulence determinants regulated in vivo, *International journal of medical microbiology*. 291 (2001) 411-417.

- [58] B. Dehghan, Synergistic Modeling of in-vitro and in-vivo data via Stochastic Kriging with Qualitative Factors (SKQ), West Virginia University, USA, 2016.
- [59] G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, B. Séraphin, A generic protein purification method for protein complex characterization and proteome exploration, *Nature biotechnology*. 17 (1999) 1030.
- [60] C. Denison, A.D Rudner, S.A. Gerber, C.E. Bakalarski, D. Moazed, S.P. Gygi, A proteomic strategy for gaining insights into protein sumoylation in yeast, *Molecular & Cellular Proteomics*. 4 (2005) 246-254.
- [61] J. Graumann, L.A. Dunipace, J.H. Seol, W.H. McDonald, J.R. Yates, B.J. Wold, R.J. Deshaies, Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast, *Molecular & Cellular Proteomics*. 3 (2004) 226-237.
- [62] J.T. Hannich, A. Lewis, M.B. Kroetz, S.J. Li, H. Heide, A. Emili, M. Hochstrasser, Defining the SUMO-modified proteome by multiple approaches in *Saccharomyces cerevisiae*, *Journal of Biological Chemistry*. 280 (2005) 4102-4110.
- [63] C. Tagwerker, K. Flick, M. Cui, C. Guerrero, Y. Dou, B. Auer, P. Kaiser, A tandem affinity tag for two-step purification under fully denaturing conditions application in ubiquitin profiling and protein complex identification combined with in vivocross-linking, *Molecular & Cellular Proteomics*. 5 (2006) 737-748.
- [64] M. Urh, S. Dan, Z. Kate, Affinity chromatography: general methods, *Methods in enzymology*, Academic Press, USA, 2009.
- [65] E.M. Phizicky, S. Fields, Protein-protein interactions: methods for detection and analysis, *Microbiological reviews*. 59 (1995) 94-123.
- [66] E. Golemis, D.A. Peter, Protein-protein interactions: a molecular cloning manual, CSHL Press, USA, 2005.
- [67] M. Ohh, R.L. Yauch, K.M. Lonergan, J.M. Whaley, R.A.O. Stemmer, D.N. Louis, O. Iliopoulos, The von Hippel-Lindau tumor suppressor protein is required for proper assembly of an extracellular fibronectin matrix, *Molecular cell*. 1 (1998) 959-968.

- [68] G. MacBeath, S.L. Schreiber, Printing proteins as microarrays for high-throughput function determination, *Science*. 289 (2000) 1760-1763.
- [69] P.O. Brown, D. Botstein, Exploring the new world of the genome with DNA microarrays, *Nature genetics*. 21 (1999) 33.
- [70] D. Pinkel, R. Seagraves, D. Sudar, S. Clark, I. Poole, D. Kowbel, S.H. Dairkee, High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays, *Nature genetics*. 20 (1998) 207.
- [71] J.R. Pollack, C.M. Perou, A.A. Alizadeh, M.B. Eisen, A. Pergamenschikov, C.F. Williams, P.O. Brown, Genome-wide analysis of DNA copy-number changes using cDNA microarrays, *Nature genetics*. 23 (1999) 41.
- [72] E. Baldelli, V. Calvert, A. Hodge, A. VanMeter, E.F. Petricoin, M. Pierobon, Reverse Phase Protein Microarrays. *Molecular Profiling: Methods and Protocols*, Springer, USA, 2017.
- [73] S.W. Michnick, P.H. Ear, C. Landry, M.K. Malleshaiah, V. Messier, Protein-fragment complementation assays for large-scale analysis, functional dissection and dynamic studies of protein–protein interactions in living cells, *Signal Transduction Protocols*. Humana Press, Totowa, 2011.
- [74] J.J. Moresco, P.C. Carvalho, J.R. Yates III, Identifying components of protein complexes in *C. elegans* using co-immunoprecipitation and mass spectrometry, *Journal of proteomics*. 73 (2010) 2198-2204.
- [75] M. Morell, A. Espargaró, F.X. Avilés, S. Ventura, Detection of transient protein–protein interactions by bimolecular fluorescence complementation: The Abl-SH3 case, *Proteomics*. 7 (2007) 1023-1036.
- [76] S. Cabantous, T.C. Terwilliger, G.S. Waldo, Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein, *Nature biotechnology*. 23 (2005) 102.
- [77] J.K. Scott, G.P. Smith, Searching for peptide ligands with an epitope library, *Science*. 249 (1990) 386-390.

- [78] S. Bass, R. Greene, J.A. Wells, Hormone phage: an enrichment method for variant proteins with altered binding properties, *Proteins: Structure, Function, and Bioinformatics*. 8 (1990) 309-314.
- [79] S.S. Sidhu, H.B. Lowman, B.C. Cunningham, J.A. Wells, Phage display for selection of novel binding peptides, *Methods in enzymology*, Academic Press, USA, 2000.
- [80] G.P. Smith, V.A. Petrenko, Phage display, *Chemical reviews*. 97 (1997) 391-410.
- [81] S.S. Sidhu, Phage display in pharmaceutical biotechnology, *Current opinion in biotechnology*. 11 (2000) 610-616.
- [82] N. Gustav, I. Sundell, I. Ylva, Interaction Analysis through Proteomic Phage Display, *BioMed Research International*. 1 (2014) 176172.
- [83] I.S. Dunn, Phase display of proteins, *Current Opinion in biotechnology*. 7 (1996) 547-553.
- [84] A.H.Y. Tong, M. Evangelista, A.B. Parsons, H. Xu, G.D. Bader, N. Pagé, B. Andrews, Systematic genetic analysis with ordered arrays of yeast deletion mutants, *Science*. 294 (2001) 2364-2368.
- [85] B. Kobe, G. Guncar, R. Buchholz, T. Huber, B. Maco, N. Cowieson, J.K. Forwood, Crystallography and protein–protein interactions: biological interfaces and crystal contacts, *Biochemical Society Transactions*. 36 (2008) 1438-1441.
- [86] Y. Urakubo, T. Ikura, N. Ito, Crystal structural analysis of protein–protein interactions drastically destabilized by a single mutation, *Protein Science*. 17 (2008) 1055-1065.
- [87] R.B. Russell, F. Alber, P. Aloy, F.P. Davis, D. Korkin, M. Pichaud, A. Sali, A structural perspective on protein–protein interactions, *Current opinion in structural biology*. 14 (2004) 313-324.
- [88] D.E. Scott, A.R. Bayly, C. Abell, J. Skidmore, Small molecules, big targets: drug discovery faces the protein–protein interaction challenge, *Nature Reviews Drug Discovery*. 15 (2016) 533.

- [89] L. Barbieri, E. Luchinat, L. Banci, Protein interaction patterns in different cellular environments are revealed by in-cell NMR, *Scientific reports*. 5 (2015) 14456.
- [90] R.A. Burton, S. Mattila, E.J. Taparowsky, C.B. Post, Protein-protein interaction analysis by nuclear magnetic resonance spectroscopy, *Protein-Protein Interactions*. Humana Press, Newyork, USA, 2004.
- [91] M. Lescot, P. Déhais, G. Thijs, K. Marchal, Y. Moreau, P.Y. Van de, S. Rombauts, PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, *Nucleic acids research*. 30 (2002) 325-327.
- [92] D.S. Wishart, C. Knox, A.C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, J. Woolsey, DrugBank: a comprehensive resource for in silico drug discovery and exploration, *Nucleic acids research*. 34 (2006) D668-D672.
- [93] W.H. Shin, C.W. Christoffer, D. Kihara, In silico structure-based approaches to discover protein-protein interaction-targeting drugs, *Methods*. 131 (2017) 22-32.
- [94] M.D. Shortridge, G. Varani, Structure based approaches for targeting non-coding RNAs with small molecules, *Current opinion in structural biology*. 30 (2015) 79-88.
- [95] N. Blom, S. Gammeltoft, S. Brunak, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites1, *Journal of molecular biology*. 294 (1998) 1351-1362.
- [96] G.M. Pelay, A. Glas, O. Koch, T.N. Grossmann, Structure-Based Design of Inhibitors of Protein–Protein Interactions: Mimicking Peptide Binding Epitopes, *Angewandte Chemie International Edition*. 54 (2015) 8896-8927.
- [97] C.T. Supuran, Advances in structure-based drug discovery of carbonic anhydrase inhibitors, *Expert opinion on drug discovery*. 12 (2017) 61-88.
- [98] X. Chen, J.C. Jeong, Sequence-based prediction of protein interaction sites with an integrative method, *Bioinformatics*. 25 (2009) 585-591.

- [99] Y.A. Huang, Z.H. You, X. Chen, K. Chan, X. Luo, Sequence-based prediction of protein-protein interactions using weighted sparse representation model combined with global encoding, *BMC bioinformatics*. 17 (2016) 184.
- [100] T. Sun, B. Zhou, L. Lai, J. Pei, Sequence-based prediction of protein protein interaction using a deep-learning algorithm, *BMC bioinformatics*. 18 (2017) 277.
- [101] X. Liu, B. Liu, Z. Huang, T. Shi, Y. Chen, J. Zhang, SPPS: a sequence-based method for predicting probability of protein-protein interaction partners, *PloS one*. 7 (2012) e30938.
- [102] T. Hamp, B. Rost, Evolutionary profiles improve protein–protein interaction prediction from sequence, *Bioinformatics*. 31 (2015) 1945-1950.
- [103] S. De, M.M. Babu, Genomic neighbourhood and the regulation of gene expression, *Current opinion in cell biology*. 22 (2010) 326-333.
- [104] B. Oliver, M. Parisi, D. Clark, Gene expression neighborhoods, *Journal of biology*. 1 (2002) 4.
- [105] A.J. Enright, I. Iliopoulos, N.C. Kyrpides, C.A. Ouzounis, Protein interaction maps for complete genomes based on gene fusion events, *Nature*. 402 (1999) 86.
- [106] S. Thieme, P. Groth, Genome fusion detection: a novel method to detect fusion genes from SNP-array data, *Bioinformatics*. 29 (2013) 671-677.
- [107] N.S. Latysheva, M.E. Oates, L. Maddox, T. Flock, J. Gough, M. Buljan, M.M. Babu, Molecular principles of gene fusion mediated rewiring of protein interaction networks in cancer, *Molecular cell*. 63 (2016) 579-592.
- [108] F. Pazos, A. Valencia, Similarity of phylogenetic trees as indicator of protein–protein interaction, *Protein engineering*. 14 (2001) 609-614.
- [109] S. Erten, X. Li, G. Bebek, J. Li, M. Koyutürk, Phylogenetic analysis of modularity in protein interaction networks, *BMC bioinformatics*. 10 (2009) 333.
- [110] F. Pazos, D. Juan, J.M. Izarzugaza, E. Leon, A. Valencia A, Prediction of protein interaction based on similarity of phylogenetic trees, *Functional Proteomics*. Humana Press, Newyork, USA, 2008.

- [111] A. Grigoriev, A relationship between gene expression and protein interactions on the proteome scale: analysis of the bacteriophage T7 and the yeast *Saccharomyces cerevisiae*, *Nucleic acids research*. 29 (2001) 3513-3519.
- [112] C. Wu, J. Zhu, X. Zhang, Integrating gene expression and protein-protein interaction network to prioritize cancer-associated genes, *BMC bioinformatics*. 13 (2012) 182.
- [113] R. Chen, Z. Zhang, Z. Xue, L. Wang, M. Fu, Y. Lu, Z. Fan, Protein-protein interaction network of gene expression in the hydrocortisone-treated keloid, *International journal of dermatology*. 54 (2015) 549-554.
- [114] N. Papanikolaou, G.A. Pavlopoulos, T. Theodosiou, I. Iliopoulos, Protein-protein interaction predictions using text mining methods, *Methods*. 74 (2015) 47-53.
- [115] V.D. Badal, P.J. Kundrotas, I.A. Vakser, Natural language processing in text mining for structural modeling of protein complexes, *BMC bioinformatics*. 19 (2018) 84.
- [116] J. Fleischer, B. Alan, Increasing Incidence within PubMed of the Use of the Misspelling, *Acta dermato-venereologica*. 96 (2016) 826-827.
- [117] R. Garofalo, J.L. Schilling, Transgender Health Accepted for Indexing in PubMed Central and Inclusion in PubMed, *Transgender Health*. 2 (2017) 80-80.
- [118] S. Orchard, S. Kerrien, S. Abbani, B. Aranda, J. Bhate, S. Bidwell, A. Chatr-Aryamontri, Protein interaction data curation: the International Molecular Exchange (IMEx) consortium, *Nature methods*. 9 (2012) 345.
- [119] I. Xenarios, D.W. Rice, L. Salwinski, M.K. Baron, E.M. Marcotte, D. Eisenberg, DIP: the database of interacting proteins, *Nucleic acids research*. 28 (2000) 289-291.
- [120] I. Xenarios, L. Salwinski, X.J. Duan, P. Higney, S.M. Kim, D. Eisenberg, DIP the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions, *Nucleic acids research*. 30 (2002) 303-305.
- [121] Gene Ontology Consortium, The gene ontology (GO) project in 2006, *Nucleic acids research*. 34 (2006) D322-D326.

- [122] H. Hermjakob, P.L. Montecchi, C. Lewington, S. Mudali, S. Kerrien, S. Orchard, H. Margalit, IntAct: an open source molecular interaction database, *Nucleic acids research*. 32 (2004) D452-D455.
- [123] L. Licata, L. Briganti, D. Peluso, L. Perfetto, M. Iannuccelli, E. Galeota, L. Castagnoli, MINT the molecular interaction database: 2012 update, *Nucleic acids research*. 40 (2011) D857-D861.
- [124] A. Chatr-Aryamontri, A. Ceol, L.M. Palazzi, G. Nardelli, M.V. Schneider, L. Castagnoli, G. Cesareni, MINT: the Molecular INTeraction database, *Nucleic acids research*. 35 (2006) D572-D574.
- [125] H.W. Mewes, A. Ruepp, F. Theis, T. Rattei, M. Walter, D. Frishman, A. Antonov, MIPS: curated databases and comprehensive secondary data resources in 2010, *Nucleic acids research*. 39 (2010) D220-D224.
- [126] H.W. Mewes, D. Frishman, U. Güldener, G. Mannhaupt, K. Mayer, M. Mokrejs, B. Weil, MIPS: a database for genomes and protein sequences, *Nucleic acids research*. 30 (2002) 31-34.
- [127] A.A. Chatr, R. Oughtred, L. Boucher, J. Rust, C. Chang, N.K. Kolas, C. Stark, The BioGRID interaction database: 2017 update, *Nucleic acids research*. 45 (2017) D369-D379.
- [128] C. Stark, B.J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, M. Tyers, BioGRID: a general repository for interaction datasets, *Nucleic acids research*. 34 (2006) D535-D539.
- [129] R.D. Finn, A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, E.L. Sonnhammer, Pfam: the protein families database, *Nucleic acids research*. 42 (2013) D222-D230.
- [130] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, *Nucleic acids research*. 28 (2000) 27-30.
- [131] D. Szklarczyk, J.H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, L.J. Jensen, The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible, *Nucleic acids research*. 1 (2016) gkw937.

- [132] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, C.J. Huerta, M. Kuhn, STRING v10: protein–protein interaction networks, integrated over the tree of life, *Nucleic acids research*. 43 (2014) D447-D452.
- [133] K. Han, B. Park, H. Kim, J. Hong, J. Park, HPID: the human protein interaction database, *Bioinformatics*. 20 (2004) 2466-2470.
- [134] S. Peri, J.D. Navarro, R. Amanchy, T.Z. Kristiansen, C.K. Jonnalagadda, V. Surendranath, N. Ibarrola, Development of human protein reference database as an initial platform for approaching systems biology in humans, *Genome research*. 13 (2003) 2363-2371.
- [135] G.D. Bader, I. Donaldson, C. Wolting, B.F. Ouellette, T. Pawson, C.W. Hogue, BIND—the biomolecular interaction network database, *Nucleic acids research*. 29 (2001) 242-245.
- [136] G.D. Bader, D. Betel, C.W. Hogue, BIND: the biomolecular interaction network database, *Nucleic acids research*. 31 (2003) 248-250.
- [137] J. Das, H. Yu, HINT: High-quality protein interactomes and their applications in understanding human disease, *BMC systems biology*. 6 (2012) 92.
- [138] Y. López, K. Nakai, A. Patil, HitPredict version 4: comprehensive reliability scoring of physical protein–protein interactions from more than 100 species, *Database*. 1 (2015).
- [139] A. Patil, K. Nakai, H. Nakamura, HitPredict: a database of quality assessed protein–protein interactions in nine species, *Nucleic acids research*. 39 (2010) D744-D749.
- [140] M.E. Smoot, K. Ono, J. Ruscheinski, P.L. Wang, T. Ideker, Cytoscape 2.8: new features for data integration and network visualization, *Bioinformatics*. 27 (2010) 431-432.
- [141] M. Franz, H. Rodriguez, C. Lopes, K. Zuberi, J. Montojo, G.D. Bader, Q. Morris, GeneMANIA update 2018, *Nucleic acids research*. 46 (2018) W60-W64.

- [142] M. Giurgiu, J. Reinhard, B. Brauner, I. Dunger-Kaltenbach, G. Fobo, G. Frishman, A. Ruepp, CORUM: the comprehensive resource of mammalian protein complexes—2019, *Nucleic acids research*. 47 (2019) D559-D563.
- [143] S. Razick, G. Magklaras, I.M. Donaldson, iRefIndex: a consolidated protein interaction database with provenance, *BMC bioinformatics*. 9 (2008) 405.
- [144] V. Altuntas, M. Gök, T. Kahveci, Stability Analysis of Biological Networks' Diffusion State, *IEEE/ACM transactions on computational biology and bioinformatics*. (2020)
- [145] A.L. Turinsky, S. Razick, B. Turner, I.M. Donaldson, S.J. Wodak, Literature curation of protein interactions: measuring agreement across major public databases, *Database*. 1 (2010).
- [146] N. Bhardwaj, H. Lu, Correlation between gene expression profiles and protein–protein interactions within and across genomes, *Bioinformatics*. 21 (2005) 2730-2738.
- [147] S. Kohler, S. Bauer, D. Horn, P.N. Robinson, Walking the interactome for prioritization of candidate disease genes, *The American Journal of Human Genetics*. 82 (2008) 949–958.
- [148] M. Cao, H. Zhang, J. Park, N.M. Daniels, M.E. Crovella, L.J. Cowen, B. Hescott, Going the distance for protein function prediction: a new distance metric for protein interaction networks, *PloS One*. 8 (2013) e76339.
- [149] H. Cho, B. Berger, J. Peng, Diffusion component analysis: unraveling functional topology in biological networks, *International Conference on Research in Computational Molecular Biology*, 2015, Poland.
- [150] H. Jeong, X. Qian, B.J. Yoon, Effective comparative analysis of protein-protein interaction networks by measuring the steadystate network flow using a markov model, *BMC Bioinformatics*. 17 (2016) 395.
- [151] H. He, D. Lin, J. Zhang, Y.P. Wang, H.W. Deng, Comparison of statistical methods for subnetwork detection in the integration of gene expression and protein interaction network, *BMC Bioinformatics*, 18 (2017) 149.

- [152] K. Guan, Important notes on lyapunov exponents, arXiv preprint. 1401 (2014) 3315.
- [153] K. Liu, H. Wang, J. Xiao, The multivariate largest lyapunov exponent as an age-related metric of quiet standing balance, Computational and mathematical methods in medicine. 1 (2015).
- [154] F. Nazarimehr, S. Jafari, S.M.R.H. Golpayegani, J. Sprott, Can lyapunov exponent predict critical transitions in biological systems? Nonlinear Dynamics. 88 (2017) 1493–1500.
- [155] J. Gao, Multiscale analysis of biological data by scale-dependent lyapunov exponent, Frontiers in physiology. 2 (2012) 110.
- [156] G. Valenza, P. Allegrini, A. Lanata, E.P. Scilingo, Dominant lyapunov exponent and approximate entropy in heart rate variability during emotional visual elicitation, Frontiers in neuroengineering. 5 (2012) 3.
- [157] Q. Han, P. Wang, Estimation of the largest lyapunov exponent of the hrv signals, Journal of biomedical engineering. 24 (2007) 732–735.
- [158] M. Gök, O.H. Koç, S. Genç, Prediction of Disordered Regions in Proteins Using Physicochemical Properties of Amino Acids, International Journal of Peptide Research and Therapeutics, 22 (2016) 31-36.
- [159] B.H. Junker, F. Schreiber, Analysis of biological networks. John Wiley & Sons vol. 2, Germany, 2011.
- [160] T. Can, O.C. Amolu, A.K. Singh, Analysis of protein-protein interaction networks using random walks, Proceedings of the 5th international workshop on Bioinformatics ACM, 2005, Hong Kong, China.
- [161] B. Schwikowski, P. Uetz, S. Fields, A network of protein-protein interactions in yeast, Nature Biotechnology. 18 (2000) 1257–1261.
- [162] A. Li, S. Horvath, Network neighborhood analysis with the multi-node topological overlap measure, Bioinformatics. 23 (2007) 222–231.

- [163] V. Altuntas, S. Altuntas, M. Gok, Automatic colony segmentation on agar surface by image processing, *Journal of Scientific & Industrial Research*. 77 (2018) 8.
- [164] Z.H. Ling, R.H. Wang, Hmm-based hierarchical unit selection combining kullback-leibler divergence with likelihood criterion, *Acoustics, Speech and Signal Processing*, 2007, Hawaii, USA.
- [165] M.B. Kennel, R. Brown, H.D. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Physical review A*. 45 (1992) 3403.
- [166] H. Abarbanel, *Analysis of observed chaotic data*. Springer Science & Business Media, New York, 2012.
- [167] R. Hegger, H. Kantz, T. Schreiber, Practical implementation of nonlinear time series methods: The tisean package, *Chaos: An Interdisciplinary Journal of Nonlinear Science*. 9 (1999) 413-435.
- [168] A. Serletis, A. Shahmoradi, D. Serletis, Effect of noise on estimation of lyapunov exponents from a time series, *Chaos Solitons & Fractals*. 32 (2007) 883–887.
- [169] O.H. Kocal, E. Yuruklu, I. Avcibas, Chaotic-type features for speech steganalysis, *IEEE Transactions on Information Forensics and Security*. 3 (2008) 651–661.
- [170] M. Newman, *Networks*. Oxford university press, Michigan, 2018.
- [171] L.C. Freeman, S.P. Borgatti, D.R. White, Centrality in valued graphs: A measure of betweenness based on network flow, *Social networks*. 13 (1991) 141–154.
- [172] S.P. Borgatti, Centrality and network flow, *Social networks*. 27 (2005) 55–71.
- [173] S.P. Borgatti, Centrality and aids, *Connections*. 18 (1995) 112–114.
- [174] D. Chen, L. Lu, M.S. Shang, Y.C. Zhang, T. Zhou, Identifying influential nodes in complex networks, *Physica a: Statistical mechanics and its applications*. 391 (2012) 1777–1787.

- [175] J. Golbeck, Analyzing the social web. Newnes, USA, 2013.
- [176] L.C. Freeman, A set of measures of centrality based on betweenness, *Sociometry*. 40 (1977) 35–41.
- [177] C. Perez, R. Germon, Graph creation and analysis for linking actors: Application to social data, *Automating Open Source Intelligence*. 1 (2016) 103–129.
- [178] A. Hagberg, P. Swart, D.S Chult, Exploring network structure, dynamics, and function using networkx, Conference: SCIPY 08, 2008, Pasadena, United States.
- [179] K.I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks, *Physical Review Letters*. 87 (2001) 278701.
- [180] D.J. Watts S.H. Strogatz, Collective dynamics of small-world networks, *nature*. 393 (1998) 440.
- [181] F.E. Sevilgen, S. Aluru, N. Futamura, Parallel algorithms for tree accumulations, *Journal of Parallel and Distributed Computing*. 65 (2005) 85-93.

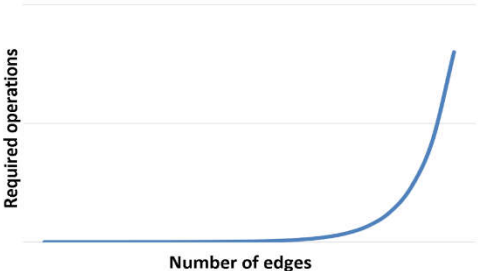


EKLER

EK A.1	Zaman Karmaşıklığı
EK A.2	Sentetik Ağ Özellikleri
EK A.3	Yöntem Karşılaştırma Detaylı Sonuçlar
EK A.4	Ağ Parametre Değerlendirme Detaylı Sonuçlar
EK A.5	Mutasyon Oranı Değerlendirme Detaylı Sonuçlar
EK A.6	Ağ Büyüklüğü Değerlendirme Detaylı Sonuçlar
EK A.7	Organizmaları Karşılaştırma Detaylı Sonuçlar
EK A.8	Mutasyon Oranı Karşılaştırma Detaylı Sonuçlar
EK A.9	Ağ Parametreleri Karşılaştırma Detaylı Sonuçlar (LE)
EK A.10	Mutasyon Etkisi Karşılaştırma Detaylı Sonuçlar (LE)
EK A.11	Düğüm Sırası Karşılaştırma Detaylı Sonuçlar (LE)
EK A.12	Organizmaları Karşılaştırma Detaylı Sonuçlar (LE)
EK A.13	Düğüm Sırası Karşılaştırma Detaylı Sonuçlar (LE)

EK A.1

Çizelge A.1 : Zaman Karmaşıklığı.

Growth curve	Edge size	Search space size
	100	$1.0e + 02$
	200	$1.9e + 04$
	300	$4.4e + 06$
	400	$1.0e + 09$
	500	$2.5e + 11$
	600	$6.3e + 13$

EK A.2

Çizelge A.2 : Sentetik Ağ Özellikleri.

Ağ	Düğüm	Kenar	Ekle	Çıkar
M1_N500_AND2_R1	500	1000	0	56
M1_N500_AND2_R2	500	1000	0	17
M1_N500_AND2_R3	500	1000	0	19
M1_N500_AND2_R4	500	1000	0	28
M1_N500_AND2_R5	500	1000	0	16
M1_N500_AND3_R1	500	1500	0	21
M1_N500_AND3_R2	500	1500	0	12
M1_N500_AND3_R3	500	1500	0	3
M1_N500_AND3_R4	500	1500	0	45
M1_N500_AND3_R5	500	1500	0	7
M1_N500_AND4_R1	500	2000	0	80
M1_N500_AND4_R2	500	2000	0	80
M1_N500_AND4_R3	500	2000	0	29
M1_N500_AND4_R4	500	2000	0	14
M1_N500_AND4_R5	500	2000	0	28
M1_N500_AND5_R1	500	2500	0	12
M1_N500_AND5_R2	500	2500	0	19
M1_N500_AND5_R3	500	2500	0	10
M1_N500_AND5_R4	500	2500	0	3
M1_N500_AND5_R5	500	2500	0	57
M1_N1000_AND2_R1	1000	2000	0	45
M1_N1000_AND2_R2	1000	2000	0	45
M1_N1000_AND2_R3	1000	2000	0	7
M1_N1000_AND2_R4	1000	2000	0	7
M1_N1000_AND2_R5	1000	2000	0	51
M1_N1000_AND3_R1	1000	3000	0	47
M1_N1000_AND3_R2	1000	3000	0	59
M1_N1000_AND3_R3	1000	3000	0	9
M1_N1000_AND3_R4	1000	3000	0	59
M1_N1000_AND3_R5	1000	3000	0	64
M1_N1000_AND4_R1	1000	4000	0	23
M1_N1000_AND4_R2	1000	4000	0	123
M1_N1000_AND4_R3	1000	4000	0	117
M1_N1000_AND4_R4	1000	4000	0	93
M1_N1000_AND4_R5	1000	4000	0	31
M1_N1000_AND5_R1	1000	5000	0	34
M1_N1000_AND5_R2	1000	5000	0	134

1. Erdős- Rényi
2. Watts-Strogatz
3. Barabasi-Albert

M1_N1000_AND5_R3	1000	5000	0	21
M1_N1000_AND5_R4	1000	5000	0	60
M1_N1000_AND5_R5	1000	5000	0	22
M1_N1500_AND2_R1	1500	3000	0	100
M1_N1500_AND2_R2	1500	3000	0	100
M1_N1500_AND2_R3	1500	3000	0	34
M1_N1500_AND2_R4	1500	3000	0	28
M1_N1500_AND2_R5	1500	3000	0	110
M1_N1500_AND3_R1	1500	4500	0	96
M1_N1500_AND3_R2	1500	4500	0	22
M1_N1500_AND3_R3	1500	4500	0	44
M1_N1500_AND3_R4	1500	4500	0	44
M1_N1500_AND3_R5	1500	4500	0	64
M1_N1500_AND4_R1	1500	6000	0	33
M1_N1500_AND4_R2	1500	6000	0	8
M1_N1500_AND4_R3	1500	6000	0	98
M1_N1500_AND4_R4	1500	6000	0	100
M1_N1500_AND4_R5	1500	6000	0	13
M1_N1500_AND5_R1	1500	7500	0	56
M1_N1500_AND5_R2	1500	7500	0	140
M1_N1500_AND5_R3	1500	7500	0	3
M1_N1500_AND5_R4	1500	7500	0	31
M1_N1500_AND5_R5	1500	7500	0	104
M1_N2000_AND2_R1	2000	4000	0	61
M1_N2000_AND2_R2	2000	4000	0	14
M1_N2000_AND2_R3	2000	4000	0	71
M1_N2000_AND2_R4	2000	4000	0	24
M1_N2000_AND2_R5	2000	4000	0	84
M1_N2000_AND3_R1	2000	6000	0	13
M1_N2000_AND3_R2	2000	6000	0	30
M1_N2000_AND3_R3	2000	6000	0	86
M1_N2000_AND3_R4	2000	6000	0	71
M1_N2000_AND3_R5	2000	6000	0	16
M1_N2000_AND4_R1	2000	8000	0	149
M1_N2000_AND4_R2	2000	8000	0	38
M1_N2000_AND4_R3	2000	8000	0	52
M1_N2000_AND4_R4	2000	8000	0	2
M1_N2000_AND4_R5	2000	8000	0	30
M1_N2000_AND5_R1	2000	10000	0	27
M1_N2000_AND5_R2	2000	10000	0	230
M1_N2000_AND5_R3	2000	10000	0	234
M1_N2000_AND5_R4	2000	10000	0	11
M1_N2000_AND5_R5	2000	10000	0	56
M2_N500_AND2_R1	500	1000	0	0
M2_N500_AND2_R2	500	1000	0	0

M2_N500_AND2_R3	500	1000	0	0
M2_N500_AND2_R4	500	1000	0	0
M2_N500_AND2_R5	500	1000	0	0
M2_N500_AND3_R1	500	1500	0	0
M2_N500_AND3_R2	500	1500	0	0
M2_N500_AND3_R3	500	1500	0	0
M2_N500_AND3_R4	500	1500	0	0
M2_N500_AND3_R5	500	1500	0	0
M2_N500_AND4_R1	500	2000	0	0
M2_N500_AND4_R2	500	2000	0	0
M2_N500_AND4_R3	500	2000	0	0
M2_N500_AND4_R4	500	2000	0	0
M2_N500_AND4_R5	500	2000	0	0
M2_N500_AND5_R1	500	2500	0	0
M2_N500_AND5_R2	500	2500	0	0
M2_N500_AND5_R3	500	2500	0	0
M2_N500_AND5_R4	500	2500	0	0
M2_N500_AND5_R5	500	2500	0	0
M2_N1000_AND2_R1	1000	2000	0	0
M2_N1000_AND2_R2	1000	2000	0	0
M2_N1000_AND2_R3	1000	2000	0	0
M2_N1000_AND2_R4	1000	2000	0	0
M2_N1000_AND2_R5	1000	2000	0	0
M2_N1000_AND3_R1	1000	3000	0	0
M2_N1000_AND3_R2	1000	3000	0	0
M2_N1000_AND3_R3	1000	3000	0	0
M2_N1000_AND3_R4	1000	3000	0	0
M2_N1000_AND3_R5	1000	3000	0	0
M2_N1000_AND4_R1	1000	4000	0	0
M2_N1000_AND4_R2	1000	4000	0	0
M2_N1000_AND4_R3	1000	4000	0	0
M2_N1000_AND4_R4	1000	4000	0	0
M2_N1000_AND4_R5	1000	4000	0	0
M2_N1000_AND5_R1	1000	5000	0	0
M2_N1000_AND5_R2	1000	5000	0	0
M2_N1000_AND5_R3	1000	5000	0	0
M2_N1000_AND5_R4	1000	5000	0	0
M2_N1000_AND5_R5	1000	5000	0	0
M2_N1500_AND2_R1	1500	3000	0	0
M2_N1500_AND2_R2	1500	3000	0	0
M2_N1500_AND2_R3	1500	3000	0	0
M2_N1500_AND2_R4	1500	3000	0	0
M2_N1500_AND2_R5	1500	3000	0	0
M2_N1500_AND3_R1	1500	4500	0	0
M2_N1500_AND3_R2	1500	4500	0	0

M2_N1500_AND3_R3	1500	4500	0	0
M2_N1500_AND3_R4	1500	4500	0	0
M2_N1500_AND3_R5	1500	4500	0	0
M2_N1500_AND4_R1	1500	6000	0	0
M2_N1500_AND4_R2	1500	6000	0	0
M2_N1500_AND4_R3	1500	6000	0	0
M2_N1500_AND4_R4	1500	6000	0	0
M2_N1500_AND4_R5	1500	6000	0	0
M2_N1500_AND5_R1	1500	7500	0	0
M2_N1500_AND5_R2	1500	7500	0	0
M2_N1500_AND5_R3	1500	7500	0	0
M2_N1500_AND5_R4	1500	7500	0	0
M2_N1500_AND5_R5	1500	7500	0	0
M2_N2000_AND2_R1	2000	4000	0	0
M2_N2000_AND2_R2	2000	4000	0	0
M2_N2000_AND2_R3	2000	4000	0	0
M2_N2000_AND2_R4	2000	4000	0	0
M2_N2000_AND2_R5	2000	4000	0	0
M2_N2000_AND3_R1	2000	6000	0	0
M2_N2000_AND3_R2	2000	6000	0	0
M2_N2000_AND3_R3	2000	6000	0	0
M2_N2000_AND3_R4	2000	6000	0	0
M2_N2000_AND3_R5	2000	6000	0	0
M2_N2000_AND4_R1	2000	8000	0	0
M2_N2000_AND4_R2	2000	8000	0	0
M2_N2000_AND4_R3	2000	8000	0	0
M2_N2000_AND4_R4	2000	8000	0	0
M2_N2000_AND4_R5	2000	8000	0	0
M2_N2000_AND5_R1	2000	10000	0	0
M2_N2000_AND5_R2	2000	10000	0	0
M2_N2000_AND5_R3	2000	10000	0	0
M2_N2000_AND5_R4	2000	10000	0	0
M2_N2000_AND5_R5	2000	10000	0	0
M3_N500_AND2_R1	500	1000	4	0
M3_N500_AND2_R2	500	1000	4	0
M3_N500_AND2_R3	500	1000	4	0
M3_N500_AND2_R4	500	1000	4	0
M3_N500_AND2_R5	500	1000	4	0
M3_N500_AND3_R1	500	1500	9	0
M3_N500_AND3_R2	500	1500	9	0
M3_N500_AND3_R3	500	1500	9	0
M3_N500_AND3_R4	500	1500	9	0
M3_N500_AND3_R5	500	1500	9	0
M3_N500_AND4_R1	500	2000	16	0
M3_N500_AND4_R2	500	2000	16	0

M3_N500_AND4_R3	500	2000	16	0
M3_N500_AND4_R4	500	2000	16	0
M3_N500_AND4_R5	500	2000	16	0
M3_N500_AND5_R1	500	2500	25	0
M3_N500_AND5_R2	500	2500	25	0
M3_N500_AND5_R3	500	2500	25	0
M3_N500_AND5_R4	500	2500	25	0
M3_N500_AND5_R5	500	2500	25	0
M3_N1000_AND2_R1	1000	2000	4	0
M3_N1000_AND2_R2	1000	2000	4	0
M3_N1000_AND2_R3	1000	2000	4	0
M3_N1000_AND2_R4	1000	2000	4	0
M3_N1000_AND2_R5	1000	2000	4	0
M3_N1000_AND3_R1	1000	3000	9	0
M3_N1000_AND3_R2	1000	3000	9	0
M3_N1000_AND3_R3	1000	3000	9	0
M3_N1000_AND3_R4	1000	3000	9	0
M3_N1000_AND3_R5	1000	3000	9	0
M3_N1000_AND4_R1	1000	4000	16	0
M3_N1000_AND4_R2	1000	4000	16	0
M3_N1000_AND4_R3	1000	4000	16	0
M3_N1000_AND4_R4	1000	4000	16	0
M3_N1000_AND4_R5	1000	4000	16	0
M3_N1000_AND5_R1	1000	5000	25	0
M3_N1000_AND5_R2	1000	5000	25	0
M3_N1000_AND5_R3	1000	5000	25	0
M3_N1000_AND5_R4	1000	5000	25	0
M3_N1000_AND5_R5	1000	5000	25	0
M3_N1500_AND2_R1	1500	3000	4	0
M3_N1500_AND2_R2	1500	3000	4	0
M3_N1500_AND2_R3	1500	3000	4	0
M3_N1500_AND2_R4	1500	3000	4	0
M3_N1500_AND2_R5	1500	3000	4	0
M3_N1500_AND3_R1	1500	4500	9	0
M3_N1500_AND3_R2	1500	4500	9	0
M3_N1500_AND3_R3	1500	4500	9	0
M3_N1500_AND3_R4	1500	4500	9	0
M3_N1500_AND3_R5	1500	4500	9	0
M3_N1500_AND4_R1	1500	6000	16	0
M3_N1500_AND4_R2	1500	6000	16	0
M3_N1500_AND4_R3	1500	6000	16	0
M3_N1500_AND4_R4	1500	6000	16	0
M3_N1500_AND4_R5	1500	6000	16	0
M3_N1500_AND5_R1	1500	7500	25	0
M3_N1500_AND5_R2	1500	7500	25	0

M3_N1500_AND5_R3	1500	7500	25	0
M3_N1500_AND5_R4	1500	7500	25	0
M3_N1500_AND5_R5	1500	7500	25	0
M3_N2000_AND2_R1	2000	4000	4	0
M3_N2000_AND2_R2	2000	4000	4	0
M3_N2000_AND2_R3	2000	4000	4	0
M3_N2000_AND2_R4	2000	4000	4	0
M3_N2000_AND2_R5	2000	4000	4	0
M3_N2000_AND3_R1	2000	6000	9	0
M3_N2000_AND3_R2	2000	6000	9	0
M3_N2000_AND3_R3	2000	6000	9	0
M3_N2000_AND3_R4	2000	6000	9	0
M3_N2000_AND3_R5	2000	6000	9	0
M3_N2000_AND4_R1	2000	8000	16	0
M3_N2000_AND4_R2	2000	8000	16	0
M3_N2000_AND4_R3	2000	8000	16	0
M3_N2000_AND4_R4	2000	8000	16	0
M3_N2000_AND4_R5	2000	8000	16	0
M3_N2000_AND5_R1	2000	10000	25	0
M3_N2000_AND5_R2	2000	10000	25	0
M3_N2000_AND5_R3	2000	10000	25	0
M3_N2000_AND5_R4	2000	10000	25	0
M3_N2000_AND5_R5	2000	10000	25	0

M1_N3000_AND3_R1	3000	9000	0	102
M1_N3000_AND3_R2	3000	9000	0	68
M1_N3000_AND3_R3	3000	9000	0	180
M1_N3000_AND3_R4	3000	9000	0	145
M1_N3000_AND3_R5	3000	9000	0	90
M1_N4000_AND3_R1	4000	12000	0	64
M1_N4000_AND3_R2	4000	12000	0	14
M1_N4000_AND3_R3	4000	12000	0	14
M1_N4000_AND3_R4	4000	12000	0	39
M1_N4000_AND3_R5	4000	12000	0	105
M1_N5000_AND3_R1	5000	15000	0	120
M1_N5000_AND3_R2	5000	15000	0	82
M1_N5000_AND3_R3	5000	15000	0	162
M1_N5000_AND3_R4	5000	15000	0	99
M1_N5000_AND3_R5	5000	15000	0	60
M1_N6000_AND3_R1	6000	18000	0	263
M1_N6000_AND3_R2	6000	18000	0	177
M1_N6000_AND3_R3	6000	18000	0	39
M1_N6000_AND3_R4	6000	18000	0	208
M1_N6000_AND3_R5	6000	18000	0	113
M2_N3000_AND3_R1	3000	9000	0	0

1. Erdős- Rényi
2. Watts-Strogatz
3. Barabasi-Albert

M2_N3000_AND3_R2	3000	9000	0	0
M2_N3000_AND3_R3	3000	9000	0	0
M2_N3000_AND3_R4	3000	9000	0	0
M2_N3000_AND3_R5	3000	9000	0	0
M2_N4000_AND3_R1	4000	12000	0	0
M2_N4000_AND3_R2	4000	12000	0	0
M2_N4000_AND3_R3	4000	12000	0	0
M2_N4000_AND3_R4	4000	12000	0	0
M2_N4000_AND3_R5	4000	12000	0	0
M2_N5000_AND3_R1	5000	15000	0	0
M2_N5000_AND3_R2	5000	15000	0	0
M2_N5000_AND3_R3	5000	15000	0	0
M2_N5000_AND3_R4	5000	15000	0	0
M2_N5000_AND3_R5	5000	15000	0	0
M2_N6000_AND3_R1	6000	18000	0	0
M2_N6000_AND3_R2	6000	18000	0	0
M2_N6000_AND3_R3	6000	18000	0	0
M2_N6000_AND3_R4	6000	18000	0	0
M2_N6000_AND3_R5	6000	18000	0	0
M3_N3000_AND3_R1	3000	9000	9	0
M3_N3000_AND3_R2	3000	9000	9	0
M3_N3000_AND3_R3	3000	9000	9	0
M3_N3000_AND3_R4	3000	9000	9	0
M3_N3000_AND3_R5	3000	9000	9	0
M3_N4000_AND3_R1	4000	12000	9	0
M3_N4000_AND3_R2	4000	12000	9	0
M3_N4000_AND3_R3	4000	12000	9	0
M3_N4000_AND3_R4	4000	12000	9	0
M3_N4000_AND3_R5	4000	12000	9	0
M3_N5000_AND3_R1	5000	15000	9	0
M3_N5000_AND3_R2	5000	15000	9	0
M3_N5000_AND3_R3	5000	15000	9	0
M3_N5000_AND3_R4	5000	15000	9	0
M3_N5000_AND3_R5	5000	15000	9	0
M3_N6000_AND3_R1	6000	18000	9	0
M3_N6000_AND3_R2	6000	18000	9	0
M3_N6000_AND3_R3	6000	18000	9	0
M3_N6000_AND3_R4	6000	18000	9	0
M3_N6000_AND3_R5	6000	18000	9	0

EK A.3

Çizelge A.3 : Yöntem Karşılaştırma Detaylı Sonuçlar.

Ağ	Değişim(hill)	Değişim (genetic)	Değişim (genetic+hill)
M1_N500_AND2	84,95	96,76	96,76
M1_N500_AND3	48,08	54,07	54,07
M1_N500_AND4	37,61	42,26	42,26
M1_N500_AND5	26,18	29,13	29,13
	49,20	55,55	55,55
M1_N1000_AND2	92,71	99,18	99,18
M1_N1000_AND3	80,89	89,71	89,71
M1_N1000_AND4	62,02	66,08	66,08
M1_N1000_AND5	45,47	49,31	49,31
	70,27	76,07	76,07
M1_N1500_AND2	98,31	100,00	100,00
M1_N1500_AND3	90,22	98,66	98,66
M1_N1500_AND4	74,61	80,48	80,48
M1_N1500_AND5	59,09	58,33	59,41
	80,56	84,37	84,64
M1_N2000_AND2	100,00	100,00	100,00
M1_N2000_AND3	97,92	100,00	100,00
M1_N2000_AND4	90,24	91,49	91,75
M1_N2000_AND5	69,52	68,45	69,56
	89,42	89,98	90,33
M2_N500_AND2	3,93	4,42	4,42
M2_N500_AND3	2,94	2,93	2,97
M2_N500_AND4	2,31	2,37	2,38
M2_N500_AND5	1,95	1,99	1,99
	2,79	2,93	2,94
M2_N1000_AND2	4,97	5,28	5,28
M2_N1000_AND3	3,26	3,27	3,28
M2_N1000_AND4	2,36	2,34	2,36
M2_N1000_AND5	2,15	2,17	2,18
	3,18	3,27	3,28
M2_N1500_AND2	5,20	5,55	5,55
M2_N1500_AND3	3,19	3,21	3,23
M2_N1500_AND4	2,69	2,73	2,73
M2_N1500_AND5	2,25	2,06	2,25
	3,33	3,38	3,44
M2_N2000_AND2	5,21	5,36	5,36
M2_N2000_AND3	3,32	3,44	3,44
M2_N2000_AND4	2,62	2,37	2,62
M2_N2000_AND5	2,35	2,10	2,35
	3,37	3,32	3,44

M3_N500_AND2	4,79	5,16	5,16
M3_N500_AND3	3,32	3,44	3,44
M3_N500_AND4	2,47	2,51	2,53
M3_N500_AND5	2,07	2,12	2,12
	3,16	3,31	3,31
M3_N1000_AND2	5,57	6,08	6,08
M3_N1000_AND3	3,51	3,64	3,66
M3_N1000_AND4	2,57	2,61	2,61
M3_N1000_AND5	2,28	2,35	2,35
	3,48	3,67	3,68
M3_N1500_AND2	5,59	5,89	5,89
M3_N1500_AND3	3,60	3,69	3,69
M3_N1500_AND4	2,71	2,80	2,80
M3_N1500_AND5	2,32	2,13	2,32
	3,56	3,63	3,68
M3_N2000_AND2	6,02	6,19	6,20
M3_N2000_AND3	3,73	3,88	3,88
M3_N2000_AND4	2,72	2,53	2,72
M3_N2000_AND5	2,31	2,08	2,31
	3,70	3,67	3,78
M3_N500_AND2	49,93	53,50	53,50
M3_N500_AND3	29,03	34,11	34,11
M3_N500_AND4	20,33	22,79	22,79
M3_N500_AND5	16,62	18,14	18,14
	28,98	32,14	32,14
M3_N1000_AND2	67,21	72,15	72,15
M3_N1000_AND3	40,72	45,26	45,26
M3_N1000_AND4	29,32	32,59	32,59
M3_N1000_AND5	24,89	27,74	27,74
	40,53	44,44	44,44
M3_N1500_AND2	81,10	88,26	88,53
M3_N1500_AND3	49,50	55,90	55,90
M3_N1500_AND4	36,37	41,27	41,27
M3_N1500_AND5	29,86	30,55	30,91
	49,20	53,99	54,15
M3_N2000_AND2	82,60	92,05	92,05
M3_N2000_AND3	59,10	66,21	66,21
M3_N2000_AND4	42,40	43,39	43,58
M3_N2000_AND5	33,29	33,40	33,52
	54,35	58,76	58,84

EK A.4**Çizelge A.4 : Ağ Parametre Değerlendirme Detaylı Sonuçlar.**

Model	Optimize CC							
	DKL(R W)	DKL(DS D)	RE(RW)	RE(DS D)	PCC(R W)	PCC(DS D)	NS(R W)	NS(DS D)
M1_N500_AND 2	0,0186	0,0124	4,6184	0,3663	0,1704	1,0069	2,2650	4,9600
M1_N500_AND 3	0,0125	0,0202	4,1044	0,2871	0,1309	0,2618	2,4900	3,7300
M1_N500_AND 4	0,0111	0,0219	3,4177	0,3107	0,1059	0,3693	2,7850	4,5800
M1_N500_AND 5	0,0102	0,0254	3,0855	0,3029	0,0904	0,3708	3,5250	5,1650
M1_N1000_AN D2	0,0019	0,0030	4,3085	0,2177	0,1784	0,0788	2,3675	4,6575
M1_N1000_AN D3	0,0023	0,0038	4,9708	0,2605	0,1144	0,2395	2,3700	3,4925
M1_N1000_AN D4	0,0035	0,0172	3,9295	0,2684	0,0897	0,3541	2,8675	3,8875
M1_N1000_AN D5	0,0052	0,0081	3,4617	0,2538	0,0869	0,3905	3,6950	4,8725
M1_N1500_AN D2	0,0009	0,0021	3,5365	0,2423	0,2166	0,7862	2,5500	4,3867
M1_N1500_AN D3	0,0011	0,0030	5,4997	0,2571	0,1245	0,1832	2,4117	3,6317
M1_N1500_AN D4	0,0011	0,0022	4,3123	0,2377	0,0892	0,3114	2,8567	3,7150
M1_N1500_AN D5	0,0010	0,0023	3,6973	0,2603	0,0728	0,3846	3,5583	4,6833
M1_N2000_AN D2	0,0004	0,0021	3,2446	0,2532	0,2095	1,0884	2,3900	4,5025
M1_N2000_AN D3	0,0007	0,0023	5,8647	0,2999	0,1368	2,7492	2,5000	3,4600
M1_N2000_AN D4	0,0022	0,0110	4,8008	0,2664	0,0969	0,4281	2,8850	3,5963
M1_N2000_AN D5	0,0007	0,0015	3,9377	0,2690	0,0737	0,4197	3,6500	4,4888
M2_N500_AND 2	0,0508	0,0090	0,5148	0,0345	0,1320	0,0295	0,7450	0,5300
M2_N500_AND 3	0,0100	0,0064	1,7000	0,0584	0,0654	0,0156	0,7700	0,3400
M2_N500_AND 4	0,0091	0,0058	2,9040	0,0868	0,0532	0,0122	0,3500	0,3950
M2_N500_AND 5	0,0087	0,0074	3,1635	0,1046	0,0416	0,0089	1,8650	0,5050
M2_N1000_AN D2	0,0220	0,0021	0,3146	0,0200	0,1311	0,0359	0,8200	0,6100
M2_N1000_AN D3	0,0031	0,0017	1,1618	0,0339	0,0667	0,0178	0,6375	0,3550
M2_N1000_AN D4	0,0019	0,0018	2,1167	0,0444	0,0489	0,0098	0,3625	0,3725

M2_N1000_AN								
D5	0,0016	0,0018	2,9967	0,0636	0,0407	0,0079	1,8500	0,4750
M2_N1500_AN								
D2	0,0079	0,0011	0,2001	0,0132	0,1293	0,0339	0,8417	0,5733
M2_N1500_AN								
D3	0,0021	0,0006	0,7754	0,0227	0,0658	0,0177	0,7100	0,3533
M2_N1500_AN								
D4	0,0015	0,0008	1,9482	0,0371	0,0504	0,0120	0,4300	0,4133
M2_N1500_AN								
D5	0,0006	0,0011	3,2656	0,0503	0,0405	0,0085	2,0617	0,5983
M2_N2000_AN								
D2	0,0064	0,0005	0,1528	0,0100	0,1320	0,0355	0,8025	0,5875
M2_N2000_AN								
D3	0,0010	0,0004	0,5851	0,0170	0,0653	0,0159	0,6225	0,3400
M2_N2000_AN								
D4	0,0004	0,0006	1,6919	0,0291	0,0511	0,0117	0,4225	0,4400
M2_N2000_AN								
D5	0,0004	0,0006	2,8849	0,0395	0,0417	0,0088	1,9962	0,5600
M3_N500_AND								
2	0,0394	0,0207	5,0391	0,3812	0,2048	0,5104	2,2800	3,5300
M3_N500_AND								
3	0,0123	0,0234	3,4388	0,4237	0,1144	0,5522	2,1000	3,8500
M3_N500_AND								
4	0,0124	0,0212	2,7333	0,4110	0,0717	0,4846	2,2400	3,9200
M3_N500_AND								
5	0,0074	0,0639	2,4237	0,3809	0,0579	0,4424	2,2250	4,0950
M3_N1000_AN								
D2	0,0072	0,0046	6,1317	0,3581	0,1936	0,6129	2,1600	4,1225
M3_N1000_AN								
D3	0,0041	0,0075	4,0001	0,3784	0,1296	0,6321	2,0925	4,0300
M3_N1000_AN								
D4	0,0026	0,0054	3,1979	0,3856	0,0934	0,6864	2,0950	4,2125
M3_N1000_AN								
D5	0,0015	0,0076	2,6695	0,3686	0,0650	0,6051	2,2575	4,1475
M3_N1500_AN								
D2	0,0038	0,0041	7,0585	0,3411	0,1828	0,7015	2,2367	4,2450
M3_N1500_AN								
D3	0,0014	0,0080	4,5088	0,3623	0,1340	0,7345	2,0983	4,2217
M3_N1500_AN								
D4	0,0011	0,0117	3,4217	0,3850	0,0887	0,7186	2,2217	4,0333
M3_N1500_AN								
D5	0,0004	0,0080	2,9402	0,3611	0,0665	0,6352	2,2933	3,8583
M3_N2000_AN								
D2	0,0028	0,0032	8,0666	0,3395	0,1701	0,7543	2,3250	4,0513
M3_N2000_AN								
D3	0,0005	0,0027	4,7526	0,3587	0,1390	0,8719	1,9663	4,3350
M3_N2000_AN								
D4	0,0006	0,0031	3,6633	0,3765	0,0909	0,7529	2,0888	4,1088
M3_N2000_AN								
D5	0,0002	0,0322	3,2118	0,3821	0,0703	0,7746	2,3575	3,9175

Model	DKL(R W)	DKL(DS D)	Optimize RW		PCC(R W)	PCC(DS D)	NS(R W)	NS(DS D)
			RE(RW)	RE(DS D)				
M1_N500_AND 2	0,0293	0,0243	8,0545	0,6263	0,0967	0,1598	1,6100	4,3950
M1_N500_AND 3	0,2196	0,4424	7,0752	2,4455	0,5140	39,0118	3,4200	4,4400
M1_N500_AND 4	0,2971	0,7071	6,0386	2,4820	0,5445	57,5428	3,1950	4,9700
M1_N500_AND 5	0,3263	5,4580	5,0922	2,2175	0,3501	55,7873	3,7400	5,4000
M1_N1000_AN D2	0,0025	0,0114	7,6827	0,4689	0,0782	0,0976	1,4625	3,7975
M1_N1000_AN D3	0,0908	0,1957	7,4810	0,7567	0,2160	5,1099	2,5925	3,5100
M1_N1000_AN D4	0,1527	1,0812	6,4512	1,0174	0,4750	27,2134	3,3800	4,6925
M1_N1000_AN D5	0,5537	11,2134	5,3839	1,6775	0,3487	55,2103	4,1675	4,9500
M1_N1500_AN D2	0,0011	0,0104	6,9509	0,3835	0,0759	0,0736	1,3600	3,3650
M1_N1500_AN D3	0,0044	0,0166	7,6264	0,5779	0,1199	0,4152	2,1533	3,1083
M1_N1500_AN D4	0,1136	1,4001	6,6501	0,9818	0,4440	29,9824	3,5067	4,4700
M1_N1500_AN D5	0,0940	1,3250	5,4118	1,0067	0,2982	40,8455	4,2917	4,8783
M1_N2000_AN D2	0,0006	0,0098	6,1820	0,3242	0,0753	0,0531	1,3775	3,4100
M1_N2000_AN D3	0,0031	0,0150	7,8187	0,5302	0,0816	0,3006	1,8863	2,4288
M1_N2000_AN D4	0,0499	0,7319	6,6738	0,8969	0,3707	35,4842	3,5313	4,0813
M1_N2000_AN D5	0,0630	1,2704	5,5109	0,8535	0,2721	44,1987	4,2925	4,7288
M2_N500_AND 2	0,0173	0,0446	3,4935	0,2595	0,1873	0,7112	1,4550	1,4950
M2_N500_AND 3	0,0116	0,0598	12,059 6	0,5237	0,1262	0,6259	1,1950	0,7050
M2_N500_AND 4	0,0101	0,0594	19,308 3	0,8039	0,0978	0,5301	0,6100	0,7850
M2_N500_AND 5	0,0094	0,0395	20,366 0	1,0324	0,0814	0,4493	0,5700	0,9900
M2_N1000_AN D2	0,0042	0,0242	1,7780	0,1391	0,1893	0,7271	1,3850	1,3800
M2_N1000_AN D3	0,0029	0,0502	7,0833	0,2870	0,1255	0,6275	1,1600	0,6650
M2_N1000_AN D4	0,0035	0,0675	15,462 3	0,4639	0,0942	0,5226	0,5900	0,8275

M2_N1000_AN			21,154					
D5	0,0035	0,0553	7	0,6865	0,0766	0,4572	0,4875	1,1575
M2_N1500_AN								
D2	0,0026	0,0165	1,2752	0,0978	0,1915	0,7234	1,3933	1,5400
M2_N1500_AN								
D3	0,0018	0,0485	5,3530	0,2086	0,1255	0,6250	1,1733	0,7267
M2_N1500_AN			13,041					
D4	0,0021	0,0611	3	0,3695	0,0935	0,5454	0,6317	0,7067
M2_N1500_AN			17,957					
D5	0,0020	0,0674	3	0,5055	0,0775	0,4626	0,4900	1,0200
M2_N2000_AN								
D2	0,0019	0,0125	0,9534	0,0727	0,1841	0,7103	1,3938	1,3938
M2_N2000_AN								
D3	0,0011	0,0490	4,3090	0,1636	0,1224	0,6181	1,1500	0,6313
M2_N2000_AN								
D4	0,0011	0,0572	9,7493	0,2675	0,0968	0,5275	0,7112	0,7138
M2_N2000_AN			16,158					
D5	0,0015	0,0654	3	0,4015	0,0764	0,4488	0,5700	0,9887
M3_N500_AND			10,027					
2	0,0569	0,0317	7	1,1201	0,0904	1,5495	2,0050	3,0900
M3_N500_AND								
3	0,0406	0,3287	6,0962	1,2316	0,2340	5,7457	2,1200	4,0650
M3_N500_AND								
4	0,0308	0,1855	4,8345	0,9152	0,2088	1,9226	2,8550	4,4700
M3_N500_AND								
5	0,0170	0,5321	4,3718	0,6702	0,1867	1,2386	2,6600	4,5750
M3_N1000_AN			11,970					
D2	0,0053	0,0142	8	0,9993	0,0463	1,1474	1,9925	3,4350
M3_N1000_AN								
D3	0,0097	0,2122	6,7276	0,9992	0,1970	3,5280	2,0175	3,8700
M3_N1000_AN								
D4	0,0092	0,3188	5,2725	0,8375	0,1942	2,5451	2,9225	4,0800
M3_N1000_AN								
D5	0,0045	0,2658	4,6479	0,6001	0,1812	1,3900	2,9000	4,2775
M3_N1500_AN			13,939					
D2	0,0018	0,0123	4	0,9625	0,0332	1,0156	2,0233	3,7233
M3_N1500_AN								
D3	0,0045	0,1949	7,2100	0,9123	0,1708	3,5075	2,1083	3,5983
M3_N1500_AN								
D4	0,0031	0,3383	5,5119	0,7884	0,1856	2,8737	2,8933	3,7783
M3_N1500_AN								
D5	0,0008	0,2548	4,6983	0,5088	0,1740	1,3245	3,0950	3,9267
M3_N2000_AN			15,094					
D2	0,0008	0,0173	9	0,8881	0,0275	0,9499	1,8650	3,7737
M3_N2000_AN								
D3	0,0045	0,1910	7,5646	0,8544	0,1607	3,5793	2,0837	3,5600
M3_N2000_AN								
D4	0,0011	0,1819	5,5486	0,6624	0,1783	2,4055	2,8813	3,7438
M3_N2000_AN								
D5	0,0017	0,3541	4,7686	0,4535	0,1691	1,2938	3,2225	3,8975

Model	DKL(R W)	DKL(DS D)	Optimize DKL		PCC(R W)	PCC(DS D)	NS(R W)	NS(DS D)
			RE(RW)	RE(DS D)				
M1_N500_AND 2	0,1989	0,1645	4,8458	0,6382	0,3522	4,8507	2,5950	5,1100
M1_N500_AND 3	0,4738	2,1688	5,4108	0,8479	0,3184	11,3982	2,8750	4,1050
M1_N500_AND 4	0,3830	1,8029	4,3472	0,8751	0,2205	20,5629	2,8350	4,4300
M1_N500_AND 5	0,8003	22,8980	3,7539	0,6352	0,1582	4,8758	2,9950	5,2650
M1_N1000_AN D2	0,0600	0,0127	3,8668	0,6212	0,3858	8,1341	2,7450	4,9750
M1_N1000_AN D3	0,2102	0,3693	5,8792	0,5809	0,2290	7,3466	2,6625	3,7725
M1_N1000_AN D4	0,4597	6,4531	4,9200	0,6370	0,2074	8,2284	2,6675	3,9350
M1_N1000_AN D5	0,4852	14,5577	3,8978	0,5426	0,1293	8,0391	2,9850	4,2925
M1_N1500_AN D2	0,0359	0,0076	3,1091	0,7727	0,4111	8,4184	2,9483	4,8583
M1_N1500_AN D3	0,1027	0,1194	6,0392	0,5190	0,2790	5,9087	2,7400	3,8983
M1_N1500_AN D4	0,2639	5,5140	5,2912	0,5962	0,2534	11,0041	2,7067	3,9050
M1_N1500_AN D5	0,3310	6,6904	3,9683	0,4738	0,1026	7,4378	2,8233	4,1917
M1_N2000_AN D2	0,0213	0,0041	3,1454	0,4492	0,3369	6,4169	2,6538	4,6838
M1_N2000_AN D3	0,0714	0,0985	6,0216	0,6748	0,3285	17,6507	3,0313	3,9813
M1_N2000_AN D4	0,2350	4,4458	5,4157	0,4732	0,2196	8,0009	2,7337	3,4163
M1_N2000_AN D5	0,4805	29,2528	4,1388	0,4896	0,1081	11,3664	2,8538	3,9025
M2_N500_AND 2	0,4939	0,0119	0,8700	0,0841	0,2864	0,4015	1,4650	1,5000
M2_N500_AND 3	0,7766	0,0089	2,0071	0,1215	0,2784	0,4729	1,0400	0,9700
M2_N500_AND 4	0,7679	0,0304	3,5514	0,1621	0,2310	0,3418	0,7850	0,8350
M2_N500_AND 5	0,1792	0,0105	4,1417	0,1530	0,0832	0,1070	1,3550	0,8950
M2_N1000_AN D2	0,2419	0,0028	0,5783	0,0498	0,2427	0,3354	1,4400	1,3475
M2_N1000_AN D3	0,4124	0,0045	1,9447	0,0861	0,2642	0,4018	1,0075	0,9025
M2_N1000_AN D4	0,5027	0,0183	4,8623	0,1895	0,2128	2,2908	0,9125	0,9125

M2_N1000_AN								
D5	0,3766	0,0615	5,3681	0,1345	0,1184	0,2102	1,2700	0,9700
M2_N1500_AN								
D2	0,1513	0,0060	0,4704	0,0664	0,2125	2,9026	1,3350	1,3083
M2_N1500_AN								
D3	0,2694	0,0064	2,4333	0,0962	0,2027	0,4267	1,1250	0,9200
M2_N1500_AN								
D4	0,3672	0,0058	4,7877	0,1365	0,2021	0,4345	0,8567	0,9550
M2_N1500_AN								
D5	0,0847	0,0038	3,7131	0,0631	0,0687	0,0753	1,5967	0,8067
M2_N2000_AN								
D2	0,1143	0,0009	0,3767	0,0289	0,1974	0,2934	1,2475	1,2975
M2_N2000_AN								
D3	0,2019	0,0091	2,2625	0,0859	0,1842	0,4130	1,0688	0,8387
M2_N2000_AN								
D4	0,2779	0,0069	4,4718	0,1215	0,1819	0,4046	0,9400	0,8738
M2_N2000_AN								
D5	0,2796	0,0025	6,5374	0,1331	0,1265	0,2540	1,1475	1,0150
M3_N500_AND								
2	0,2503	1,2642	5,5984	0,5526	0,3008	1,6694	2,4500	3,6400
M3_N500_AND								
3	0,7141	30,5263	4,1811	0,8312	0,2670	8,4115	2,1300	3,5700
M3_N500_AND								
4	0,7602	40,9287	3,7586	0,9089	0,2450	10,2484	2,1950	3,5050
M3_N500_AND								
5	0,7767	42,2854	3,2740	0,8548	0,1887	11,9383	1,8850	3,7300
M3_N1000_AN								
D2	0,1083	0,1434	5,8610	0,4117	0,3380	1,5961	2,4575	3,9725
M3_N1000_AN								
D3	0,3233	18,1436	4,7810	0,5524	0,1714	3,2142	2,1300	3,3125
M3_N1000_AN								
D4	0,4537	44,0666	4,0005	0,6569	0,1657	10,5825	2,1300	3,2125
M3_N1000_AN								
D5	0,4271	38,8770	3,5052	0,5937	0,1212	8,0744	1,9875	3,3875
M3_N1500_AN								
D2	0,0709	0,2392	6,7499	0,4071	0,2914	1,6486	2,2600	3,8967
M3_N1500_AN								
D3	0,2068	9,4525	5,2175	0,5034	0,1785	3,2335	2,1967	2,9600
M3_N1500_AN								
D4	0,3500	45,3513	4,2989	0,5559	0,1414	8,2762	2,1250	2,9267
M3_N1500_AN								
D5	0,4383	63,5364	3,5195	0,5409	0,1187	8,4331	1,8150	2,8117
M3_N2000_AN								
D2	0,0490	0,0977	6,6322	0,3448	0,2789	1,3823	2,3050	3,5837
M3_N2000_AN								
D3	0,1695	8,2377	5,3789	0,4594	0,1492	2,2249	2,1650	3,0775
M3_N2000_AN								
D4	0,2485	45,5941	4,4439	0,4948	0,1464	7,5534	2,1600	2,8600
M3_N2000_AN								
D5	0,2849	49,8736	3,6524	0,5251	0,0980	9,5506	1,8863	2,8638

EK A.5

Çizelge A.5 : Mutasyon Oranı Değerlendirme Detaylı Sonuçlar.

Optimize CC									
Model	Hata Oranı	DKL(R W)	DKL(D SD)	RE(RW)	RE(DS D)	PCC(R W)	PCC(D SD)	NS(R W)	NS(DS D)
M1_N1000_		0,001	0,0022	2,5444	0,136	0,060	0,1107	1,375	2,0175
AND3	0,5	36	4	3	99	93	4	00	0
M1_N1000_		0,002	0,0037	4,9991	0,262	0,115	0,2414	2,442	3,5100
AND3	1	39	5	0	38	95	5	50	0
M1_N1000_		0,022	0,0275	9,5049	0,589	0,252	0,6294	4,377	5,9150
AND3	2	62	2	4	92	26	0	50	0
M1_N1000_		0,007	0,0285	17,358	1,464	0,569	6,2761	7,730	9,3600
AND3	4	88	3	29	73	74	2	00	0
M2_N1000_		0,001	0,0006	0,5830	0,017	0,035	0,0100	0,307	0,1675
AND3	0,5	53	2	5	65	18	9	50	0
M2_N1000_		0,002	0,0017	1,1717	0,033	0,066	0,0182	0,627	0,3550
AND3	1	23	8	8	88	98	1	50	0
M2_N1000_		0,003	0,0026	2,1319	0,061	0,126	0,0401	1,105	0,6350
AND3	2	20	0	9	22	83	1	00	0
M2_N1000_		0,003	0,0046	4,1459	0,114	0,253	0,0814	2,350	1,2150
AND3	4	94	9	9	49	67	1	00	0
M3_N1000_		0,003	0,0042	2,2232	0,207	0,061	0,3093	1,032	2,3650
AND3	0,5	05	0	2	59	16	6	50	0
M3_N1000_		0,003	0,0075	4,0227	0,389	0,126	0,6571	2,032	4,0275
AND3	1	90	3	0	81	16	6	50	0
M3_N1000_		0,004	0,0132	7,2059	0,723	0,255	1,2893	3,870	6,7775
AND3	2	74	3	9	92	12	3	00	0
M3_N1000_		0,006	0,0224	13,281	1,645	0,473	2,8936	6,740	10,385
AND3	4	47	4	20	55	04	5	00	00
Optimize RW									
Model	Hata Oranı	DKL(R W)	DKL(D SD)	RE(RW)	RE(DS D)	PCC(R W)	PCC(D SD)	NS(R W)	NS(DS D)
M1_N1000_		0,056	0,1217	3,9696	0,350	0,072	0,4252	1,227	1,9275
AND3	0,5	33	2	7	47	91	3	50	0
M1_N1000_		0,090	0,1957	7,4809	0,756	0,215	5,1098	2,592	3,5100
AND3	1	84	4	6	67	98	7	50	0
M1_N1000_		0,084	0,1742	13,221	1,710	0,542	18,618	5,137	6,1400
AND3	2	86	7	99	18	13	34	50	0
M1_N1000_		0,078	0,1510	21,908	2,677	1,076	27,051	9,155	10,057
AND3	4	94	9	41	30	33	84	00	50
M2_N1000_		0,001	0,0132	3,9594	0,153	0,060	0,3101	0,547	0,3725
AND3	0,5	72	4	1	92	28	9	50	0
M2_N1000_		0,002	0,0501	7,0832	0,287	0,125	0,6275	1,160	0,6650
AND3	1	95	6	9	01	54	2	00	0
M2_N1000_		0,006	0,1727	11,698	0,509	0,274	1,2467	2,312	1,1750
AND3	2	67	0	38	26	74	6	50	0
M2_N1000_		0,014	0,4137	17,089	0,788	0,573	2,1666	4,285	2,3975
AND3	4	33	3	21	78	46	5	00	0
M3_N1000_		0,004	0,1250	3,6458	0,528	0,094	1,8018	1,080	2,2650
AND3	0,5	47	6	3	81	53	5	00	0

M3_N1000_		0,009	0,2121	6,7275	0,999	0,196	3,5279	2,017	3,8700
AND3	1	68	9	8	18	98	9	50	0
M3_N1000_		0,027	0,3054	11,776	1,739	0,426	5,8367	3,852	6,4950
AND3	2	23	8	69	68	73	9	50	0
M3_N1000_		0,040	0,4076	19,423	2,908	0,858	9,2590	7,182	10,622
AND3	4	85	7	27	22	72	2	50	50
Optimize DKL									
Model	Hata Oranı	DKL(R W)	DKL(D SD)	RE(RW)	RE(DS D)	PCC(R W)	PCC(D SD)	NS(R W)	NS(DS D)
M1_N1000_		0,201	0,4527	3,0797	0,286	0,130	2,3250	1,615	2,3250
AND3	0,5	85	6	3	82	68	0	00	0
M1_N1000_		0,210	0,3693	5,8791	0,580	0,229	4,3465	2,662	3,7725
AND3	1	16	0	9	85	05	6	50	0
M1_N1000_		0,220	0,3206	10,700	1,233	0,431	5,8725	4,527	5,8725
AND3	2	45	9	10	71	04	0	50	0
M1_N1000_		0,227	0,1945	18,447	2,060	0,685	9,5050	7,810	9,5050
AND3	4	89	3	94	48	49	0	00	0
M2_N1000_		0,283	0,0010	0,6387	0,031	0,123	0,4550	0,442	0,4550
AND3	0,5	23	6	4	45	55	0	50	0
M2_N1000_		0,412	0,0045	1,9447	0,086	0,264	0,4017	1,007	0,9025
AND3	1	40	0	2	12	20	5	50	0
M2_N1000_		0,417	0,0376	6,8001	0,284	0,371	1,7475	2,242	1,7475
AND3	2	78	1	0	88	45	0	50	0
M2_N1000_		0,419	0,0912	10,825	0,447	0,578	3,1525	4,012	3,1525
AND3	4	62	5	58	85	47	0	50	0
M3_N1000_		0,296	15,723	2,5345	0,301	0,108	1,9350	1,192	1,9350
AND3	0,5	30	20	7	17	73	0	50	0
M3_N1000_		0,323	14,143	4,7810	0,552	0,171	3,2142	2,130	3,3125
AND3	1	27	57	4	36	38	1	00	0
M3_N1000_		0,329	11,532	8,7320	1,048	0,319	5,2500	3,985	5,2500
AND3	2	13	48	8	41	45	0	00	0
M3_N1000_		0,325	9,8364	15,518	1,976	0,649	10,300	7,225	10,300
AND3	4	22	1	74	87	96	00	00	00

EK A.6

Çizelge A.6 : Ağ Büyüklüğü Değerlendirme Detaylı Sonuçlar.

		Optimize CC							
		DKL(R W)	DKL(DS D)	RE(RW)	RE(DS D)	PCC(R W)	PCC(DS D)	NS(RW)	NS(DS D)
ER	ER : 1000	0,0023	0,0038	4,9708	0,2605	0,1144	0,2395	2,3700	3,4925
	ER : 2000	0,0007	0,0023	5,8647	0,2999	0,1368	2,7492	2,5000	3,4600
	ER : 3000	0,0003	0,0033	6,1216	0,3007	0,1353	2,9286	2,4633	3,3892
	ER : 4000	0,0001	0,0034	5,9147	0,2908	0,1488	3,9963	2,6163	3,3663
	ER : 5000	0,0002	0,0046	5,7291	0,2738	0,1485	3,1698	2,5160	3,3525
	ER : 6000	0,0001	0,0049	5,4575	0,2733	0,1502	3,2905	2,5138	3,3458
WS	WS : 1000	0,0031	0,0017	1,1618	0,0339	0,0667	0,0178	0,6375	0,3550
	WS : 2000	0,0010	0,0004	0,5851	0,0170	0,0653	0,0159	0,6225	0,3400
	WS : 3000	0,0011	0,0002	0,4295	0,0122	0,0675	0,0174	0,6992	0,3717
	WS : 4000	0,0002	0,0001	0,3212	0,0092	0,0681	0,0172	0,7338	0,4044
	WS : 5000	0,0001	0,0001	0,2760	0,0079	0,0685	0,0194	0,8000	0,4315
	WS : 6000	0,0010	0,0001	0,2418	0,0070	0,0701	0,0217	0,8400	0,4521
BA	BA : 1000	0,0041	0,0075	4,0001	0,3784	0,1296	0,6321	2,0925	4,0300
	BA : 2000	0,0005	0,0027	4,7526	0,3587	0,1390	0,8719	1,9663	4,3350
	BA : 3000	0,0003	0,0024	5,2253	0,3451	0,1369	0,8582	2,0533	4,4617
	BA : 4000	0,0006	0,0201	5,7854	0,3392	0,1509	1,0456	2,1125	4,4994
	BA : 5000	0,0001	0,0093	6,1738	0,3312	0,1536	0,9967	2,2285	4,3635
	BA : 6000	0,0001	0,0043	6,4180	0,3274	0,1565	1,0641	2,2292	4,3112
		Optimize RW							
		DKL(R W)	DKL(DS D)	RE(RW)	RE(DS D)	PCC(R W)	PCC(DS D)	NS(RW)	NS(DS D)

ER	ER : 1000	0,0108	0,0957	7,4810	0,7567	0,2160	5,1099	2,5925	3,5100
	ER : 2000	0,0031	0,0150	7,8187	0,5302	0,0816	0,3006	1,8863	2,4288
	ER : 3000	0,0004	0,0102	7,6882	0,4315	0,0767	0,2171	1,6983	2,5375
	ER : 4000	0,0002	0,0112	7,5718	0,3880	0,0751	0,1823	1,6297	2,5648
	ER : 5000	0,0002	0,0126	7,5222	0,3508	0,0758	0,1215	1,5925	2,5425
	ER : 6000	0,0001	0,0131	7,4259	0,3332	0,0801	0,0996	1,6833	2,6958
WS	WS : 1000	0,0029	0,0502	6,0833	0,2870	0,1255	0,6275	1,1600	0,8650
	WS : 2000	0,0011	0,0490	4,3090	0,1636	0,1224	0,6181	1,1500	0,8313
	WS : 3000	0,0005	0,0217	2,3251	0,0887	0,1227	0,5007	1,2242	0,9042
	WS : 4000	0,0002	0,0137	1,5947	0,0611	0,1192	0,4474	1,2438	0,9477
	WS : 5000	0,0001	0,0101	1,2418	0,0489	0,1186	0,4420	1,2050	0,9125
	WS : 6000	0,0001	0,0060	0,9615	0,0367	0,1141	0,3775	1,2875	0,9708
BA	BA : 1000	0,0097	0,2122	6,7276	0,9992	0,1970	3,5280	2,0175	3,8700
	BA : 2000	0,0045	0,1210	6,9446	0,8544	0,1607	3,5793	2,0838	3,5600
	BA : 3000	0,0010	0,0418	7,3698	0,7028	0,1414	2,9073	2,0433	3,5650
	BA : 4000	0,0003	0,0370	7,6632	0,6493	0,1215	2,5685	1,8667	3,5000
	BA : 5000	0,0001	0,0123	7,7166	0,6081	0,1105	2,2337	1,9200	3,5150
	BA : 6000	0,0001	0,0021	7,8896	0,5704	0,1075	2,2168	1,9708	3,5542
Optimize DKL									
		DKL(R W)	DKL(DS D)	RE(RW)	RE(DS D)	PCC(R W)	PCC(DS D)	NS(RW)	NS(DS D)
ER	ER : 1000	0,2102	0,2693	5,8792	0,5809	0,2290	9,3466	2,7625	3,7725
	ER : 2000	0,0714	0,0985	6,0216	0,6748	0,3285	17,6507	3,0313	3,9813
	ER : 3000	0,0477	0,0480	5,9288	0,5952	0,2788	18,5352	2,8742	3,7067
	ER : 4000	0,0304	0,0262	5,7156	0,4723	0,2533	15,2472	2,9008	3,5766

	ER : 5000	0,0224	0,0183	5,9220	0,3883	0,1945	12,1878	2,8575	3,5183
	ER : 6000	0,0209	0,0103	5,4296	0,3410	0,1903	15,4825	2,8135	3,5438
W S	WS : 1000	0,4124	0,0045	1,9447	0,0861	0,2642	0,4018	1,0075	0,9025
	WS : 2000	0,2019	0,0091	2,2625	0,0859	0,1842	0,4130	1,0688	0,8388
	WS : 3000	0,0894	0,0013	0,9298	0,0338	0,1256	0,1970	1,1050	0,8917
	WS : 4000	0,0406	0,0004	0,5315	0,0195	0,1017	0,1256	1,0688	0,8563
	WS : 5000	0,0385	0,0005	0,5101	0,0188	0,1055	0,1550	1,1017	0,9125
	WS : 6000	0,0243	0,0004	0,5713	0,0164	0,0938	0,1350	1,2188	0,9083
BA	BA : 1000	0,3233	17,1436	4,7810	0,5524	0,1714	3,2142	2,1300	3,3125
	BA : 2000	0,1695	9,2377	5,3789	0,4594	0,1492	2,2249	2,1650	3,0775
	BA : 3000	0,0858	13,6962	5,8029	0,3964	0,1507	1,5891	2,1392	3,3717
	BA : 4000	0,0483	12,5362	6,1516	0,3760	0,1528	1,3615	2,1609	3,3484
	BA : 5000	0,0368	11,0857	6,3647	0,3446	0,1555	1,1464	2,2608	3,4550
	BA : 6000	0,0325	11,3952	6,4357	0,3507	0,1567	1,4708	2,2042	3,5854

EK A.7

Çizelge A.7 : Organizmaları Karşılaştırma Detaylı Sonuçlar.

Organizma	DKL(R W)	DKL(D SD)	RE(R W)	RE(DS D)	PCC(RW)	PCC(D SD)	NS(R W)	NS(DS D)
<i>Helicobacter pylori</i> (1)	0,06	0,11	10,50	1,02	0,09	1,46	2,63	3,12
<i>Mus musculus</i> (2)	0,01	0,04	9,22	0,40	0,07	0,64	2,49	5,46
<i>Rattus norvegicus</i> (3)	0,14	0,51	23,76	1,25	0,20	1,06	6,69	29,40
<i>Plasmodium falciparum</i> (4)	0,03	0,02	8,95	0,81	0,16	2,23	2,82	3,25
<i>Caenorhabditis elegans</i> (5)	0,00	0,14	10,39	0,76	0,10	11,20	2,93	8,43
<i>Arabidopsis thaliana</i> (6)	0,00	0,19	13,24	0,84	0,21	23,25	2,10	11,85
<i>Saccharomyces cerevisiae</i> (7)	0,01	0,23	6,41	3,11	0,58	43,45	3,37	10,49
<i>Homo sapiens</i> (8)	0,00	0,10	9,63	0,97	0,22	25,90	2,68	9,34

Organizma	Veri Tabanı
<i>Helicobacter pylori</i> (1)	String
<i>Mus musculus</i> (2)	BioGrid
<i>Rattus norvegicus</i> (3)	BioGrid
<i>Plasmodium falciparum</i> (4)	BioGrid
<i>Caenorhabditis elegans</i> (5)	BioGrid
<i>Arabidopsis thaliana</i> (6)	BioGrid
<i>Saccharomyces cerevisiae</i> (7)	BioGrid
<i>Homo sapiens</i> (8)	BioGrid

EK A.8

Çizelge A.8 : Mutasyon Oranı Karşılaştırma Detaylı Sonuçlar.

Organizma	Hata Oranı	DKL(R W)	DKL(D SD)	RE(R W)	RE(D SD)	PCC(R W)	PCC(D SD)	NS(R W)	NS(D SD)
<i>Helicobacter pylori</i> (1)	0,5	0,04	0,01	6,55	0,62	0,04	0,81	1,45	2,49
	1	0,06	0,11	10,50	1,02	0,09	1,46	2,63	3,12
	2	0,06	0,14	18,21	1,93	0,26	3,11	4,41	4,65
	4	0,09	0,31	27,94	4,49	1,02	30,23	8,72	10,56
<i>Mus musculus</i> (2)	0,5	0,01	0,02	6,28	0,25	0,03	0,37	1,93	4,28
	1	0,01	0,04	9,22	0,40	0,07	0,64	2,49	5,46
	2	0,02	0,12	12,52	0,62	0,18	1,35	4,07	6,91
	4	0,04	0,50	15,82	1,13	0,61	12,03	8,43	10,86
<i>Rattus norvegicus</i> (3)	0,5	0,08	0,39	18,56	0,76	0,12	0,78	0,59	27,49
	1	0,14	0,51	23,76	1,25	0,20	1,06	6,69	29,40
	2	0,10	0,80	33,35	2,48	0,58	2,97	12,86	31,29
	4	0,12	14,17	40,78	5,22	1,35	16,37	32,87	31,94
<i>Plasmodium falciparum</i> (4)	0,5	0,04	0,01	5,43	0,50	0,07	1,63	1,57	1,96
	1	0,03	0,02	8,95	0,81	0,16	2,23	2,82	3,25
	2	0,04	0,16	15,35	2,46	0,55	32,01	5,40	6,28
	4	0,04	0,21	24,98	4,64	1,13	48,53	10,19	10,78
<i>Caenorhabditis elegans</i> (5)	0,5	0,00	0,01	6,34	0,37	0,04	0,56	1,63	7,09
	1	0,00	0,14	10,39	0,76	0,10	11,20	2,93	8,43
	2	0,01	0,16	15,45	1,27	0,31	21,13	5,01	10,46
	4	0,00	0,20	22,99	4,40	1,14	57,30	9,33	14,59
<i>Arabidopsis thaliana</i> (6)	0,5	0,00	0,18	10,39	0,45	0,07	10,03	1,09	10,95
	1	0,00	0,19	13,24	0,84	0,21	23,25	2,10	11,85
	2	0,00	0,21	17,76	1,95	0,55	42,96	4,22	14,98
	4	0,00	0,24	25,05	4,03	1,24	58,22	7,44	18,35

<i>Saccharomyces cerevisiae</i> (7)	0,5	0,01	0,35	4,61	2,54	0,44	39,66	2,31	8,03
	1	0,01	0,23	6,41	3,11	0,58	43,45	3,37	10,49
	2	0,01	0,22	9,13	4,22	0,77	51,41	5,19	13,33
	4	0,01	0,38	13,40	5,68	1,14	54,92	8,21	18,88
<i>Homo sapiens</i> (8)	0,5	0,00	0,09	6,81	0,44	0,06	7,22	1,66	8,37
	1	0,00	0,10	9,63	0,97	0,22	25,90	2,68	9,34
	2	0,00	0,12	14,28	2,51	0,63	49,01	4,96	12,41
	4	0,00	0,16	22,44	4,54	1,32	60,33	8,13	16,78



EK A.9

Çizelge A.9 : Ağ Parametreleri Karşılaştırma Detaylı Sonuçlar (LE).

Ağ	RW LE	DSD LE
M1_N500_AND2	51,17749109	21,07498985
M1_N500_AND3	40,53327863	20,91472377
M1_N500_AND4	41,85645421	19,02714546
M1_N500_AND5	38,65142892	17,49571401
M1_N1000_AND2	52,04490619	23,20296728
M1_N1000_AND3	45,39962438	24,22688948
M1_N1000_AND4	40,76850985	20,01545297
M1_N1000_AND5	36,13739531	19,09837483
M1_N1500_AND2	52,64768618	24,74330241
M1_N1500_AND3	47,98226709	25,65741266
M1_N1500_AND4	40,023611	20,80787971
M1_N1500_AND5	35,40719842	19,35361341
M1_N2000_AND2	52,54477252	24,58303633
M1_N2000_AND3	49,55292551	28,07772729
M1_N2000_AND4	39,87904182	22,19240059
M1_N2000_AND5	35,40229777	19,88634974
M2_N500_AND2	46,28080907	23,8378796
M2_N500_AND3	47,63931998	25,97362384
M2_N500_AND4	50,69596952	25,75734595
M2_N500_AND5	50,9049712	25,20313383
M2_N1000_AND2	46,64656201	24,14877908
M2_N1000_AND3	47,82219645	26,08176279
M2_N1000_AND4	50,80047036	25,64244831
M2_N1000_AND5	51,33603716	24,75030198
M2_N1500_AND2	46,60302	23,71622328
M2_N1500_AND3	47,75252922	25,91504858
M2_N1500_AND4	50,58276028	25,96911805
M2_N1500_AND5	51,32297456	25,12202962
M2_N2000_AND2	46,60084289	24,11160631
M2_N2000_AND3	47,87444687	26,17300503
M2_N2000_AND4	50,16693403	25,94996845
M2_N2000_AND5	50,9049712	25,20651317
M4_N500_AND2	46,87506663	19,08716275
M4_N500_AND3	46,22617916	16,51049488
M4_N500_AND4	42,07894185	15,44018669
M4_N500_AND5	43,1792293	15,16269938
M4_N1000_AND2	46,94559788	18,17541874
M4_N1000_AND3	46,80453539	17,02582846
M4_N1000_AND4	41,69807312	16,47085384
M4_N1000_AND5	41,82502936	15,94891342
M4_N1500_AND2	46,7716208	17,41342978
M4_N1500_AND3	46,29200833	17,74377181
M4_N1500_AND4	41,86264603	17,26807928
M4_N1500_AND5	40,72474191	16,8760734
M4_N2000_AND2	47,6650166	17,28019182
M4_N2000_AND3	46,43071978	17,8450767
M4_N2000_AND4	41,93082623	17,35947391
M4_N2000_AND5	40,00532319	16,95315321

EK A.10**Çizelge A.10 : Mutasyon Etkisi Karşılaştırma Detaylı Sonuçlar (LE).**

		RW	RW Mutant	DSD	DSD Mutant
m1	Exp 1	100,00	100,00	64,13	97,80
	Exp 2	100,00	100,00	26,57	39,58
	Exp 3	62,50	62,26	2,21	2,85
	Exp 4	0,00	0,00	0,00	0,00
	Exp 5	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00
m2	Exp 1	100,00	100,00	100,00	100,00
	Exp 2	99,98	99,99	33,67	40,38
	Exp 3	93,78	95,98	18,23	21,47
	Exp 4	0,21	0,14	0,00	0,00
	Exp 5	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00
m4	Exp 1	100,00	100,00	63,28	57,59
	Exp 2	99,77	99,79	2,03	2,61
	Exp 3	65,48	65,66	0,00	0,00
	Exp 4	0,00	0,00	0,00	0,00
	Exp 5	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00

EK A.11

Çizelge A.11 : Düğüm Sırası Karşılaştırma Detaylı Sonuçlar (LE).

		rw	rw obc	rw occ	rw ora	rw olc	rw odc	rw olc
m1	Exp 1	100,00	100,00	100,00	100,00	100,00	100,00	100,00
	Exp 2	100,00	100,00	100,00	100,00	100,00	100,00	100,00
	Exp 3	62,50	74,14	73,24	62,42	64,86	74,43	74,20
	Exp 4	0,00	0,00	0,00	0,00	0,00	0,01	0,00
	Exp 5	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00	0,00	0,00	0,00
m2	Exp 1	100,00	100,00	100,00	100,00	100,00	100,00	100,00
	Exp 2	99,98	100,00	100,00	100,00	100,00	99,98	100,00
	Exp 3	93,78	98,67	98,39	98,59	97,53	78,70	98,56
	Exp 4	0,21	0,03	0,00	0,04	0,01	0,00	0,02
	Exp 5	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00	0,00	0,00	0,00
m4	Exp 1	100,00	100,00	99,94	100,00	100,00	100,00	100,00
	Exp 2	99,77	99,75	99,50	99,98	99,90	99,83	99,78
	Exp 3	65,48	69,97	64,42	65,19	70,00	71,40	69,83
	Exp 4	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 5	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00	0,00	0,00	0,00
		dsd	dsd obc	dsd occ	dsd ora	dsd olc	dsd odc	dsd olc
m1	Exp 1	64,34	64,34	64,34	64,34	64,34	64,34	64,34
	Exp 2	28,98	28,98	28,98	28,98	28,98	28,98	28,98
	Exp 3	2,40	2,85	2,82	2,40	2,50	2,86	2,85
	Exp 4	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 5	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00	0,00	0,00	0,00
m2	Exp 1	100,00	100,00	100,00	100,00	100,00	81,00	100,00
	Exp 2	31,45	31,85	31,26	31,44	31,64	31,55	31,37
	Exp 3	19,89	20,93	20,87	20,91	20,68	6,69	20,90
	Exp 4	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 5	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00	0,00	0,00	0,00
m4	Exp 1	62,53	62,53	62,49	62,53	62,53	62,53	62,53
	Exp 2	2,21	2,21	2,20	2,20	2,22	2,21	2,21
	Exp 3	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 4	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 5	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Exp 6	0,00	0,00	0,00	0,00	0,00	0,00	0,00

EK A.12

Çizelge A.12 : Organizmaları Karşılaştırma Detaylı Sonuçlar (LE).

Organizma	RW	DSD
<i>Helicobacter pylori</i> (1)	26,86	26,90
<i>Mus musculus</i> (2)	26,92	25,78
<i>Rattus norvegicus</i> (3)	24,94	14,35
<i>Plasmodium falciparum</i> (4)	26,77	26,84
<i>Caenorhabditis elegans</i> (5)	26,72	24,37
<i>Arabidopsis thaliana</i> (6)	27,11	22,73
<i>Saccharomyces_cerevisiae</i> (7)	26,51	23,38
<i>Homo sapiens</i> (8)	26,83	23,93

EK A.13

Çizelge A.13 : Düğüm Sırası Karşılaştırma Detaylı Sonuçlar (LE).

Organizma	obc		odc		ora		occ		olc		oclc	
	rw	dsd	rw	dsd	rw	dsd	rw	dsd	rw	dsd	rw	dsd
Helicobacter pylori (1)	26, 99	26,9 7	26, 86	26,9 0	26, 86	26,9 0	27, 11	26,8 7	26, 92	27,1 1	26, 32	26,9 6
Mus musculus (2)	26, 59	26,0 7	26, 92	26,0 5	26, 92	25,7 8	27, 11	25,6 7	26, 65	25,2 7	26, 76	26,1 8
Rattus norvegicus (3)	24, 63	14,2 1	24, 94	17,8 7	24, 94	14,3 5	24, 79	14,8 8	24, 93	14,7 3	24, 62	14,6 2
Plasmodium falciparum (4)	26, 80	26,5 6	26, 77	26,8 4	26, 77	26,8 4	26, 59	26,3 7	26, 78	26,6 5	27, 11	26,7 8
Caenorhabditis elegans (5)	26, 69	23,9 8	26, 72	24,3 7	26, 72	24,3 7	26, 54	24,6 1	26, 63	24,1 7	26, 48	23,8 4
Arabidopsis thaliana (6)	27, 28	23,0 9	27, 11	23,2 8	27, 11	22,7 3	26, 57	22,9 7	26, 79	22,9 6	27, 11	22,6 4
Saccharomyces_cerevisiae (7)	26, 77	23,5 9	26, 51	24,1 9	26, 51	23,3 8	26, 61	23,6 5	26, 92	23,6 8	26, 77	23,2 8
Homo sapiens (8)	27, 27	24,0 9	26, 83	25,0 2	26, 83	23,9 3	27, 11	23,7 6	26, 71	23,7 9	26, 87	23,8 4

ÖZGEÇMİŞ



Ad Soyad: Volkan Altuntaş
Doğum Yeri ve Tarihi: Zonguldak 1982
E-Posta: valtuntas@gmail.com

Lisans: Gebze Teknik Üniversitesi

Yüksek Lisans: Yalova Üniversitesi

Mesleki Deneyim ve Ödüller:

- Öğretim Görevlisi, Bursa Teknik Üniversitesi
- AR-GE ve Proje Yöneticisi, Yılmaz Holding,
- AR-GE ve Proje Yöneticisi, Durmazlar Holding
- Proje Yöneticisi, STOK Endüstriyel
- Yazılım Yöneticisi, İstihkam Eğitim Okulu
- AR-GE Mühendisi, ENMOS Endüstriyel Otomasyon
- Stajyer, Pi NET Yazılım
- Stajyer, Microsoft Türkiye
- Bölüm Birincisi, Gebze Teknik Üniversitesi
- Üstün Başarı Ödülü, Türk Eğitim Vakfı

Yayın ve Patent Listesi:

- V. Altuntas, M. Gök, O.H. Koçal, Response of Lyapunov Exponents to Biological Networks' Diffusion State, International Journal of Applied Mathematics and Computer Science (2020) (**sci-exp**) (**kabül**)
- V. Altuntas, M. Gök, T. Kahveci, Stability Analysis of Biological Networks Diffusion State, IEEE/ACM Transactions on Computational Biology and Bioinformatics (2020) (**sci-exp**)
- S., Altuntas, A., Çınar, V., Altuntaş, Modelling of Listeria monocytogenes growth and survival in presence of royal jelly: a promising antibiofilm agent, Journal of Food and Nutrition Research, (2020) (**sci-exp**)
- V. Altuntas, S. Altuntas, M. Gök, Automatic Colony Segmentation on Agar Surface by Image Processing, Journal of Scientific and Industrial Research, (2018) (**sci-exp**)

- V. Altuntas, M. Gök, Protein - Protein Interaction Detection Methods, Databases and Data Reliability, Avrupa Bilim ve Teknoloji Dergisi (2020) (TR)
- V., Altuntas, S., Altuntas, M., Gök, Örüntü Tanıma Teknikleri ile Agar Yüzeyi Üzerinde Koloni Morfoloji Sınıflandırması, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, (2018) (TR)
- S., Altuntas, M., Korukluoglu, V., Altuntaş, Probiyotik Escherichia coli Nissle 1917 Suşu, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, (2017) (TR)
- M., Gök, V., Altuntas, Regresyon Analizi, RapidMiner ile Veri Madenciliği, (2017)
- V. Altuntas, M. Gok, Effects of Cluster Coefficient Based Mutations on Average Cluster of Synthetic Networks, IconData (2020)
- V. Altuntas, M. Gok, T. Aktaş, Implementation of Multiprocessing Architects on Network Motif Detection Algorithms, ICAT (2018)
- V. Altuntas, M. Gok, The stability and fragility of biological networks: Eukaryotic model organism Saccharomyces cerevisiae, UBMK (2017)
- V. Altuntas, M. Gok, Görüntü İşleme ve Makine Öğrenmesi ile Diyabetik Retinopati Tanısı, IMSEC (2017)
- V. Altuntas, M. Gok, Yapay Sinir Ağları ile Parmak İzi Tanıma, IMSEC (2017)
- S. Altuntas, V. Altuntas, Lactobacillus bulgaricus Büyüme Evresinin Tanımlanması ve Sınıflandırılması, IMSEC (2017)
- V. Altuntas, M. Cetin, Analysis of satellite images with segmentation algorithms, IAC (2017)
- V. Altuntas, Waste Water Bacteria Microscopic Image Analysis and Classification System , ICRES (2017)
- V., Altuntas Iris Identification System with Image Analysis, ICRES (2017)
- V. Altuntas, S. Altuntas, M. Gök, Machine Learning Techniques for Colony Classification, ICRES (2016)
- Yükseköğretim Kurumları tarafından destekli bilimsel araştırma projesi, Araştırmacı, 2016, Koloni Morfoloji Tahmininde Yeni Yapay Öğrenme Modelleri Geliştirilmesi (7.000 TL)
- TÜBİTAK PROJESİ, Yürütücü, 2019, Otomatik Koloni Tespit ve Sayım Cihazı (199.983 TL)
- Yükseköğretim Kurumları tarafından destekli bilimsel araştırma projesi, Araştırmacı, 2020, Sentetik ve Biyolojik Ağların Yeni Tekniklerle Analizi ve Güvenirliklerinin Belirlenmesi (28.792 TL)
- Protein Etkileşim Ağları ve Ağ Güvenilirlik Analizleri (Doktora)
- Koloni Morfoloji Tahmininde Yeni Yapay Öğrenme Modelleri Geliştirilmesi (Yüksek Lisans)
- Lineer Denklem Sistemlerinin Paralel ve Dağıtık Sistemler ile Çözümü (Lisans)

TEZDEN TÜRETİLEN YAYINLAR/SUNUMLAR

Makale:

- [1] V. Altuntas, M. Gök, T. Kahveci, Stability Analysis of Biological Networks Diffusion State, IEEE/ACM Transactions on Computational Biology and Bioinformatics. (2020) (**sci-exp**)
- [2] V. Altuntas, M. Gök, O.H. Koçal, Response of Lyapunov Exponents to Biological Networks' Diffusion State, International Journal of Applied Mathematics and Computer Science (2020) (**sci-exp**)(**kabül**)
- [3] V. Altuntas, M. Gök, Protein - Protein Interaction Detection Methods, Databases and Data Reliability, Avrupa Bilim ve Teknoloji Dergisi (2020) (**TR**)

Bildiri:

- [1] V. Altuntas, M. Gök, Effects of Cluster Coefficient Based Mutations on Average Cluster of Synthetic Networks, International Conference on Data Science and Applications (ICONDATA), 2020, İstanbul, Turkey.
- [2] V. Altuntas, M. Gök, T. Aktaş, Implementation of Multiprocessing Architects on Network Motif Detection Algorithms, International Conference on Advanced Technologies (ICAT), 2018, Antalya, Turkey.
- [3] V. Altuntas, M. Gök, The stability and fragility of biological networks: Eukaryotic model organism *Saccharomyces cerevisiae*, International Conference on Computer Science and Engineering (UBMK), 2017, Antalya, Turkey.