



**KARMAŞIK AĞLARDA DÜĞÜM AĞIRLIKLI
BAĞLANTI TAHMİN YÖNTEMLERİNİN
GELİŞTİRİLMESİ**

Emrah ÖZKAYNAK

**2020
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

**Tez Danışmanı
Prof. Dr. Oğuz FINDIK**

**KARMAŐIK AĐLARDA DÜĐÜM AĐIRLIKLI BAĐLANTI TAHMİN
YÖNTEMLERİNİN GELİŐTİRİLMESİ**

Emrah ÖZKAYNAK

**T.C.
Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar MühendisliĐi Anabilim Dalında
Doktora Tezi
Olarak Hazırlanmıştır**

**Tez Danışmanı
Prof. Dr. OĐuz FINDIK**

**KARABÜK
Aralık 2020**

Emrah ÖZKAYNAK tarafından hazırlanan “KARMAŞIK AĞLARDA DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİN YÖNTEMLERİNİN GELİŞTİRİLMESİ” başlıklı bu tezin Doktora Tezi olarak uygun olduğunu onaylarım.

Prof. Dr. Oğuz FINDIK
Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışma, jürimiz tarafından Oy Birliği ile Bilgisayar Mühendisliği Anabilim Dalında Doktora tezi olarak kabul edilmiştir. 23/12/2020

<u>Ünvanı, Adı SOYADI (Kurumu)</u>	<u>İmzası</u>
Başkan : Doç. Dr. İlker TÜRKER (KBÜ)
Üye : Prof. Dr. Oğuz FINDIK (KBÜ)
Üye : Doç. Dr. Ergin YILMAZ (BEÜ)
Üye : Dr. Öğr. Üyesi Okan ERKAYMAZ (BEÜ)
Üye : Dr. Öğr. Üyesi Ümit ATİLA (KBÜ)

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Doktora derecesini onamıştır.

Prof. Dr. Hasan SOLMAZ
Lisansüstü Eğitim Enstitüsü Müdürü

“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

Emrah ÖZKAYNAK

ÖZET

Doktora Tezi

KARMAŞIK AĞLARDA DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİN YÖNTEMLERİNİN GELİŞTİRİLMESİ

Emrah ÖZKAYNAK

Karabük Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:

Prof. Dr. Oğuz FINDIK

Aralık 2020, 105 sayfa

Bağlantı tahmini, karmaşık ağlarda gelecekte ortaya çıkabilecek yeni bağlantıların tahmin edilmesinde kullanılmaktadır. Geçmişte birçok bağlantı tahmin yöntemi önerilmiştir. Önerilen yöntemlerin birçoğu ağın anlık topolojik yapısına ait bilgileri kullanmaktadır. Özellikle düğümler arasındaki ortak komşulara ait bilgiler kullanılarak benzerlik ölçümü yapan yöntemler, başarı oranlarının diğer yöntemlere göre yüksek olmasından dolayı tercih edilmektedir. Ayrıca dinamik ağlarda zaman içerisinde ağın topolojik yapısında meydana gelen değişiklikleri dikkate alarak düğümler arasında benzerlik ölçümü yapan yöntemler de önerilmiştir. Ancak hem statik hem de dinamik ağlar için önerilen yöntemler, tahmin işlemlerine düğümlerin ağdaki gücünü dahil etmemektedir. Düğümlerin zaman içerisinde ağda içerisindeki etkileşimlerinde meydana gelen değişimlerin ortaya çıkarılması düğümün gücünü tespit etmede önemlidir. Tahmin işlemlerinin topolojik faktörlerle sınırlandırılması tahminlerin başarısını olumsuz etkilemektedir. Özellikle dinamik ağlar için etkili olan

ve topolojik bilgilerle ifade edilemeyen faktörlerin tahmin sürecine dahil edilmesi önemlidir. Düğümlerin ağdaki gücünü hesaplarken, topolojik bilgilerle birlikte zaman periyodu gibi düğümün ağdaki gücünü ölçebilecek faktörlerinde dikkate alınması tahmin başarısını artırıcı bir etken olmaktadır. Bu tez çalışmasında, bahsedilen problemlerin üstesinden gelmek için karmaşık ağlarda düğümlerin ağırlıklandırıldığı ve ağırlıklandırılmış düğümlerle etkili bağlantı tahmini yapabilen bir yöntem geliştirilmiştir. Geliştirilen yöntem hem ağdaki güçlü düğümleri tespit etmede hem de iki düğüm arasındaki bağlantı olasılığını hesaplamada ortak komşuların ağırlıklarını dikkate almaktadır. Düğümlerin ağırlıkları hesaplanırken zaman periyodu ve düğümlerin ağırlık gelişimine etkisini ifade eden faktörler dikkate alınmıştır. Ayrıca zaman periyodu ile birlikte merkezilik ölçütleri kullanılarak tahmin işlemlerinde karşılaştırmalar yapılmıştır. Geliştirilen yöntem, Avustralya Açık Tenis Turnuvalarından ve IEEE Xplore'daki yayınlardan oluşturulan gerçek dünya ağlarında test edilmiştir. Farklı zaman periyotlarından oluşturulan ağlarda yapılan deneysel çalışmalarda geliştirilen yöntemin ortak komşuların ilişkilerine dayalı mevcut bağlantı tahmin yöntemlerinden daha başarılı olduğu görülmüştür. Deneysel çalışmaların bir diğer sonucu da dinamik ağlarda merkezi düğümlerin tespiti için geliştirilen yöntemin geleneksel merkezilik ölçütlerinden daha başarılı olmasıdır.

Anahtar Sözcükler : Karmaşık ağlar, sosyal ağlar, bağlantı tahmini, düğüm ağırlıklandırma.

Bilim Kodu : 92429

ABSTRACT

Ph. D. Thesis

DEVELOPMENT OF NODE WEIGHTED LINK PREDICTION METHODS IN COMPLEX NETWORKS

Emrah ÖZKAYNAK

**Karabük University
Institute of Graduate Programs
Department of Computer Engineering**

Thesis Advisor:

Prof. Dr. Oğuz FINDIK

December 2020, 105 pages

Link prediction is used to predict new connections that may arise in the future in complex networks. Many link estimation methods have been proposed in the past. Most of the proposed methods use instantaneous topological information of the network. Especially methods that measure similarity between nodes using information about common neighbors are preferred because their success rates are higher than other methods. In addition, methods that measure similarity between nodes in dynamic networks, taking into account the changes in the topological structure of the network over time are also proposed. However, the proposed methods for both static and dynamic networks do not adequately predict the power of the nodes in the network. Revealing the changes that occur in the interaction of nodes within the network over time is important in determining the strength of the node. Limiting the prediction processes to topological factors negatively affects the success of the predicts. Especially in dynamic networks, it is important to include factors that are effective in

the expansion of the network and cannot be expressed with topological information into the prediction process. When calculating the power of nodes in the network, considering the factors that can measure the power of the node in the network, such as the time period together with topological information, is a factor that increases the prediction success. In this thesis, in order to overcome the mentioned problems, a method in which nodes are weighted in complex networks and that can effectively predict connections with weighted nodes has been developed. The developed method considers the weights of common neighbors both in detecting strong nodes in the network and in calculating the connection probability between two nodes. While calculating the weights of the nodes, the time period and factors that express the effect of nodes on the development of the network were considered. In addition, comparisons were made in prediction processes by using centrality criteria with the time period. The developed method has been tested in real world networks created from Australian Open Tennis Tournaments and publications in IEEE Xplore. It has been observed that the method developed in the experimental studies performed on networks created from different time periods is more successful than the existing link prediction methods. Another result of experimental studies is that the method developed for the detection of central nodes in dynamic networks is more successful than traditional centrality measurements.

Key Word : Complex networks, social networks, link prediction, node weighting.

Science Code : 92429

TEŐEKKÜR

Bu tez alıőmasının planlanmasında, araőtırılmasında, yürütülmesinde ve oluşumunda ilgi ve desteęini esirgemeyen, engin bilgi ve tecrübelerinden yararlandıęım, yönlendirme ve bilgilendirmeleriyle alıőmamı bilimsel temeller ışığında őekillendiren sayın hocam Prof. Dr. Oęuz FINDIK'a sonsuz teőekkürlerimi sunarım.

Tez alıőmamın her aşamasında görüşleri ve önerileriyle katkılarını esirgemeyen tez izleme komitesindeki deęerli hocalarım Doę. Dr. Ergün YILMAZ ve Doę. Dr. İlker TÜRKER'e çok teőekkür ederim.

Sevgili aileme manevi hiçbir yardımını esirgemedен yanımda oldukları için tüm kalbimle teőekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER	ix
ŞEKİLLER DİZİNİ.....	xii
ÇİZELGELER DİZİNİ	xiv
SİMGELER VE KISALTMALAR DİZİNİ	xv
BÖLÜM 1	1
GİRİŞ	1
BÖLÜM 2	7
KARMAŞIK AĞLAR.....	7
2.1. KARMAŞIK AĞ ANALİZİ.....	8
2.2. ÇİZGE (GRAF).....	9
2.3. KARMAŞIK AĞ ANALİZİNDE TEMEL MERKEZİLİK ÖLÇÜTLERİ ...	10
2.3.1. Derece Merkeziliği	10
2.3.2. Yakınlık Merkeziliği.....	11
2.3.3. Arasındalık Merkeziliği.....	12
2.3.4. Özvektör Merkeziliği.....	12
BÖLÜM 3	13
BAĞLANTI TAHMİNİ.....	13
3.1. BAĞLANTI TAHMİNİ PROBLEMİ	14
3.2. BAĞLANTI TAHMİNİ YÖNTEMLERİ	15
3.2.1. Benzerlik Tabanlı Yaklaşımlar	16
3.2.1.1. Yerel Benzerlik Tabanlı Yaklaşımlar	16
3.2.1.2 Yol Tabanlı Yaklaşımlar.....	21

	<u>Sayfa</u>
3.2.1.3. Yarı Yerel Benzerlik Tabanlı Yaklaşımlar	22
3.2.2. Olasılıksal Yaklaşımlar.....	24
3.2.3. İlişkisel Modeller	24
3.2.4. Öğrenmeye Dayalı Yaklaşımlar	24
3.3. BAĞLANTI TAHMİNİ BAŞARISINI ÖLÇMEDE KULLANILAN DEĞERLENDİRME ÖLÇÜTLERİ	25
 BÖLÜM 4	 29
ÇOK KRİTERLİ KARAR VERME ANALİZİ.....	29
4.1. LOGARİTMİK KONSEPT YAKLAŞIMI (APLOCO).....	30
 BÖLÜM 5	 34
VERİ SETİNİN OLUŞTURULMASI.....	34
5.1. VERİ SETİ VE ÖNİŞLEME.....	34
5.2. ZAMAN PERİYODUNA BAĞLI OLARAK AĞLARIN OLUŞTURULMASI	35
 BÖLÜM 6	 38
KOMŞULUK İLİŞKİLERİNE DAYALI BAĞLANTI TAHMİNİ İÇİN GELİŞTİRİLEN DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİNİ YÖNTEMİ	38
6.1. DÜĞÜM AĞIRLIKLANDIRMA İŞLEMİ.....	41
6.2. DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİN İŞLEMİ.....	43
6.2.1. Düğüm Ağırlıklı Ortak Komşular (NwCN).....	44
6.2.2. Düğüm Ağırlıklı Jaccard İndeks (NwJI).....	44
6.2.3. Düğüm Ağırlıklı Tercihli Bağlantı İndeksi (NwPA).....	45
6.2.4. Düğüm Ağırlıklı Adamic Adar İndeksi (NwAA).....	45
6.2.5. Düğüm Ağırlıklı Kaynak Paylaştırma İndeksi (NwRA)	46
6.2.6. Düğüm Ağırlıklı Sorenson İndeks (NwSI).....	46
6.3. DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİNİ YÖNTEMİ İÇİN GERÇEKLEŞTİRİLEN DENEYSEL ÇALIŞMA.....	47
6.3.1. Deneysel Çalışmada Kullanılan Ağırlıklandırma Kriterlerinin Belirlenmesi	48
6.3.1.1. Tenis Ağlarının Ağırlıklandırılmasında Kullanılan Kriterler	48

	<u>Sayfa</u>
6.3.1.2. Bilimsel İşbirliği Ağlarının Ağırlıklandırılmasında Kullanılan Kriterler.....	50
6.3.2. 2000-2003 Arası Bağlantı Tahmin Sonuçları.....	52
6.3.3. 2000-2007 Arası Bağlantı Tahmin Sonuçları.....	52
6.3.4. 2000-2014 Arası Bağlantı Tahmin Sonuçları.....	59
BÖLÜM 7	60
DÜĞÜM AĞIRLIKLIL BAĞLANTI TAHMİNİNDE MERKEZİLİK ÖLÇÜTLERİ VE ZAMAN FAKTÖRÜNÜN BİRLİKTE KULLANIMI	60
7.1. 2000-2003 ARASI BAĞLANTI TAHMİN SONUÇLARI	60
7.2. 2000-2007 ARASI BAĞLANTI TAHMİN SONUÇLARI	63
7.3. 2000-2014 ARASI BAĞLANTI TAHMİN SONUÇLARI	63
BÖLÜM 8	79
DÜĞÜM AĞIRLIKLARININ MERKEZİLİK ÖLÇÜTÜ OLARAK KULLANILMASI	79
8.1. DENEYSEL ÇALIŞMA İÇİN AĞLARIN OLUŞTURULMASI	79
8.2. 2000-2003 ARASI AĞLARDAN ELDE EDİLEN MERKEZİLİK SONUÇLARI	80
8.3. 2000-2010 ARASI AĞLARDAN ELDE EDİLEN MERKEZİLİK SONUÇLARI	83
8.4. 2000-2017 ARASI AĞLARDAN ELDE EDİLEN MERKEZİLİK SONUÇLARI	86
BÖLÜM 9	90
SONUÇLAR	90
KAYNAKLAR	92

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1. Tenis turnuvaları ağı.	7
Şekil 2.2. Farklı bağlantı türleri.	10
Şekil 3.1. Ağda yeni bağlantıların oluşumu.	15
Şekil 4.1. APLOCO'nun uygulama adımları.	30
Şekil 5.1. 2000-2003 Tenis turnuvaları ağı.	36
Şekil 5.2. 2000-2007 Tenis turnuvaları ağı.	36
Şekil 5.3. 2000-2010 Tenis turnuvaları ağı.	37
Şekil 5.1. Geliştirilen yöntemin uygulama adımları.	40
Şekil 6.1. 2000-2003 Arası tenis ağından elde edilen AUC sonuçları.	53
Şekil 6.2. 2000-2003 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları. ...	54
Şekil 6.3. 2000-2007 Arası tenis ağından elde edilen AUC sonuçları.	55
Şekil 6.4. 2000-2007 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları. ...	56
Şekil 6.5. 2000-2014 Arası tenis ağından elde edilen AUC sonuçları.	57
Şekil 6.6. 2000-2014 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları. ...	58
Şekil 7.1. 2000-2003 Arası tenis ağından elde edilen AUC sonuçları.	61
Şekil 7.2. 2000-2003 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları. ...	62
Şekil 7.3. 2000-2007 Arası oluşturulan tenis ağından elde edilen AUC sonuçları. ...	64
Şekil 7.4. 2000-2007 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları. ...	65
Şekil 7.5. 2000-2014 Arası oluşturulan tenis ağından elde edilen AUC sonuçları. ...	66
Şekil 7.6. 2000-2014 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları. ...	67
Şekil 7.7. 2000-2003 Arası ağlarda NwCN yönteminin AUC sonuçları.	69
Şekil 7.8. 2000-2003 Arası ağlarda NwJI yönteminin AUC sonuçları.	70
Şekil 7.9. 2000-2003 Arası ağlarda NwAA yönteminin AUC sonuçları.	70
Şekil 7.10. 2000-2003 Arası ağlarda NwPA yönteminin AUC sonuçları.	71
Şekil 7.11. 2000-2003 Arası ağlarda NwSI yönteminin AUC sonuçları.	71
Şekil 7.12. 2000-2003 Arası ağlarda NwRA yönteminin AUC sonuçları.	72
Şekil 7.13. 2000-2007 Arası ağlarda NwCN yönteminin AUC sonuçları.	72
Şekil 7.14. 2000-2007 Arası ağlarda NwJI yönteminin AUC sonuçları.	73
Şekil 7.15. 2000-2007 Arası ağlarda NwAA yönteminin AUC sonuçları.	73

Sayfa

Şekil 7.16. 2000-2007 Arası ağlarda NwPA yönteminin AUC sonuçları.....	74
Şekil 7.17. 2000-2007 Arası ağlarda NwSI yönteminin AUC sonuçları.	74
Şekil 7.18. 2000-2007 Arası ağlarda NwRA yönteminin AUC sonuçları.	75
Şekil 7.19. 2000-2014 Arası ağlarda NwCN yönteminin AUC sonuçları.	75
Şekil 7.20. 2000-2014 Arası ağlarda NwJI yönteminin AUC sonuçları.....	76
Şekil 7.21. 2000-2014 Arası ağlarda NwAA yönteminin AUC sonuçları.	76
Şekil 7.22. 2000-2014 Arası ağlarda NwPA yönteminin AUC sonuçları.....	77
Şekil 7.23. 2000-2014 Arası ağlarda NwSI yönteminin AUC sonuçları.	77
Şekil 7.24. 2000-2014 Arası ağlarda NwRA yönteminin AUC sonuçları.	78

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 3.1. Karmaşıklık matrisi.....	25
Çizelge 5.1. Deneysel çalışmalarda kullanılan ağlarda düğüm ve bağlantı bilgileri.	35
Çizelge 5.2. Yıl bazında zaman periyoduna bağlı olarak oluşturulan ağlar.....	35
Çizelge 8.1. Zaman periyoduna bağlı olarak oluşturulan ağlar.	80
Çizelge 8.2. 2000-2003 Yılları düğüm merkeziliği sıralaması.	81
Çizelge 8.3. 2000-2003 Yılları ağırlıklandırılmış düğüm sıralaması.....	82
Çizelge 8.4. 2000-2010 Yılları düğüm merkeziliği sıralaması.	84
Çizelge 8.5. 2000-2010 Yılları ağırlıklandırılmış düğüm sıralaması.....	85
Çizelge 8.6. 2000-2017 Yılları düğüm merkeziliği sıralaması.	87
Çizelge 8.7. 2000-2017 Yılları ağırlıklandırılmış düğüm sıralaması.....	88

SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

G	: graf
V	: düğüm kümesi
E	: bağlantı kümesi
$(C_i)^D$: i düğümünün derece merkeziliği
$(C_i)^C$: i düğümünün yakınlık merkeziliği
$(C_k)^B$: i ve j düğümleri arasındaki n düğümünün arasındalık merkeziliği
$M(v)$: v düğümünün bağlı olduğu düğümler kümesi
λ	: özvektör katsayısı
$\alpha_{v,t}$: komşuluk matrisi
X_v	: düğümün özdeğeri
X_t	: düğümün özvektörü
$\Gamma(x)$: düğümünün komşularının kümesi
S_{xy}	: x ve y düğümlerinin benzerlik skoru
$ \Gamma(x) $: x düğümünün komşularının sayısı
$w(x, y)$: x ve y düğümleri arasındaki ağırlık
$w_n(z)$: z düğümünün ağırlığı

KISALTMALAR

- CN : Common Neighbor (Ortak Komşular)
- JI : Jaccard Index (Jaccard İndeks)
- SI : Sorenson Index (Sorenson İndeks)
- PA : Preferential Attachment Index (Tercihli Bağlantı İndeksi)
- AA : Adamic – Adar Index (Adamic – Adar İndeks)
- RA : Resource Allocation Index (Kaynak Paylaştırma İndeksi)
- HP : Hub Promoted Index (Hub Promoted İndeksi)
- HD : Hub Depressed Index (Hub Depressed İndeksi)
- LHN : Leicht-Holme-Newman Index (Leicht-Holme-Newman İndeksi)
- KI : Katz Index (Katz İndeksi)
- GLHN: Global Leicht-Holme-Newman Index (Global Leicht-Holme-Newman İndeksi)
- SR : SimRank Index (SimRank İndeksi)
- FI : FriendLink Index (FriendLink İndeksi)
- PFP : PropFlow Prediction Index (PropFlow Tahmin İndeksi)
- NWCN: Node Weigthed Common Neighbor (Düğüm Ağırlıklı Ortak Komşular)
- NWJI : Node Weigthed Jaccard Index (Düğüm Ağırlıklı Jaccard İndeks)
- NWSI : Node Weigthed Sorenson Index (Düğüm Ağırlıklı Sorenson İndeks)
- NWPA: Node Weigthed Preferential Attachment Index (Düğüm Ağırlıklı Tercihli Bağlantı İndeksi)
- NWAA: Node Weigthed Adamic – Adar Index (Düğüm Ağırlıklı Adamic – Adar İndeks)
- NWAA: Node Weigthed Resource Allocation Index (Düğüm Ağırlıklı Kaynak Paylaştırma İndeksi)

BÖLÜM 1

GİRİŞ

Ağ bilimi, birbiriyle dolaylı ya da doğrudan ilişki içerisinde olan varlıkların oluşturduğu karmaşık sistemleri anlamada ve modellemede farklı disiplinlerin bir araya geldiği alandır. Teknolojik gelişmelerle birlikte varlıklar arasındaki ilişkilerin ortaya çıkarılmasında, analiz edilmesinde, modellenmesinde ağ bilimi yaygın olarak kullanılmaya başlanmıştır [1]. Özellikle zaman içerisinde genişlemeye devam eden ağların oluşturulmaya başlanması ile birlikte ağları oluşturan varlıkları ve aralarındaki ilişkileri doğru ve anlaşılabilir şekilde analiz etmek büyük önem kazanmıştır [2]. Aslında karmaşık ağ bilimi, aralarında doğrudan ya da dolaylı bir şekilde ilişki olan her türlü sistemi belli kurallar ve disiplinler çerçevesinde analiz edebilmeye olanak sağlamaktadır [3]. Karmaşık ağ biliminin ilgilendiği önemli konulardan birisi de varlıklar arasında gelecekte olabilecek muhtemel ilişkileri ortaya çıkarmaktır [4]. Karmaşık ağlarda bağlantı tahmini problemi, ağdaki eksik bağlantıların tespit edilmesi ve ağın mevcut yapısından elde edilen bilgiler kullanılarak gelecekte oluşabilecek bağlantıların ortaya çıkarılmasına dayanmaktadır [5,6]. Son yıllarda iletişimdeki teknolojik gelişmelerle birlikte sosyal ağların yaygınlaşması, gelecekte oluşabilecek bağlantıların tahmin edilebilmesi için bağlantı tahmin çalışmalarına olan ilgiyi arttırmıştır [7]. Bağlantı tahmini günümüzde bir çok alanda kullanıcıların ihtiyacına yönelik tavsiye sistemi olarak çalışan filtreleme sistemlerinde [8,9], e-ticaret önerisi sistemlerinde [10], arkadaşlık öneri sistemlerinde [11,12], bibliyografik çalışmalarda [13], protein-protein etkileşim ağları ve metabolik ağlar gibi pek çok ağlarda [14,15] ve tıbbi çalışmalarda [16,17] yaygın olarak kullanılmaktadır. Ayrıca dinamik ağlarda, ağın gelecekte değişebilecek yapısını tahmin etmede ve modellemede bağlantı tahmini kullanılabilir [18–20]. Bağlantı tahminin önemli kullanım alanlarından birisi de ağdaki sahte bağlantıların tespit edilmesine yönelik yapılan çalışmalardır [21].

Ağda mevcut olmayan bağlantıların tahmininde ağın topolojik yapısı, düğümlerin özellikleri kullanılmaktadır. Burada önemli olan ağın analizinin iyi yapılması ve tahmin işleminde kullanılacak özelliklerin uygun şekilde tespit edilmesidir. Çünkü bir ağda gelecekte düğümler arasında yeni bağlantıların oluşma olasılığı kadar ağa sonradan yeni düğümlerin dahil olması veya düğümler arasındaki mevcut bağlantılarda kopmalar olması ihtimal dahilindedir. Özellikle dinamik ağlarda ağın yapısı zaman içerisinde değişebilmekte ve bu durum bağlantı tahmininin başarısını olumsuz etkilemektedir [12]. Bağlantı tahmininde ağın analizi yapılırken düğümler arasındaki bağlantıların oluşma şekli, bağlantıların türü ve ağırlıklarının ortaya çıkarılması kadar uygulanan tahmin işleminin de doğruluğu dikkate alınmalıdır [22]. Karmaşık ağlarda bağlantı tahmini için yapılan çalışmaların büyük bir çoğunluğu ağın topolojik özelliklerini analiz ederek tahmin yapan çalışmalardır [4]. Topolojik özellikleri kullanarak komşuluk tabanlı, yol tabanlı gibi düğümler arasında benzerlik ölçümü yapan yöntemler kullanılmaktadır [23–27].

Ağırlıksız ağlarda ortak komşuluk ilişkilerine dayalı ortak komşu sayısı, düğümler arasındaki yol mesafesi, Jaccard İndeksi, Adamic-Adar İndeksi, Sorensen İndeks, Salton İndeks, Tercihli Bağlanma ve Kaynak Paylaşırma İndeksi gibi yöntemler düğümler arasındaki benzerlik oranlarını hesaplayarak bağlantı tahmini yapmaktadırlar [24,28–33]. Bunların dışında ağırlıksız ağlarda düğümlerin derece bilgileri kullanılarak tahmin işlemi de yapılmıştır [34]. Önerilen bu yöntemler, ağın topolojik yapısını analiz ederek düğümlerin ortak komşularının özelliklerini kullanarak düğümler arasında benzerlik ölçümü yapılmasını sağlar [35].

Ağırlıklı ağlarda bağlantı tahmini, düğümler arasındaki benzerlik oranlarının hesaplanmasında düğüm çiftlerinin ortak komşularıyla aralarındaki ağırlık değerleri dikkate alınarak yapılmaktadır [36]. Düğümler arasındaki ağırlık değerlerinin tahmin işlemlerine dahil edilmesi bağlantı tahmininin başarısını artıran bir etken olarak kendini göstermektedir [37]. Ağırlıklı ağlarda bağlantı tahmini için önerilen komşuluk tabanlı yöntemler, ağırlıksız ağlar için önerilen yöntemlere bağlantı ağırlıklarının dahil edilmesiyle revize edilmiştir. Bu bağlamda Ağırlıklı Ortak Komşular, Ağırlıklı Adamic-Adar, Ağırlıklı Jaccard İndeks, Ağırlıklı Kaynak İndeksi, Ağırlıklı Sorenson İndex ve Ağırlıklı Salton İndex yöntemleri ağırlıklı ağlarda kullanılmak üzere önerilen

yöntemlerdir [37–39]. Bunların dışında, ağırlıklı ve yönlü ağlarda denetimli tahmin yapan [40], dinamik ağlarda zamansal değişimi dikkate alarak zamansal ağırlıklı tahmin yapan [41], düğümler arasındaki ağırlıkların öğrenme otomatası ile zamana bağlı olarak güncellenerek tahmin yapan [42], multipleks ağlarda özellik çıkarımına dayalı bağlantı tahmini yapan [43], benzerlik oranlarının hesaplanmasında düğümler arasındaki ağırlığı ve zaman periyodunu dahil ederek makine öğrenmesi ile tahmin yapan [44] yöntemler de ağırlıklı ağlarda uygulanmış ve bağlantı tahmin işlemlerinin başarısını önemli ölçüde arttırmışlardır.

Ortak komşuluk ilişkilerini inceleyerek bağlantı tahmini yapan yöntemlerin temel problemi, iki düğüm arasındaki ortak komşulara eşit davranması ve çoğunlukla sadece sayısal fazlalığı dikkate almasıdır. Ancak özellikle dinamik ağlarda zamanla etkinliği artan veya azalan düğümlerin ağına gelişimine olan etkisi tahmin işlemlerine yeterince yansıtılmamaktadır. Aslında ortak komşuların sayılarından ziyade ortak komşuların ağdaki gücü iki düğümün birbirine bağlanma olasılığını etkileyen önemli bir faktördür. Zaman periyoduna bağlı olarak, geçmişte ağda çok fazla bağlantısı olan bir düğüm zamanla bu etkinliğini kaybetmiş olabilir. Ancak bu durum geleneksel komşuluk tabanlı yöntemlerde dikkate alınmamaktadır. Bu nedenle, geçmişte ağda etkinliği fazla olan bir düğümün zaman içerisinde etkinliğini kaybetmiş olmasına rağmen başka düğümlerle bağlantı olasılığı geleneksel komşuluk tabanlı yöntemlerde yüksek çıkabilmektedir. Ayrıca düğümlerin zaman içerisinde ağına gelişimine olan etkilerinin bağlantı tahmin işlemlerinde dikkate alınmaması bağlantı tahmin başarısını etkilemektedir. Ağırlıklı ağlarda, düğümler arasındaki bağlantı ağırlıklarının tahmin işlemlerine dahil edilmesi her ne kadar başarıyı arttırmış olsa da ortak düğümlerin ağ içerisindeki etkinliği, devamlılığı ve gücü yeterince kullanılmamaktadır. Ayrıca, dinamik ağların zaman periyoduna bağlı olarak değişen yapıları, düğümlerin ağdaki konumlarında meydana gelen değişiklikler yeni bağlantıların oluşumunda önemli bir rol oynamaktadır.

Bu tez çalışmasında ortak komşuluk ilişkilerine dayalı bağlantı tahmin yöntemlerinin bahsedilen problemlerinin çözümüne yönelik olarak, düğümler arasındaki ortak komşuların ağdaki gücünü ve etkisini dikkate alan düğüm ağırlıklı bağlantı tahmin yöntemi geliştirilmiştir. Bu yöntemin temelini düğümlerin ağdaki gücünü hesaplayan

düğüm ağırlıklandırma işlemi ve elde edilen ağırlık değerlerinin komşuluk tabanlı yöntemlerde kullanılması oluşturmaktadır. Düğüm ağırlıklandırma işlemi, düğümler arasındaki bağlantı ağırlıklarından farklı olarak düğümlerin kendi ağırlıklarının çok kriterli karar verme yöntemi ile hesaplanması işlemidir. Bu yöntem, hem ağın topolojik bilgilerinin dışında ağın genişlemesine etkisi olan faktörlerin hem de ağ merkezilik ölçütlerinin ağırlıklandırma işleminde kullanılmasına olanak sağlamaktadır. Bu yöntem ile aralarında bağlantı olasılığı hesaplanacak düğümlerin ortak komşularının sayısından ziyade ortak komşuların ağırlıkları dikkate alınmaktadır. Bu sayede dinamik ağlar gibi zaman içerisinde genişlemeye devam eden ağlarda zaman faktörü dahil olmak üzere ağın genişlemesinde etkisi olan her türlü faktörün tahmin işlemlerine dahil edilmesi mümkün olmaktadır. Geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin başarısını değerlendirmek için tenis müsabakalarından ve ortak yazarlı bilimsel yayınlardan, farklı zaman periyotlarında ağlar oluşturulmuştur. Oluşturulan ağlar için belirlenen ağırlıklandırma kriterleri kullanılarak düğümlerin ağırlıkları hesaplanmıştır. Elde edilen ağırlık değerleri ile bağlantı tahmin işlemleri gerçekleştirilmiştir. AUC metriği kullanılarak elde edilen deneysel sonuçlar, geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin geleneksel komşuluk tabanlı bağlantı tahmin yöntemlerinden daha başarılı olduğunu göstermektedir.

Tez çalışmasının karmaşık ağlar ve bağlantı tahmini alanlarına katkıları şu şekilde özetlenebilir:

- Düğümlerin ağdaki gücünü tespit etmek için hem topolojik bilgiler hem de zaman faktörü gibi topolojik bilgilerle ifade edilemeyen faktörler kullanılabilir.
- İki düğüm arasındaki bağlantı ağırlığından farklı olarak düğümlerin ağdaki ağırlıkları hesaplanabilir.
- Düğümlerin ağırlıkları hesaplanırken zaman periyodu etkin bir şekilde kullanılmaktadır.
- Hesaplanan düğüm ağırlıkları kullanılarak gelecekte ağda etkin olabilecek düğümlerin tahmin edilmesi mümkün olmaktadır.
- Ağda merkezi düğümlerin tespitinde elde edilen düğüm ağırlıkları kullanılarak zaman içerisinde ağdaki merkeziliğin değişimi analiz edilebilir.

- İki düğüm arasındaki bağlantı olasılığını hesaplarken ortak komşuların ağırlıkları kullanılabilir.

Tez çalışmasının ikinci bölümünde, karmaşık ağ bilimi daha geniş bir şekilde ele alınmaktadır. Karmaşık ağ biliminin uygulama alanlarından, karmaşık ağ analizine dayalı çalışmalardan bahsedilmekte, graf teorisi ve karmaşık ağ analizinde kullanılan merkezilik ölçütleri açıklanmaktadır.

Üçüncü bölümde, bağlantı tahmini problemi ele alınmaktadır. Bağlantı tahmin yöntemlerinden genel olarak bahsedilmekte, mevcut komşuluk tabanlı yöntemler detaylı bir şekilde açıklanmaktadır. Ayrıca bağlantı tahmin başarısını ölçmede kullanılan değerlendirme ölçütleri de anlatılmaktadır.

Dördüncü bölümde, çok kriterli karar verme analizinden bahsedilerek özellikle, tez çalışmasında kullanılan Logaritmik Konsept Yaklaşımı (APLOCO) yöntemi detaylı bir şekilde açıklanmaktadır. APLOCO yöntemi tez çalışmasında düğümlerin ağırlıklandırılması aşamasında kullanılmaktadır.

Beşinci bölümde, geliştirilen yöntemin performansını değerlendirmek için yapılan deneysel çalışmalarda kullanılan veri setleri, zaman periyoduna bağlı ağların oluşturulması, ağırlıklandırma kriterlerinin belirlenmesi aşamaları anlatılmaktadır.

Altıncı bölümde, geliştirilen düğüm ağırlıklı bağlantı yöntemi anlatılmaktadır. Yöntemin uygulama adımları, düğüm ağırlıklandırma işlemi, düğüm ağırlıklı bağlantı tahmin işlemi, geliştirilen yöntemle ait matematiksel modeller ve geliştirilen düğüm ağırlıklı bağlantı tahmini yöntemi başarısı geleneksel komşuluk tabanlı yöntemler ile karşılaştırılmakta, elde edilen sonuçlar gösterilmekte ve detaylı bir şekilde açıklanmaktadır.

Yedinci bölümde, düğüm ağırlıklı bağlantı tahmin yönteminde merkezilik ölçütleri ve zaman faktörünün birlikte kullanımına dair gerçekleştirilen deneysel çalışma anlatılmakta, yöntemlerin başarısı geleneksel komşuluk tabanlı yöntemler ile karşılaştırılmakta, elde edilen sonuçlar gösterilmekte ve yorumlanmaktadır.

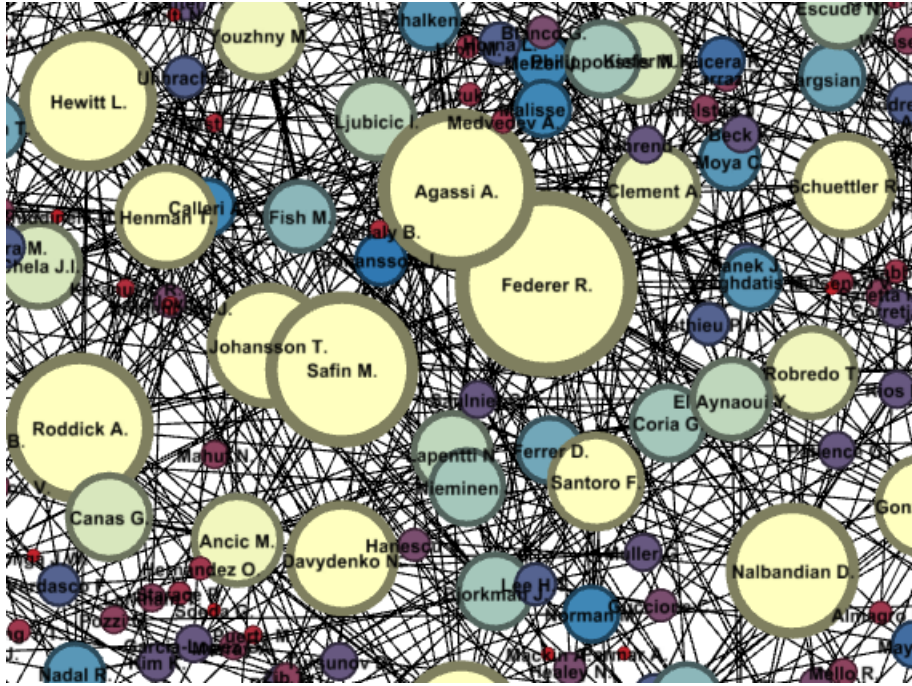
Sekizinci bölümde, düğüm ağırlıklarının karmaşık ağlarda merkezilik ölçütü olarak kullanılması üzerine gerçekleştirilen deneysel çalışma anlatılmakta, sonuçları gösterilmekte ve yorumlanmaktadır.

Dokuzuncu bölümde, tez çalışmasının amacına uygun olarak geliştirilen yöntem, gerçekleştirilen deneysel çalışmaların sonuçları ışığında genel olarak değerlendirilmekte ve nihai hedefin sonuçları vurgulanmaktadır.

BÖLÜM 2

KARMAŞIK AĞLAR

Ağlar, günümüzde karmaşık sistemlerin içerdiği unsurların birbirleriyle olan ilişkilerini graf yapıları ile temsil etmede kullanılmaktadır [45]. Şekil 2.1’de gösterildiği gibi bir ağı oluşturan unsurlar köşeler ve kenarlardır. Köşeler varlıkları temsil ederken kenarlar varlıklar arasındaki ilişkilerdir.



Şekil 2.1.Tenis turnuvaları ağı [1].

Karmaşık ağlar, içerdiği köşe ve bağlantı sayıları bakımından oldukça büyük ağlardır. Karmaşık ağlar gerçek dünya ağlarının modellenmesi olarak da ifade edilebilir [46]. Karmaşık ağların yapısında köşeler düğümleri, kenarlar ise bağlantıları ifade etmektedir [47]. Gerçek dünya ağlarının graflarla gösterimi yaygın olarak kullanılmaktadır. İletişim teknolojisindeki gelişmeler ile özellikle sosyal ağların yaygınlaşması karmaşık ağların kullanımının önemini arttırmıştır.

Karmaşık ağ tanımı aslında, iki şekilde ifade edilebilir. Birincisi modellenen ağın çok büyük olmasından dolayı karmaşık ağ olarak ifade edilmesidir. Bir diğeri ise ağdaki varlıklar arasındaki ilişkilerin topolojik olarak karmaşık olmasından dolayı karmaşık ağ olarak ifade edilmesidir [48]. Dilsel birliktelikleri ifade eden ağlar, elektrik şebekelerini ifade eden ağlar, sinir ağları, bilimsel işbirliği ağları, spor ağları, ulaşım ağları, internet ağı, arkadaşlık ağları [49–57], gibi pek çok gerçek dünya ağı karmaşık ağların kullanıldığı alanlara örnek gösterilebilir.

Karmaşık ağlar yapısal olarak genellikle dinamiktir. Örneğin arkadaşlık ağları, bilimsel işbirliği ağları, spor ağları, ulaşım ağları, internet ağı gibi karmaşık ağlar zamanla genişlemeye devam ederler. Ağların sürekli genişlemeye devam eden dinamik yapısından dolayı bu ağların analizlerinde Giriş bölümünde bahsedildiği gibi çeşitli zorluklar ortaya çıkmaktadır.

2.1. KARMAŞIK AĞ ANALİZİ

Karmaşıklık teorisi dünyadaki herşeyin birbiriyle dolaylı ya da doğrudan bağlantılı olduğu temeline dayanmaktadır. Bu teoriye göre aslında dünyadaki herşey bir ağın parçasıdır [48]. Karmaşık ağ bilimi, ağı oluşturan varlıkları, varlıklar arasındaki ilişkileri ve ağdaki veri akışını belli kurallar ve disiplinler çerçevesinde analiz edebilmeye imkan sağlamaktadır [58]. Günümüzde karmaşık ağ analizi çalışmaları yaygın olarak sosyal ağların analizi üzerine yoğunlaşmış olsa da hemen hemen her alandan ilgi görmektedir [59]. Karmaşık ağ analizi teknikleri ilk başlarda sosyal ağ analizi üzerinde kullanılırken günümüzde spor, ulaşım, eğitim, bankacılık, sosyoloji, biyoloji, gibi daha karmaşık yapıya sahip alanlarda da uygulanmaktadır [60–64].

Gerçek dünya ağlarının incelenmesi ile ilgili tarihsel sürece bakıldığında, İsviçreli matematikçi Leonhard Euler'in 1736'da çizge teorisiyle çözdüğü Königsberg Köprüsü problemi bu alanda bilinen ilk çalışma olarak dikkat çekmektedir [65]. 1950'den sonra Erdős–Rényi tarafından önerilen rassal ağlar bu alanda çalışmaların yaygınlaşmasında önemli rol oynamıştır [66]. Rassal ağlar, ağların ortak özelliklerini ve olasılık dağılımlarını göstermesine yönelik çözümleriyle yeni araştırmaların önünü açmıştır [67–71]. Özellikle bilgisayar ve iletişim teknolojisinin hızla gelişmesiyle daha büyük

ve karmaşık gerçek dünya ağlarının analizi ilgi gören araştırma konuları arasında yerini almıştır. Karmaşık ağlarla ilgili çalışmalar gerçek dünya ağlarının ortak özelliklerinin ortaya çıkarılmasına katkı sağlamıştır [72]. Elde edilen bu özellikler sonraki dönemlerde geliştirilerek, karmaşık yapıları anlamamız için yapılan çalışmalara katkı sunmuştur [73].

Hesaplamalı ağ analizi çalışmalarına bakıldığında, temel olarak karmaşık ağların istatistiksel analizi [74], topluluk algılama ve düğümlerin sınıflandırılması [75–77], ağ dinamiklerinin zaman içerisindeki değişimi [78–81], bilgi yayılımı [82–85], ağlar üzerinde veri madenciliği [86–88] ve ağların görselleştirilmesi [89–93] üzerine yoğunluk kazandığı görülmektedir.

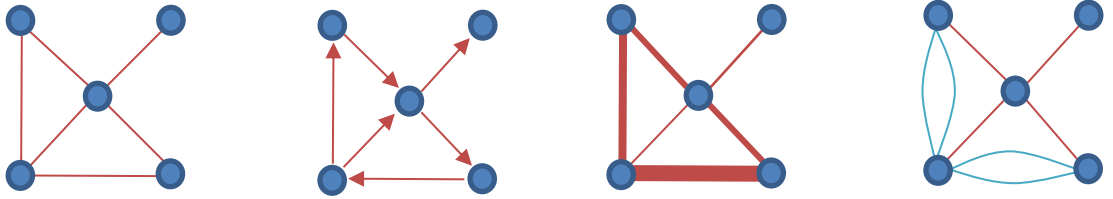
2.2. ÇİZGE (GRAF)

Karmaşık ağların analizinde en önemli kısım yapısal analizdir. Çünkü yapısal analiz ağın oluşumu ile ilgili en önemli kanıtları içermektedir. Ağların yapısal analizinde çizge teorisinden yararlanılmaktadır. Çizge teorisinin temelinde sosyoloji ve matematik kökenli çalışmalar yer alsa da günümüzde telekomünikasyon, bilişim, biyoloji gibi pek çok alanın ağ analizinde çizge teorisi önemli bir yere sahiptir [46]. Ağların gösteriminde kullanılan çizgeler, düğümlerin ve düğümler arasındaki bağlantıların kolay anlaşılabilir şekilde ifade edilmesini sağlar. Özellikle anlaşılması zor olan ve matematiksel olarak ifade edilen ağların çizge olarak gösterilmesi anlaşılabilirliği kolaylaştırmaktadır.

$$G = (V, E) \quad (2.1)$$

Eşitlik 2.1’de G Grafi, V düğüm kümesini ve E bağlantı kümesini ifade etmektedir. Burada modellenecek varlıklar (Vertex) (V), varlıklar arasındaki ilişkiler ise (Edges) (E) olarak temsil edilir [46]. Örneğin bir tenis ağında sporcular düğümleri, sporcular arasındaki müsabakalar bağlantıları temsil eder. Ya da bir bilimsel iş birliği ağında yazarlar düğümleri ifade ederken yazarlar arasındaki ortak yayınlar bağlantıları ifade eder [94].

Ağları graflarla temsil ederken, düğümlerin ve düğümler arasındaki bağlantıların konumları ya da grafiksel boyutları önemli değildir. Burada önemli olan düğümler arasındaki bağlantıların doğru bir şekilde tanımlanmasıdır. Şekilde 2.2’de olduğu gibi bağlantıları gösterirken farklı ifade şekilleri vardır. Düğümler arasındaki bağlantılar oluşurken yön bilgisi yok ise oluşan bağlantı şekli yönsüzdür. Örneğin iki sporcu arasındaki müsabakayı temsil eden bir bağlantı yönsüzdür. Eğer bağlantı oluşurken bir düğümden diğer düğüme bağlantı yapılmışsa yönlü bağlantıdır. Örneğin bir iletişim ağında a kişisinden b kişisine yapılan aramayı temsil eden bağlantı türü yönlü bağlantıdır [95]. Bir başka bağlantı gösterim türü de ağırlıklı bağlantılardır. Düğümler arasındaki bağlantının gücü ifade edilmek isteniyorsa ağırlıklı çizge ile bağlantı gösterilir [96]. Örneğin şehirlerarası ulaşım ağında iki şehir arasındaki yolun uzunluğu ağırlık olarak kabul edilebilir ve ağırlıklı bağlantı ile ifade edilebilir. Düğümler arasında ilişkinin türü birden fazla farklı olduğu durumlarda bağlantıyı çoklu bağlantı ile göstermek mümkündür.



a) Yönsüz bağlantı b) Yönlü Bağlantı c) Ağırlıklı Bağlantı d) Çoklu Bağlantı

Şekil 2.2. Farklı bağlantı türleri.

Yönlü bağlantılar içeren ağlarda bağlantıların yönü tek taraflı ya da çift taraflı olabilir. Burada önemli olan ilişki türüne göre bağlantı yönünün doğru bir şekilde ifade edilmesidir.

2.3. KARMAŞIK AĞ ANALİZİNDE TEMEL MERKEZİLİK ÖLÇÜTLERİ

2.3.1. Derece Merkeziliği

Ağda merkezi düğümleri ortaya çıkarmak için kullanılan ve en basit ölçüt olan derece merkeziliği, ağdaki bir düğüme gerçekleşen bağlantı sayısı ile hesaplanır.

Hesaplanması basit olmasına rağmen, düğümün ağdaki konumunu gösterebilen önemli bir ölçüttür. Çoğu karmaşık ağda, bir düğüm ne kadar çok bağlantıya sahipse o kadar önemli ve güçlüdür. Aslında, en yüksek dereceye sahip düğüm, ağın en aktif üyesi olarak yorumlanabilir. Bağlantıların yönlü olduğu ağlarda, derece içi bağlantıların sayısı ve derece dışı bağlantıların sayısı ayrı ayrı hesaplanır [97].

$$(C_i)^D = \sum_{j=1}^n a_{ij} \quad (2.2)$$

Burada i , derece merkeziliği hesaplanacak düğüm, j ağdaki diğer düğümlerdir. Eğer i düğümü, komşu matrisine göre herhangi bir j düğümüne bağlıysa, o zaman $a_{ij} = 1$, aksi takdirde $a_{ij} = 0$ 'dır. Bu şekilde, i düğümün derecesi, bağlı olduğu düğümlerin toplamı olarak ifade edilir.

2.3.2. Yakınlık Merkeziliği

Ağda merkezi düğümleri ortaya çıkarmak için kullanılan başka bir merkezilik ölçütü olan yakınlık merkeziliği, ağdaki herhangi bir düğümün, diğer tüm düğümlere olan en kısa yol uzunluklarının (jeodezik mesafe) ortalaması ile hesaplanır. Yakınlık merkeziliği, düğümün ağdaki diğer düğümlere olan yakınlığını ölçen bir metriktir. Ağdaki düğümlerin tamamına en yakın olan düğüm, ağda bilginin en hızlı aktığı düğümdür. Eğer ağdaki bağlantılar yönlü ise, en kısa yolların ortalaması hesaplanırken bu hususlar dikkate alınmalıdır [98].

$$(C_i)^C = (l_i)^{-1} = \frac{n-1}{\sum d_{ij}} \quad (2.3)$$

Burada d_{ij} , i 'den j 'ye en kısa mesafeyi gösterir. Bu mesafe i ve j düğümleri arasındaki en kısa yolun uzunluğu olarak tanımlanır. l_i , yakınlık merkeziliği hesaplanacak düğümün diğer düğümlere ortalama mesafesidir.

2.3.3. Arasındalık Merkeziliği

Ağdaki merkezi düğümleri belirlemede kullanılan ölçütlerden birisi olan arasındalık merkeziliği, hesaplaması karmaşık gibi görünse de bir düğümden kaç defa geçildiği bilgisini vermesi açısından önemli bir ölçüttür. Özellikle ağdaki toplulukların kesiştiği noktada bulunan düğümlerin tespit edilmesinde etkili bir merkezilik ölçütüdür. Bu merkezilik ölçütüne göre, üzerinden en fazla geçilen düğüm ağdaki en merkezi düğüm bilgisini vermektedir [97].

$$(C_k)^B = \sum_{i=1}^n \sum_{j=1}^n \frac{G_{ikj}}{G_{ij}}, i \neq j \neq k \quad (2.4)$$

Burada C_k , i ve j düğümleri arasında bulunan n düğümünün arasındalık merkeziliğini temsil ederken, G_{ij} , i düğümünden j düğümüne en kısa yolların sayısını temsil eder. G_{ikj} , i ile j düğümleri arasındaki ve ayrıca k düğümü üzerinden geçen en kısa yolların sayısı olarak kabul edilir.

2.3.4. Özvektör Merkeziliği

Ağ merkeziliği ölçümünde kullanılan bir diğer ölçüt olan özvektör merkeziliği, bir düğümün bağlı olduğu düğümlerin sayısından ziyade bağlı olduğu düğümlerin kalitesini dikkate almaktadır. Özvektör merkeziliğine göre ağ üzerindeki bir düğüm için anahtar düğümlere olan bağlantıların etkisi, diğer herhangi bir sıradan bağlantıdan daha fazla olabilir. Bir düğümün bağlı olduğu düğümlerin daha merkezi olması, düğümün daha merkezi bir konumda olacağını gösterir. Bu ölçüyü hesaplarken, komşuların merkeziliğinin toplamı hesaba katılır [99].

$$X_v = \frac{1}{\lambda} \sum_{t \in M(v)} X_t \quad X_t = \frac{1}{\lambda} \sum_{t \in G} \alpha_{v,t} X_t \quad (2.5)$$

Burada, $M(v)$, v düğümünün bağlı olduğu düğümler kümesini, G ağdaki tüm düğümlerin kümesini, λ özvektör katsayısını, $\alpha_{v,t}$ komşu matrisi, X_v düğümünün özdeğerini ve X_t v düğümünün özvektörünü ifade eder.

BÖLÜM 3

BAĞLANTI TAHMİNİ

Gerçek dünyada varlıklar arasındaki birlikteliklerin çeşitlenmesi, artması ve bununla beraber varlıklar arasındaki birlikteliklerin ortaya çıkarılması çok büyük yapıda ağların oluşmasına neden olurken oluşan bu ağlardan elde edilen kullanışlı veriler, karmaşık ağ analizi için de farklı amaçlarla çalışmalar yapılmasının önünü açmıştır. Karmaşık ağlar, varlıkların farklı amaçlarla bir araya gelmesiyle ya da getirilmesiyle oluşturdukları ve günümüzde sosyal ağların yaygınlaşması ile birlikte bilgi akışının çok yoğun olduğu ağlardır. Dinamik yapıya sahip karmaşık ağlar göz önüne alındığında varlıklar arasındaki bağlantıların zaman içerisinde kopması ya da yeni bağlantıların oluşması mümkün olmaktadır. Karmaşık ağ analizinde varlıkların gelecekte ağdaki ilişkilerinin tahmin edilmesi önemli bir problem olarak görünmektedir. Bu noktada bağlantı tahmininin görevi, karmaşık bir ağda gelecekte ortaya çıkacak yeni bağlantıları ya da ağda mevcut olan gizli bağlantıları tespit etmek olarak tanımlanabilir [94].

Bağlantı tahmini, varlıkları arasındaki ilişkileri verilen bir ağın, mevcut graf yapısına bakarak varlıklar arasında gelecekte meydana gelebilecek yeni bağlantıları tespit etmektedir. Bu bağlamda bağlantı tahmin modelleri, ağın topolojik yapısından elde edilen bilgileri düğümlerin öznitelikleriyle birleştirerek düğümler arasında benzerlik ölçümüne dayalı tahminler yapmaktadırlar. Tüm modeller düğümler arasındaki yüksek benzerliğin gelecekte daha yüksek bir bağlantı olasılığı ile sonuçlandığı hipotezine dayanmaktadırlar [100].

Bağlantı tahmin çalışmalarının çok çeşitli uygulamaları mevcuttur. Sosyal ağlarda, kullanıcı-kullanıcı ya da kullanıcı-içerik önerilerinin analizi [101–103], biyoinformatik alanında, PPI (protein protein etkileşimi) ağının yeniden yapılandırılması ve mevcut gürültünün azaltılması [104–106], hiper bağlantı tahmini

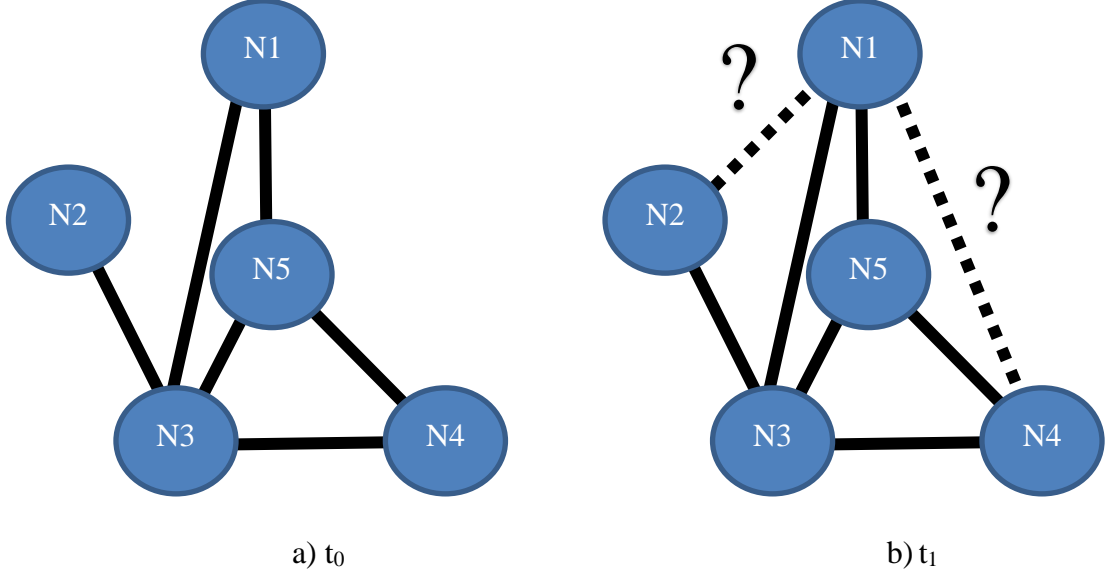
[107], ulaşım ağlarının tahmini [108], terör kampanyalarının ve faaliyetlerinin takibi [109] gibi farklı kullanım alanları örnek verilebilir.

3.1. BAĞLANTI TAHMİNİ PROBLEMİ

Bağlantı tahmininin en klasik tanımı, karmaşık bir ağın t_0 zamanındaki anlık yapısı göz önüne alınarak, t_0 zamanından t_1 zamanına kadar geçecek sürede ağda oluşacak yeni bağlantıların tahmin edilmesi işlemidir [4]. Gelecekte oluşabilecek yeni bağlantıların tahmin edilmesi problemi, ağın anlık topolojik yapısının analiz edilmesine veya zaman periyoduna bağlı olarak ağdaki bağlantı değişimlerinin analizine dayanmaktadır. Her iki durum için de ağın yapısının iyi analiz edilmesi gerekmektedir. Çünkü bağlantı tahmin probleminin çözümü için ağın yapısından elde edilen bilgiler kullanılmaktadır. Ağdan ne kadar çok bilgi elde edilir ve bağlantı tahmin sürecine dahil edilirse tahmin sonuçları o kadar başarılı olur [18].

Matematiksel olarak, anlık t_0 zamanında düğümleri ve bağlantıları $G = (V, E)$ şeklinde ifade edilen karmaşık bir ağda bağlantı tahmini, $t_0 \longrightarrow t_1$ aralığında bağlantılar kümesinde (E) meydana gelebilecek değişimi tahmin etmeyi amaçlar. Bu değişim, Şekil 3.1'de gösterildiği gibi ağda yeni bağlantıların oluşmasından, mevcut bağlantılarda kopmalar olmasından, gizli bağlantıların ortaya çıkmasından kaynaklı olabilir. t_0 anında, düğümler arasında N1-N3, N1-N4, N1-N5, N2-N3, N3-N4, N3-N5 ve N4-N5 şeklinde bağlantılar bulunmaktadır. Bu noktada, bağlantı tahmin problemine yönelik geliştirilen yöntemlerin amacı $t_0 \longrightarrow t_1$ aralığında geçen sürede aralarında bağlantı olmayan düğümlerden hangilerinin birbiriyle bağlantı kurabileceğini doğru bir şekilde tahmin etmektir. Neticede $t_0 \longrightarrow t_1$ arasındaki zaman periyodunda aralarında bağlantı olmayan bütün düğümler de birbiriyle bağlantı kurabilir ya da hiçbir düğüm de birbiriyle bağlantı kurmayabilir. Bağlantı tahmin probleminin çözümünü zorlaştıran kısım ise işte bu belirsizliktir. Özellikle sosyal ağlarda insanların birbiriyle bağlantı kurmasının belli bir periyodu, kuralı ve zorunluluğu yoktur. Bağlantılar tamamen tercihe bağlı olarak gerçekleştiği için bağlantı tahminin çözümü de zorlaşmaktadır. Bu noktada bağlantı tahmin probleminin çözümü için t_0 anında ağın yapısal özelliklerinin ve düğümlerin özniteliklerinin iyi bir şekilde analiz edilmesi çok

önemlidir. Ağın yapısına dair düğümleri birbirinden ayırt edebilecek ne kadar çok bilgi elde edilebilirse bağlantı tahmin işlemlerinin başarısı da o kadar artar.



Şekil 3.1. Ağda yeni bağlantıların oluşumu.

3.2. BAĞLANTI TAHMİNİ YÖNTEMLERİ

Bağlantı tahmini probleminin çözümüne yönelik son yıllarda gerçekleştirilen çalışmalar bu problemin karmaşık ağların analizinde önemli bir konu olduğunu göstermektedir [110]. Bu konuda yapılan ilk çalışmalara bakıldığında, geleneksel veri madenciliği yöntemlerinin uygulandığı görülmektedir [111,112]. Daha sonraki dönemlerde, ağ yapısından elde edilen bilgilerin kullanılmasıyla birlikte graf tabanlı modeller [113] ve yerel ağ bilgisini kullanan modeller kullanılmaya başlanmıştır [29]. Düğümlerden elde edilen öznitelik bilgilerinin ağın yapısal özellikleriyle birlikte kullanılmasıyla daha başarılı sonuçlar veren yöntemler üzerinde çalışmalar yoğunlaşmıştır [114]. Bağlantı tahmini yöntemleri, ağın yapısal özelliklerinin türüne ve düğüm özniteliklerin durumuna göre farklı başlıklar altında kategorize edilmektedir.

3.2.1. Benzerlik Tabanlı Yaklaşımlar

Benzerlik tabanlı yaklaşımlar, ağın topolojik yapısına odaklanan en eski ve uygulaması en basit bağlantı tahmin yöntemleridir. Bu yaklaşıma dayalı önerilen yöntemler, ağın yapısal özelliklerini kullanarak düğümler arasındaki benzerlik oranını hesaplayarak gelecekte oluşabilecek bağlantıları tahmin etmeye çalışırlar. Benzerlik tabanlı yaklaşımlar üç ana kategori altında incelenebilir. Bunlar, yerel benzerlik tabanlı yaklaşımlar, yarı yerel benzerlik tabanlı yaklaşımlar ve yol tabanlı yaklaşımlardır [110]. Benzerlik tabanlı bağlantı tahmin yaklaşımlarında kullanılan bazı ölçütlerin ne ifade ettiğine bakılacak olursa;

- $\Gamma(x)$, karmaşık ağda x düğümünün komşularının kümesi.
- S_{xy} , karmaşık ağda x ve y düğümlerin benzerlik skoru.
- $|\Gamma(x)|$, karmaşık ağda x düğümünün derecesi (komşularının sayısı).
- $w(x, y)$, karmaşık ağda x ve y düğümlerinin arasındaki ağırlık.

3.2.1.1. Yerel Benzerlik Tabanlı Yaklaşımlar

Yerel benzerliğe dayalı yaklaşımlar, düğüm çiftlerinin ortak komşularının yapısal özelliklerini kullanarak gelecekte bir bağlantı oluşturup oluşturmayacakları varsayımına dayanmaktadır. Ortak komşuluk ilişkileri için önerilen yöntemler, tüm ağ topolojisini dikkate almak yerine sadece komşulukla ilgili bilgileri kullandıklarından dolayı hızlı sonuç verebilmektedirler [4,110]. Sadece komşuluk ilişkilerine dair bilgilerin dahil edilmesi başarı ölçütlerini etkileyen olumsuz yönlerden birisidir.

Ortak Komşular (CN)

CN, bağlantı tahmini için en basit yöntemlerden birisi olmasına rağmen tahmin başarısındaki verimliliği nedeniyle en çok kullanılan yöntemlerden birisidir. CN'nin bağlantı tahminindeki yaklaşımı, iki düğümün ortak komşularının fazlalığı üzerinedir. Yani CN'ye göre gelecekte iki düğümün bağlanma olasılığı, ortak komşularının sayısıyla doğru orantılıdır. İki düğüm arasında ne kadar fazla ortak komşu var ise bu iki düğümün gelecekte bağlanma olasılığı o kadar yüksektir [115].

$$S_{xy} = |\Gamma(x) \cap \Gamma(y)| \quad (3.1)$$

CN'nin ağırlıklı ağlar için önerilen yaklaşımında iki düğümün bağlanma olasılığı, ortak komşuların düğümlerle aralarındaki ağırlıkların toplamına dayalı olarak hesaplanmaktadır. Yine burada da iki düğüm arasındaki ortak komşularla ağırlıkların toplamı ne kadar fazla ise gelecekte düğümlerin bağlanma ihtimali de o kadar yüksek olur sonucu çıkarılmaktadır [37].

$$S_{xy}^{WCN} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} w(x, z) + w(y, z) \quad (3.2)$$

Jaccard İndeks (JI)

Veri madenciliğinde de kullanılan JI ölçütü, bağlantı tahmininde de başarılı bir şekilde uygulanan yöntemlerden birisidir. JI ölçütü, ortak komşuların sayısını kullanmakla birlikte tüm komşuların toplam sayısını dikkate alarak normalleştirme işlemi yapar ve iki düğüm arasındaki gelecekte bağlanma olasılığını hesaplar [30].

$$S_{xy} = \frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x) \cup \Gamma(y)|} \quad (3.3)$$

JI'nın ağırlıklı ağlar için önerilen yaklaşımında ise iki düğümün bağlanma olasılığı, ortak komşuların düğümlerle aralarındaki ağırlıkları toplamının tüm komşuların ağırlıkları toplamına oranı ile hesaplanmaktadır [39].

$$S_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{w(x, z) + w(y, z)}{\sum_{a \in \Gamma(x)} w(a, x) + \sum_{a \in \Gamma(y)} w(b, y)} \quad (3.4)$$

Sorenson İndeks (SI)

JI'ye benzer bir ölçüt olan SI, iki düğüm arasındaki ortak komşuların sayısına önem vermekle birlikte, düğümlerin dereceleri toplamını da dikkate almaktadır. Burada

amaç, dereceleri düşük düğümlerin gelecekte bağlantı kurma olasılığının daha yüksek olabileceği varsayımdır [32].

$$S_{xy} = \frac{2|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x)| + |\Gamma(y)|} \quad (3.5)$$

SI'nın ağırlıklı ağlar için önerilen yaklaşımında ise JI'ye benzer bir şekilde hesaplanmaktadır. Hesaplamadaki tek fark ortak komşuların düğümlerle aralarındaki ağırlıkları toplamının iki katının hesaplama işlemine dahil edilmesidir [39].

$$S_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{2(w(x, z) + w(y, z))}{\sum_{a \in \Gamma(x)} w(a, x) + \sum_{b \in \Gamma(y)} w(b, y)} \quad (3.6)$$

Tercihli Bağlantı İndeksi (PA)

Genellikle ölçeksiz büyüyen ağlarda tercih edilen PA ölçütü, ağa yeni dahil olan bir düğümün yüksek dereceli düğümlere bağlanma olasılığının daha kuvvetli olduğunu varsaymaktadır. Bir düğümün derecesi ne kadar yüksek ise gelecekte bağlantılarını artırma olasılığı da o kadar yüksektir. PA ölçütüne göre iki düğüm arasında bağlantı oluşma ihtimalini düğümlerin ağdaki komşu sayıları ile doğru orantılıdır [31].

$$S_{xy} = \Gamma(x) * \Gamma(y) \quad (3.7)$$

PA'nın ağırlıklı ağlar için önerilen yaklaşımında ise düğümlerin komşu sayılarından farklı olarak komşuların ağırlıklı toplamı dikkate alınmaktadır. Düğümlerin komşu sayıları kadar komşularıyla aralarındaki ağırlıkların kuvveti de başka düğümlere bağlanma olasılığını doğrudan etkilemektedir [39].

$$S_{xy} = \sum_{a \in \Gamma(x)} w(a, x) * \sum_{b \in \Gamma(y)} w(b, y) \quad (3.8)$$

Adamic – Adar İndeks (AA)

AA ölçütü, ağda bağlantı sayısı düşük ortak komşulara önem vererek bağlantı olasılığını hesaplamaktadır. $\Gamma(x)$ ve $\Gamma(y)$, ortak komşusu z düğümünün kaç düğümlle bağlantısının olduğu, yani z düğümünün derecesi bağlantı olasılığında dikkate alınır [28].

$$S_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{\log (|\Gamma(z)|)} \quad (3.9)$$

Ağırlıklı ağlarda AA ölçütü, Jaccard ölçütünün yeniden formüle edilmesiyle elde edilmiştir [37]. Ağırlıklı ağlar için önerilen AA ölçütüne göre düğüm çiftlerinin aralarında özel ya da güçlü bir ortak komşuya sahip olmak, gelecekte bu düğüm çiftlerinin birbirine bağlanma olasılığını artırmaktadır.

$$S_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{w(x, z) + w(y, z)}{\log \log (1 + \sum_{c \in \Gamma(z)} w(z, c))} \quad (3.10)$$

Kaynak Paylaştırma İndeksi (RA)

Karmaşık ağlarda yaygın olarak kullanılan RA ölçütü, birbiriyle doğrudan bağlantısı olmayan düğüm çiftleri arasındaki bağlantı gücünü ölçer. Yani birbiri ile bağlantısı olmadığı halde ortak komşular üzerinden iletişim halinde olan düğüm çiftlerinin benzerliklerini hesaplar. Bu düğüm çiftleri arasındaki benzerlik, birbirlerinden aldıkları kaynaklara göre hesaplanır [29]. Eşitlik 3.11’de k_z , $\Gamma(x)$ ve $\Gamma(y)$ ‘nin ortak komşusu olan z ’nin komşularının sayısını ifade etmektedir

$$S_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{k_z} \quad (3.11)$$

Hub Promoted İndeksi (HP)

HP ölçütü, metabolik ağlarda birleşen veya ayrışan maddelerin benzerliğini ölçmek için önerilen bir yöntemdir [116].

$$S_{xy}^{HP} = \frac{|\Gamma(x) \cap \Gamma(y)|}{\min(|\Gamma(x)|, |\Gamma(y)|)} \quad (3.12)$$

Eşitlik 3.12'de gösterildiği gibi HP, hem $\Gamma(x)$ hem de $\Gamma(y)$ 'nin ortak komşularının $\Gamma(x)$ ve $\Gamma(y)$ 'nin minimum derecelerine oranıyla hesaplanır. Burada, derecesi az olan düğümler ile merkezi düğümlerin bağlantı oluşumu desteklenmektedir [100].

Hub Depressed İndeksi (HD)

HD ölçütü, HP ölçütünün tam aksine ortak komşuların düğüm çiftlerinin maksimum derecelerine oranıyla hesaplanır. Burada, yüksek dereceli düğümlerin merkezi düğümlerle bağlantı olasılığı desteklenmektedir [5].

$$S_{xy}^{HD} = \frac{|\Gamma(x) \cap \Gamma(y)|}{\max(|\Gamma(x)|, |\Gamma(y)|)} \quad (3.13)$$

Leicht-Holme-Newman İndeksi (LHN)

SI ölçütüne benzerlik gösteren LHN ölçütü, $\Gamma(x)$ ve $\Gamma(y)$ 'nin ortak komşularının bu düğüm çiftinin aralarındaki bağlantı sayısının beklenen değerine oranıyla hesaplanmaktadır. Bu hesaplama işlemi Eşitlik 3.14'de gösterildiği gibi $\Gamma(x)$ ve $\Gamma(y)$ 'nin ortak komşularının sayısının $\Gamma(x)$ ve $\Gamma(y)$ 'nin komşularının sayısının çarpımına oranı ile gerçekleşmektedir [117].

$$S_{xy}^{LHN} = \frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x)| \cdot |\Gamma(y)|} \quad (3.14)$$

3.2.1.2 Yol Tabanlı Yaklaşımlar

Yol tabanlı yaklaşımlar, komşuluk tabanlı yaklaşımların aksine düğüm çiftleri arasındaki benzerliği sıralamak için ağın tüm topolojisini kullanır. Bu neden sadece ortak komşuları olan düğümlerin benzerliğini ölçmekle sınırlı değildirler. Ağın tüm topolojisinin dikkate alınması bağlantı tahmini için esneklik sağlasa da zaman karmaşıklığını artıran bir duruma yol açmaktadır. Yol tabanlı yaklaşımlar, düğüm çiftleri arasındaki tüm yolları benzerlik ölçümünde kullanmaktadırlar [110].

Katz İndeksi (KI)

Katz tarafından önerilen bu ölçüt, düğüm çiftleri arasındaki tüm yolları dikkate alır. Düğüm çifti arasındaki en kısa yollara daha fazla önem vererek en kısa yolların toplamına göre benzerlik ölçümü yapar [118].

$$S_{xy}^{KI} = \sum_{i=1}^{\infty} \beta^i \cdot |A_{zy}^{(i)}| \quad (3.15)$$

β burada, ($\beta > 0$) olmak üzere serbest bir parametredir ve aynı zamanda "sönümlenme faktörü" olarak da adlandırılır. KI ölçütüne göre daha yüksek uzunluklara sahip yollar, bağlanma olasılığına daha az katkıda bulunur ve aynı şekilde kısa uzunluklara sahip yollar, bağlanma olasılığına daha fazla katkıda bulunurlar [5].

Global Leicht-Holme-Newman İndeksi (GLHN)

KI ölçütüne benzer bir yaklaşımda olan *GLHN* ölçütü, KI ölçütünden farklı olarak düğüm çiftleri arasındaki yolların çokluğu ya da azlığı ile ilgilenir. *GLHN* ölçütüne göre düğüm çiftleri arasındaki yolların sayısı ne kadar fazlaysa bu düğüm çiftleri arasındaki benzerlik oranı da o kadar yüksektir [117].

$$S_{xy}^{GLHN} = \beta_1 (I - \beta_2 A)^1 \quad (3.16)$$

Burada β_1 ve β_2 , ($\beta > 0$) olmak üzere serbest parametrelerdir ve β_2 en kısa yollar için daha yüksek bir benzerlik sonucu elde edilmesini sağlar.

SimRank İndeksi (SR)

SR ölçütü, "iki nesne benzer nesnelere ilişkiliyse benzerdir" hipotezinden yola çıkılarak düğüm çifti arasındaki benzerliği komşularının yollarına olan benzerliğe bakarak hesaplar [119].

$$S_{xy}^{SR} = \gamma \cdot \frac{\sum_{z_1 \in \Gamma(x)} \sum_{z_2 \in \Gamma(y)} S_{(z_1, z_2)}^{SR}}{|\Gamma(x)| \cdot |\Gamma(y)|} \quad (3.17)$$

γ burada, ($\gamma \in [0, 1]$) bozulma faktörü olarak adlandırılır ve benzer düğüm çiftlerinin etkisinin, benzerlikleri hesaplanan düğüm çiftlerinden uzaklaştıkça benzerliğin ne kadar hızla azaldığını kontrol eder. Ağdaki bütün yolları dolaşmasından dolayı ve hesaplama karmaşıklığından dolayı büyük ölçekli ağlarda uygulanabilirliği zayıftır [6,120].

3.2.1.3. Yarı Yerel Benzerlik Tabanlı Yaklaşımlar

Yol tabanlı yaklaşımların tüm ağ yapısına dair bilgileri kullanabilme verimliliği ile yerel benzerlik tabanlı yaklaşımların daha az zamanda benzerlik ölçümü yapabilmesi arasındaki dengeyi kurabilmek adına bağlantı tahmininde kullanılmak üzere yarı yerel tabanlı benzerlik yaklaşımlar ortaya çıkmıştır. Aslında yarı yerel benzerlik tabanlı ölçütler düğüm çiftleri arasındaki benzerliği hesaplamada ortak komşulara dair bilgileri yol bilgileri ile birlikte kullanmaktadırlar. Bazı yarı yerel benzerlik yaklaşımı ölçütler, tüm ağın topolojisini dikkate alsa da zaman karmaşıklığından dolayı uygulamada çok fazla tercih edilmemektedirler [110].

Yerel Yol İndeksi (LPI)

LPI ölçütü, yol tabanlı yaklaşım olan KI ve yerel benzerlik yaklaşımı olan CN'ye çok benzemekle birlikte, yol uzunluğunu yalnızca en yakın komşuların bilgilerini

kullanarak değil, aynı zamanda sonraki 2. ve 3. en yakın komşuları kullanarak daha geniş bir perspektifle ele alır [29,121].

$$S_{xy}^{LP} = A^2 + \beta A^3 \quad (3.18)$$

β burada, $-1 < \beta < 1$ olmak üzere, yol uzunluğunu 2 ve 3 düğüm mesafe içinde komşuların önemini ayarlamak için serbest kullanılan bir parametredir. Düğüm çiftleri arasındaki yol mesafelerinin kısa olması benzerlik oranlarını artırmaktadır.

FriendLink İndeksi (FL)

FL ölçütüne göre düğüm çiftleri arasındaki benzerlik, değişen yol uzunluklarına gelen düğümler arasındaki yolların normalleştirilmiş sayıları ile hesaplanmaktadır [122].

$$S_{xy}^{FL} = \sum_{l=1}^L \frac{1}{l-1} \cdot \frac{|A_{x,y}^l|}{\prod_2^l |V| - j} \quad (3.19)$$

Burada V , ağdaki yolların sayısıdır. FL ölçütü performans ve hızı nedeniyle yaygın olarak kullanılan ölçütlerden birisidir.

PropFlow Tahmin Endeksi (PFP)

PFP ölçütü, rastgele yürüyüş teorisinden [6] esinlenerek benzerlik ölçümü yapılacak düğüm çiftleri arasında rastgele yürüyerek en kısa adımda birbirine ulaşma olasılığının hesaplanmasına dayanmaktadır [123]. Yürüyüş adımı parametre olarak belirlenen bu ölçütte, bağlantılar ağırlıklara göre seçilir.

$$S_{xy}^{PFP} = S_{ax}^{PFP} \frac{w_{xy}}{\sum_{z \in \Gamma(x)} w_{xy}} \quad (3.20)$$

PFP'nin diğer yaklaşımlardan en önemli üstünlüğü, yönlendirilmiş, yönlendirilmemiş, ağırlıklı, ağırlıksız, seyrek ve yoğun ağlarda kullanılabilmesidir.

3.2.2. Olasılıksal Yaklaşımlar

Olasılıksal yaklaşımlar, Bayes kurallarını kullanan denetimli modellerdir. Bu modellerden bazılarının en önemli dezavantajı, büyük ağlar için yavaş ve maliyetli olmalarıdır [7]. Pek çok olasılıksal yaklaşıma dayalı bağlantı tahmin yöntemi önerilmiştir. Bunların içerisinde ağda hiyerarşik yapılar oluşturarak daha sonra araştırılan bu hiyerarşik yapıların ortalamasını alarak bağlantı olasılığı olan düğüm çiftlerini temsil eden Hiyerarşik Yapı Modeli [116,124], birbirine çok bağlı düğümleri bir blok ya da topluluk olarak ele alan ve düğüm çiftlerinin bağlanma olasılığını yalnızca ait oldukları blok grubuna bağlı olarak tahmin eden Stokastik Blok Modelleri [125–130], ağın zaman içerisinde evrimini belirli zaman dilimlerinde ele alarak düğümlerin birbiri üzerindeki etkilerini bağlanma olasılığına dahil eden Ağ Evrim Modeli [131], genel bağlantı tahmin yaklaşımlarının çözümsüz kaldığı problemlerin çözümü için problemin özelliğine göre uyarlanabilir Yerel Olasılık Modeli [132] olasılıksal yaklaşımlarda yaygın olarak kullanılan modellerdir.

3.2.3. İlişkisel Modeller

İlişkisel modeller, daha önce bahsedilen olasılıksal modellerin varlıklar arasındaki ilişkiyi yalnızca düğüm niteliklerini ya da yalnızca bağlantı niteliklerini birleştirmelerinden kaynaklı eksikliği ortadan kaldırmak için ortak olasılık dağılımını modellemede düğümlerin ve bağlantıların niteliklerinin birlikte kullanımını içeren modellerdir [7]. İlişkisel modeller, karmaşık problemlerin çözümünde yapısal bilgilerin dışında mantıksal yaklaşımları da dahil ederler. Genellikle ağdaki eksik bağlantıların ortaya çıkarılmasında kullanılan bu modeller, İlişkisel Bayes Ağları ve İlişkisel Markov Ağları şeklinde bağlantı tahmini için önerilen modellerdir [133–135]. İlişkisel Bayes Ağları ve İlişkisel Markov Ağlarını baz alarak önerilen farklı ilişkisel bağlantı tahmin modelleri de mevcuttur [119,136–138].

3.2.4. Öğrenmeye Dayalı Yaklaşımlar

Öğrenmeye dayalı yaklaşımlar, düğüm özniteliklerinin elde edilerek bağlantı tahmin süreçlerine dahil edilmesine yönelik önerilen modelleri kapsamaktadır. Öğrenmeye

dayalı yaklaşımlarda makine öğrenmesi, denetimli öğrenme algoritmaları gibi yöntemlerin düğüm çiftlerinin gelecekte bağlanma olasılıklarının hesaplanmasında etkin bir şekilde kullanımı hedeflenir [12,139]. Bu modeller bağlantı tahmin problemini bir sınıflandırma görevi olarak görür ve sınıflandırma metotlarını kullanarak etkili ve yüksek performanslı bağlantı tahmini gerçekleştirmeye çalışırlar [140]. Bağlantı tahmini için, Destek Vektör Makineleri, K En Yakın Komşular, Lojistik Regresyon, Radial Tabanlı Fonsiyon, Çok Katmanlı Algılayıcı ve Naive Bayes yöntemleri yaygın olarak kullanılan bazı denetimli öğrenme yöntemleridir [12,139–143].

3.3. BAĞLANTI TAHMİNİ BAŞARISINI ÖLÇMEDE KULLANILAN DEĞERLENDİRME ÖLÇÜTLERİ

Bağlantı tahmini yöntemlerinin başarılarını ölçmede farklı değerlendirme ölçütleri kullanılabilir. Bu ölçütler içerisinde Area Under the ROC Curve (AUC), Precision ve F1 Ölçütü en yaygın kullanılan değerlendirme ölçütleridir. Bağlantı tahmini sonuçları Karışıklık Matrisi (Confusion Matrix)'ne dönüştürülerek yapılan tahmin işleminin başarısı yukarıda verilen ölçütlerinden herhangi bir tanesi veya hepsi ile ölçülerek değerlendirilmektedir. Hata matrisi olarak da bilinen ve Çizelge 3.1.'de gösterilen karmaşıklık matrisini oluştururken tahmin işleminden elde edilen sonuç verileri kullanılmaktadır [144].

Çizelge 3.1. Karmaşıklık matrisi.

		VAR OLAN DURUM	
		POZİTİF	NEGATİF
TAHMİN	POZİTİF	DP	YP
	NEGATİF	YN	DN

Çizelge 3.1’de gösterilen karmaşıklık matrisinde kullanılan bazı terimlere bakıldığında,

DP (Doğru Pozitif): Uygulanan yöntem tarafından pozitif olarak işaretlenen ve deney ağında bulunan bağlantı sayısı.

YP (Yanlış Pozitif): Uygulanan yöntem tarafından negatif olarak işaretlenen ve deney ağında bulunan bağlantı sayısı.

DN (Doğru Negatif): Uygulanan yöntem tarafından pozitif olarak işaretlenen ve deney ağında bulunmayan bağlantı sayısı.

YN (Yanlış Negatif): Uygulanan yöntem tarafından negatif olarak işaretlenen ve deney ağında bulunmayan bağlantı sayısı.

Karmaşıklık matrisinde var olan durum, deney ağında işaretlenmiş durumlardır. Var olan durum için işaretlenen durumlar kesinlikle doğru olarak kabul edilir. Deney ağı için gerçekleştirilen pozitif bir tahmin durumu, var olan ağda da pozitif bir durum olarak işaretlenmişse karmaşıklık matrisinde DP olarak değerlendirilir. Ya da deney ağında bulunan negatif bir durum için tahmin edilen durum negatif ise karmaşıklık matrisinde DN olarak değerlendirilir. Benzer gösterimler YP ve YN içinde değerlendirilir. Deney ağında var olan negatif bir durum tahmin işleminde pozitif olarak işaretlenmişse karmaşıklık matrisinde YP olarak değerlendirilirken deney ağında var olan pozitif bir durum tahmin işleminde negatif olarak işaretlenmişse YN olarak değerlendirilir. Karmaşıklık matrisine bakılarak farklı hesaplama ölçütleri oluşturulmuştur. Bunlar, Recall (Anma), Accuracy (Doğruluk), Precision (Kesinlik), F1 Ölçütü ve Area Under the ROC Curve (AUC) ölçütleridir.

Recall (Anma): Duyarlılık olarak da bilinen bu ölçüt, tahmin işleminde pozitif durumların işaretlenmesindeki başarıyı ölçmek için kullanılmaktadır. Bu ölçme işlemi pozitif olarak işaretlenen tahminlerin olması gereken tüm durumların sayısına oranıyla hesaplanmaktadır.

$$Anma = \frac{DP}{DP + YN} \quad (3.21)$$

Precision (Kesinlik): Tahmin işlemi sonucunda doğru olarak işaretlenen pozitif tahminlerin olması gereken tüm pozitiflere olan oranıyla hesaplanmaktadır.

$$Kesinlik = \frac{DP}{DY + YP} \quad (3.22)$$

Accuracy (Doğruluk): Tahmin işleminde doğru olarak gerçekleşen tahminlerin deney ağındaki tüm tahminlere oranıyla hesaplanan doğruluk ölçütü, doğru tahminlerin başarısını doğrulamak için kullanılmaktadır.

$$Doğruluk = \frac{(DP + DN)}{(DP + DN + YN + YP)} \quad (3.23)$$

F1- Ölçütü: Kesinlik ve Anma ölçütlerinin harmonik ortalaması olan bu ölçüt, kesinlik ve anma arasında dengeli bir ölçüm yapmaktadır.

$$F1 = \frac{2 * Kesinlik * Anma}{Kesinlik + Anma} \quad (3.24)$$

Area Under the ROC Curve (AUC): Receiver operator characteristics curve (ROC) altında kalan alan olarak tanımlanan AUC, bağlantı tahmin yaklaşımlarının başarısını ölçmede kullanılan en yaygın değerlendirme ölçütüdür. AUC değeri, bir deney ağındaki iki düğüm arasında rastgele seçilen gerçekte var olan bir bağlantının için hesaplanan benzerlik puanının, rastgele seçilen ve eksik olan bir bağlantı için hesaplanan benzerlik puanından daha yüksek olma olasılığıdır [145].

$$AUC = \frac{n' + 0.5n''}{n} \quad (3.25)$$

Burada n , toplam karşılaştırma sayısını ifade ederken n' , gerçekte var olan seçili bağlantıların benzerlik puanlarının tahmin edilen bağlantılar için hesaplanan benzerlik

puanlarından fazla olanlarının sayısını, n ' ise, eşit veya daha küçük benzerlik puanına sahip olanların sayısını ifade etmektedir.

BÖLÜM 4

ÇOK KRİTERLİ KARAR VERME ANALİZİ

Çok kriterli karar verme analizi (ÇKKVA), birden çok özelliğe dayalı alternatiflerin performansını ölçmek için yapılandırılmış bir yaklaşımdır [146]. Bu yaklaşıma dayalı önerilen yöntemler, öznitelik olarak da bilinen birden fazla kriterin aynı anda değerlendirildiği sorunlar için karar analiz sürecini kolaylaştırmaktadır [147]. Bu karar analizi yöntemleri, her bir kriter için göreceli önemin veya ağırlığın dahil edilmesini sağlar. Ağırlıklar, uygulanacak alternatiflerin performansını seçilen kriterlere göre sıralamak için kullanılır. Bu yöntemler, özellikle karmaşık durumlarda karar verme süreçlerinin şeffaflığına, denetlenebilirliğine ve analitik değerlendirme sürecinin geliştirilebilirliğine etki etmektedir.[148].

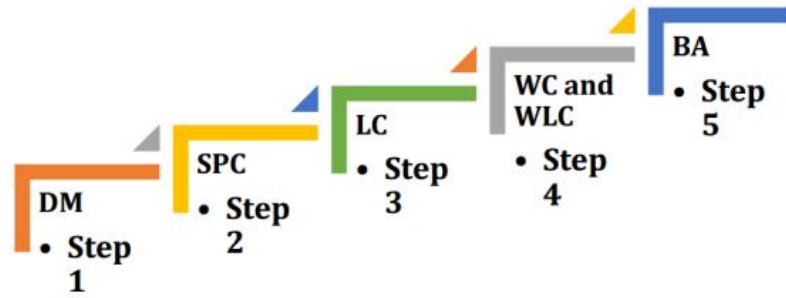
Karar verme, birden fazla kriter temelinde en iyi alternatifi belirleme ve karar vericilerin tercihlerini dikkate alma sürecidir. Bu süreç, problemin belirlenmesi, alternatiflerin tanıtılması, alternatiflerin değerlendirilmesi ve en iyi alternatifin belirlenmesi olarak dört aşamadan oluşmaktadır [149]. Karar verme sürecini desteklemek için kullanılan ÇKKVA yöntemleri, alternatiflerin tespitinde hem esnek bir şekilde hem de anlaşılır bir yapıda birden çok kriterin göz önünde bulundurulmasına imkan tanır [150]. Bu yöntemler, karar verme sürecini destekleyen matematiksel araçları içeren bir yönelem dalıdır [151]. ÇKKVA kullanılarak, alternatifleri sıralamak, alternatiflerin sayısını azaltmak, en iyi alternatifi belirlemek mümkün olabilmektedir [152].

ÇKKVA yöntemlerinin uygulama alanlarına bakıldığında, aynı alanda hizmet veren ticari kuruluşlar için hem nicel hem de nitel faktörleri içeren çok kriterli karar verme işlemi uygulanarak en iyi alternatifi belirleme [153], deniz mühendisliğinde mürettebatın risk dürtüsünü etkileyen faktörleri dikkate alarak en iyi güvenlik önleminin alınması [154], farklı konum özelliklerinin birer kriter olarak kabul edilip

en iyi konum planlaması [152] , şehir planlamasında en iyi şehir merkezinin seçimi için tüm kriterlerin birlikte kullanılarak seçim yapılması [155], bedene uygun en iyi giysinin eşleştirilmesinde beden memnuniyetine dair faktörlerin birer kriter olarak ele alınarak en iyi beden-giysi eşleştirilmesinin gerçekleştirilmesi [156] gibi farklı alanlarda uygulanmıştır. Bunu dışında, sağlık yönetimi [157], bilgi yönetimi [158], otomotiv endüstrisi [159] gibi alanlar da ÇKKVA yöntemlerinin yaygın olarak kullanıldığı alanlar içerisinde yer almaktadır. Kullanım alanları geniş olan ÇKKVA için önerilen yöntemlerin sayısı oldukça fazla olmasına rağmen bir kaç yöntem yaygın olarak tercih edilmektedir [160–162].

4.1. LOGARİTMİK KONSEPT YAKLAŞIMI (APLOCO)

Çok kriterli karar verme yöntemlerinden biri olan APLOCO, derin öğrenme yöntemlerinden olan çok katmanlı algılayıcıyı kullanmaktadır. APLOCO, herhangi bir yöntemle belirlenen kriterlerin kullanılmasıyla olayların ya da durumların ağırlıklarını hesaplayabilmektedir. Karar verme aşamasında hesaplanan ağırlık değerleri kullanılabilir. Şekil 4.1’de gösterildiği gibi APLOCO’nun uygulanması 5 adımdan oluşmaktadır [152].



Şekil 4.1. APLOCO’nun uygulama adımları[152].

Adım 1: Karar matrisinin oluşturulması (DM)

Karar Matrisi (X_{ij}), alternatiflerden ve bu alternatiflerden elde edilen kriterlerden oluşmaktadır. Karar matrisi, kriter sayısı ve alternatif sayısı yeniden değerlendirilmeden geçirilerek oluşturulur.

$$X_{ij} = \begin{matrix} & \begin{matrix} \text{Alternatives} \\ [A_1 A_2 \dots A_r] \end{matrix} \\ \begin{matrix} \text{Criteria} \\ [C_1 \\ C_2 \\ \dots \\ \dots \\ \dots \\ C_c] \end{matrix} & \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1r} \\ X_{21} & X_{22} & \dots & X_{2r} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ X_{c1} & X_{c2} & \dots & X_{cr} \end{bmatrix} \end{matrix} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1r} \\ X_{21} & X_{22} & \dots & X_{2r} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ X_{c1} & X_{c2} & \dots & X_{cr} \end{bmatrix} \quad (4.1)$$

Eşitlik 4.1’de sütunlar alternatifleri ifade ederken satırlar, alternatiflerin ağırlıklandırılmasında kullanılacak kriterleri ifade etmektedir. Burada kullanılacak kriter sayısında herhangi bir sınırlandırma bulunmamaktadır. Alternatiflerin ağırlıklarını hesaplamada farklılıkları ortaya koyabilecek kriterlerin fazla olması ağırlıklandırma sonuçlarının daha belirgin farklarda oluşmasında etkili olmaktadır.

Step 2: Başlangıç noktası kriter değerlerinin hesaplanması (SPC)

Bu adımda kriter değerleri, maksimum ve minimum değerlerinden herhangi birisiyle yeniden hesaplanır. Eğer kriter değeri maksimum olsun isteniyorsa ilgili satırdaki kriterin değerleri içerisinde maksimum olan değer maksimum değer olarak atanır. Eğer kriter değeri minimum olacak ise ilgili satırdaki kriter değerleri içerisinde minimum olan değer minimum değer olarak atanır. Ayarlanan kriter değeri maksimum değerlerden seçilmiş ise maksimum değerden ilgili satırdaki bütün değerler çıkarılır. Ayarlanan kriter değeri minimum ise ilgili satırdaki değerlerden minimum değer çıkarılır. Ve bu şekilde yeniden hesaplama işlemi tamamlanmış olur. Eğer kriter değerleri bulunduğu satırdaki minimum kriter değerine göre yeniden hesaplanmak istenirse Eşitlik 4.2 kullanılır. Aynı şekilde kriter değerleri bulunduğu satırdaki maksimum kriter değerine göre yeniden hesaplanmak istenirse Eşitlik 4.3 kullanılır. Örneğin karar matrisindeki kriter değerleri buldukları satırlardaki minimum kriter değerlerine Eşitlik 4.3 kullanılarak Eşitlik 4.4’te gösterildiği gibi karar matrisi yeniden hesaplanır.

$$P_{ij} = \begin{cases} \max x_{ij} - x_{ij} & \text{if } P_{ij} \text{ is the maximum starting point criterion.} \\ x_{ij} - \min x_{ij} & \text{if } P_{ij} \text{ is the minimum starting point criterion.} \end{cases} \quad (4.2)$$

$$(4.3)$$

$$X_{ij} = \begin{bmatrix} X_{11} - \min x_{ij} & X_{12} - \min x_{ij} & \dots & X_{1r} - \min x_{ij} \\ X_{21} - \min x_{ij} & X_{22} - \min x_{ij} & \dots & X_{2r} - \min x_{ij} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ X_{c1} - \min x_{ij} & X_{c2} - \min x_{ij} & \dots & X_{cr} - \min x_{ij} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1r} \\ X_{21} & X_{22} & \dots & X_{2r} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ X_{c1} & X_{c2} & \dots & X_{cr} \end{bmatrix} \quad (4.4)$$

Adım 3: Logaritmik dönüşüm matrisini oluşturma (LC)

Bu adımda, bir önceki adımda maksimum veya minimum kriter değerlerine göre yeniden hesaplanan karar matrisindeki her bir kritere +2 değer eklenir. Daha sonra elde edilen yeni değerlerin tersine çevrilmiş doğal logaritması hesaplanarak LC değerleri elde edilir. Kriter değerlerine +2 değerinin eklenmesinin sebebi, bir önceki adımda hesaplanan kriter değerlerinden 0 ya da negatif olanların ters logaritmasının tanımsız olmasının önüne geçmektir.

$$\ln(x) = \log_e(x) \text{ and } L_{ij} = \frac{1}{\ln(p_{ij} + 2)} \text{ for } (i = 1 \dots \dots c \text{ and } j = 1 \dots \dots r) \quad (4.5)$$

$$X_{ij} = \begin{bmatrix} \frac{1}{\ln(p_{11} + 2)} & \frac{1}{\ln(p_{12} + 2)} & \dots & \frac{1}{\ln(p_{1r} + 2)} \\ \frac{1}{\ln(p_{21} + 2)} & \frac{1}{\ln(p_{22} + 2)} & \dots & \frac{1}{\ln(p_{2r} + 2)} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \frac{1}{\ln(p_{c1} + 2)} & \frac{1}{\ln(p_{c2} + 2)} & \dots & \frac{1}{\ln(p_{cr} + 2)} \end{bmatrix} = \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1r} \\ l_{21} & l_{22} & \dots & l_{2r} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ l_{c1} & l_{c2} & \dots & l_{cr} \end{bmatrix} \quad (4.6)$$

Eşitlik 4.4'de gösterildiği gibi kriter değerine +2 eklenerek tersi alınmış doğal logaritma işlemi uygulanmaktadır. Bu işlemin karar matrisinin tamamına uygulanmış hali Eşitlik 4.5'de görülmektedir.

Adım 4: Kriter ağırlıklarının belirlenmesi ve ağırlıklı logaritmik dönüşüm matrisinin hesaplanması (WLC)

Bu adımda alternatiflerin ağırlıklandırılmasında kullanılan kriterlerin ağırlık katsayıları belirlenerek bu katsayılar göre ağırlıklı logaritmik dönüşüm matrisi hesaplanmaktadır.

$$T_{ij} = \begin{bmatrix} l_{11}w_1 & l_{12}w_1 & \cdots & l_{1r}w_1 \\ l_{21}w_2 & l_{22}w_2 & \cdots & l_{2r}w_2 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ l_{c1}w_n & l_{c2}w_n & \cdots & l_{cr}w_n \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1r} \\ t_{21} & t_{22} & \cdots & t_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ t_{c1} & t_{c2} & \cdots & t_{cr} \end{bmatrix} \quad (4.7)$$

Eşitlik 4.6'da w_1, w_2, \dots, w_n kriterlerin etki değerlerini ifade ederken, bu etki değerleri bir önceki adımda elde edilen karar matrisine uygulanarak ağırlıli logaritmik dönüşüm matrisi elde edilmektedir.

Adım 5: En iyi alternatifin tespiti (BA)

Bu adımda, Eşitlik 4.8 ve 4.9 kullanılarak, her bir satırdaki kriterlerin maksimum değerleri optimal çözüm değeri (β_j) olarak belirlenir ve toplam puan elde edildikten sonra her bir kriterin puanı (β_{sj}) hesaplanır. Eşitlik 4.10 ve 4.11'de her bir alternatif için nihai puanlar (θ_i), alternatiflerin karşılaştırma değerlerinin toplam puanlarının (α_{si}), toplanan optimum çözüm değerlerine (β_{sj}) oranlanmasıyla hesaplanır. Eşitlik 4.11'dan elde edilen puanlar, 0 ile 1 arasındadır. Daha sonra θ_i değerleri en büyükten en küçüğe doğru sıralanır ve birinci dereceden alternatif en uygun alternatif olarak kabul edilir [aploco].

$$B_{sj} = \{ \max t_{ij} \} \text{ and } B_{ij} = \{ t_1, t_2, t_3, \dots, t_n \} \quad (4.8)$$

$$B_{si} = \sum_{i=1}^n (l_1, l_2, l_3, \dots, l_n) \quad (4.9)$$

$$a_{sj} = \sum_{j=1}^n (t_1, t_2, t_3, \dots, t_n) \quad (4.10)$$

$$0 \leq \theta_i \leq 1 \text{ and } j = 1, 2, \dots, r. \theta_i = \frac{\alpha_i}{\beta_i} \quad (4.11)$$

BÖLÜM 5

VERİ SETİNİN OLUŞTURULMASI

Geliştirilen yöntemin hem karmaşık ağlarda bağlantı tahmininde hem de karmaşık ağlarda düğüm merkeziliği tespitindeki başarısını test etmek ve mevcut yöntemlerle karşılaştırmalar yapmak üzere üç farklı deneysel çalışma gerçekleştirilmiştir. Deneysel çalışmalarda hem topolojik bilgilerle ifade edilemeyen faktörler hem de topolojik faktörler ağırlıklandırma kriterleri olarak kullanılmış ve yöntemin başarısı test edilmiştir.

Uygulamaların birincisinde, geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin hem topolojik bilgilerin hemde topolojik bilgilerle ifade edilemeyen faktörlerin dahil edildiği kriterler kullanılarak bağlantı tahminindeki başarısı araştırılmış ve mevcut yöntemlerle karşılaştırmalar yapılmıştır. Uygulamaların ikincisinde geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin, karmaşık ağlarda kullanılan merkezilik ölçütleri ve zaman faktörünün dahil edildiği kriterler kullanılarak bağlantı tahminindeki başarısı araştırılmış ve mevcut yöntemlerle karşılaştırmalar yapılmıştır. Uygulamaların üçüncüsü, düğüm ağırlıklandırma işleminin karmaşık ağlarda merkezi düğümlerin belirlenmesindeki başarısının araştırılması ve düğüm ağırlıklandırma işlemi ile merkezi düğümlerin ortaya çıkarılmasına yönelik gerçekleştirilmiştir. Yapılan deneysel çalışmalarda ağların görselleştirilmesi ve merkezilik ölçütlerinin hesaplanmasında Gephi [93] kullanılmıştır.

5.1. VERİ SETİ VE ÖNİŞLEME

Deneysel çalışmalarda geliştirilen yöntemin gerçek dünya ağlarındaki başarısını test etmek için yaygın olarak kullanılan veri setlerinden topolojik özelliklerin dışında özellik çıkarımı mümkün olmadığı için iki farklı alandan veriler toplanarak veri setleri

oluşturulmuştur. Bunlardan birincisi 2000-2017 yılları arasında Avustralya Açık Tenis Turnuvalarındaki turnuva yılları, sporcular, karşılaşmalar ve turlardan oluşan veri setidir [164]. Toplanan verilerden düğümler için sporcular, bağlantılar için ise sporcular arasındaki karşılaşmalar kullanılmıştır. Bir diğer veri seti ise 2000-2017 yılları arasındaki IEEE Xplore' [165] çevrimiçi kütüphanesinde bulunan en az iki yazarlı yayınlardaki, yazar, yayınlanma yılından oluşmaktadır. Toplanan verilerden düğümler için yazarlar, bağlantılar için ortak yazarlı yayınlar kullanılmıştır. Oluşturulan ağlar için kullanılan düğüm ve bağlantı bilgileri Çizelge 5.1'de verilmiştir.

Çizelge 5.1. Deneysel çalışmalarda kullanılan ağlarda düğüm ve bağlantı bilgileri.

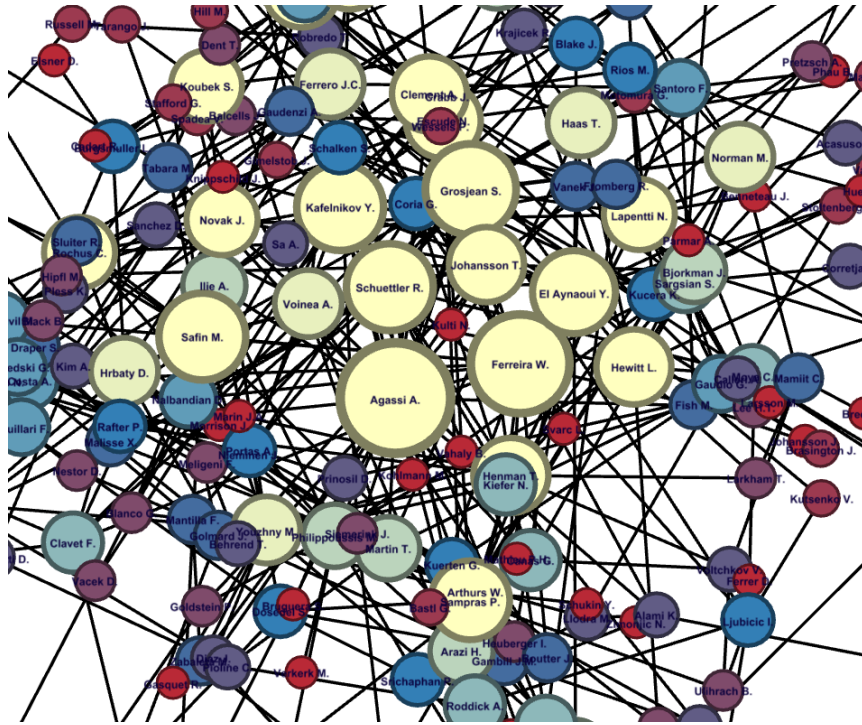
	Avustralya Açık Tenis Turnuvası	Ortak Yazarlı Bilimsel Yayımlar
Düğümler	553	16892
Bağlantılar	2156	61584

5.2. ZAMAN PERİYODUNA BAĞLI OLARAK AĞLARIN OLUŞTURULMASI

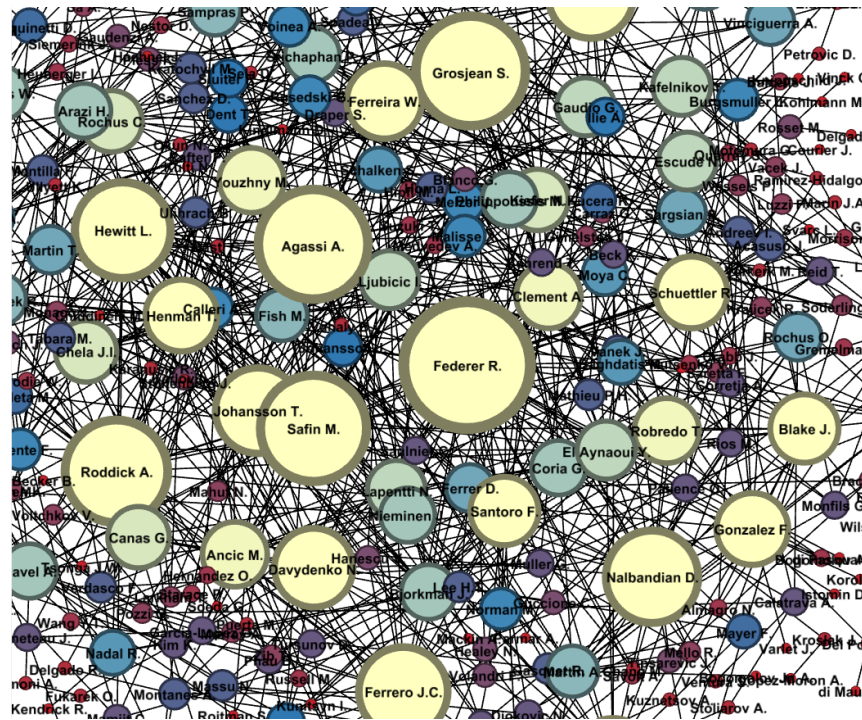
Çizelge 5.2'de gösterildiği gibi eğitim ağları farklı zaman periyotlarına bölünmüş ve ağların oluşturulma yıllarından sonraki üç yıl test ağı olarak kullanılmıştır. Şekil 5.1, 5.2, ve 5.3'te görüldüğü gibi ağların farklı zaman periyotlarına ayrılmasındaki amaç hem düğüm ağırlıklandırma işleminin zaman periyoduna bağlı olarak ağdaki merkezi düğümleri tespit etme performansını hem de geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin ağı büyüklüğüne göre tahmin başarısını ölçmek olmuştur.

Çizelge 5.2. Yıl bazında zaman periyoduna bağlı olarak oluşturulan ağlar.

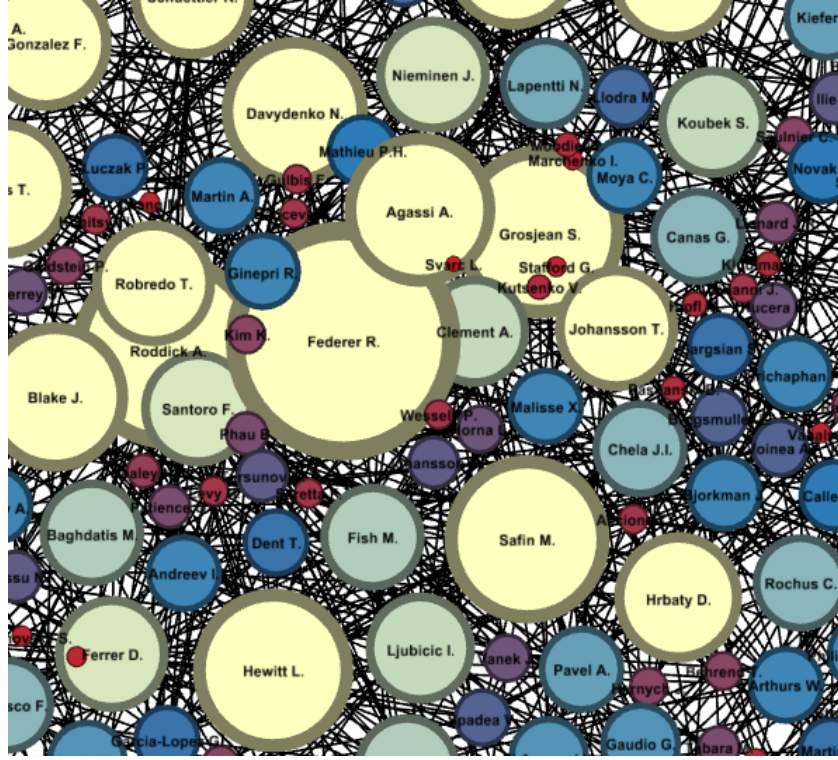
	Ağlar		
	1	2	3
Eğitim Ağı (Yıl)	2000-2003	2000-2010	2000-2014
Test Ağı (Yıl)	2004, 2005, 2006	2011, 2012, 2013	2015, 2016, 2017



Şekil 5.1. 2000-2003 Tenis turnuvaları ağı.



Şekil 5.2. 2000-2007 Tenis turnuvaları ağı.



Şekil 5.3. 2000-2010 Tenis turnuvaları ağı.

BÖLÜM 6

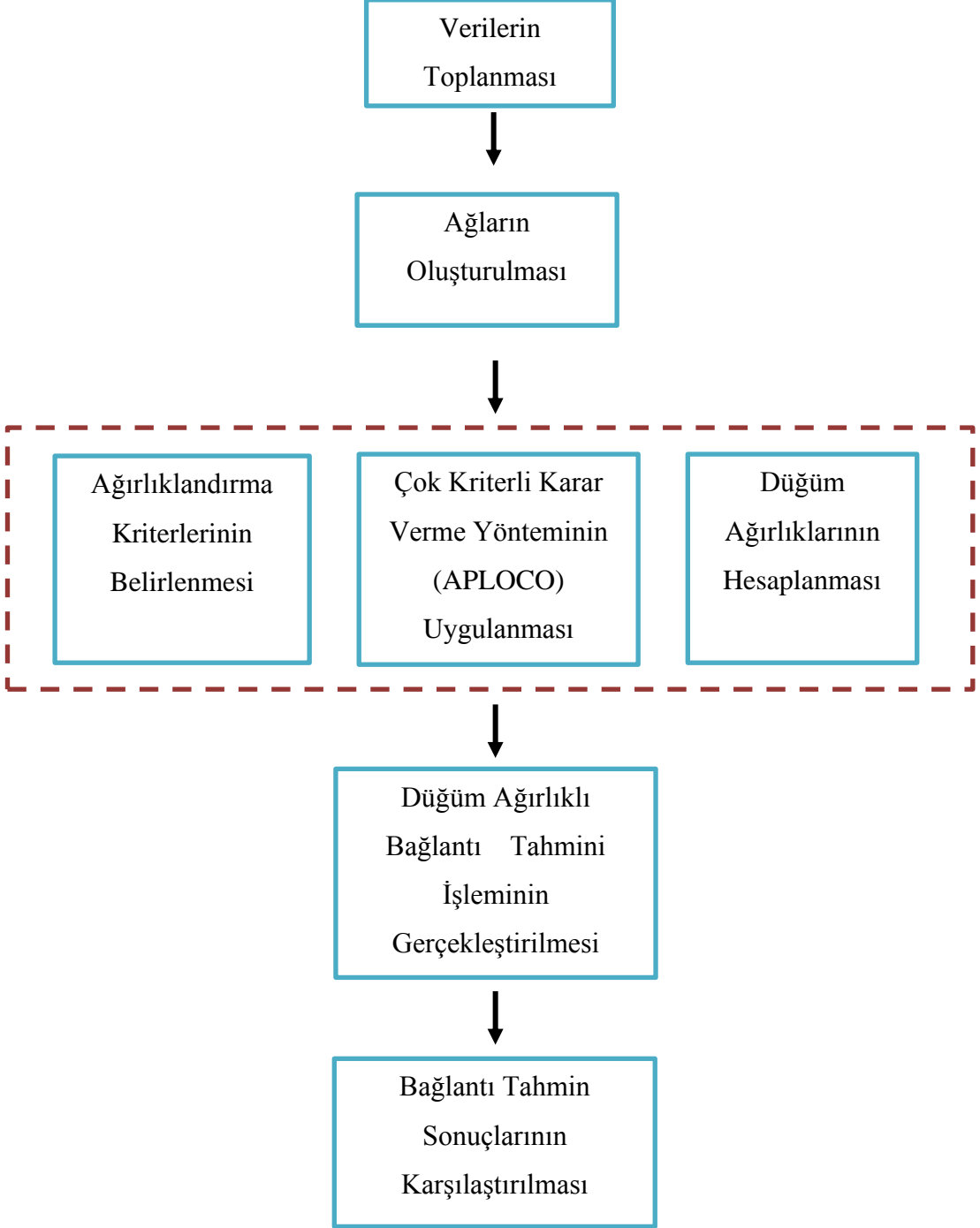
KOMŞULUK İLİŞKİLERİNE DAYALI BAĞLANTI TAHMİNİ İÇİN GELİŞTİRİLEN DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİNİ YÖNTEMİ

Komşuluk tabanlı bağlantı tahmini çalışmalarına bakıldığında, zaman içerisinde ağda meydana gelen değişimin ve ağdaki düğümlerin gücünün tahmin süreçlerine yeterince dahil edilmemesi bağlantı tahmini işlemlerinin başarısını etkileyen önemli faktörler olarak görülmektedir. Düğüm çiftleri arasındaki benzerlik hesaplamalarında ortak komşuların hepsine eşit davranılması, ağın yapısal özellikleri dışındaki faktörlerin göz ardı edilmesine sebep olmaktadır. Örneğin bir spor ağında sporcuların turnuvalara katılım zamanları, turnuvadaki müsabaka sayıları, başarıları ve deneyimleri gibi faktörler de ağın genişlemesinde etkilidir. Başka bir örnek verilecek olursa, bilimsel işbirliği ağında, yazarların zamana bağlı olarak gerçekleştirdikleri yayınlar, yayın sayıları, farklı kişilerle yapmış oldukları çalışmalar bilimsel işbirliği ağının genişlemesinde etkilidir. Ya da sosyal ağlarda bir kişinin ağa katılımından itibaren zaman içerisindeki paylaşımları, başkalarıyla etkileşimleri, paylaşımlarının yayılımı gibi faktörler ağın genişlemesinde etkilidir. Özellikle zaman periyodu içerisinde genişlemeye devam eden dinamik ağlarda bağlantı tahmini yapılırken zaman periyodunun ve ağın genişlemesine etkisi olan faktörlerin dikkate alınması tahmin sürecinin başarısını artıran etkenlerdir. Karmaşık bir ağın zaman içerisinde gelişimi gözlemlendiğinde, düğümlerin farklı zaman periyotlarında ağdaki konumu, etkinliği ve gücünün değiştiği görülmektedir. Bu nedenlerle düğüm çiftleri arasındaki benzerliği hesaplarken, bütün düğümlerin ağ içerisindeki ağırlıklarının hesaplanması ve hesaplanan bu ağırlıkların bağlantı tahmini sürecine dahil edilmesi önemlidir. Ayrıca, başarılı bir tahmin süreci için bağlantı tahmini yapılırken ortak komşulara eşit davranılması ya da düğümler arasındaki bağlantı ağırlıklarının ayırt edici faktör olarak dikkate alınması yeterli olmamaktadır. Zaman periyodu içerisinde ağda önemi ve gücü azalmış veya artmış düğümlerin ortaya çıkarılması bağlantı tahmini sürecini olumlu

etkileyecektir. Çünkü zaman içerisinde ağda önemi azalan bir düğümün ya da önemini tamamen kaybeden bir düğümün gelecekte yeni bağlantılar kurma olasılığı azalacaktır. Aynı şekilde zaman içerisinde ağda etkileşimi, gücü artan bir düğümün de gelecekte yeni bağlantılar kurma olasılığı artacaktır. Ayrıca iki düğüm arasındaki ortak komşuların gücü de bu iki düğümün gelecekte birbiriyle bağlantı kurma olasılığını etkilemektedir. Güçlü ortak komşulara sahip düğümlerin birbirine bağlanma olasılığı yüksektir. Aynı şekilde ağda etkinliği olmayan ortak komşulara sahip düğümlerin birbirine bağlanma olasılığı da düşüktür. Düğümlerin ağdaki gücünü tespit etmede genellikle topolojik metrikler kullanılmaktadır. Ancak topolojik metrikler, ağın genel yapısını gözlemleyerek değerlendirme yaptıkları için düğümlerin ağda zaman içerisindeki değişimlerini gözleme de yetersiz kalmaktadırlar. Örneğin geçmişte güçlü bağlantılar kuran bir düğüm zaman içerisinde bağlantı kurma gücünü kaybedebilir. Ya da geçmiş bağlantıları zayıf olan bir düğümün zaman içerisinde bağlantı kurma gücü artabilir. Topolojik metrikler bu farklılıkları dikkate almakta yetersiz kalmaktadırlar. Aslında zaman içerisinde ağda aktivitelerini artıran bir düğümün daha güçlü olması muhtemeldir. Bu noktada düğümün ağdaki gücünü tespit ederken topolojik metriklerin dışındaki faktörleri de dahil etmek daha etkili bir sonuç elde etmeyi mümkün kılar.

Bu tez çalışmasında komşuluk ilişkilerine dayalı bağlantı tahmin yöntemlerinin topolojik metriklerle sınırlandırılmış tahmin işlemlerindeki başarılarını artırmak için düğümlerin ağırlıklandırılmasına dayalı bağlantı tahmini yöntemi geliştirilmiştir. Geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin mevcut yöntemlerden en önemli farkı zaman periyodu içerisinde ağda etkinliği artan ya da azalan düğümleri tespit ederken topolojik metriklerin dışında zaman faktörünü de dahil ederek farklı kriterleri dikkate alabilmesidir. Ayrıca geliştirilen yöntem karmaşık ağlarda merkezi düğümlerin tespitinde etkili çözümler sunmaktadır. Karmaşık ağlarda kullanılan geleneksel merkezilik ölçütlerinin kısıtlılıklarını ortadan kaldırarak zamana dayalı merkezilik analizi yapmayı mümkün kılmaktadır. Uygulama adımları Şekil 5.1’de gösterildiği gibi düğüm ağırlıklı bağlantı tahmin yöntemi ile düğümlerin ağdaki ağırlıkları hesaplanabilmekte ve bu ağırlıklar kullanılarak komşuluk ilişkilerine dayalı bağlantı tahmini işlemi yapılabilmektedir. Geliştirilen yöntem iki kısımdan oluşmaktadır. Birinci kısım, ağdaki düğümlerin gücünü hesaplayan düğüm

ağırlıklandırma işlemi, ikinci kısım ise düğüm ağırlıklarının kullanıldığı düğüm ağırlıklı bağlantı tahmin işlemidir.



Şekil 5.1. Geliştirilen yöntemin uygulama adımları.

6.1. DÜĞÜM AĞIRLIKLANDIRMA İŞLEMİ

Karmaşık ağlarda, merkezilik ölçütleri kullanılarak bir düğümün ağdaki merkeziliğini tespit etmek mümkündür [94,97–99]. Ancak özellikle dinamik ağlarda düğümün ağdaki ağırlığını ifade etmek için bu merkezilik ölçütlerinden elde edilen bilgiler yeterli olmamaktadır. Bunun nedeni ise geleneksel merkezilik ölçütlerinin ağın sadece topolojik yapısına bakarak hesaplama yapmalarıdır. Ancak zamanla yapısal değişime uğrayan dinamik ağlarda bir düğümün merkeziliği ağdaki faaliyetlerine bağlı olarak değişebilmektedir. Geleneksel merkezilik ölçütlerinin bu kısıtlılığını ortadan kaldırmak düğümlerin zaman içerisinde ağda değişen konumlarını tespit edebilmekle mümkündür. Düğümlerin zaman içerisinde konumlarının değişimini göstermek içinse düğümlerin merkeziliğini ifade eden faktörleri dikkate almak gerekmektedir. Bu noktada, düğüm ağırlıklandırma işlemi zaman periyoduna bağlı olarak düğümlerin ağdaki konumunu tespit edebilmeyi mümkün kılmaktadır. Ayrıca ağın genişlemesinde ve düğümlerinin ağdaki gücünü hesaplamada etkisi olan ancak topolojik bilgilerle ifade edilemeyen faktörler de dikkate alınmaktadır. Düğüm ağırlıklandırma işlemi ağdaki hem düğüm merkeziliği hem de düğümlerin gücünü göstermektedir.

Burada bahsedilen düğüm ağırlığı, bir düğümün ağdaki gücünü, etkisini gösteren bir metriktir. Bu kavram, iki düğüm arasındaki bağlantının ağırlığını ifade etmemektedir. Bu metrik zaman periyodu dahil olmak üzere düğümün ağdaki gücünü ortaya çıkarabilecek hem topolojik faktörlerin hem de topolojik olarak ifade edilemeyen faktörlerin dahil edilmesiyle hesaplanmaktadır.

Geliştirilen yöntemde düğümlerin ağırlıkları hesaplanırken öncelikle ağın yapısal özelliklerinin haricinde düğümlerin etkinliğini ortaya çıkarabilecek faktörler belirlenir. Bu faktörler, özellikle zaman periyodu ve ağın gelişimine bağlı olarak düğümleri birbirinden ayırt edebilecek kriterler kullanılarak belirlenir. Kriter belirleme sürecinde herhangi bir kısıtlama yoktur. Burada önemli olan husus, belirlenen kriterlerin matematiksel olarak hesaplanabilir olması ve düğümler arasındaki etkinliği ayırt edici özelliğe sahip olmasıdır. Örneğin düğümün ağda zaman içerisindeki hareketliliğini ifade eden Eşitlik 6.1'deki denklem bir kriter olarak kullanılabilir. Ya da düğümün ağın genişlemesine olan etkisini gösteren Eşitlik 6.2'deki gibi denklem

bir kriter olarak kullanılabilir. Bir başka örnek verilecek olursa Eşitlik 6.3'de gösterildiği gibi düğümün derece bilgisi [97] ya da Eşitlik 6.4'de gösterildiği gibi düğümün kümelenme katsayısı [163] da bir kriter olarak kullanılabilir. Düğüm ağırlıklandırma işleminin en önemli avantajı da kriter belirlemede herhangi bir sayısal ya da matamtiksel bir kısıtlama koymamasıdır.

$$Zaman = \frac{Son\ Aktivite\ Zamanı - ilk\ Aktivite\ Zamanı}{Zaman\ Peryodu\ Bitişi - Zaman\ Peryodu\ Başlangıcı + 1} \quad (6.1)$$

$$Deneyim = \frac{Aktivite\ Sayısı}{Ağda\ Toplam\ Aktivite\ Sayısı} \quad (6.2)$$

$$C_D(j) = \sum_{j=1}^n A_{ij} \quad (6.3)$$

$$C_i = \frac{number\ of\ triangles\ connectes\ node\ i}{number\ of\ triples\ centered\ around\ node\ i} \quad (6.4)$$

Düğümün ağırlıklandırılması için kullanılacak kriterlerin belirlenmesinden sonraki adımda belirlenen kriterler çok kriterli karar verme analizi (ÇKKVA) işlemine tabi tutulur ve düğümlerin ağırlıkları hesaplanır. Düğümlerin ağırlıklandırılması işleminde kullanılacak ÇKKVA yönteminin seçimi de önemlidir. Çünkü düzgün dağılım gösteren veriler üzerinde kullanılan pek çok ÇKKVA yöntemi mevcuttur. Ancak karmaşık bir ağda, düğümler arasındaki ilişkiler ve bu ilişkilerin oluşumunda etkili olan faktörler düzgün dağılım göstermeyebilir. Bu nedenle ÇKKVA yöntemlerinde düzgün dağılım göstermeyen verilere uygulanabilir bir yöntemin seçilmesi gerekmektedir. Bölüm 4.1'de açıklanan ve ÇKKVA yöntemlerinden birisi olan APLOCO, geliştirilen yöntemde düğümlerin ağırlıklandırılması için kullanılmaktadır.

Eşitlik 4.1'de verilen karar matrisi düğümlerin ağırlıklandırılması işlemi için belirlenen kriterleri bir araya getirmektedir. Karar matrisinde her satır kriterleri ifade ederken her sütun düğümlerin kriter değerlerini ifade etmektedir. Eşitlik 4.2. ve Eşitlik 4.3'teki yöntemlerden birisi seçilerek karar matrisindeki kriter değerleri maksimum

veya minimum değere göre normalize edilerek yeniden hesaplanır ve Eşitlik 4.4 elde edilir. Eşitlik 4.4'teki normalize edilmiş kriter değerlerinin ters logaritması alınarak Eşitlik 4.6 elde edilir. Bundan sonraki aşamada her bir kriter için ağırlık etkisi belirlenerek yeniden hesaplama yapılır. Örneğin beş kriterli bir ağırlık hesaplamasında kriterlerin etkisi eşitse Eşitlik 4.7'de w_1, w_2, w_3, w_4, w_5 katsayıları 0,2 olarak alınmalıdır. Son olarak APLOCO'nun 5. adımındaki işlemler yapıldıktan sonra her bir düğüm için ağırlık değerleri hesaplanır.

Elde edilen düğüm ağırlıkları bağlantı tahmini için düğüm çiftleri arasındaki benzerliklerin hesaplanmasında kullanılmaktadır. Ayrıca zaman faktörü gibi topolojik bilgilerle ifade edilemeyen faktörlerin dahil edilmesiyle elde edilen bu ağırlıklar, bir düğümün ağdaki konumu hakkında bilgi vermesinden dolayı karmaşık ağlarda merkezilik ölçütü olarak da kullanılabilir.

6.2. DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİN İŞLEMİ

Ortak komşuların benzerliğinde kullanılan geleneksel komşuluk tabanlı yöntemlerin tahmin başarısını artırmak için geliştirilen düğüm ağırlıklı bağlantı tahmini, iki düğüm arasındaki bağlantı olasılığının hesaplanmasında ortak komşuların eşit muamele görmesini ortadan kaldırır. Ortak komşuluk ilişkilerine göre düğüm çiftlerinin benzerliğini hesaplayan geleneksel yöntemlerle gerçekleştirilen tahmin işlemlerinin başarısı, zaman periyodunu veya farklı etkenleri dikkate almamasından dolayı tatmin edici bir başarı gösterememektedir. Geliştirilen yöntemde, iki düğüm arasındaki ortak komşulara dayalı benzerlik hesaplamaları yapılırken, ağdaki ortak komşuların gücü dikkate alınır. İki düğüm arasında ağda yüksek verimliliğe sahip ortak komşuların varlığı, bu iki düğümün gelecekte birbirine bağlanma olasılığını artıran bir unsurdur. Geliştirilen yöntemde, Bölüm 6.1'de anlatılan düğüm ağırlıklandırma işlemi, ağdaki ortak komşuların verimliliğini hesaplamak için kullanılır. Düğüm ağırlıklandırma işleminin zaman periyodunu kriter olarak kullanabilmesi, düğümün hem ağda zaman içerisinde değişen konumunun hem de ağdaki gücünün tahmin işlemine yansıtılmasında etkilidir.

Geliştirilen yöntem ile bağlantı tahmininde kullanılan komşuluk ilişkilerine dayalı geleneksel yöntemler, matematiksel olarak yeniden modellenmiştir.

6.2.1. Düğüm Ağırlıklı Ortak Komşular (NwCN)

Geleneksel CN yönteminde, ortak komşuların sadece sayısı dikkate alınırken Eşitlik 6.5'te verilen NwCN yönteminde ortak komşuların ağırlıkları dikkate alınmaktadır. x ve y düğümleri arasındaki benzerlik hesaplanırken ortak komşuların ağırlıkları toplamı hesaplanır. Bu yöntem ile ağda ortak komşuların sayısal çoğunluğundan ziyade ortak komşuların ağırlıkları önem kazanmaktadır. Güçlü ortak komşuları olan düğümlerin gelecekte bağlanma olasılığı daha yüksektir.

$$NwCN_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} wn(z) \quad (6.5)$$

Burada, $\Gamma(x)$ ve $\Gamma(y)$, bağlantı olasılığı olan x ve y düğümlerinin komşularının kümesini, $wn(z)$, x ve y 'nin ortak komşusu olan z düğümünün ağırlığını belirtir.

6.2.2. Düğüm Ağırlıklı Jaccard İndeks (NwJI)

Bağlantı tahmininde kullanılan komşuluk tabanlı JI, benzerlik hesabı yapılacak düğüm çiftlerinin ortak komşularının sayısının tüm komşulara olan oranının hesaplanmasıyla elde edilirken, Eşitlik 6.6'da verilen NWJI, düğüm çiftlerinin ortak komşularının ağırlıkları toplamının tüm komşuların ağırlıklarının toplamına oranıyla hesaplanmaktadır. Burada temel farklılık benzerlik ölçümü yapılan düğüm çiftinin ortak komşularının gücünün tüm komşulara olan oranını dikkate almasıdır. Bu oran ne kadar yüksek çıkarsa bu iki düğümün gelecekte birbirine bağlanma olasılığı da yüksek olacaktır.

$$NwJI_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{wn(z)}{\sum_{a \in |\Gamma(x)|} wn(a) + \sum_{b \in |\Gamma(y)|} wn(b)} \quad (6.6)$$

Burada, $\Gamma(x)$ ve $\Gamma(y)$, bağlantı olasılığı olan x ve y düğümlerinin komşularının kümesini, $wn(z)$, x ve y 'nin ortak komşusu olan z düğümünün ağırlığını, $wn(a)$, x düğümünün komşularının ağırlığını ve $wn(b)$, ise y düğümünün komşularının ağırlığını ifade etmektedir.

6.2.3. Düğüm Ağırlıklı Tercihli Bağlantı İndeksi (NwPA)

Bağlantı tahmininde kullanılan PA, Eşitlik 3.7'de gösterildiği gibi düğüm çiftleri arasında bağlanma olasılığını düğüm çiftlerinin komşularının sayısı ile doğru orantılı olarak hesaplamaktadır. Bu noktada komşuları fazla olan düğümlerin bağlanma olasılıkları daha yüksek ihtimal olarak değerlendirilmektedir. Eşitlik 6.7'de verilen NwPA ise düğümlerin bağlanma olasılığını komşularının ağırlıkları toplamıyla doğru orantılı olarak hesaplamaktadır. Yani komşu sayılarının fazlalığı kadar komşularının ağırlıkları daha fazla düğüm çiftlerinin gelecekte birbiriyle bağlantı kurma olasılığı daha yüksek ihtimal dahilinde değerlendirilmektedir.

$$NwPA_{xy} = \sum_{a \in |\Gamma(x)|} wn(a) * \sum_{b \in |\Gamma(y)|} wn(b) \quad (6.7)$$

Burada, $\Gamma(x)$ ve $\Gamma(y)$, bağlantı olasılığı olan x ve y düğümlerinin komşularının kümesini, $wn(a)$, x düğümünün komşularının ağırlığını ve $wn(b)$, ise y düğümünün komşularının ağırlığını ifade etmektedir.

6.2.4. Düğüm Ağırlıklı Adamic Adar İndeksi (NwAA)

Bağlantı tahmininde kullanılan geleneksel AA ölçütü, ortak komşularının bağlantı sayıları düşük olanlara daha çok önem vererek gelecekte bağlantı kurma ihtimallerinin daha yüksek olduğunu değerlendirmektedir. Eşitlik 6.8'de verilen NwAA ise düğüm çiftleri arasında bağlantı olasılığını hesaplarken ortak komşuların bağlı olduğu düğümlerin ağırlıkları toplamının önemini ortaya çıkarmaktadır. Bağlantı sayısı az olan düğümlerin bağlanma olasılığının aksine komşularının ağırlıkları toplamı daha az olan düğümlerin bağlanma olasılığına yüksek ihtimal vermektedir.

$$NwAA_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{wn(z)}{\log(1 + \sum_{c \in |\Gamma(z)|} wn(c))} \quad (6.8)$$

Burada, $\Gamma(x)$ ve $\Gamma(y)$, bağlantı olasılığı olan x ve y düğümlerinin komşularının kümesini, $wn(z)$ x ve y düğümlerinin ortak komşusu olan z düğümünün ağırlığını ve $wn(c)$ ise z düğümünün komşularının ağırlığını ifade etmektedir.

6.2.5. Düğüm Ağırlıklı Kaynak Paylaştırma İndeksi (NwRA)

Komşularının komşuları üzerinden değerlendirme yaparak düğüm çiftleri arasındaki bağlantı olasılığını hesaplayan RA ölçütü, düğümlerin ortak komşuları üzerinden birbiriyle iletişim kurabileceğini değerlendirirken, Eşitlik 6.9'da verilen NwRA ölçütü, düğüm çiftleri arasındaki ortak komşuların komşularının ağırlıklarını dikkate almaktadır. Bu ölçüte göre hesaplanan benzerlik, düğüm çiftlerinin birbirlerinden aldıkları kaynakların ağırlığına göre hesaplanmaktadır.

$$NwRA_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{1}{\sum_{b \in |\Gamma(z)|} wn(k_b)} \quad (6.9)$$

Burada, $\Gamma(x)$ ve $\Gamma(y)$, bağlantı olasılığı olan x ve y düğümlerinin komşularının kümesini, z , x ve y düğümlerinin ortak komşularını, $wn(k_b)$ ise z düğümünün komşularının ağırlığını ifade etmektedir.

6.2.6. Düğüm Ağırlıklı Sorenson İndeks (NwSI)

Bağlantı tahmininde kullanılan geleneksel SI, ortak komşuların sayısal çoğunluğuna önem vermekle birlikte düğüm derecesi az olan düğüm çiftlerinin bağlanma olasılığına daha yüksek ihtimal verirken, Eşitlik 6.10'da verilen NwSI ölçütü, hem ortak komşuların ağırlıkları dikkate almakta hem de komşuların ağırlıkları toplamına önem vermektedir.

Eşitlik 6.10'da $\Gamma(x)$ ve $\Gamma(y)$, bağlantı olasılığı olan x ve y düğümlerinin komşularının kümesini, $wn(z)$, x ve y 'nin ortak komşusu olan z düğümünün ağırlığını, $wn(a)$, x

düğümünün komşularının ağırlığını ve $wn(b)$ ise y düğümünün komşularının ağırlığını ifade etmektedir.

$$NwSI_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{2 * Wn(z)}{\sum_{a \in |\Gamma(x)|} wn(a) + \sum_{b \in |\Gamma(y)|} wn(b)} \quad (6.10)$$

Yeniden matematiksel olarak modellenen geleneksel bağlantı tahmin yöntemlerine bakıldığında, geleneksel bağlantı tahmin yöntemlerinin bağlantı olasılıklarını hesaplarken ortak komşuların baskın olma durumuna verdikleri önemin aksine geliştirilen yöntemde ise bağlantı olasılıkları hesabında ortak komşuların ağırlıklarının önemini dikkate aldığı görülmektedir.

6.3. DÜĞÜM AĞIRLIKLI BAĞLANTI TAHMİNİ YÖNTEMİ İÇİN GERÇEKLEŞTİRİLEN DENEYSSEL ÇALIŞMA

Bu tez çalışmasının temel amacı olan ve ortak komşuluk ilişkilerine dayalı bağlantı tahmin yöntemlerinin başarısını arttırmak için Bölüm 6.1 ve 6.'de detaylı bir şekilde açıklanan düğüm ağırlıklı bağlantı tahmini yönteminin başarısını değerlendirmek için Bölüm 5.1 ve 5.2'de açıklanan veri setleri ve bu veri setlerinden oluşturulan ağlar kullanılarak ortak komşuluk ilişkilerine dayalı ağırlıksız ve kenar ağırlıklı ağlar için önerilen geleneksel yöntemlerle karşılaştırma işlemleri gerçekleştirilmiştir. Kenar ağırlıklı ağlarda ağırlık değerleri tenis turnuvaları ağları için sporcular arasındaki müsabaka sayıları, bilimsel işbirliği ağları için yazarların ortak makale sayıları olarak kullanılmıştır. Her veri seti için de oluşturulan ağlarda ağırlıklandırma kriterleri kullanılarak düğümlerin ağırlıklandırma işlemi gerçekleştirilmiştir. Düğümler arasında gelecekte oluşması muhtemel bağlantı olasılıklarını hesaplamak için elde edilen düğüm ağırlıkları kullanılarak geliştirilen bağlantı tahmin yöntemleri uygulanmıştır. Deneysel çalışmaların başarısını ölçmek için Bölüm 3.2'de anlatılan değerlendirme ölçütlerinden en yaygın olan AUC metriği kullanılmıştır.

6.3.1. Deneysel Çalışmada Kullanılan Ağırlıklandırma Kriterlerinin Belirlenmesi

Düğüm ağırlıklarını hesaplamak için gerçekleştirilen ÇKKVA’da kullanılacak kriter belirleme işlemi ağın topolojik bilgilerinin dışında ağın gelişimini daha net ortaya çıkarabilecek verilerin tespit sürecidir. Hem tenis hem de ortak yazar ağı için kriterlerin belirlenmesinde zaman periyodu dahil olmak üzere güçlü düğümleri ortaya çıkarabilecek faktörler tespit edilmiştir. Kriterlerin belirlenme süreci uygulanacak ağların yapısına göre farklılıklar gösterebilmektedir.

6.3.1.1. Tenis Ağlarının Ağırlıklandırılmasında Kullanılan Kriterler

Zaman Kriteri: Zaman içerisinde genişlemeye devam eden ağlarda, düğümlerin ağdaki konumları zaman periyoduna bağlı olarak değişkenlik gösterebilmektedir. Bu durum gelecekte yeni bağlantıların oluşmasında önemli bir faktördür. Zaman periyodunun başında ağda etkin olan bir düğümün zaman periyodunun devamında etkinliğini azaltması ağdaki önemini doğrudan etkilemektedir. Bu nedenle zaman kriteri düğüm ağırlıklarının hesaplanmasında önemli bir kriter olarak kullanılabilir. Tenis turnuvaları ağında kullanılan zaman kriteri, Eşitlik 6.11’de gösterildiği gibi bir sporcunun zaman periyodu aralığında turnuvaya katılım yılının zaman periyoduna oranıyla hesaplanmaktadır.

$$Zaman = \frac{TSKY - TİK Y}{ZASY - ZAIY + 1} \quad (6.11)$$

Burada $TSKY$, sporcunun hesaplanmak istenen zaman aralığı içerisinde son katıldığı turnuvaya ait yılını, Turnuvaya $TİK Y$ sporcunun hesaplanmak istenen zaman aralığı içerisinde ilk katıldığı turnuvaya ait yılını, $ZASY$ hesaplanmak istenen zaman aralığının bitiş yılını ve $ZAIY$ ise hesaplanmak istenen zaman aralığının başlangıç yılını ifade etmektedir. Eşitlik 6.11’deki +1 değeri hesaplanmak istenen zaman aralığının aynı yıllara denk gelmesi durumunda 0 durumunun oluşmasını önlemek için kullanılmaktadır.

Deneyim Kriteri: Tenis turnuvalarında bir sporcunun turnuvalardaki deneyimi ağdaki diğer düğümlerle olan bağlantısını etkileyen bir unsurdur. Bir sporcu belirli zaman periyodu içerisinde turnuvalarda ne kadar çok müsabaka yaparsa ağdaki etkinliği ve gücü o kadar artar. Tenis turnuvaları için belirlenen deneyim kriteri Eşitlik 6.12’de gösterildiği gibi sporcunun turnuvalarda oynadığı müsabaka sayısının turnuvalarda oynanan toplam müsabaka sayısına oranı ile hesaplanmaktadır.

$$Deneyim = \frac{STMS}{TTMS} \quad (6.12)$$

Burada *STMS*, belirli zaman aralığında turnuvada oynadığı maç sayısını, *TTMS* ise belirli zaman aralığında turnuvada gerçekleşen toplam müsabaka sayısını ifade etmektedir.

Başarı Kriteri: Tenis turnuvalarında sporcunun turnuvadaki başarısı turnuva ağındaki etkinliği ölçmede kullanılabilir önemli faktörlerden birisidir. Bir sporcunun turnuvada kazandığı maç sayısı onun ağdaki gücünün de göstergesidir. Çünkü sporcu ne kadar çok müsabaka kazanırsa ağdaki bağlantı sayısını artırmaktadır ve daha fazla sporcuyla bağlantı kurmaktadır. Tenis turnuvalarında başarı kriteri Eşitlik 6.13’te gösterildiği gibi kazanılan toplam müsabaka sayısının toplam müsabaka sayısına oranıyla hesaplanmaktadır.

$$Başarı = \frac{KTMS}{OTMS} \quad (6.31)$$

Burada *KTMS*, sporcunun belirli zaman aralığında turnuvalarda kazandığı müsabaka sayısını, *OTMS* ise sporcunun belirli spor aralığında oynadığı toplam müsabaka sayısını ifade etmektedir.

Tur Kriteri: Düğüm ağırlıklandırma işleminde kullanılacak kriterleri belirlemede veri setinin durumuna göre düğümlerin ağdaki ağırlığını ortaya çıkarabilecek faktörler kullanılabilir. Tur kriteri de tenis turnuvalarından oluşturulan bir ağda düğümlerin ağırlıklarını hesaplamak için kullanılabilir önemli bir kriterdir. Tenis turnuvalarının kendi içerisinde turlardan oluştuğu göz önüne alındığında, bir

sporçunun katıldığı her turnuvada oynadığı tur sayısı sporçunun ağdaki gücünü ifade etmede önemli bir faktördür. Tenis turnuvalarında tur kriteri Eşitlik 6.14'te gösterildiği gibi sporçunun belirli yıllar arasında katıldığı turnuvalarda oynadığı tur sayılarının toplamının turnavadaki toplam tur sayısına oranı ile hesaplanmaktadır.

$$Tur = \sum_{Yıl=i}^j \frac{TOTS_{Yıl}}{TTTS_{Yıl}} \quad (6.14)$$

Burada i , düğüm ağırlıkları hesaplanacak olan ağın zaman periyodunun başlangıç yılını, j düğüm ağırlıkları hesaplanacak olan ağın zaman periyodunun bitiş yılını, $TOTS$ sporçunun yıl bazında katıldığı turnuvada oynadığı tur sayısını, $TTTS$ ise yıl bazında turnuvada sporcuların maksimum oynayacağı tur sayısını ifade etmektedir.

6.3.1.2. Bilimsel İşbirliği Ağlarının Ağırlıklandırılmasında Kullanılan Kriterler

Zaman Kriteri: Ortak yazarlı bilimsel işbirliği ağları zamana bağlı olarak genişleyen ve dinamik bir özelliğe sahip olan ağlar oldukları için düğümlerin zaman içerisinde ağdaki etkinlikleri ve güçleri değişkenlik göstermektedir. Zaman periyodu içerisinde ortak yayınlara katılımı azalan ya da sona eren bir yazarın ağın içerisinde etkinliğinin de azalması beklenir. Ya da zaman periyodu içerisinde ortak yayınlara katılımı artan bir yazarın ağın içerisinde etkinliğinin de artması beklenir. Bu nedenle zaman kriteri, düğümlerin ağırlıklarının hesaplanmasında önemli bir faktör olarak kullanılmaktadır. Bilimsel işbirliği ağlarında kullanılan zaman kriteri Eşitlik 6.15'de gösterildiği gibi belirli bir zaman aralığında yazarın son yaptığı ortak yayın ile ilk yaptığı ortak yayın yılının farkının, düğüm ağırlıklandırılması yapılan ağın zaman periyoduna oranı ile hesaplanmaktadır.

$$Zaman = \frac{SYT - İYT}{ZASY - ZAIY + 1} \quad (6.15)$$

Burada SYT , yazarın en son yaptığı ortak yayın yılını, $İYT$ yazarın ilk ortak yayın yılını, $ZASY$ düğüm ağırlıklandırılması yapılan ağda gerçekleşen son ortak yayın yılını, $ZAIY$ ise ilk ortak yayın yılını ifade etmektedir. Eşitlik 6.15'teki +1 değeri

hesaplanan zaman aralığı aynı yıla denk geldiğinde 0 sonucunu önüne geçmek için kullanılmaktadır.

Deneyim Kriteri: Bilimsel işbirliğinde yazarın deneyimi, ağda komşu olduğu yazarların birbirine bağlanmasında etkilidir. Bir yazar ne kadar çok ortak yayın yaparsa ortak yayın yaptığı yazarlarında birbiriyle yayın yapma olasılığı artar. Bu durum yazarın ağdaki gücünün ve etkinliğinin göstergesidir. Bilimsel işbirliği ağlarında kullanılan deneyim kriteri Eşitlik 6.16'da gösterildiği gibi belirli zaman aralığında yazarın yayınladığı ortak makale sayısının aynı zaman aralığında yayınlanan ortak yazarlı toplam makale sayısına oranı hesaplanmaktadır.

$$Deneyim = \frac{YMS}{TMS} \quad (6.16)$$

Burada YMS , yazarın ortak yayınladığı makale sayısını ifade ederken TMS ise, yayınlanan ortak yazarlı makale sayısını ifade etmektedir.

Etkinlik Kriteri: Bilimsel işbirliğinde zaman periyoduna bağlı olarak yazarın yıl bazında yaptığı ortak yayınlar ağdaki etkinliğini gözlemlemede önemli bir faktördür. Ağ içerisinde etkin olan bir yazar komşularını birbirlerine bağlamada da önemli bir rol oynayabilmektedir. Bu nedenle bilimsel işbirliği ağı için belirlenen etkinlik kriteri düğüm ağırlıklandırmasında önemlidir.

$$Etkinlik = \sum_{Yil=i}^j \frac{YMS_{Yil}}{TMS_{Yil}} \quad (6.17)$$

Eşitlik 6.17.'de i , düğüm ağırlıklandırılması yapılacak ağda zaman periyodunun başlangıç yılını, j zaman periyodunun bitiş yılını, YMS , yazarın yıl bazında yaptığı ortak yayın sayısını, TMS ise, yıl bazında ağda yapılan ortak yayınların sayısını ifade etmektedir.

6.3.2. 2000-2003 Arası Bağlantı Tahmin Sonuçları

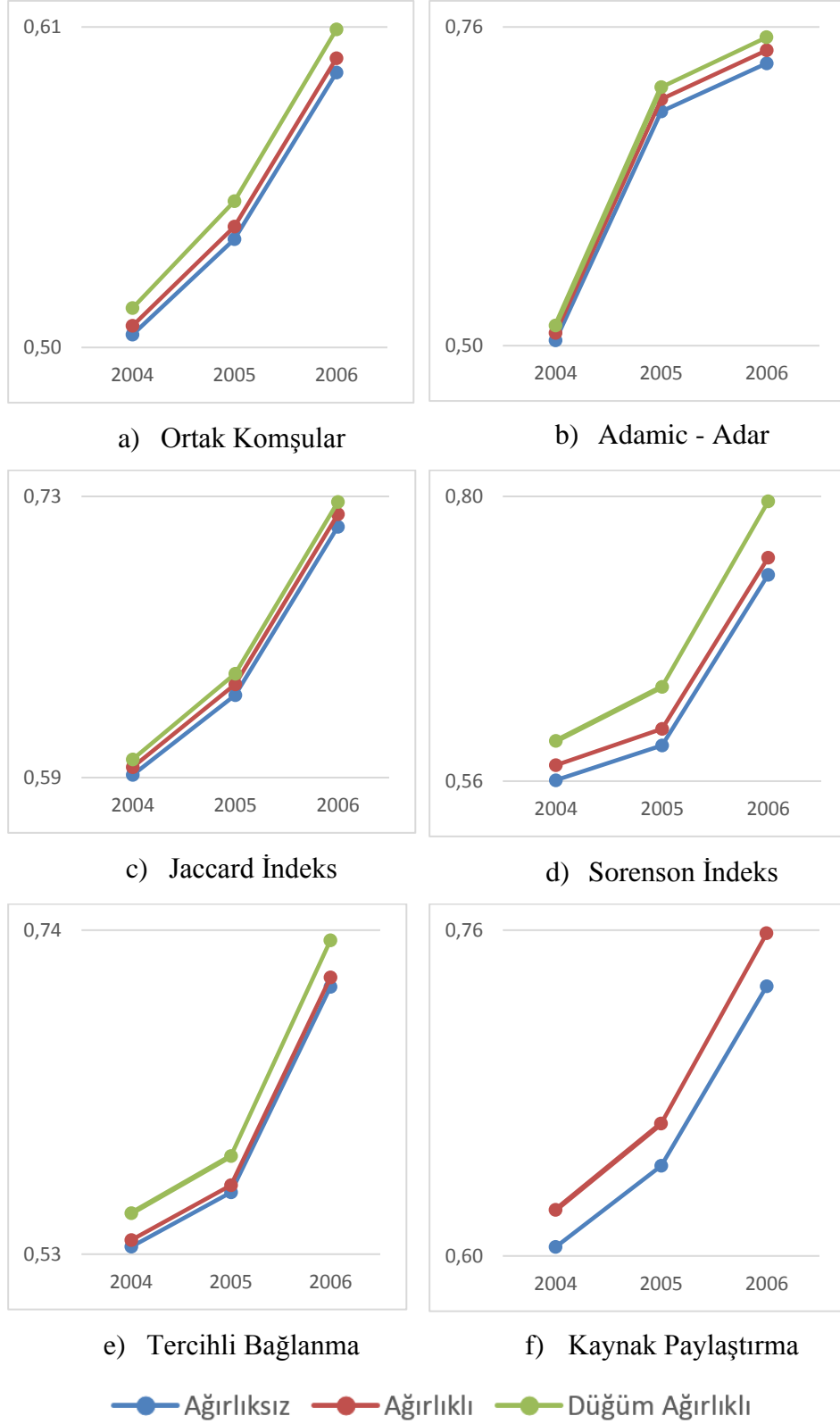
Tenis turnuvaları ve bilimsel işbirliği verilerinden 2000-2003 yılları arasında oluşturulan ağ eğitim için kullanılırken bağlantı tahmin işlemleri 2004, 2005 ve 2006 yıllarında oluşturulan ağlar için gerçekleştirilmiştir. Tahmin sonuçlarına bakıldığında Şekil 6.1 ve 6.2’de görüldüğü gibi düğüm ağırlıklı bağlantı tahmini yönteminin geleneksel komşuluk ilişkilerine dayalı yöntemlerden daha başarılı olduğu görülmektedir.

Tenis turnuvalarından oluşturulan ağlarda en iyi tahminlerin 2004 yılı için NwRA, 2005 yılı için NwAA ve 2006 yılı için NwSI tarafından gerçekleştirildiği görülmektedir. Bilimsel işbirliği ağlarının tahmin sonuçlarına bakıldığında ise en başarılı tahminler 2004 yılı için NwCN, 2005 yılı NwRA ve 2006 yılı için ise NwPA tarafından gerçekleştirildiği görülmektedir.

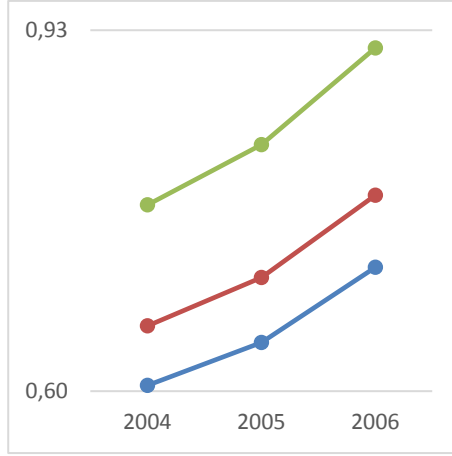
6.3.3. 2000-2007 Arası Bağlantı Tahmin Sonuçları

Şekil 6.3 ve 6.4’te tenis turnuvaları ve bilimsel işbirliği verilerinden 2000-2007 yılları arasında oluşturulan ağ eğitim için kullanılırken, bağlantı tahmin işlemleri 2008, 2009 ve 2010 yıllarından oluşturulmuştur. Şekil 6.3 ve 6.4’teki tahmin sonuçlarına bakıldığında geliştirilen Düğüm Ağırlıklı Bağlantı Tahmin yönteminin geleneksel yöntemlerin tamamından daha başarılı olduğu görülmektedir. Eğitim için kullanılan ağın zaman periyoduna bağlı olarak genişlemesinin tahmin sonuçlarını olumlu etkilediği görülmektedir. Geliştirilen Düğüm Ağırlıklı Bağlantı Tahmin yönteminin, geniş zaman aralığına sahip eğitim ağlarında daha başarılı tahmin işlemi gerçekleştirebildiği anlaşılmaktadır. Zaman periyodu geniş bir eğitim ağından alınan bilgileri kullanarak daha başarılı tahmin yapılabilmektedir.

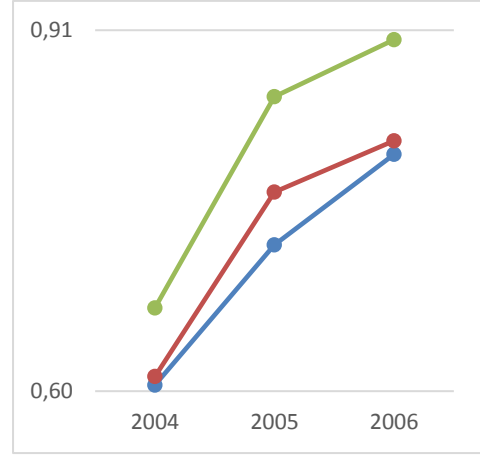
Tenis turnuvalarından oluşturulan ağlarda en başarılı tahminlerin 2008 yılı için NwPA, 2009 yılı için NwAA ve 2010 yılı için ise NwJI tarafından gerçekleştirildiği görülmektedir. Bilimsel işbirliği ağlarının tahmin sonuçlarına bakıldığında ise en başarılı tahminler 2008 yılı için NwRA, 2009 ve 2010 yılları için ise NwPA tarafından gerçekleştirilmiştir.



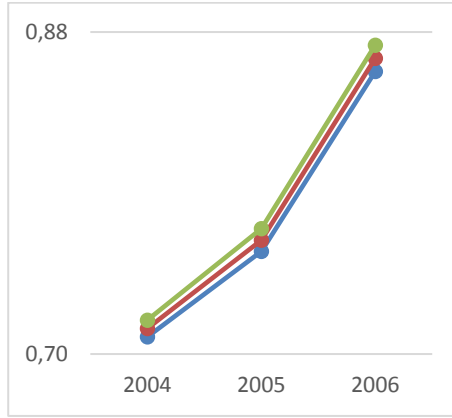
Şekil 6.2. 2000-2003 Arası tenis ağından elde edilen AUC sonuçları.



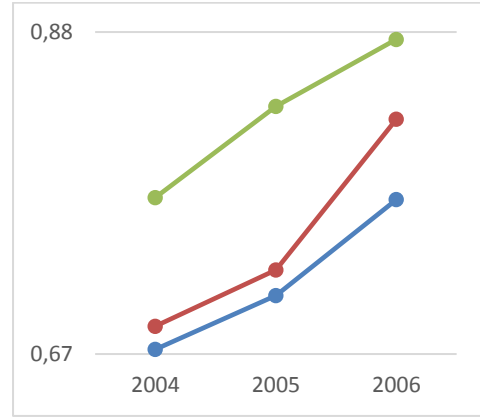
a) Ortak Komşular



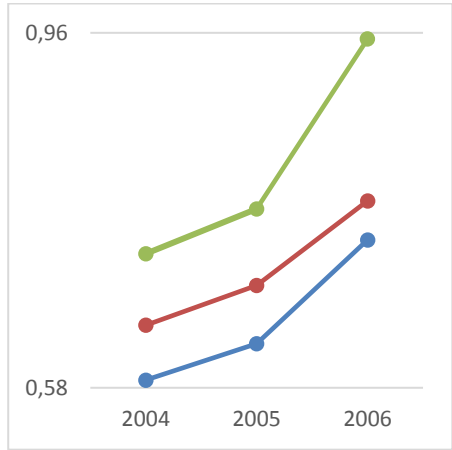
b) Adamic - Adar



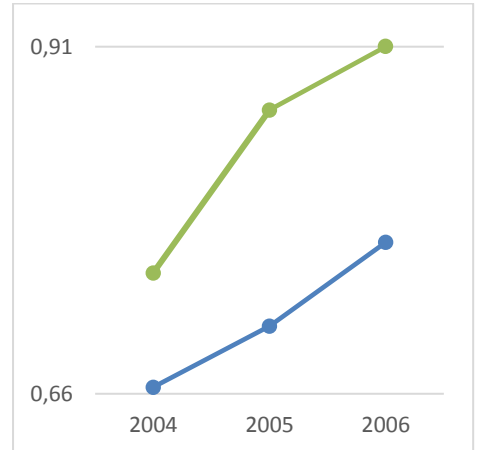
c) Jaccard İndeks



d) Sorenson İndeks



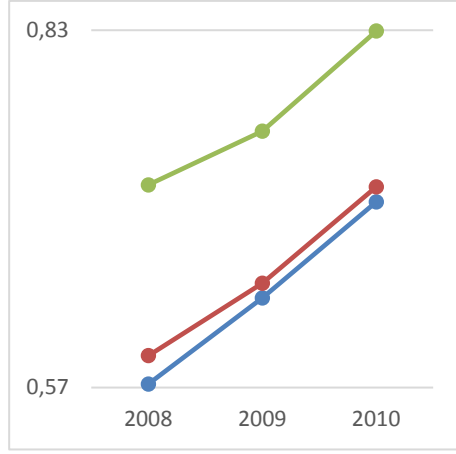
e) Tercihli Bağlanma



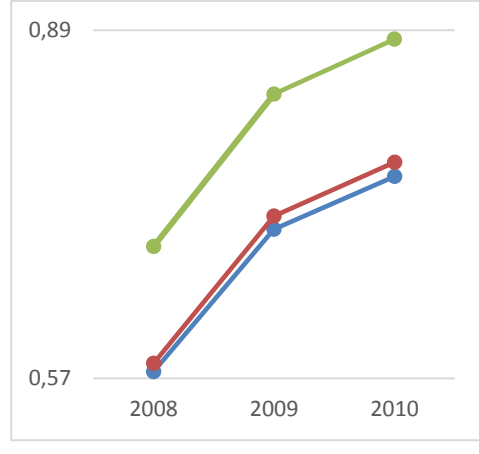
f) Kaynak Paylaşırma

—●— Ağırlıksız —●— Ağırlıklı —●— Düğüm Ağırlıklı

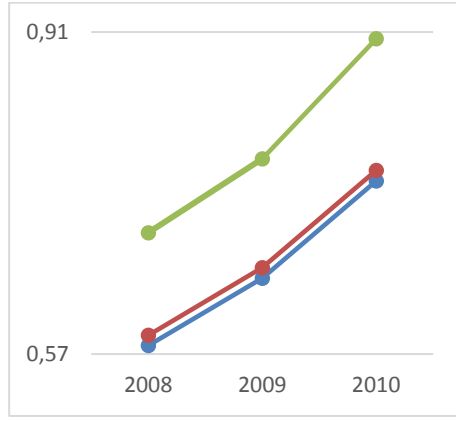
Şekil 6.3. 2000-2003 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları.



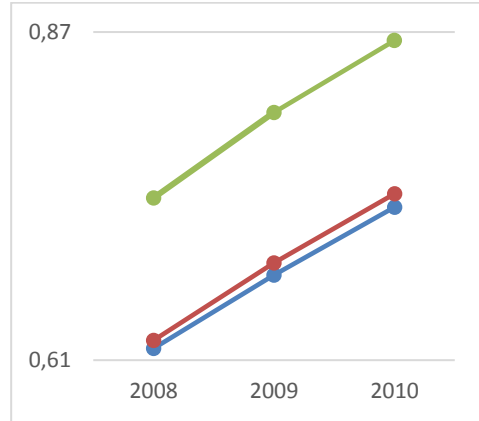
a) Ortak Komşular



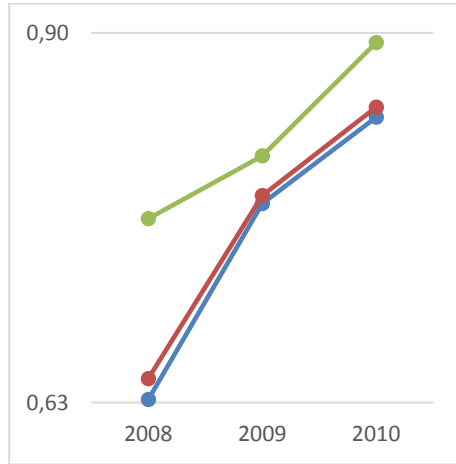
b) Adamic - Adar



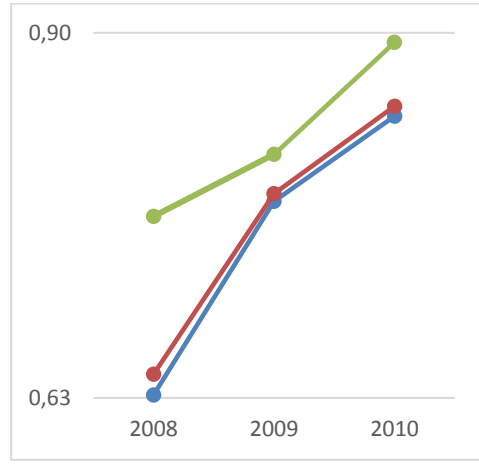
c) Jaccard İndeks



d) Sorenson İndeks



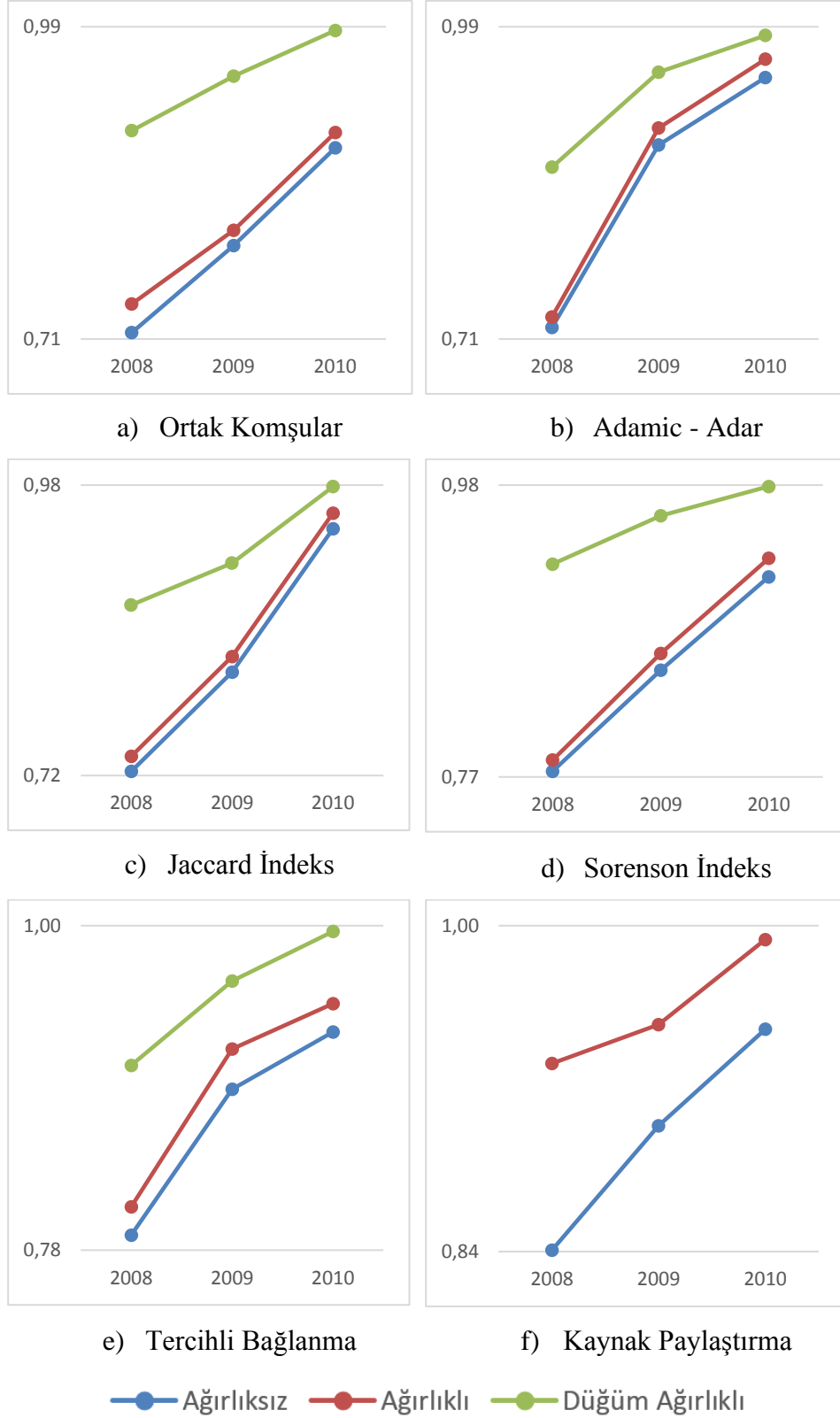
e) Tercihli Bağlanma



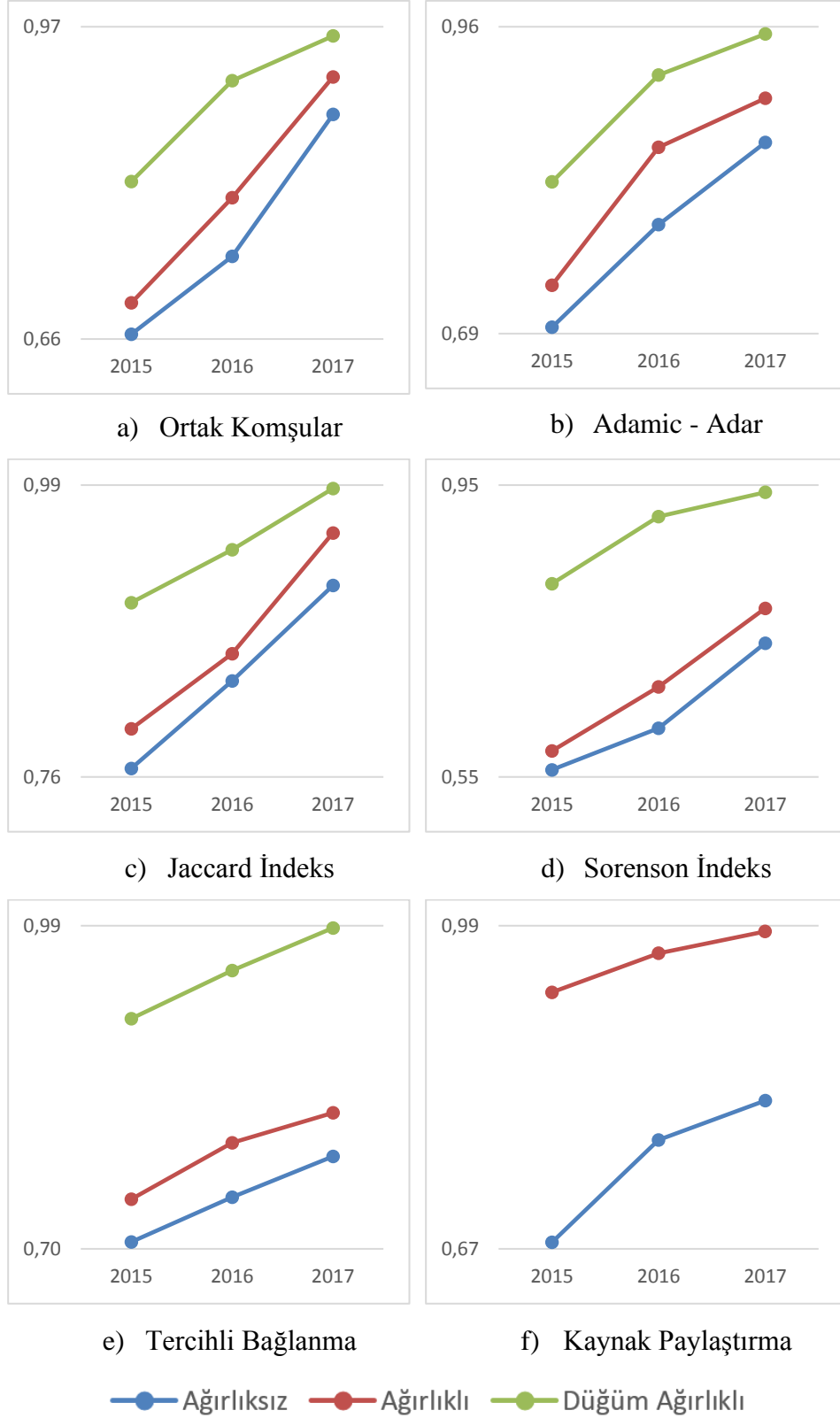
f) Kaynak Paylaştırma

—●— Ağırlıksız —●— Ağırlıklı —●— Düğüm Ağırlıklı

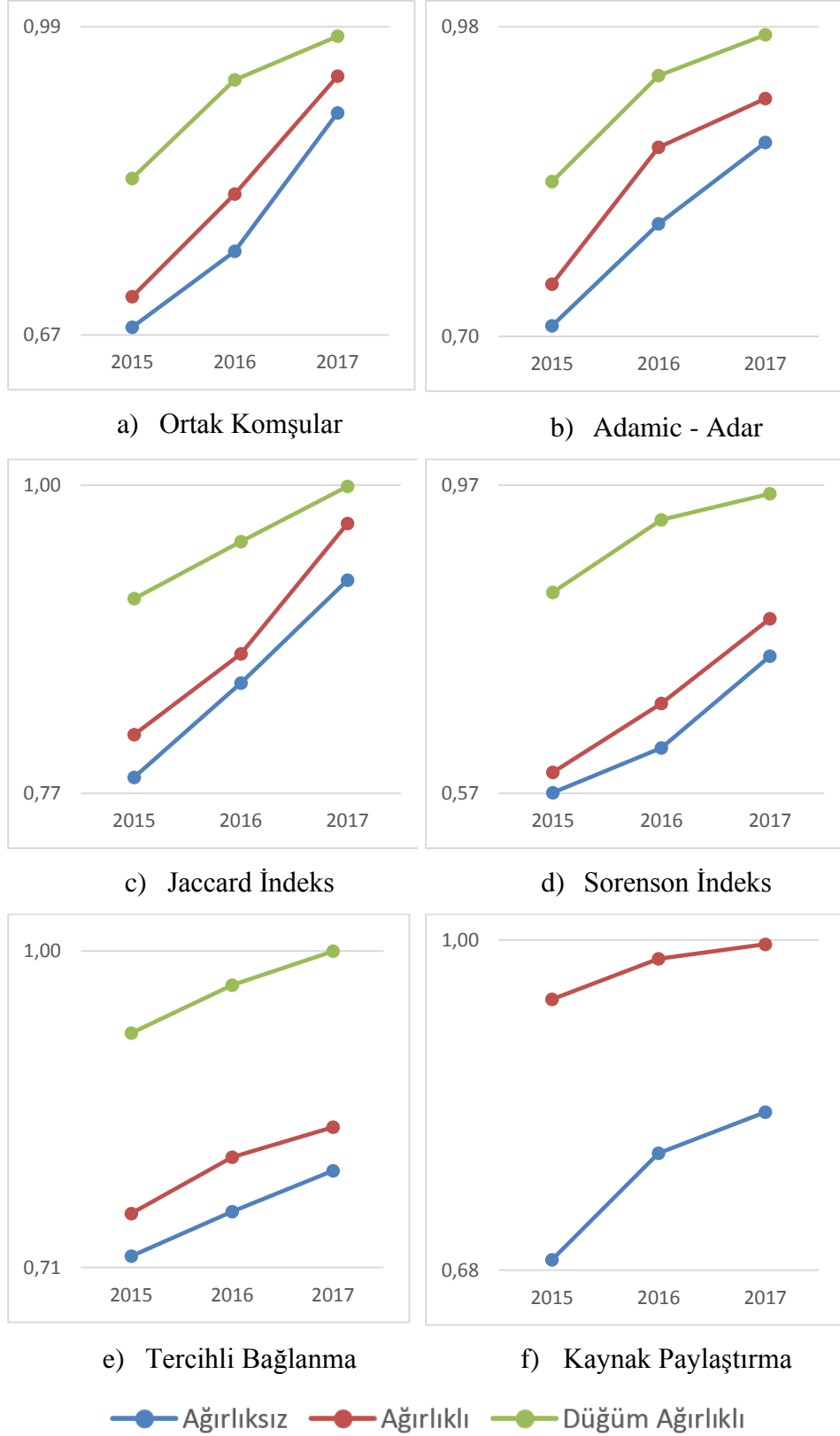
Şekil 6.4. 2000-2007 Arası tenis ağından elde edilen AUC sonuçları.



Şekil 6.5. 2000-2007 Arası bilimsel işbirliği açısından elde edilen AUC sonuçları.



Şekil 6.6. 2000-2014 Arası tenis ağından elde edilen AUC sonuçları.



Şekil 6.7. 2000-2014 Arası bilimsel işbirliği ağından elde edilen AUC sonuçları.

6.3.4. 2000-2014 Arası Bağlantı Tahmin Sonuçları

Şekil 6.5 ve 6.6’da tenis turnuvaları ve bilimsel işbirliği verilerinden 2000-2014 yılları arasında oluşturulan ağ eğitim için kullanılırken, bağlantı tahmin işlemleri 2015, 2016 ve 2017 yıllarından oluşturulan ağlardır. Her iki veri seti için de sonuçlara bakıldığında, zaman periyodunun genişlemesinin tahmin işlemlerinin başarısına olan etkisi daha net olarak ortaya çıkmaktadır. Geliştirilen yöntemin, ağırlıklı ve ağırlıksız ağlarda kullanılan komşuluk ilişkilerine dayalı yöntemlerden başarılı olmasının en önemli sebebi tahmin işlemlerinde zaman periyodunu dikkate almasıdır. Her iki veri seti içinde en iyi tahminler NwRA ve NwPA tarafından gerçekleştirilmiştir.

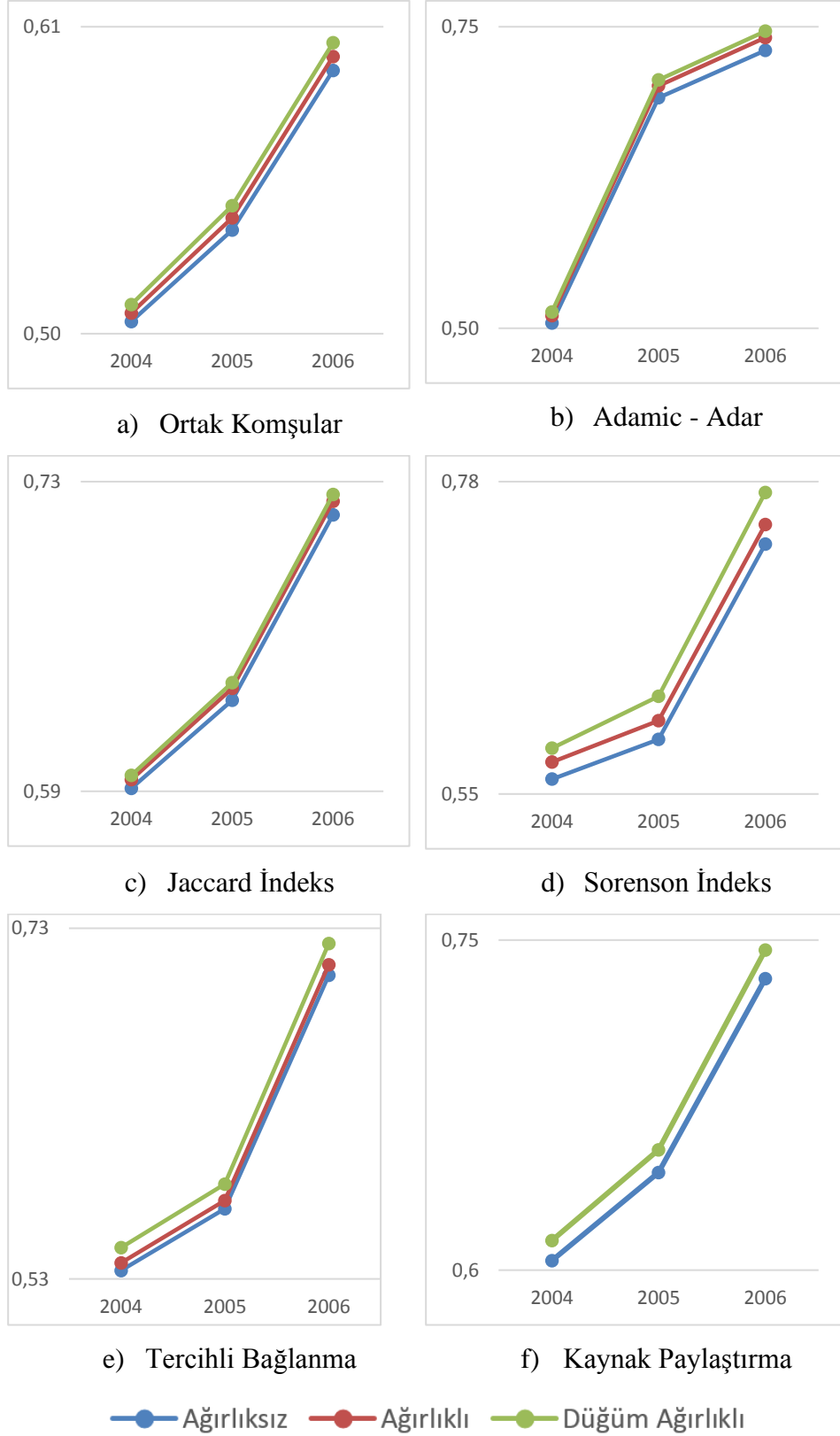
BÖLÜM 7

DÜĞÜM AĞIRLIKLIL BAĞLANTI TAHMİNİNDE MERKEZİLİK ÖLÇÜTLERİ VE ZAMAN FAKTÖRÜNÜN BİRLİKTE KULLANIMI

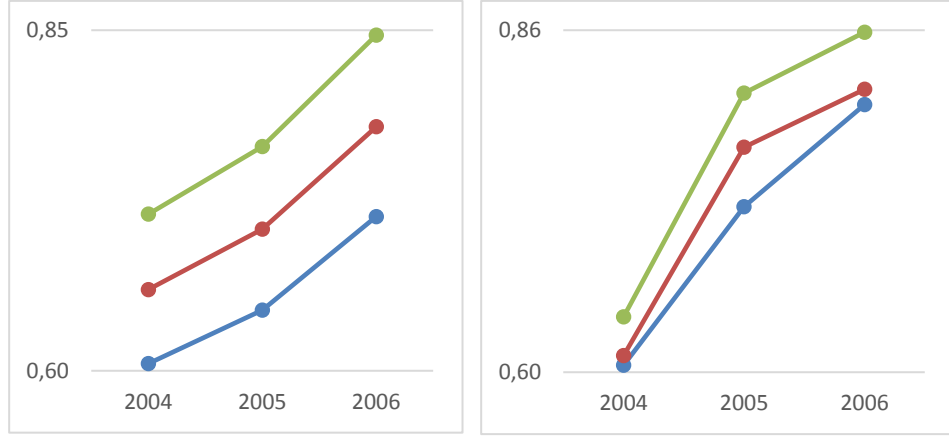
Düğüm ağırlıklı bağlantı tahmin yönteminde düğüm ağırlıklarının hesaplanması işlemi, verisetinin içeriğinden elde edilen kriterler kullanılarak gerçekleştirilebilmektedir. Bu kriterler, ağın topolojik yapısı dışında düğümlerin gücünü ortaya çıkarabilecek faktörlerden oluşabilmektedir. Tez çalışmasının bu kısmında, geliştirilen yöntemin düğümlerin ağırlıklandırılması aşamasında her ağ için ortak kriterler, ağın topolojik yapısından elde edilen ağ merkeziliği ölçütlerinden ve düğümlerin ağdaki faaliyetlerini zamansal olarak ifade eden zaman faktöründen oluşmaktadır. Zaman faktörü, düğümlerin zaman periyodu içerisinde ağdaki hareketini ölçmek adına önemli bir kriterdir. Diğer kriterler ise ağ merkezilik ölçütleri olan düğüm merkeziliği, yakınlık merkeziliği, arasındalık merkeziliği ve özvektör merkeziliği olarak belirlenmiştir.

7.1. 2000-2003 ARASI BAĞLANTI TAHMİN SONUÇLARI

2000-2003 arasında oluşturulan ağlar için 2004, 2005 ve 2006 yıllarında yapılan bağlantı tahmin işlemlerinin sonuçlarına bakıldığında, Şekil 7.1 ve 7.2’de görüldüğü gibi merkezilik ölçütleri ve zaman periyodunun ağırlıklandırma kriteri olarak kullanıldığı düğüm ağırlıklı bağlantı tahmin yöntemleri geleneksel bağlantı tahmin yöntemlerinden daha başarılıdır. Bölüm 6.3.2.’de yapılan deneysel çalışmaya benzer şekilde AUC metriğine göre genel olarak başarının düşük olduğu görülmektedir. Bunun temel nedeni de yine kullanılan eğitim ağının 2000-2003 yıllarından oluşmasından dolayı tahmin işlemlerinde kullanılacak bilginin yeterli olmamasından kaynaklanmaktadır. Ancak buna rağmen düğümlerin merkezilik ölçütleri ve zaman periyoduna göre ağırlıklandırılmasıyla gerçekleştirilen tahmin işleminin tatmin edici bir başarı gösterdiği görülmektedir.

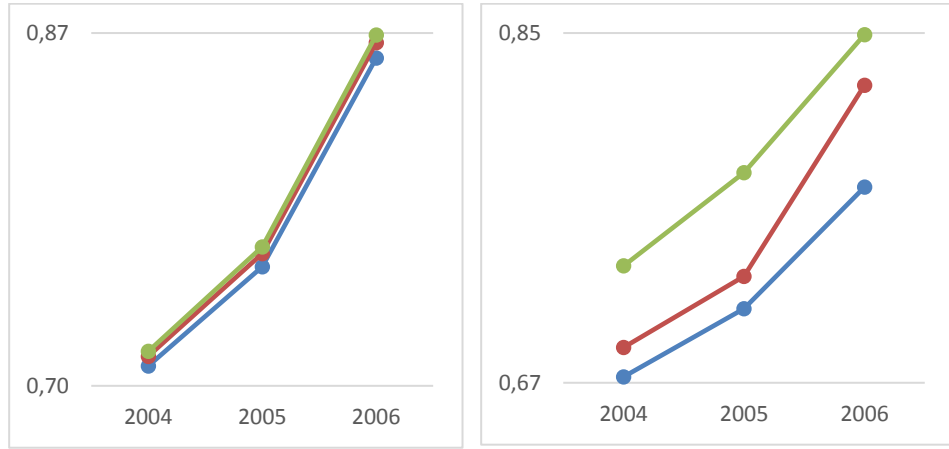


Şekil 7.1. 2000-2003 Arası tenis ağından elde edilen AUC sonuçları.



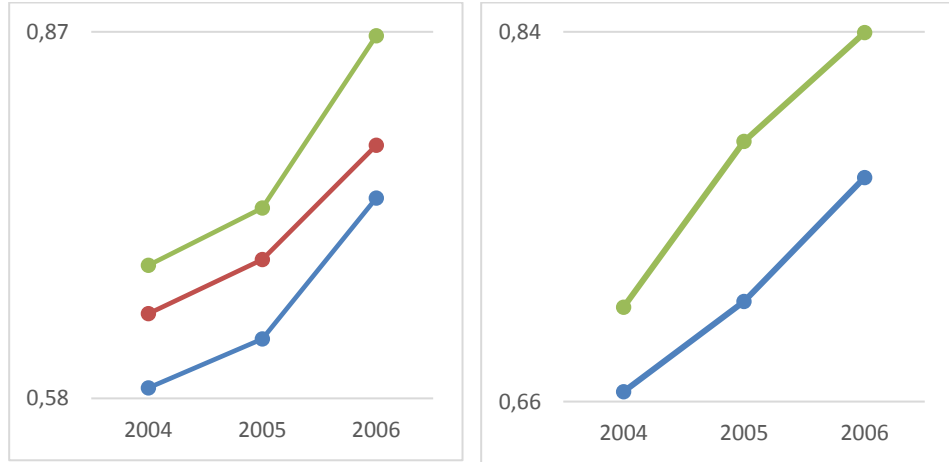
a) Ortak Komşular

b) Adamic - Adar



c) Jaccard İndeks

d) Sorenson İndeks



e) Tercihli Bağlanma

f) Kaynak Paylaşırma

—●— Ağırlıksız —●— Ağırlıklı —●— Düğüm Ağırlıklı

Şekil 7.2. 2000-2003 Arası bilimsel işbirliği açısından elde edilen AUC sonuçları.

Özellikle bilimsel işbirliği ağlarında bu başarı daha net görülmektedir. Tahmin sonuçlarına bakıldığında, tenis turnuvalarından oluşturulan ağlarda en iyi tahminlerin 2004 yılı için NwRA, 2005 yılı için NwAA ve 2006 yılı için NwSI tarafından gerçekleştirildiği görülmektedir. Bilimsel işbirliği ağlarının tahmin sonuçlarına bakıldığında ise en başarılı tahminlerin 2004 yılı için NwSI, 2005 yılı NwAA ve 2006 yılı için ise NwJI tarafından gerçekleştirildiği görülmektedir.

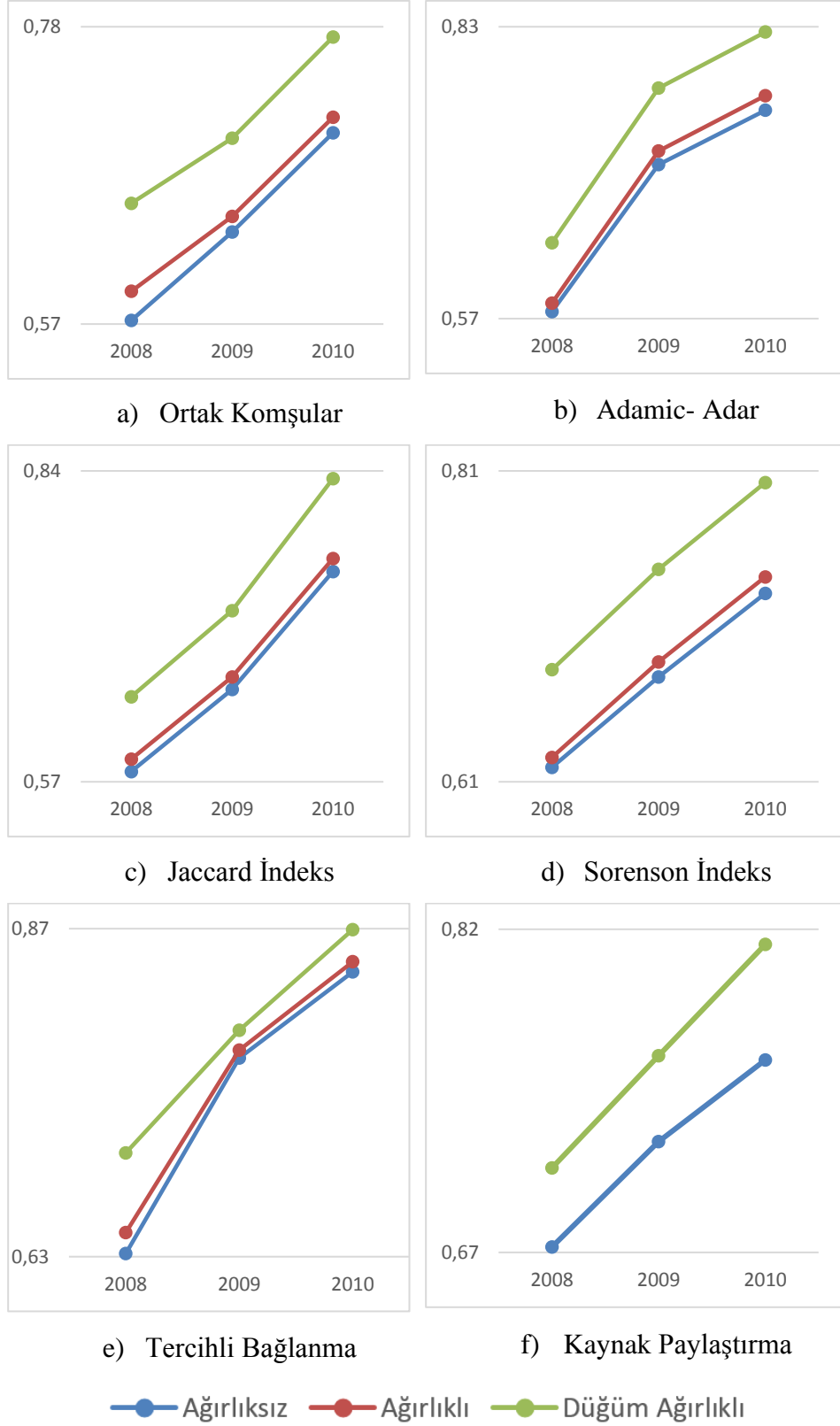
7.2. 2000-2007 ARASI BAĞLANTI TAHMİN SONUÇLARI

2000-2007 arasında oluşturulan ağlar için 2008, 2009 ve 2010 yıllarında yapılan bağlantı tahmin işlemlerinin sonuçlarına bakıldığında, Şekil 7.3 ve 7.4'te görüldüğü gibi merkezilik ölçütlerinin ve zaman periyodunun ağırlıklandırma kriteri olarak kullanıldığı düğüm ağırlıklı bağlantı tahmin yöntemlerinin geleneksel bağlantı tahmin yöntemlerinden daha başarılı sonuçlar verdiği görülmektedir. Elde edilen bağlantı tahmin sonuçları eğitim ağının genişlemesinin tahmin başarısına olan olumlu etkisini göstermektedir. Düğüm ağırlıklarının merkezilik ölçütleri ve zaman periyoduna bağlı kriterler kullanılarak hesaplandığı düğüm ağırlıklı bağlantı tahmin sonuçlarının hem ağırlıksız hem de kenar ağırlıklı ağlar için önerilen yöntemlerden daha üstün bir başarıya sahip olduğu burada da görülmektedir. Hatta eğitim ağının genişlemesiyle birlikte düğüm ağırlıklı bağlantı tahmininin başarısı diğer yöntemlerden çok daha belirgin bir şekilde ortaya çıkmaktadır.

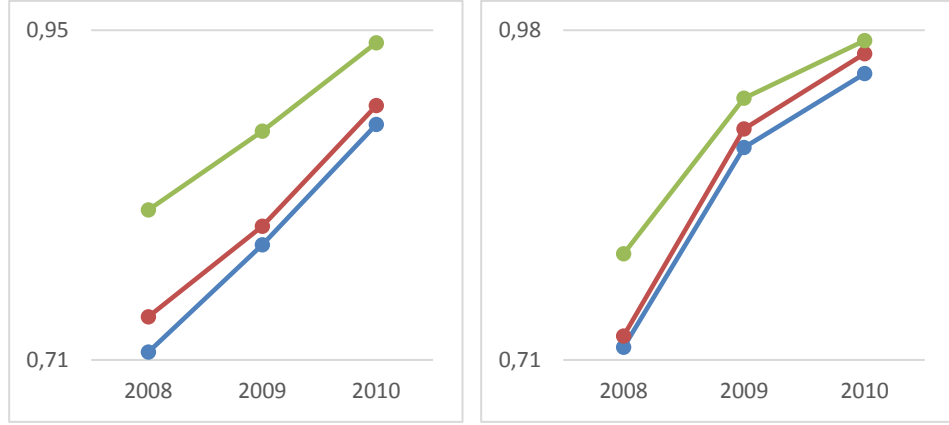
Tahmin sonuçlarına bakıldığında, tenis turnuvalarından oluşturulan ağlarda en iyi tahminlerin her iki veri seti için de 2008 yılı için NwRA, 2009-2010 yılı için NwPA tarafından gerçekleştirildiği görülmektedir.

7.3. 2000-2014 ARASI BAĞLANTI TAHMİN SONUÇLARI

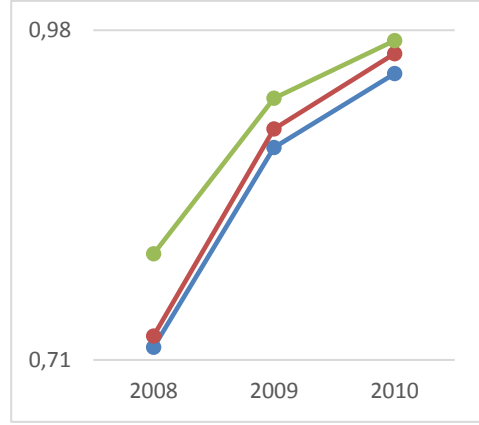
2000-2014 arasında oluşturulan ağlar için 2015, 2016 ve 2017 yıllarında yapılan bağlantı tahmin işlemlerinin sonuçlarına bakıldığında, Şekil 7.5 ve 7.6'da görüldüğü gibi merkezilik ölçütlerinin ve zaman periyodunun ağırlıklandırma kriteri olarak kullanıldığı düğüm ağırlıklı bağlantı tahmin yöntemlerinin başarısının geleneksel bağlantı tahmin yöntemlerinden daha yüksek olduğu görülmektedir.



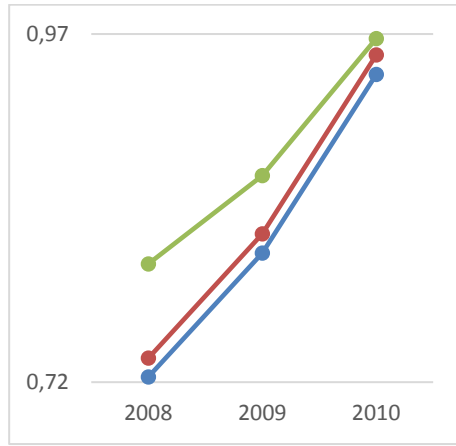
Şekil 7.3. 2000-2007 Arası oluşturulan tenis ağından elde edilen AUC sonuçları.



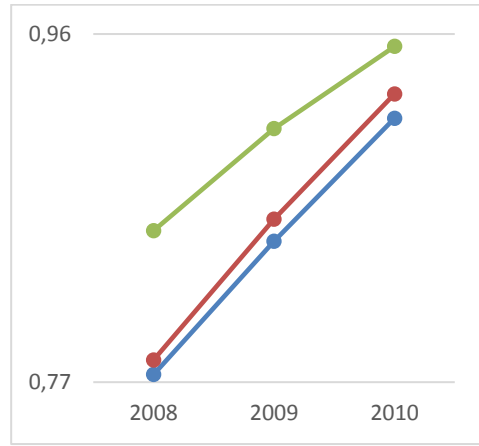
a) Ortak Komşular



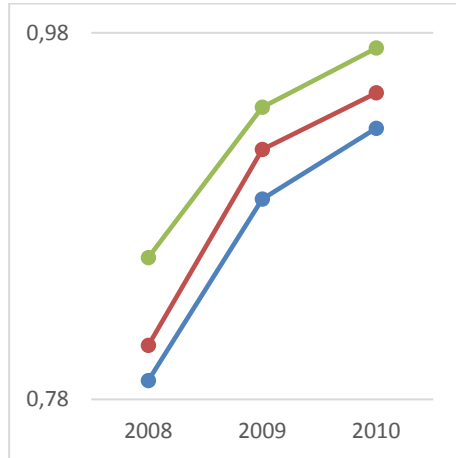
b) Adamic - Adar



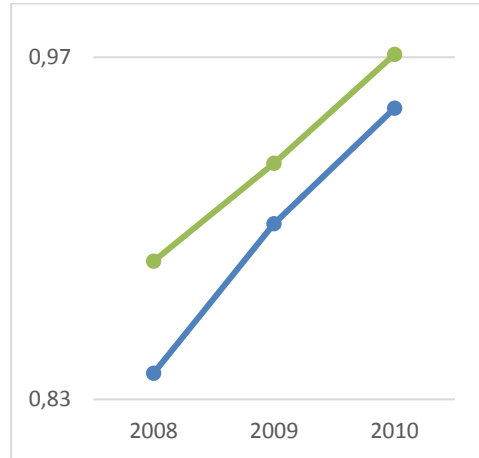
c) Jaccard İndeks



d) Sorenson İndeks



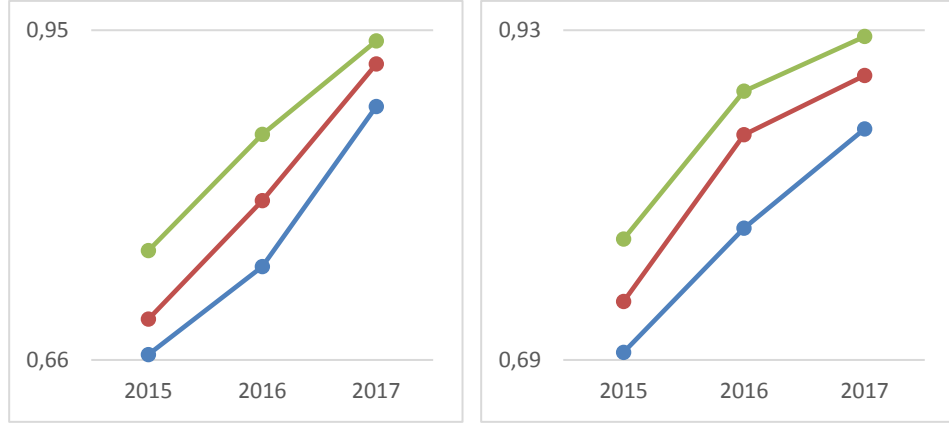
e) Tercihli Bağlanma



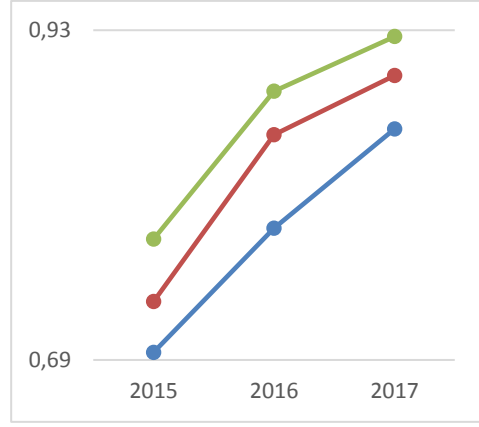
f) Kaynak Paylaşırma

—●— Ağırksız —●— Ağırıklı —●— Düğüm Ağırıklı

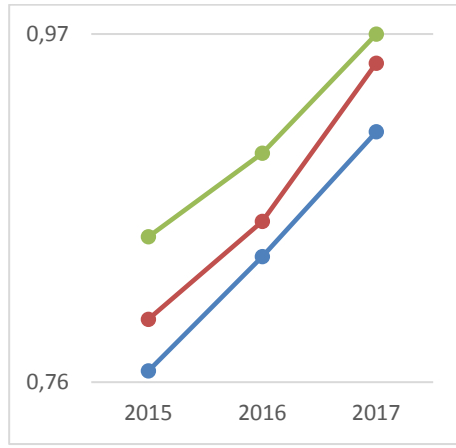
Şekil 7.4. 2000-2007 Arası bilimsel işbirliği açısından elde edilen AUC sonuçları.



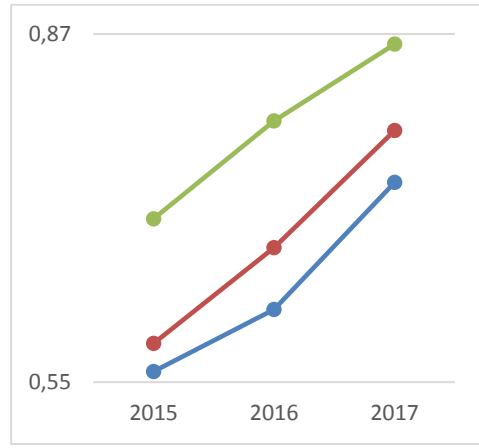
a) Ortak Komşular



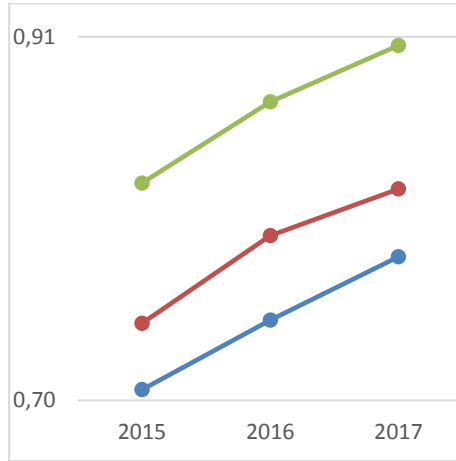
b) Adamic - Adar



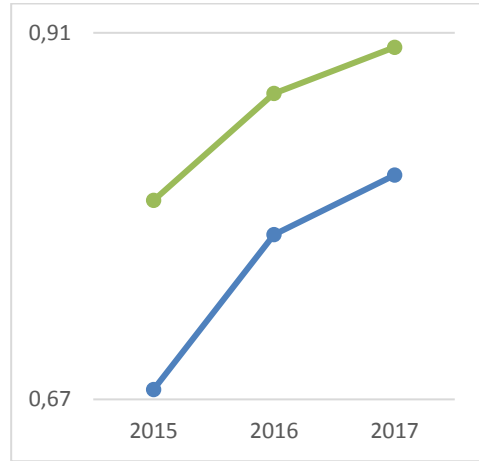
c) Jaccard İndeks



d) Sorenson İndeks



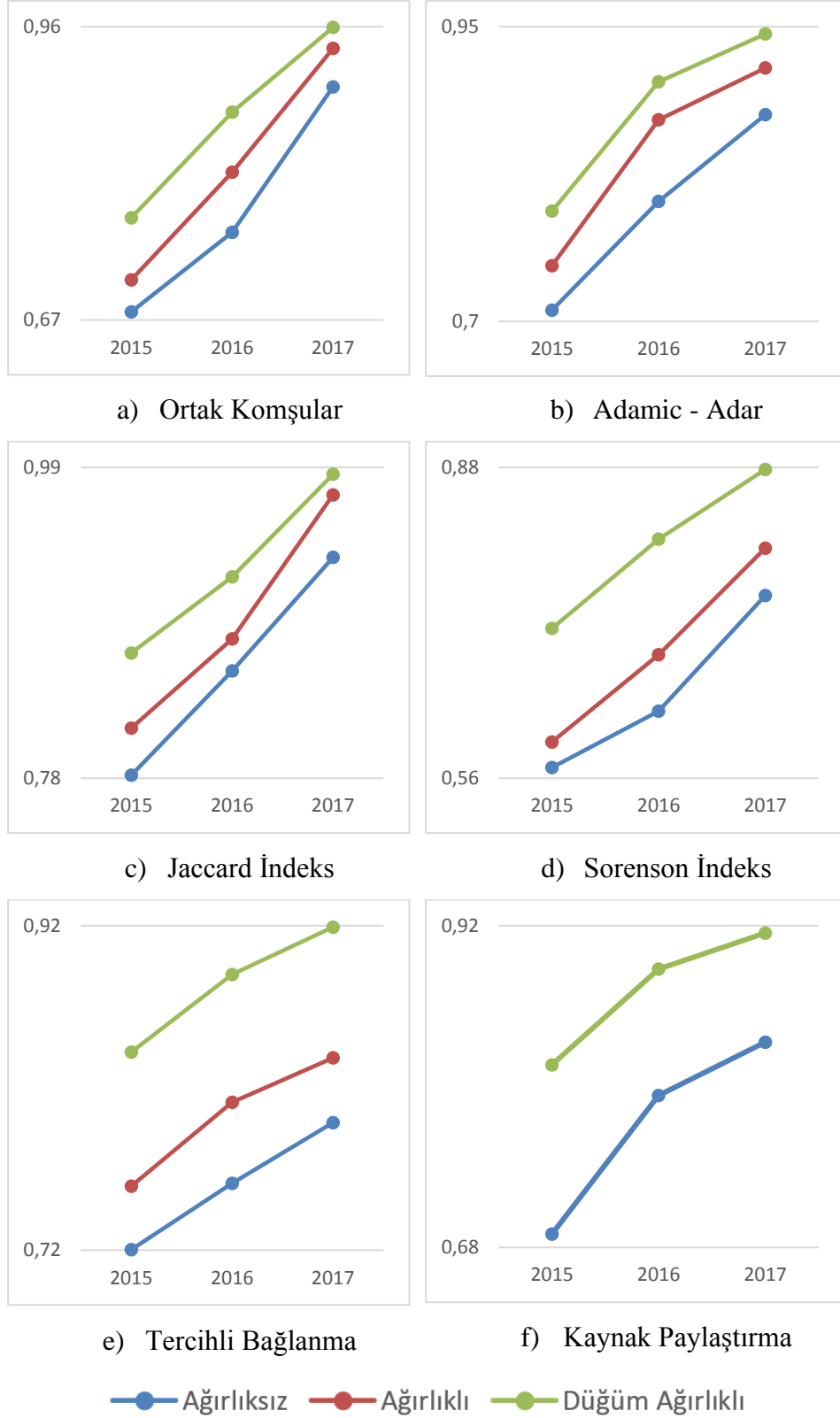
e) Tercihli Bağlanma



f) Kaynak Paylaşırma

—●— Ağırksız —●— Ağırıklı —●— Dğüm Ağırıklı

Şekil 7.5. 2000-2014 Arası oluşturulan tenis ağından elde edilen AUC sonuçları.



Şekil 7.6. 2000-2014 Arası bilimsel işbirliği açısından elde edilen AUC sonuçları.

Bağlantı tahmin sonuçlarına bakıldığında eğitim ağının genişlemesiyle birlikte tahmin başarılarındaki artış daha net olarak ortaya çıkmaktadır. Düğüm ağırlıklarının merkezilik ölçütleri ve zaman periyoduna bağlı kriterler kullanılarak hesaplandığı düğüm ağırlıklı bağlantı tahmininin geleneksel bağlantı tahmin yöntemlerinden önceki deneysel sonuçlarda olduğu gibi yine daha başarılı olduğu görülmektedir.

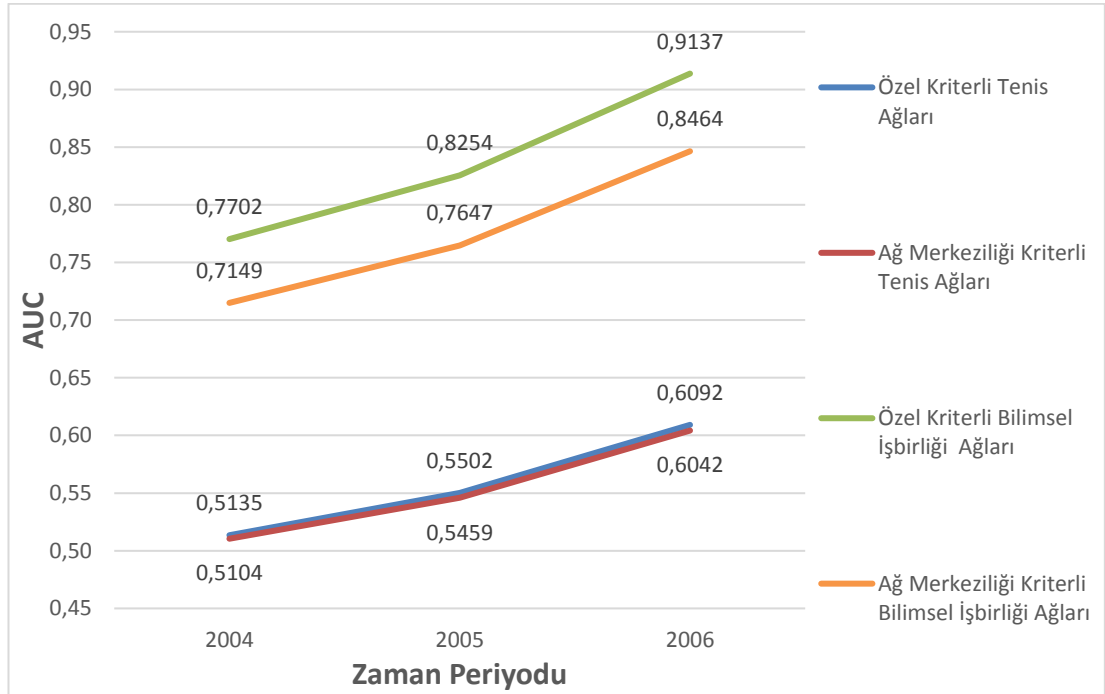
Tenis turnuvalarından oluşturulan ağlarda en iyi tahminlerin her iki veri seti için de NwJI tarafından gerçekleştirildiği görülmektedir.

Geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin başarısını değerlendirmek için Bölüm 6 ve 7’de açıklanan deneysel çalışmalarda farklı ağırlıklandırma kriterleri kullanılmış ve bu kriterlere göre düğüm ağırlıkları hesaplanarak bağlantı tahmin işlemleri gerçekleştirilmiştir. Bölüm 6’da ağırlıklandırma kriteri olarak tenis turnuvaları veri setinde zaman, deneyim, başarı, tur ve özvektör merkeziliği kriterleri, bilimsel işbirliği veri setinde ise zaman, deneyim, etkinlik ve özvektör merkeziliği kriterleri kullanılmıştır. Burada kriter belirleme süreci verisetinin içerisinde düğümlerin ağdaki gücünün ortaya çıkarabilecek faktörlerin belirlenmesi şeklinde gerçekleşmiştir. Bölüm 7’de gerçekleştirilen deneysel çalışmada ise düğüm ağırlıklandırma işlemi için ağların topolojik olarak yapılarından elde edilebilen ağ merkeziliği ölçütleri ve bunlara ek olarak zaman periyodu faktörü kriter olarak kullanılmıştır. Buradaki hedef düğüm ağırlıklı bağlantı tahmin yönteminde düğümlerin ağırlıklandırılması için kullanılacak kriterleri genelleştirerek yöntemin, zaman bilgisi içeren tüm ağlarda kullanılabilirliğini göstermek olmuştur. Her iki farklı kriter belirleme süreçlerinden elde edilen düğüm ağırlıklarının kullanıldığı bağlantı tahmin işlemlerinde de geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin mevcut bağlantı tahmin yöntemlerinden daha başarılı olduğu görülmektedir.

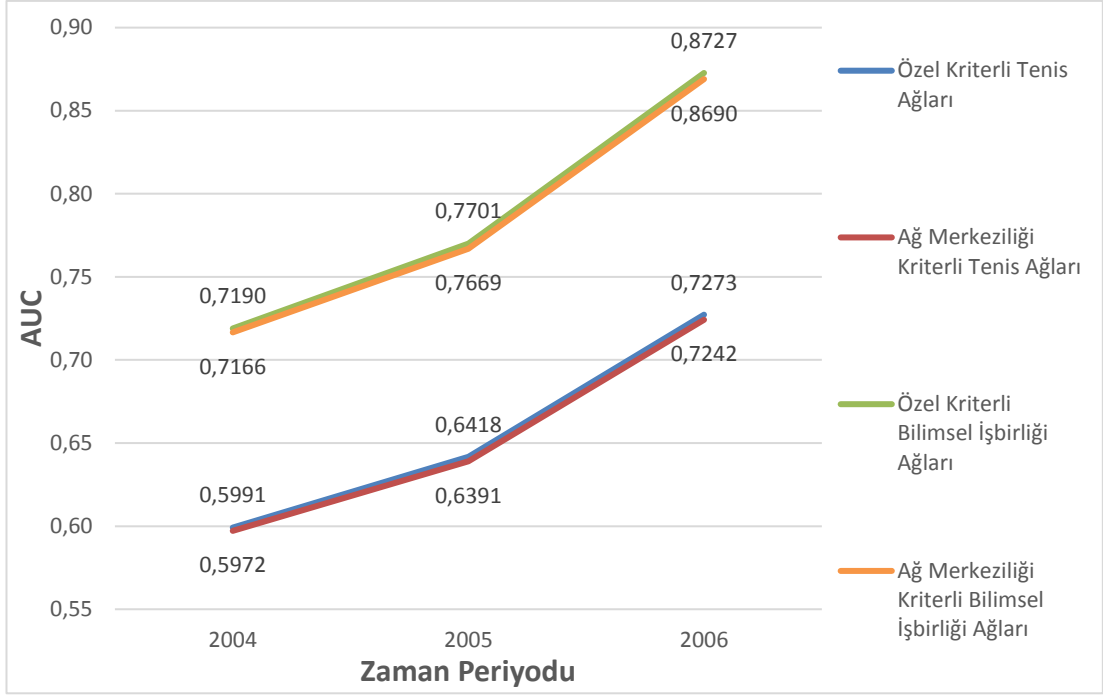
Genel olarak her iki kriter belirleme sürecine bağlı olarak elde edilen tahmin sonuçları başarılı olsa da kullanılan kriterlere göre gerçekleştirilen tahmin işlemlerinin AUC sonuçları karşılaştırıldığında Şekil 7.7 - 7.24’te görüldüğü gibi özel kriterler seçilerek yapılan düğüm ağırlıklandırma işlemlerine bağlı tahmin sonuçlarının daha başarılı oldukları görülmektedir. Ayrıca yalnızca topolojik bilgilerin ve zaman bilgilerinin elde

edilebildiği ağlarda da Bölüm 7'deki deneysel sonuçlarda görüldüğü gibi düğüm ağırlıklı bağlantı tahmin yöntemi mevcut yöntemlerden daha başarılı olmuştur.

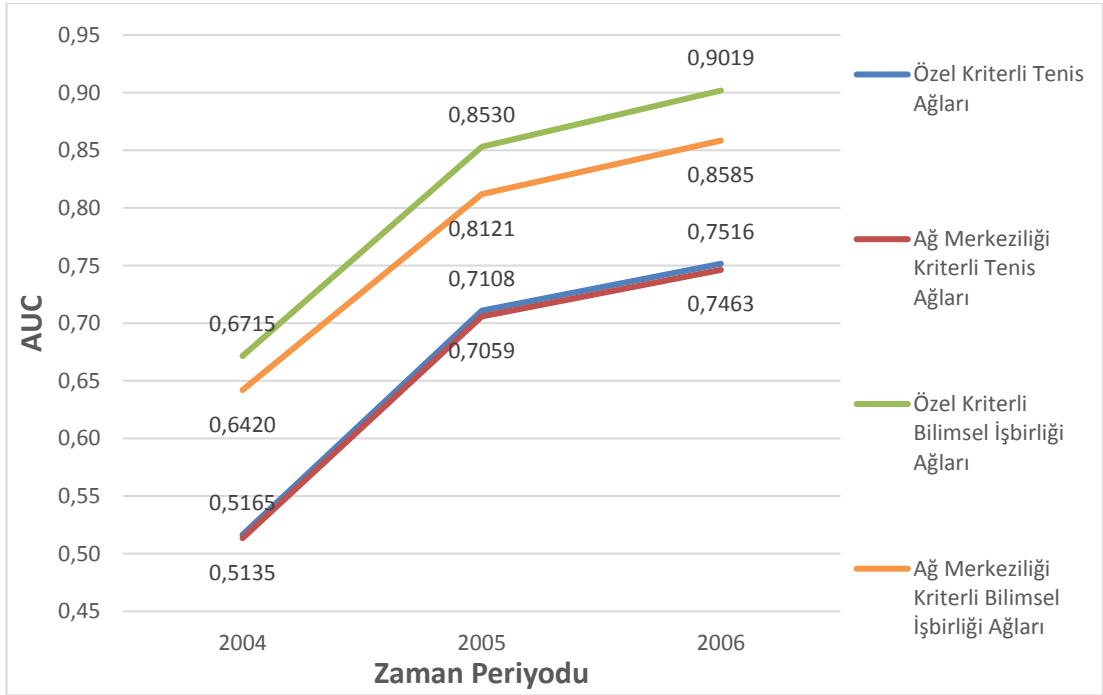
Yapılan deneysel çalışmalardan elde edilen bir diğer sonuç ise bağlantı tahmin işlemlerinde kullanılacak eğitim ve test verilerinin seçiminin önemi olmuştur. Şekil 7.1 ve 7.2'a bakıldığında bağlantı tahmin sonuçlarının çok tatmin edici bir başarı gösteremediği görülmektedir. Bunun sebebi ise çalışmalarda kullanılan eğitim ve test ağlarının zaman periyodunun dar olması başarılı bir bağlantı tahmini yapılmasının önüne geçmektedir. Şekil 7.3, 7.4, 7.5 ve 7.6'da görüldüğü gibi eğitim ve test için kullanılan verilerden oluşan ağlar geniş zaman periyodu içerisinde olursa bağlantı tahmin işlemleri de buna paralel olarak tatmin edici bir başarı göstermektedir. Özellikle, geliştirilen yöntemin başarısının geniş zaman periyoduna bağlı olarak oluşturulan ağlarda yüksek olduğu tespit edilmiştir. Bu yüzden bağlantı tahmin işlemlerinde kullanılacak eğitim ve test verilerinin başarılı bir tahmin işlemine uygun olacak şekilde belirlenmesi önemlidir.



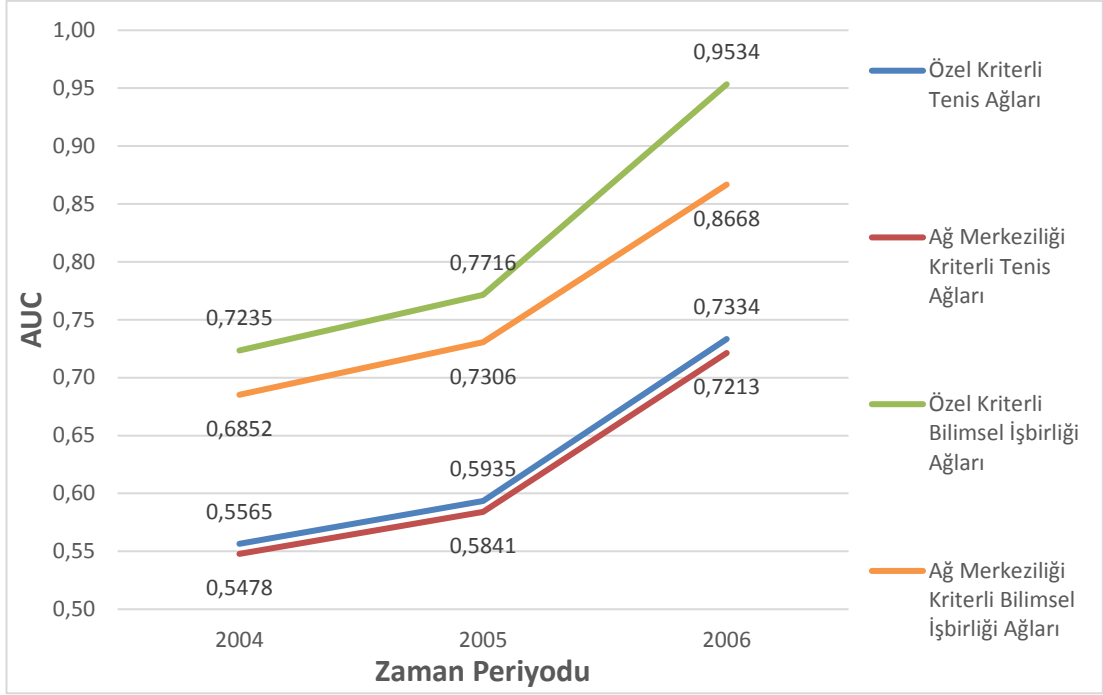
Şekil 7.7. 2000-2003 Arası ağlarda NwCN yönteminin AUC sonuçları.



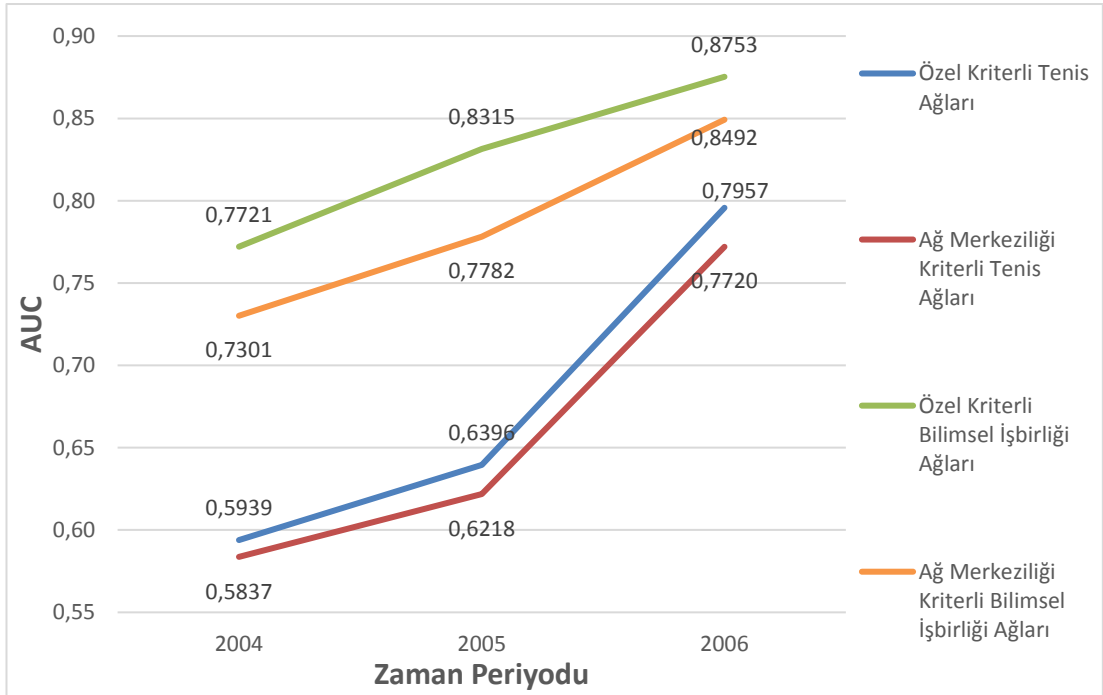
Şekil 7.8. 2000-2003 Arası ağlarda NwJI yönteminin AUC sonuçları.



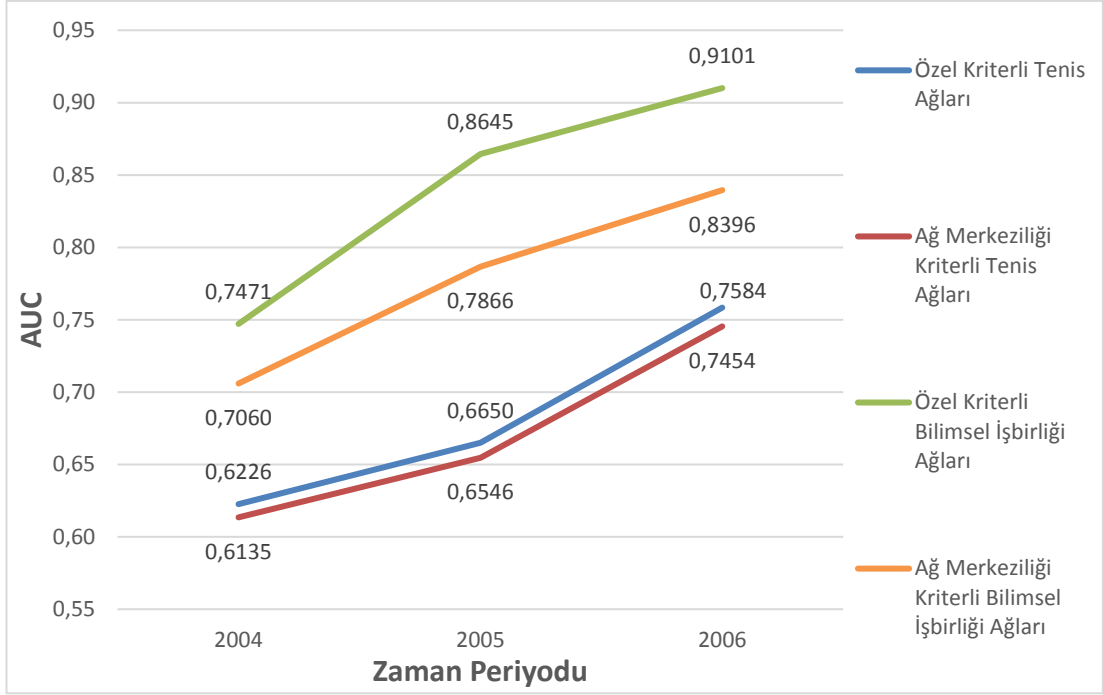
Şekil 7.9. 2000-2003 Arası ağlarda NwAA yönteminin AUC sonuçları.



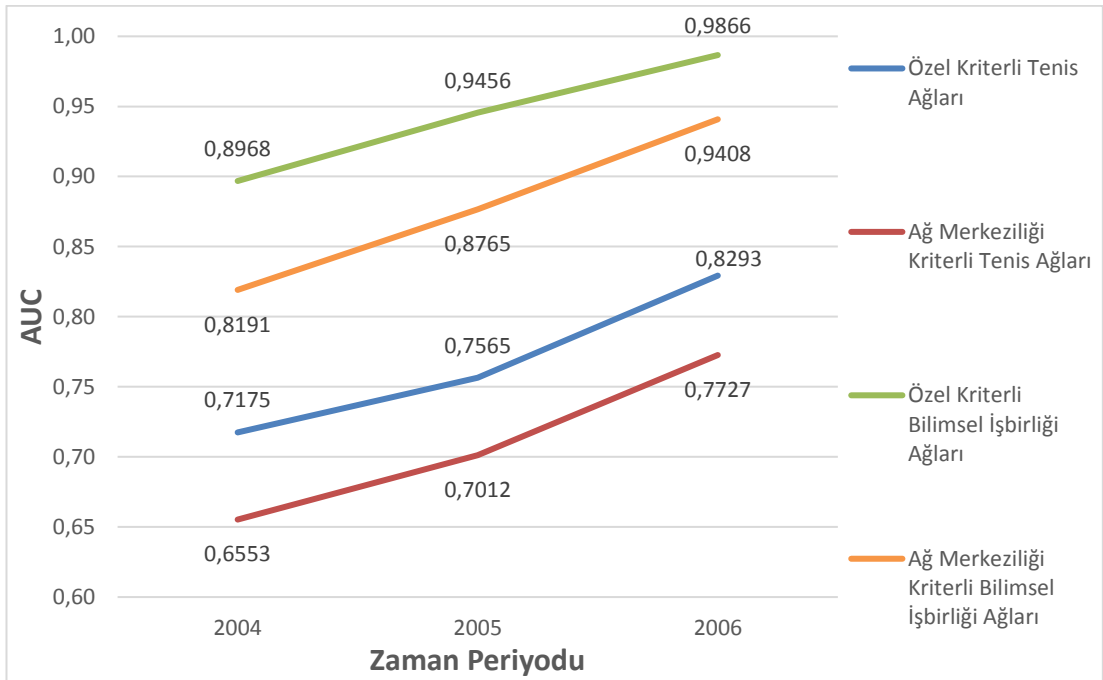
Şekil 7.10. 2000-2003 Arası ağlarda NwPA yönteminin AUC sonuçları.



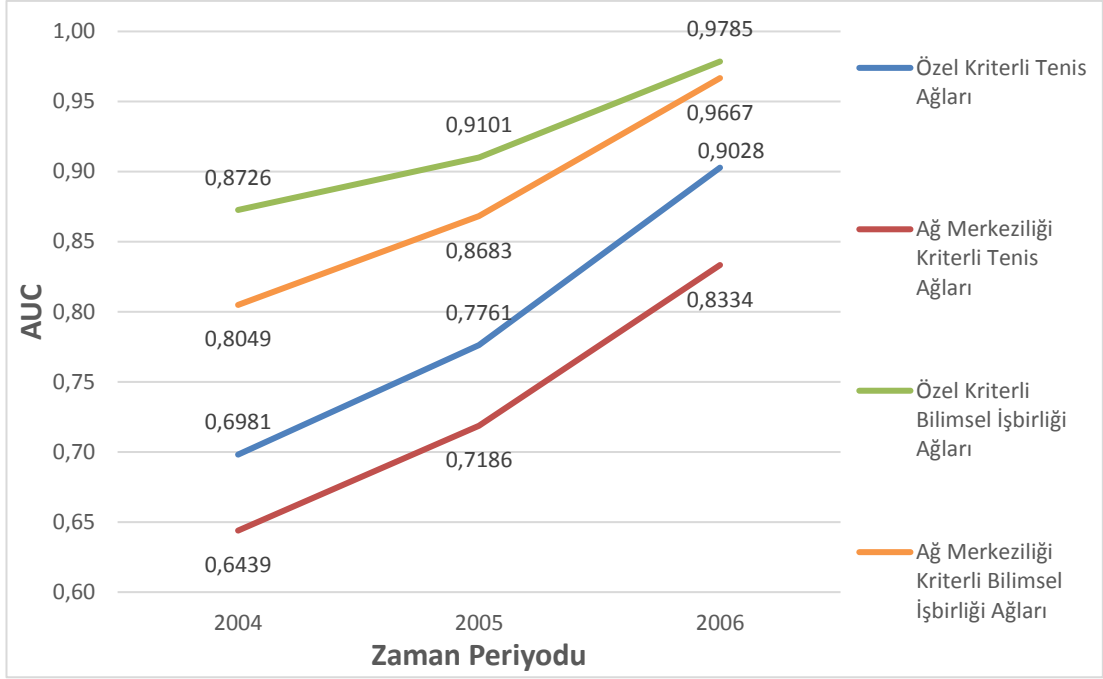
Şekil 7.11. 2000-2003 Arası ağlarda NwSI yönteminin AUC sonuçları.



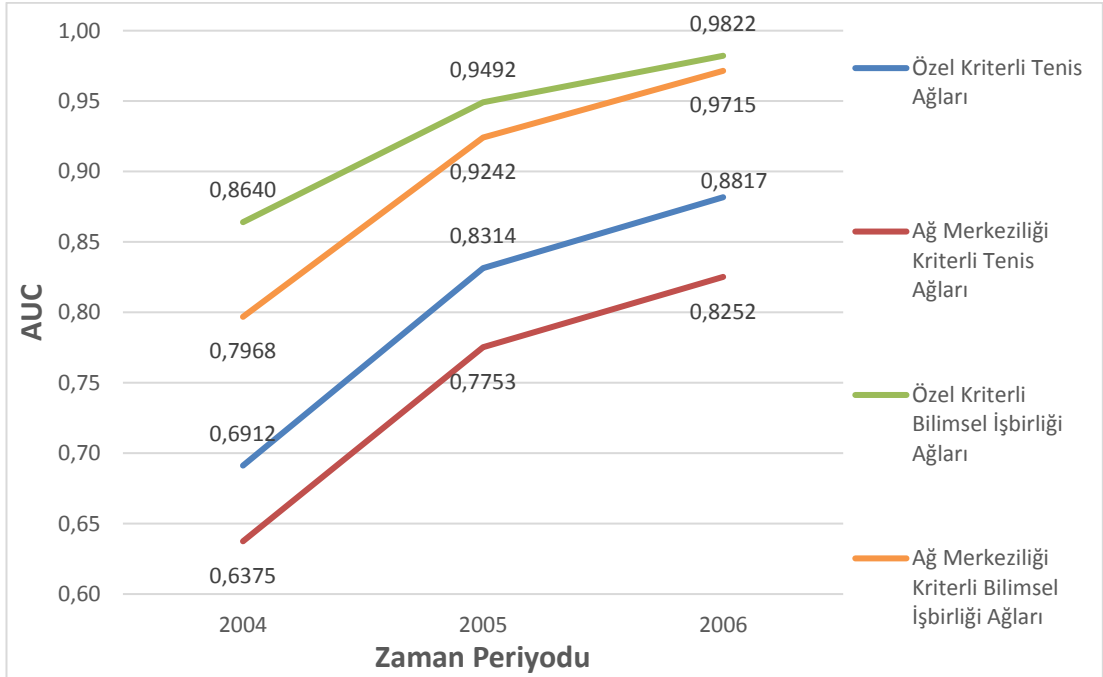
Şekil 7.12. 2000-2003 Arası ağlarda NwRA yönteminin AUC sonuçları.



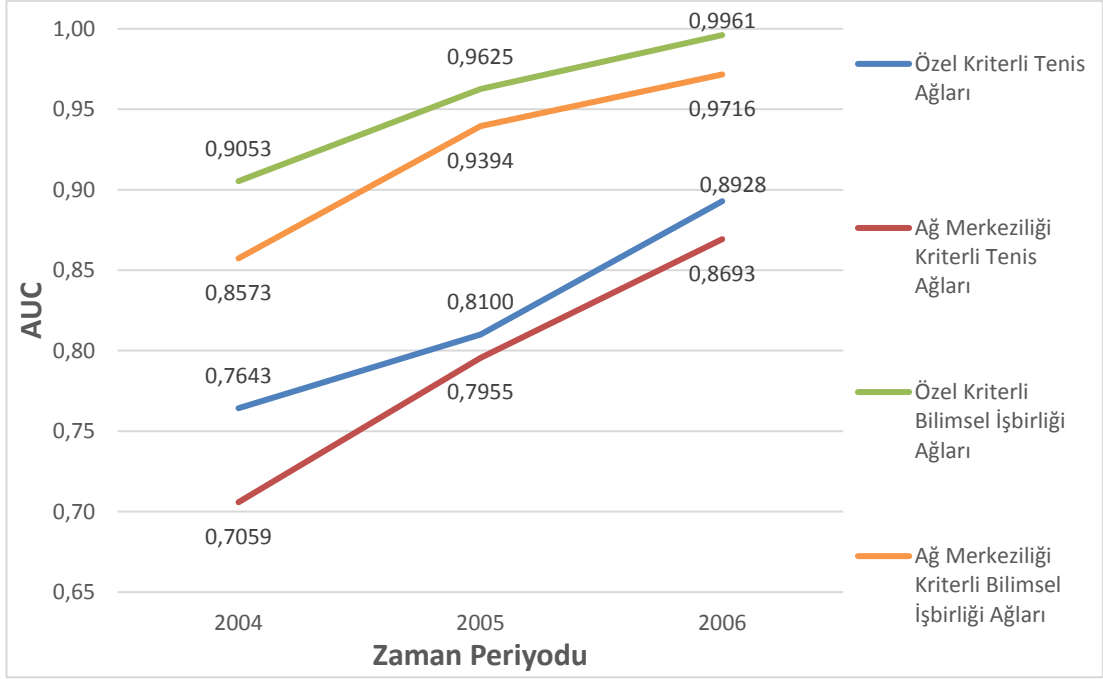
Şekil 7.13. 2000-2007 Arası ağlarda NwCN yönteminin AUC sonuçları.



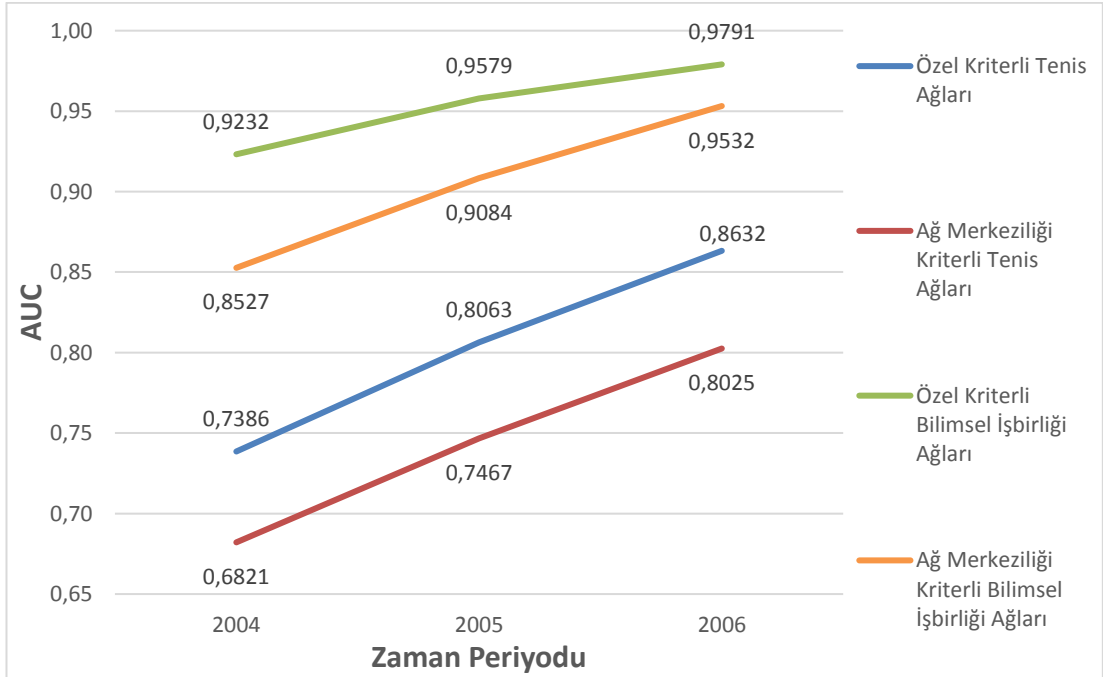
Şekil 7.14. 2000-2007 Arası ağlarda NwJI yönteminin AUC sonuçları.



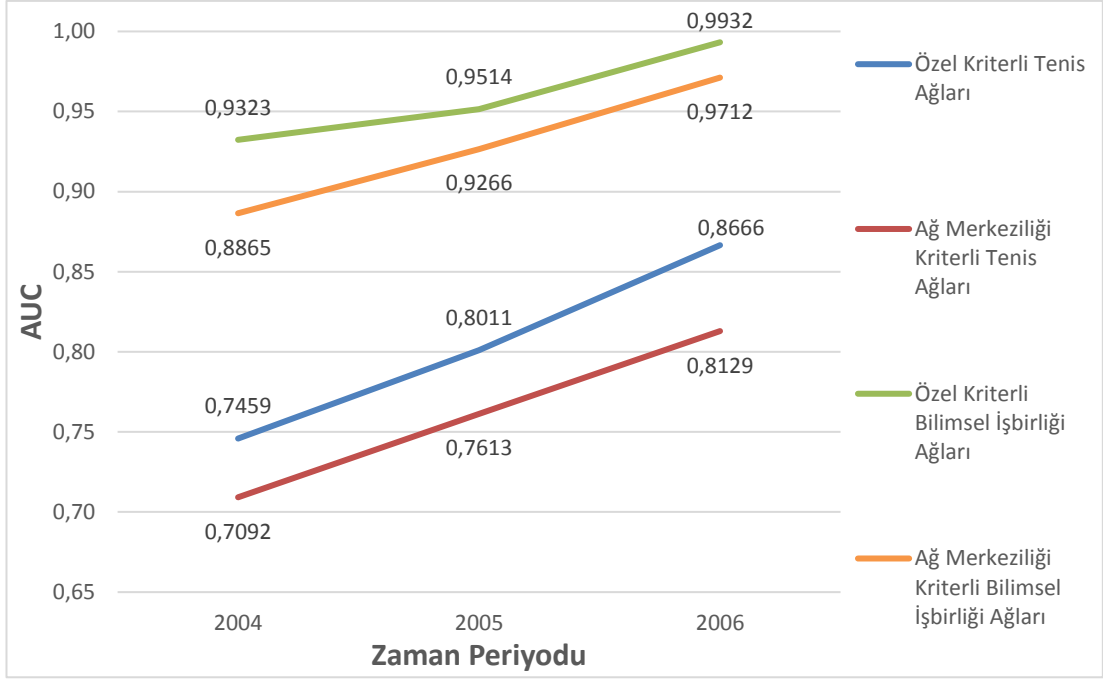
Şekil 7.15. 2000-2007 Arası ağlarda NwAA yönteminin AUC sonuçları.



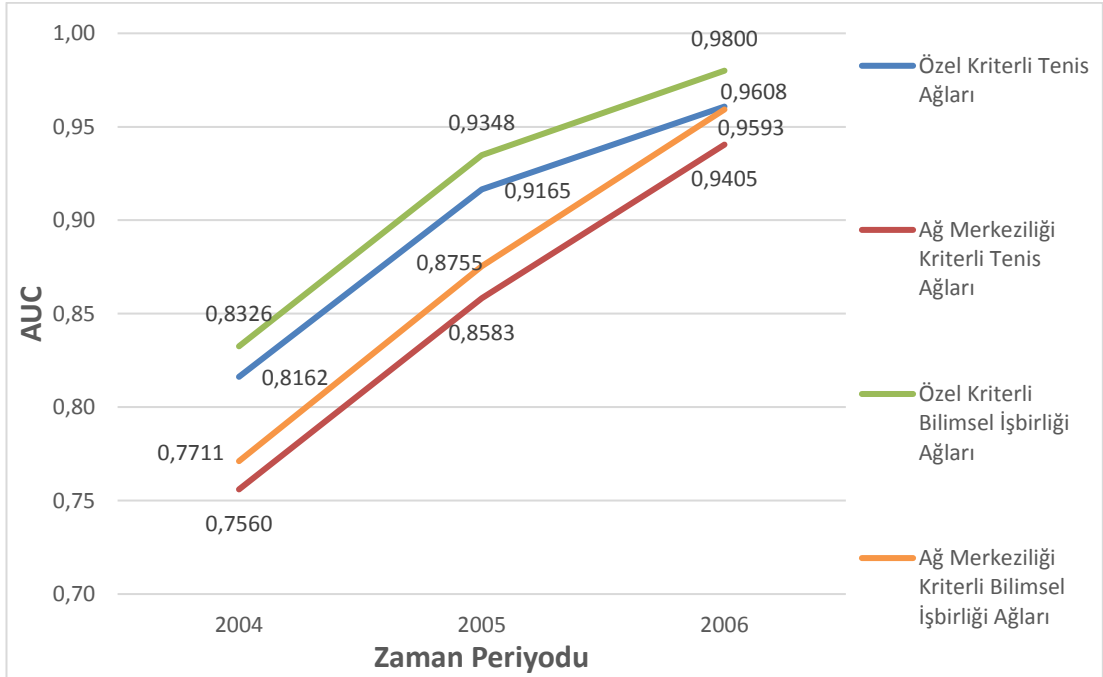
Şekil 7.16. 2000-2007 Arası ağlarda NwPA yönteminin AUC sonuçları.



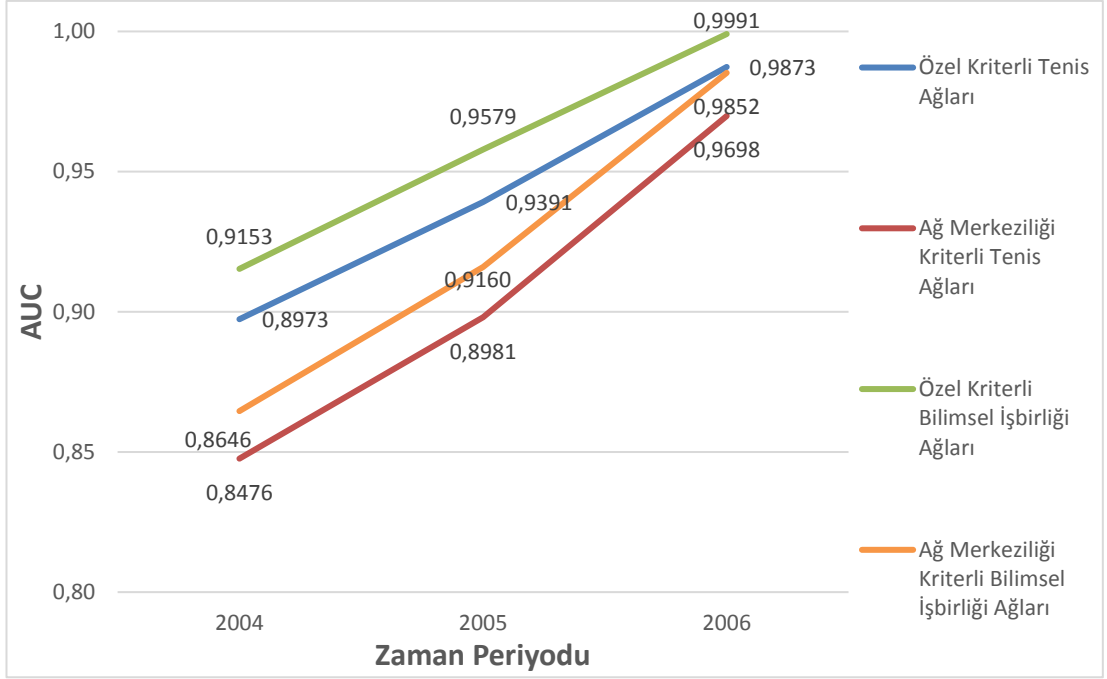
Şekil 7.17. 2000-2007 Arası ağlarda NwSI yönteminin AUC sonuçları.



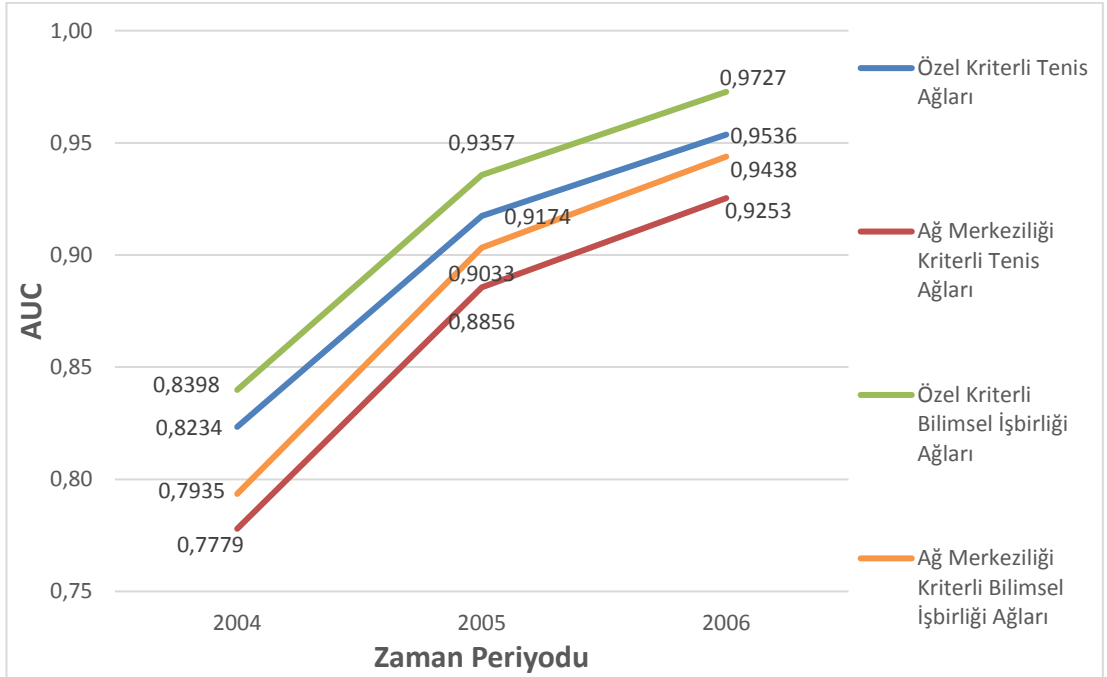
Şekil 7.18. 2000-2007 Arası ağlarda NwRA yönteminin AUC sonuçları.



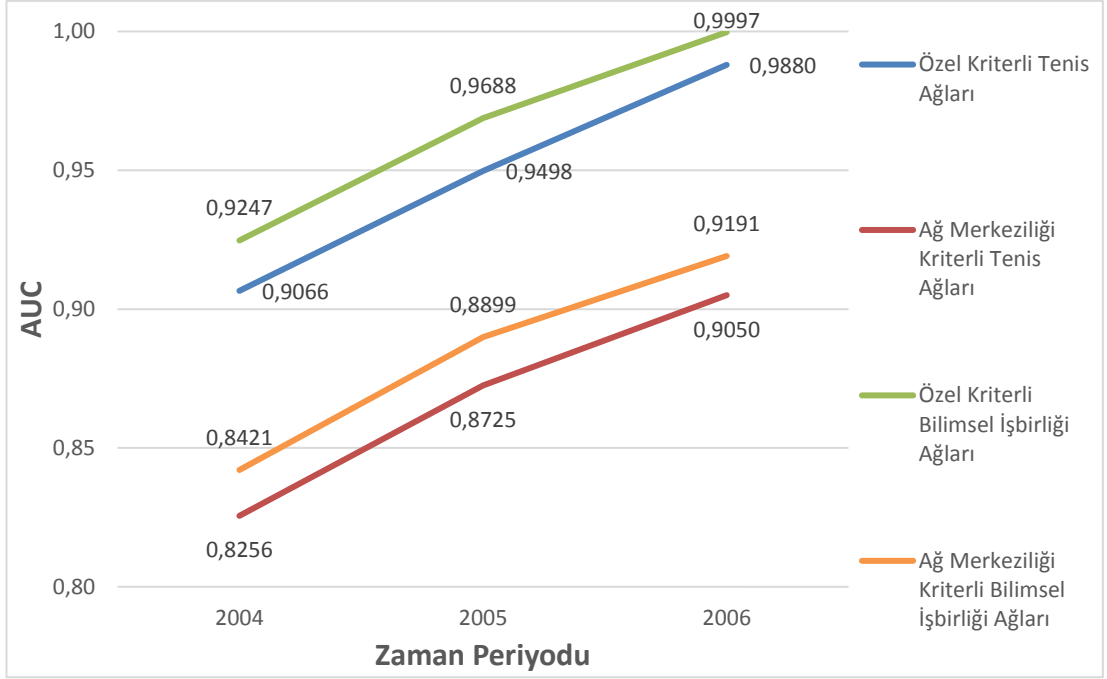
Şekil 7.19. 2000-2014 Arası ağlarda NwCN yönteminin AUC sonuçları.



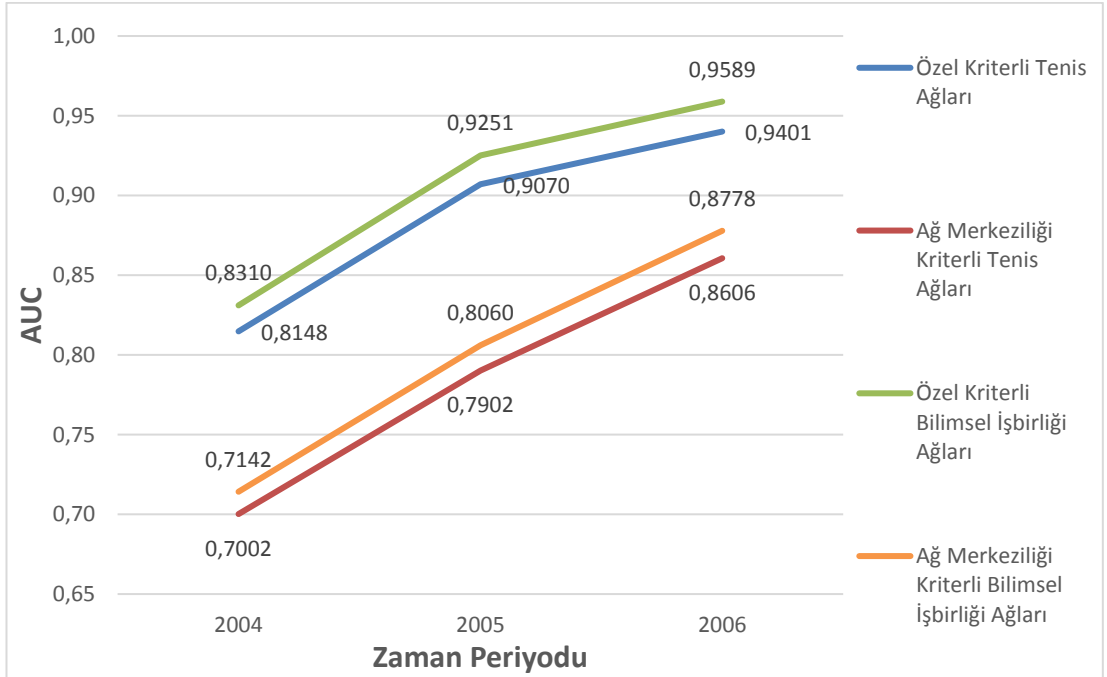
Şekil 7.20. 2000-2014 Arası ağlarda NwJI yönteminin AUC sonuçları.



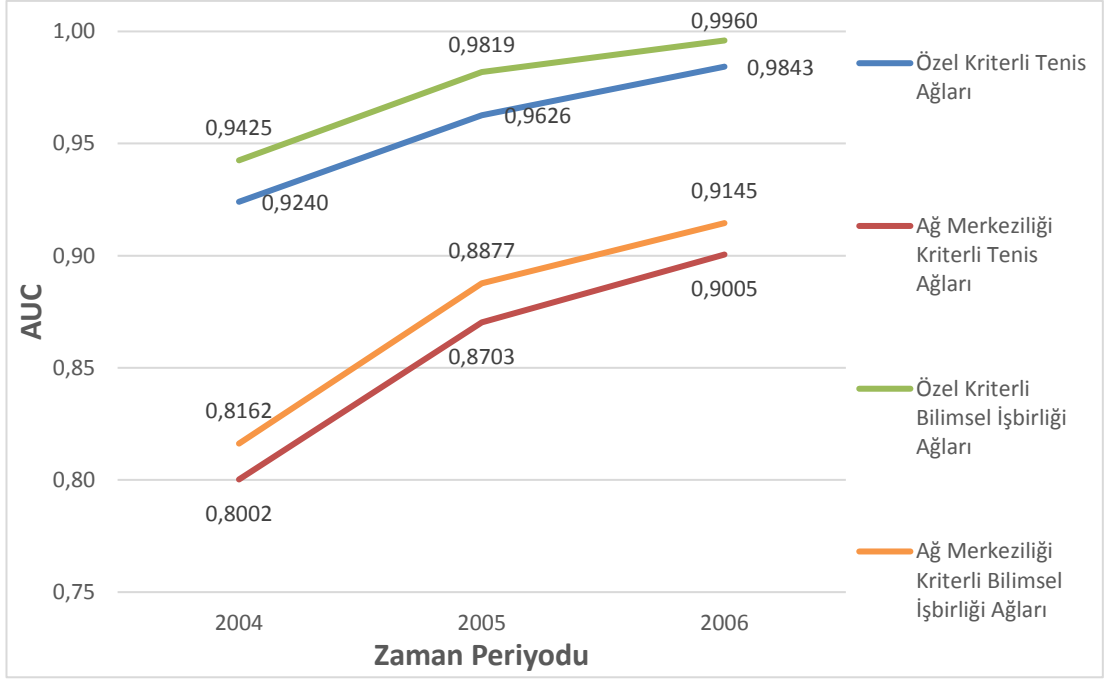
Şekil 7.21. 2000-2014 Arası ağlarda NwAA yönteminin AUC sonuçları.



Şekil 7.22. 2000-2014 Arası ağlarda NwPA yönteminin AUC sonuçları.



Şekil 7.23. 2000-2014 Arası ağlarda NwSI yönteminin AUC sonuçları.



Şekil 7.24. 2000-2014 Arası ağlarda NwRA yönteminin AUC sonuçları.

BÖLÜM 8

DÜĞÜM AĞIRLIKLARININ MERKEZİLİK ÖLÇÜTÜ OLARAK KULLANILMASI

Karmaşık ağlarda yaygın olarak kullanılan geleneksel ağ merkeziliği ölçütleri [97–99], ağın genel olarak topolojik yapısını dikkate alır ve topolojik bilgileri kullanarak ağda mevcut merkezi düğümleri tespit eder. Ancak zaman periyoduna bağlı olarak genişleyen ağlarda düğümlerin zaman içerisinde konumları değişkenlik göstermektedir. Ağda genel olarak etkin olan merkezi düğümlerin tespit edilmesinde özellikle zaman faktörünün dikkate alınması önemlidir. Düğüm ağırlıklandırma işlemi düğümlerin ağdaki konumunu belirlemede zaman faktörü dahil olmak üzere topolojik bilgilerle ifade edilemeyen faktörlerin etkisini dahil etmesi açısından önemli bir avantaj sağlamaktadır. Özellikle dinamik ağlarda, düğümlerin ağdaki konumunun değişimine etki eden faktörler değişkenlik gösterebilir. Bu nedenle ağdan ağa değişkenlik gösterebilecek faktörler düğüm ağırlıklandırma işleminde kullanılabilir.

8.1. DENEYSEL ÇALIŞMA İÇİN AĞLARIN OLUŞTURULMASI

Geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin bir parçası olan düğüm ağırlıklandırma işleminin karmaşık ağlarda, ağ merkeziliğinin tespitinde kullanılabilirliğini tespit etmek için geleneksel ağ merkeziliği ölçütleri ile karşılaştırmalar yapılmıştır. Deneysel çalışma için Bölüm 5.1’de verilen Avustralya Açık Tenis Turnuvalarına ait veriler kullanılmıştır. Tenis turnuvaları verileri ile zaman periyoduna bağlı olarak Çizelge 8.1’de verilen ağlar oluşturulmuş ve hem geleneksel merkezilik ölçütleri hem de düğüm ağırlıklandırma işlemi uygulanarak merkezi düğümler tespit edilmiştir. Geleneksel merkezilik ölçütlerinin hesaplanmasında Gephi[kaynak] uygulaması kullanılmıştır.

Düğümlerin ağırlıklandırılması işleminde Bölüm 6.3.1.'de açıklanan kriterlerden zaman, deneyim, başarı ve tur kriterleri kullanılmıştır. Tespit edilen kriterlere bağlı olarak APLOCO yöntemi uygulanmış ve düğüm ağırlıkları hesaplanmıştır.

Çizelge 8.1. Zaman periyoduna bağlı olarak oluşturulan ağlar.

Ağlar		
Ağ 1	Ağ 2	Ağ 3
2000-2003	2000-2010	2000-2017

8.2. 2000-2003 ARASI AĞLARDAN ELDE EDİLEN MERKEZİLİK SONUÇLARI

2000-2003 yılları arasında oynanan müsabakalardan oluşturulan ağda hesaplanan derece merkeziliğine göre sonuçlar Çizelge 8.2 ve düğüm ağırlığına göre merkezilik sonuçları Çizelge 8.3'te verilmiştir. Ayrıca bu düğümlerin hem geleneksel merkezilik ölçütlerine göre hem de düğüm ağırlıklandırma işlemi sonucuna göre sıralama bilgileri de Çizelge 8.2 ve 8.3'te görülmektedir.

Çizelge 8.2'te görüldüğü gibi zaman periyoduna bağlı olarak ağda aktif olan düğümlerin sıralaması, geleneksel merkezilik ölçütlerinin hesaplamalarına göre benzerlik göstermektedir. Bununla birlikte, geliştirilen düğüm ağırlıklandırma yöntemi ile diğer ölçütler arasındaki farkı gösteren iki durum vardır.

Birincisi, dar zaman aralığında derece merkeziliğine göre yapılan hesaplamada aynı dereceye sahip çok sayıda düğüm bulunması hangi düğümün daha önemli olduğunun belirlenmesini zorlaştırmaktadır. Çizelge 8.2'de görüldüğü gibi düğüm ağırlıklandırma yöntemi ile aynı dereceye sahip düğümler arasında en etkili düğümü bulmak mümkün olmaktadır.

Çizelge 8.2. 2000-2003 Yılları düğüm merkeziliği sıralaması.

	Sporcular	Derece Merkeziliği	Derece Merkeziliği Sırası	Arasındalık Merkeziliği	Arasındalık Merkeziliği Sırası	Yakındalık Merkeziliği	Yakındalık Merkeziliği Sırası	Özvektör Merkeziliği	Özvektör Merkeziliği Sırası	Düğüm Ağırlıklı Merkezilik	Düğüm Ağırlıklı Merkezilik Sırası
1	Agassi A.	21	1	0,02341	1	0,40067	1	1	1	1	1
2	Ferreira W.	18	2	0,01819	2	0,38436	2	0,79536	2	0,95957	2
3	Grosjean S.	16	3	0,01399	6	0,36589	8	0,69451	3	0,95737	5
4	Kafelnikov Y.	15	4	0,01497	4	0,37942	4	0,65618	4	0,95810	3
5	Safin M.	15	4	0,01563	3	0,37639	5	0,65615	5	0,95689	6
6	Schuettler R.	15	4	0,01287	7	0,37519	6	0,62722	6	0,95772	4
7	El Aynaoui Y.	14	5	0,01244	8	0,36085	10	0,56049	9	0,95520	7
8	Sampras P.	13	6	0,01437	5	0,36251	9	0,52604	10	0,83519	130
9	Clement A.	12	7	0,00977	11	0,36875	7	0,62220	12	0,83470	131
10	Escude N.	12	7	0,00915	13	0,35066	12	0,51350	11	0,95269	16
11	Federer R.	12	7	0,01183	10	0,34005	18	0,45128	14	0,95460	8
12	Henman T.	12	7	0,00820	16	0,34859	13	0,50104	12	0,83279	132
13	Hewitt L.	12	7	0,00774	18	0,34153	16	0,42067	16	0,95287	14
14	Johansson T.	12	7	0,00943	12	0,35811	12	0,56640	8	0,83579	129
15	Vinciguerra A.	12	7	0,01204	9	0,34705	14	0,46144	13	0,95269	13
16	Koubek S.	11	8	0,00612	23	0,33053	30	0,34856	28	0,95197	29
17	Lapentti N.	11	8	0,00801	17	0,33286	27	0,37349	23	0,95171	26
18	Novak J.	11	8	0,00548	26	0,33333	25	0,41682	17	0,95448	24
19	Rochus C.	11	8	0,00704	20	0,33956	20	0,42090	15	0,95171	19
20	Ferrero J.C.	10	9	0,00856	14	0,33810	21	0,32816	33	0,95320	20

İkinci fark ise geleneksel merkezilik ölçütleri ile elde edilen sonuçlarda önemli düğümler arasında gösterilen Sampras S., Clement A., Henman T., Johansson T. gibi sporcuların geliştirilen yöntemde sıralamanın gerisinde olmasıdır.

Bu durumun nedenine bakıldığında düğüm ağırlıklandırma yöntemi, merkezi düğümleri tespit ederken zaman kriteri olmak üzere düğümlerin ağdaki aktivitesi, etkinliği ve sürekliliği gibi kriterleri dikkate alarak daha hassas bir hesaplama yapılmasına imkân vermektedir. Geleneksel merkezilik ölçütleri, yalnızca topolojik bilgileri kullanarak düğümlerin merkeziliğini hesapladığından, zaman periyodunun başında ağda çok aktif olan düğümün daha sonraki zamanlarda pasif hale gelmesini

dikkate almaz. Çizelge 8.2’de görüldüğü gibi bu sorun geliştirilen yöntemle çözülmüştür.

Çizelge 8.3. 2000-2003 Yılları ağırlıklandırılmış düğüm sıralaması.

	Sporcular	Düğüm Ağırlıklı Merkezilik	Düğüm Ağırlıklı Merkezilik Sırası	Derece Merkeziliği	Derece Merkeziliği Sırası	Arasındalık Merkeziliği	Arasındalık Merkeziliği Sırası	Yakındalık Merkeziliği	Yakındalık Merkeziliği Sırası	Özvektör Merkeziliği	Özvektör Merkeziliği Sırası
1	Agassi A.	1	1	21	1	0,40067	1	1	1	1	1
2	Ferreira W.	0,95957	2	18	2	0,38436	2	0,79536	2	0,95957	2
3	Kafelnikov Y.	0,95810	3	15	4	0,36589	4	0,69451	3	0,95737	4
4	Schuetzler R.	0,95772	4	15	4	0,37942	7	0,65618	5	0,95810	6
5	Grosjean S.	0,95737	5	16	3	0,37519	6	0,62722	7	0,95772	3
6	Safin M.	0,95689	6	15	4	0,37639	3	0,65610	4	0,95689	5
7	El Aynaoui Y.	0,95520	7	14	5	0,36085	8	0,56050	9	0,95520	9
8	Federer R.	0,95460	8	12	7	0,34005	10	0,45128	17	0,9546	14
9	Novak J.	0,95448	9	11	8	0,34153	26	0,42067	24	0,95287	17
10	Roddick A.	0,95424	10	8	11	0,34705	36	0,46144	27	0,95269	29
11	Ferrero J.C.	0,95320	11	10	9	0,35066	14	0,51350	20	0,95269	33
12	Youzhny M.	0,95302	12	10	9	0,33333	28	0,41682	33	0,95448	30
13	Ancic M.	0,95290	13	4	15	0,33053	53	0,34856	78	0,95197	143
14	Hewitt L.	0,95287	14	12	7	0,33286	18	0,37349	15	0,95171	16
15	Nalbandian D.	0,95282	15	7	12	0,33956	50	0,42090	38	0,95171	35
16	Escude N.	0,95269	16	12	7	0,33810	13	0,32816	11	0,95320	11
17	Vinciguerra A.	0,95269	17	12	7	0,32596	9	0,34276	13	0,95302	13
18	Hrbaty D.	0,95216	18	10	9	0,32914	35	0,39151	30	0,95216	19
19	Moya C.	0,95213	19	8	11	0,33380	22	0,35984	37	0,95073	36
20	Koubek S.	0,95197	20	11	8	0,325517	23	0,331624	29	0,95196	28

Geliştirilen yönteme göre belirlenen en etkili 20 düğüm ve bu düğümlerin geleneksel merkezilik ölçütlerindeki sırası verilen Çizelge 8.3’te geliştirilen yöntem, zaman periyodu içerisinde etkin olan düğümlerin ortaya çıkarılmasında geleneksel merkezilik ölçüleriyle benzer sonuçlar vermektedir. Zaman aralığının başında ağda herhangi bir aktivitesi olmayan, ancak daha sonra ağda aktif olan veya ağa daha sonra katılarak güçlü etkileşimleri olan düğümlerin, geleneksel merkezilik ölçütlerine göre önemli düğümler arasında gösterilmediği anlaşılmaktadır. Kullanılan veri setine bakıldığında

Roddick A., Youzhny M., Nalbandian D., Hrbaty D., Moya C. ve Koubek S. gibi sporcuların ağı daha sonra katıldıkları ve zaman periyodunun son dönemlerinde aktif oldukları görülmektedir.

Geliştirilen yöntemin, zaman periyodu içinde sürekli etkileşim halinde olan düğümlere önem vermesi, bu düğümlerin aktif düğümler olarak tanımlanmasını mümkün kılmaktadır. Bir diğer dikkat çekici sonuç ise Ancic M. isimli sporcunun zaman periyodunun son yılında ağı katılması ve ağı dahil olduğu andan itibaren ağıda güçlü etkileşimlere sahip olmasından dolayı etkili bir düğüm olarak gösterilmesidir. Geleneksel merkezilik ölçütleri, merkeziliği yalnızca topolojik bilgileri kullanarak hesapladıkları için zaman periyodunun son yılında ağı dahil olan bu düğümün ağıdaki etkinliğini dikkate almamaktadır. Bununla birlikte, geliştirilen yöntem, bu düğümün zaman periyodunun son döneminde ağı dahil edildiğini ve ağıda aktif olduğunu dikkate alır. Aslında geliştirilen yöntem ağıdaki düğümlerin gücünü hesaplarken, gelecekte zaman içinde aktif olabilecek düğümleri öne çıkararak analiz yapılmasını mümkün kılmaktadır.

8.3. 2000-2010 ARASI AĞLARDAN ELDE EDİLEN MERKEZİLİK SONUÇLARI

2000-2010 yılları arasında oynanan müsabakalardan oluşturulan ağıda hesaplanan derece merkeziliğine göre sonuçlar Çizelge 8.4 ve düğüm ağırlıklığına göre merkezilik sonuçları Çizelge 8.5'te verilmiştir. Ayrıca bu düğümlerin hem geleneksel merkezilik ölçütlerine göre hem de düğüm ağırlıklandırma işlemi sonucuna göre sıralama bilgileri de Çizelge 8.4 ve 8.5'te görülmektedir.

Çizelge 8.4'te görüldüğü gibi zaman periyodu içerisinde ağıda sürekli aktif olan düğümlerin sıralaması geleneksel merkezilik ölçütleri arasında benzerlik göstermektedir. Ancak, geleneksel merkezilik ölçütlerinin önemli düğümler arasında gösterdiği Safin S., Agassi A., Nalbandian D., Hrbaty D., Johansson T. gibi bazı düğümlerin geliştirilen düğüm ağırlıklandırma yönteminde önemli düğümler arasında olmadığı görülmektedir. Bunun nedeni geliştirilen yöntemin zaman periyoduna göre düğüm hareketlerini dikkate alarak merkezilik hesaplaması yapmasıdır. Burada, ağıda

güçlü düğümlerin etkinliğinin zaman içinde azalmasını dikkate alarak merkezilik hesaplaması gerçekleştiren düğüm ağırlıklandırma yönteminin geleneksel merkezilik ölçütlerinden üstünlüğü görülmektedir.

Çizelge 8.4. 2000-2010 Yılları düğüm merkeziliği sıralaması.

	Sporcular	Derece Merkeziliği	Derece Merkeziliği Sırası	Arasındalık Merkeziliği	Arasındalık Merkeziliği Sırası	Yakındalık Merkeziliği	Yakındalık Merkeziliği Sırası	Özvektör Merkeziliği	Özvektör Merkeziliği Sırası	Düğüm Ağırlıklı Merkezilik	Düğüm Ağırlıklı Merkezilik Sırası
1	Federer R.	48	1	0,05948	1	0,44820	1	1	1	1,03977	1
2	Roddick A.	39	2	0,03989	2	0,42723	2	0,75637	2	0,99590	2
3	Grosjean S.	34	3	0,03190	4	0,40977	6	0,54749	8	0,99171	6
4	Safin M.	33	4	0,03226	3	0,42590	3	0,69479	3	0,94179	129
5	Hewitt L.	32	5	0,02892	5	0,41306	5	0,64125	4	0,99158	7
6	Agassi A.	30	6	0,02354	8	0,40254	9	0,50818	12	0,79992	253
7	Nalbandian D.	30	6	0,02423	6	0,41473	4	0,63909	5	0,93965	130
8	Blake J.	29	7	0,02108	9	0,40492	8	0,58210	6	0,99061	11
9	Davydenko N.	29	7	0,02096	10	0,40612	7	0,52361	9	0,99081	10
10	Ferrero J.C.	27	8	0,02052	11	0,39217	18	0,41635	19	0,98981	16
11	Haas T.	27	8	0,02376	7	0,39443	16	0,51804	10	0,99101	9
12	Gonzalez F.	26	9	0,01745	13	0,40176	11	0,55366	7	0,98991	15
13	Hrbaty D.	25	10	0,01678	16	0,38519	24	0,39971	22	0,93706	131
14	Nadal R.	25	10	0,01628	18	0,39367	17	0,47442	17	0,99522	3
15	Youzhny M.	25	10	0,01898	12	0,39980	13	0,48560	15	0,98807	20
16	Johansson T.	24	11	0,01630	17	0,38555	22	0,39636	23	0,89263	173
17	Robredo T.	24	11	0,01526	22	0,39105	19	0,47965	16	0,98721	25
18	Schuettler R.	24	11	0,01622	19	0,39902	14	0,50617	13	0,98732	23
19	Ferrer D.	22	12	0,01301	26	0,38375	25	0,38776	25	0,98833	18
20	Nieminen J.	22	12	0,01722	14	0,38957	20	0,41016	21	0,98719	26

Kullanılan veri setine bakıldığında Nadal R., Djokovic N., Tsonga JW, Baghdatis M., Cilic M., Kubot L., Murray A. gibi sporcuların ağa 2004 yılında katılmış olmaları ve ağda çok fazla etkileşim kurmuş olmalarına rağmen geleneksel merkezilik ölçütlerinin bu durumu yeterince tespit edemedikleri Çizelge 8.4'te görülmektedir. Ancak geliştirilen yöntemin bu durumu dikkate alarak merkezilikleri belirlediği görülmektedir.

Ayrıca, ağdaki güçlü etkileşimleri nedeniyle geliştirilen yöntemde önemli düğümler arasında yer alan, 2000-2003 zaman aralığı içerisinde ağa 2002 yılında dahil olan ve Çizelge 8.3'te gösterilen Roddick A.'nın gelecekte daha güçlü etkileşimler

kurabileceği görülmektedir. Çizelge 8.4 ve 8.5'te görüldüğü gibi bu düğüm hem derece merkeziliğine hem de geliştirilen düğüm ağırlıklı merkezilik yöntemine göre yapılan sıralamalarda önemli düğümler arasında yer almaktadır. Bu sonuç, ağa zaman periyodunun belli bir kısmından sonra dahil olan ve ağda sürekli etkileşim halinde olan düğümlerin gelecekte daha güçlü etkileşimlere sahip olabileceğinin tespitini de teyit etmektedir.

Çizelge 8.5. 2000-2010 Yılları ağırlıklandırılmış düğüm sıralaması.

	Sporcular	Düğüm Ağırlıklı Merkezilik	Düğüm Ağırlıklı Merkezilik Sırası	Derece Merkeziliği	Derece Merkeziliği Sırası	Arasındalık Merkeziliği	Arasındalık Merkeziliği Sırası	Yakındalık Merkeziliği	Yakındalık Merkeziliği Sırası	Özvektör Merkeziliği	Özvektör Merkeziliği Sırası
1	Federer R.	1,03977	1	48	1	0,05948	1	0,44820	1	1	1
2	Roddick A.	0,99590	2	39	2	0,03989	2	0,42723	2	0,75637	2
3	Nadal R.	0,99522	3	25	10	0,01628	18	0,39367	17	0,47442	17
4	Djokovic N.	0,99273	4	21	13	0,01602	20	0,37880	30	0,38820	24
5	Tsonga J.W.	0,99271	5	18	16	0,01131	34	0,38555	23	0,41521	20
6	Grosjean S.	0,99170	6	34	3	0,03190	4	0,40977	6	0,54749	8
7	Hewitt L.	0,99158	7	32	5	0,02892	5	0,41306	5	0,64125	4
8	Baghdatis M.	0,99131	8	20	14	0,01200	31	0,4017	10	0,49843	14
9	Haas T.	0,99101	9	27	8	0,02376	7	0,39443	16	0,51804	10
10	Davydenko N.	0,99081	10	29	7	0,02096	10	0,40612	7	0,52361	9
11	Blake J.	0,99061	11	29	7	0,02108	9	0,40492	8	0,58210	6
12	Cilic M.	0,99049	12	14	20	0,00728	56	0,36084	50	0,27558	39
13	Kubot L.	0,99013	13	4	30	0,00050	202	0,30354	212	0,06522	196
14	Murray A.	0,98996	14	14	20	0,00662	62	0,36468	43	0,25901	47
15	Gonzalez F.	0,98991	15	26	9	0,01745	13	0,4017	11	0,55366	7
16	Ferrero J.C.	0,98980	16	27	8	0,02052	11	0,39217	18	0,41635	19
17	Del Petro J.M.	0,98914	17	11	23	0,00431	86	0,34771	85	0,22800	60
18	Ferrer D.	0,98833	18	22	12	0,01301	26	0,38375	18	0,38776	19
19	Kohlschreiber P.	0,98819	19	15	19	0,00742	55	0,35553	64	0,23462	56
20	Youzhny M.	0,98807	20	25	10	0,01898	12	0,39980	13	0,48560	15

8.4. 2000-2017 ARASI AĞLARDAN ELDE EDİLEN MERKEZİLİK SONUÇLARI

2000-2017 yılları arasında oynanan müsabakalardan oluşturulan ağda hesaplanan derece merkeziliğine göre sonuçlar Çizelge 8.6 ve düğüm ağırlığına göre merkezilik sonuçları Çizelge 8.7'de verilmiştir. Ayrıca bu düğümlerin hem geleneksel merkezilik ölçütlerine göre hem de düğüm ağırlıklandırma işlemi sonucuna göre sıralama bilgileri de Çizelge 8.6 ve 8.7'de görülmektedir.

Çizelge 8.6'de görüldüğü gibi belirlenen zaman periyodu içerisinde ağda sürekli aktif olan düğümlerin sıralaması geleneksel merkezilik ölçütleri ile benzerlik göstermektedir.

Ancak zaman periyodunun başında güçlü etkileşim içinde olan ve zaman periyodunun sonunda etkileşimleri azalan veya sona eren Roddick A., Grosjean S., Davydenko N., Safin M., Nalbandian D. gibi sporcuların geleneksel merkezilik ölçütlerine göre hala güçlü düğümler olarak tespit edildiği görülmektedir. Ancak geliştirilen düğüm ağırlıklandırma yönteminde bu düğümler zayıf düğümler olarak gösterilmektedir. Ağları oluşturulduğu veri setine bakıldığında bu sporcuların 2010 'dan sonra turnuvalara katılımında azalma olduğu ya da katılımlarının sona erdiği anlaşılmaktadır.

Bununla birlikte geleneksel merkezilik ölçütlerinin yalnızca topolojik bilgilerle merkezilik hesaplaması yapmalarından dolayı bu sporcular hala önemli düğümler arasında sıralanmaktadır. Bu durum geliştirilen yöntemin zaman periyoduna bağlı olarak etkin düğümleri analiz etmede geleneksel merkezilik ölçütlerinden daha başarılı olduğunu göstermektedir.

Geliştirilen düğüm ağırlıklı merkezilik ölçütü, ağdaki sürekliliği ve etkileşimi azalan veya biten bu düğümlerin önemli düğümlerden çıkarılıp zayıf düğümler arasına dahil edilmesini sağlar. Bu süreç, düğüm etkileşimlerinin sürekliliğinin, zaman dilimine bağlı olarak genişleyen ağlarda güçlü ve zayıf düğümlerin tespitinde önemli bir faktör

olduğunu göstermektedir. Geliştirilen yöntem, ağdaki düğümün etkileşim sürekliliğine bağlı olarak ağın merkeziliğinin anlık olarak belirlenmesinde etkilidir. Çizelge 8.7’de geliştirilen yaklaşıma göre belirlenen en etkili 20 düğüm ve bu düğümlerin diğer merkezilik ölçülerindeki sıralaması gösterilmektedir. Çizelge 8.7’de görüldüğü gibi geliştirilen yöntem, ağda aktif olan düğümlerin ortaya çıkarılmasında geleneksel merkezilik ölçütlerine benzer sonuçlar vermektedir.

Çizelge 8.6. 2000-2017 Yılları düğüm merkeziliği sıralaması.

	Sporcular	Derece Merkeziliği	Derece Merkeziliği Sırası	Arasındalık Merkeziliği	Arasındalık Merkeziliği Sırası	Yakındalık Merkeziliği	Yakındalık Merkeziliği Sırası	Özvektör Merkeziliği	Özvektör Merkeziliği Sırası	Düğüm Ağırlıklı Merkezilik	Düğüm Ağırlıklı Merkezilik Sırası
1	Federer R.	71	1	0,11981	1	0,45937	1	1	1	1,04191	1
2	Ferrer D.	51	2	0,05485	2	0,42484	3	0,69245	4	0,99362	7
3	Djokovic N.	48	3	0,03754	7	0,42097	4	0,75037	3	1,00072	2
4	Murray A.	46	4	0,04392	4	0,41748	6	0,64049	5	0,99646	4
5	Berdych T.	45	5	0,04638	3	0,41250	8	0,58064	10	0,99378	6
6	Nadal R.	45	5	0,03702	8	0,42813	2	0,76217	2	0,99807	3
7	Roddick A.	43	6	0,04117	6	0,41969	5	0,59859	8	0,90419	192
8	Wawrinka S.	42	7	0,02983	14	0,40915	11	0,6364	6	0,99393	5
9	Hewitt L.	40	8	0,04237	5	0,41006	9	0,55660	11	0,99124	11
10	Tsonga J.W.	39	9	0,03408	10	0,40855	12	0,60250	7	0,96070	9
11	Lopez F.	35	10	0,03673	9	0,39402	21	0,39508	26	0,98926	23
12	Grosjean S.	34	11	0,03396	11	0,38154	41	0,30524	45	0,85334	232
13	Youzhny M.	34	11	0,03206	12	0,40615	13	0,46516	15	0,98859	26
14	Davydenko N.	34	11	0,02764	17	0,40526	14	0,42941	20	0,95691	133
15	Baghdatis M.	34	11	0,02048	31	0,41436	7	0,59814	9	0,99149	10
16	Safin M.	33	12	0,03048	13	0,40946	10	0,45579	17	0,83467	257
17	Monfils G.	33	12	0,02964	15	0,38128	43	0,36312	31	0,98958	20
18	Robredo T.	33	12	0,02378	24	0,40144	17	0,45307	18	0,98899	24
19	Nalbandian D.	32	13	0,02710	20	0,40497	15	0,47078	14	0,90153	199
20	Simon G.	31	114	0,02722	18	0,39798	18	0,46054	16	0,99017	13

Kullanılan veri setine bakıldığında 2009 yılında Kyrgios N., Janowicz J., Pospisil V., Jaziri M. gibi sporcular ağa katılmış ve ağa katıldıkları andan itibaren ağdaki etkileşim içerisinde olmalarına rağmen geleneksel merkezilik ölçütleri bu düğümleri ağın zayıf

düğümleri olarak tanımlamıştır. Geliştirilen düğüm ağırlıklandırma yönteminde bu düğümler, ağa dahil oldukları andan itibaren belirlenen kriterlere göre ağdaki etkileşimlerinin sürekliliğinden dolayı önemli düğümler olarak tespit edilmiştir.

Ayrıca 2004 yılında ağa dahil olan Nadal R., 2005 yılında ağa dahil olan Djokovic N. ve 2007 yılında ağa dahil olan Tsonga JW, Çizelge 8.5'te görüldüğü gibi 2000-2010 döneminde ağ içindeki güçlü etkileşimleri nedeniyle geliştirilen yöntemde önemli düğümler arasında gösterilmiş ve bu düğümlerin gelecekte daha güçlü etkileşimler kurabilecek bir düğümler oldukları tespit edilmiştir.

Çizelge 8.7. 2000-2017 Yılları ağırlıklandırılmış düğüm sıralaması

	Sporcular	Düğüm Ağırlıklı Merkezilik	Düğüm Ağırlıklı Merkezilik Sırası	Derece Merkeziliği	Derece Merkeziliği Sırası	Arasındalık Merkeziliği	Arasındalık Merkeziliği Sırası	Yakındalık Merkeziliği	Yakındalık Merkeziliği Sırası	Özvektör Merkeziliği	Özvektör Merkeziliği Sırası
1	Federer R.	1,04191	1	71	1	0,11981	1	0,45937	1	1	1
2	Djokovic N.	1,00072	2	48	3	0,03754	7	0,42097	4	0,75037	3
3	Nadal R.	0,99807	3	45	5	0,03702	8	0,42813	2	0,76217	2
4	Murray A.	0,99646	4	46	4	0,04392	4	0,41748	6	0,64049	5
5	Wawrinka S.	0,99393	5	42	7	0,02983	14	0,40915	11	0,6364	6
6	Berdych T.	0,99378	6	45	5	0,04638	3	0,41250	8	0,58064	10
7	Ferrer D.	0,99362	7	51	2	0,05485	2	0,42484	3	0,69245	4
8	Raonic M.	0,99211	8	30	15	0,01877	33	0,40261	16	0,50313	12
9	Tsonga J.W.	0,99201	9	29	16	0,01620	43	0,39124	27	0,47845	13
10	Baghdatis M.	0,99149	10	34	11	0,02048	31	0,41436	7	0,59814	9
11	Hewitt L.	0,99124	11	40	8	0,04237	5	0,41006	9	0,55660	11
12	Kyrgios N.	0,99100	12	11	34	0,00654	101	0,33718	144	0,12406	135
13	Simon G.	0,99017	13	31	14	0,02722	18	0,39798	18	0,46054	16
14	Kohlschreiber P.	0,98995	14	30	15	0,01864	34	0,38579	34	0,35805	33
15	Janowicz J.	0,98980	15	11	34	0,00578	108	0,33474	154	0,12728	131
16	Verdasco F.	0,98975	16	31	14	0,02105	27	0,39346	24	0,43210	19
17	Gasquet R.	0,98973	17	31	14	0,02456	22	0,39348	23	0,42167	22
18	Pospisil V.	0,98969	18	7	38	0,00064	259	0,32153	208	0,09760	172
19	Jaziri M.	0,98959	19	7	38	0,00239	175	0,30812	267	0,05841	244
20	Monfils G.	0,98958	20	33	12	0,02964	15	0,38128	43	0,36312	31

Çizelge 8.6 ve 8.7’de görüldüğü üzere bu düğümler hem geleneksel merkezilik ölçütleri hem de geliştirilen düğüm ağırlıklı merkezilik ölçütüne göre yapılan sıralamalarda önemli düğümler arasındadır. Bu durum, zaman periyodunun belli bir bölümünde ağa katılan ve ağda sürekli etkileşim halinde olan düğümlerin gelecekte daha güçlü etkileşimlere sahip olabileceğinin tespitini de teyit etmektedir.

Üç farklı zaman periyoduna bağlı olarak oluşturulan ağlarda yapılan deneysel çalışmalar neticesinde geliştirilen düğüm ağırlıklı merkezilik ölçütünün, zaman periyodunun başından sonuna kadar sürekli etkileşim halinde olan düğümlerin belirlenmesinde mevcut geleneksel merkezilik ölçütleriyle aynı sonuçları hesapladığı görülmektedir. Deneysel çalışmaların bir başka sonucu, geliştirilen yöntemin, zaman periyodunun başında ağda güçlü etkileşimleri olan ancak etkileşimleri zamanla azalan veya sona eren düğümleri tespit etmede geleneksel merkezilik ölçütlerinden daha başarılı olmasıdır. Ayrıca deneysel çalışmalar, geliştirilen yöntemin düğümlerin zaman içinde güçlü ve sürekli bağlantılar oluşturmasının ağ merkeziliği üzerindeki etkilerinin analizinde geleneksel merkezilik ölçütlerinden daha başarılı olduğunu göstermektedir. Deneysel sonuçlarda geliştirilen yöntemin bağlantı sayısı az olsa bile ilerde güçlü bağlantılar kurabilen düğümleri tespit edebildiği de görülmektedir.

BÖLÜM 9

SONUÇLAR

Bu tez çalışmasında karmaşık ağlarda bağlantı tahmin problemlerinin çözümü için komşuluk tabanlı bağlantı tahmininde kullanılmak üzere düğüm ağırlıklandırması ve ağırlıklandırılmış düğümlerle bağlantı tahmini gerçekleştiren Düğüm Ağırlıklı Bağlantı Tahmin Yöntemi geliştirilmiştir. Düğüm ağırlıklı bağlantı tahmini ile iki düğümün benzerlik oranı hesaplanırken ortak komşuların sayısından ziyade ortak komşuların ağdaki güçleri (ağırlıkları) dikkate alınmaktadır. Farklı zaman periyotlarından oluşturulan ağlar üzerinde, bağlantı tahmin çalışmalarında kullanılan geleneksel komşuluk tabanlı yöntemler ve geliştirilen düğüm ağırlıklı bağlantı tahmini yöntemi ile bağlantı tahmin işlemleri gerçekleştirilmiştir. Yapılan bağlantı tahmin işlemlerinin başarısı AUC metriği kullanılarak ölçülmüştür. Deneysel çalışmalarda düğüm ağırlıklandırma işleminde iki farklı kriter belirleme süreci uygulanmıştır. Birincisinde, kriterler verisetinin içerdiği bilgiler incelenerek belirlenmiş, ikincisinde ise ağların topolojik bilgileri kullanılarak elde edilmiştir. Her iki kriter belirleme sürecinde de zaman periyodu bir kriter olarak dahil edilmiştir. Elde edilen deneysel sonuçlarda geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin, geleneksel bağlantı tahmin yöntemlerine göre daha başarılı olduğu görülmüştür. Ayrıca deneysel çalışmalar kriter belirleme süreci için ağ merkeziliği ölçütleri olan düğüm merkeziliği, yakınlık merkeziliği, arasındalık merkeziliği ve özvektör merkeziliği ölçütlerinin zaman periyodu ile birlikte düğümlerin ağırlıklandırılmasında kriter olarak kullanılabileceklerini göstermiştir. Deneysel çalışmalardan çıkarılan bir başka sonuç ise düğümlerin birbirine bağlanma şeklinin, bağlantı tahmin işlemlerinde başarıyı etkileyen bir faktör olduğudur. Tenis turnuvalarında düğümler arasındaki bağlantılar kura ya da üst turlara geçildikçe oluşurken, bilimsel işbirliği ağında ise düğümler arasındaki bağlantılar tercihli bağlanma şeklinde gerçekleşmektedir. Bu durum bağlantı tahmininin başarısını etkileyen unsurlardır. Bu sonuçlar ışığında geliştirilen düğüm ağırlıklı bağlantı tahmin yönteminin, hem ağın gelişimini etkileyen başta

zaman periyodu olmak üzere topolojik bilgilerle ifade edilemeyen faktörler dikkate alınarak yapılacak bağlantı tahmin çalışmalarında hem de ağ merkeziliği ölçütleri dahil edilerek yapılacak bağlantı tahmin çalışmalarında kullanılabileceği görülmektedir. Çalışmanın bir diğer katkısı ise geliştirilen düğüm ağırlıklandırma işleminin ağ merkeziliğini tespit etmede bir ölçüt olarak kullanılabilmesidir. Düğüm ağırlıklandırma işlemi, zaman periyodunu ve ağın genişlemesine etki eden faktörleri dikkate alarak düğümlerin ağırlık değerlerini hesaplamakta ve ağdaki güçlü düğümleri tespit etmektedir. Zaman karmaşıklığı açısından önerilen yöntem, literatürdeki komşuluk tabanlı yöntemlerin algoritmasından farklı bir yapıda olmadığından dolayı zaman maliyeti mevcut yöntemlere benzerlik göstermiştir.

Çalışmaların devamında, geliştirilen yöntemin başarısı farklı alanlarda gerçek dünya ağları kullanılarak tespit edilmeye devam edilecektir. Ayrıca geliştirilen düğüm ağırlıklandırma yönteminin özellikle dinamik ağlarda merkezi düğümlerin tespitindeki başarısı da araştırılmaya devam edilecektir.

KAYNAKLAR

1. Findık O. and Özkaynak, E., "Complex Network Analysis Of Players In Tennis Tournaments", *Icatces 2018 Proceeding Book*, Karabük, 383–388 (2018).
2. Subbaraj, K. and Sundan, B., "MutatedSocioAgentSim (MSAS): semisupervised modelling of multiagent simulation to predict and detect the mutation in a camouflaged social network", *Turkish Journal Of Electrical Engineering & Computer Sciences*, 26 (2): 961–973 (2018).
3. Strogatz, S. H., "Exploring complex networks", *Nature*, 410 (6825): 268–276 (2001).
4. Liben-Nowell, D. and Kleinberg, J., "The link-prediction problem for social networks", *Journal Of The American Society For Information Science And Technology*, 58 (7): 1019–1031 (2007).
5. Lü, L. and Zhou, T., "Link prediction in complex networks: A survey", *Physica A: Statistical Mechanics And Its Applications*, 390 (6): 1150–1170 (2011).
6. Wang, P., Xu, B., Wu, Y., and Zhou, X., "Link prediction in social networks: the state-of-the-art", *Science China Information Sciences*, 58 (1): 1–38 (2015).
7. Hasan, M. Al and Zaki, M. J., "A Survey of Link Prediction in Social Networks", *Social Network Data Analytics*, Springer US, Boston, MA, 243–275 (2011).
8. Resnick, P. and Varian, H. R., "Recommender systems", *Communications Of The ACM*, 40 (3): 56–58 (1997).
9. Lü, L., Medo, M., Yeung, C. H., Zhang, Y.-C., Zhang, Z.-K., and Zhou, T., "Recommender systems", *Physics Reports*, 519 (1): 1–49 (2012).
10. Huang, Z., Li, X., and Chen, H., "Link prediction approach to collaborative filtering", *Proceedings Of The 5th ACM/IEEE-CS Joint Conference On Digital Libraries - JCDL '05*, 141–142 (2005).
11. Kleinberg, J., "Analysis of large-scale social and information networks", *Philosophical Transactions Of The Royal Society A: Mathematical, Physical And Engineering Sciences*, 371 (1987): 20120378 (2013).
12. Hasan, M., Chaoji, V., Salem, S., and Zaki, M., "Link Prediction Using Supervised Learning", *SDM'06: Workshop on Link Analysis Counterterrorism and Security*, (2006).

13. Hwang, S.-Y., Wei, C.-P., and Liao, Y.-F., "Coauthorship networks and academic literature recommendation", *Electronic Commerce Research And Applications*, 9 (4): 323–334 (2010).
14. Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J.-F., Dricot, A., Vazquez, A., Murray, R. R., Simon, C., Tardivo, L., Tam, S., Svrvzikapa, N., Fan, C., de Smet, A.-S., Motyl, A., Hudson, M. E., Park, J., Xin, X., Cusick, M. E., Moore, T., Boone, C., Snyder, M., Roth, F. P., Barabasi, A.-L., Tavernier, J., Hill, D. E., and Vidal, M., "High-Quality Binary Protein Interaction Map of the Yeast Interactome Network", *Science*, 322 (5898): 104–110 (2008).
15. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., and Barabási, A.-L., "The large-scale organization of metabolic networks", *Nature*, 407 (6804): 651–654 (2000).
16. Folino, F. and Pizzuti, C., "Link Prediction Approaches for Disease Networks", *ITBAM*, 99–108 (2012).
17. Gül, S., Kaya, M., and Kaya, B., "Predicting links in weighted disease networks", (2016). *In Computer and Information Sciences (ICCOINS)*, Scimago, 77-81 (2016).
18. Findik, O. and Özkaynak, E., "Link Prediction on Networks Created from UEFA European Competitions", *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 207–217 (2020).
19. Wang, W.-Q., Zhang, Q.-M., and Zhou, T., "Evaluating network models: A likelihood analysis", *EPL (Europhysics Letters)*, 98 (2): 28004 (2012).
20. Zhang, Q.-M., Xu, X.-K., Zhu, Y.-X., and Zhou, T., "Measuring multiple evolution mechanisms of complex networks", *Scientific Reports*, 5 (1): 10350 (2015).
21. Guimerà, R. and Sales-Pardo, M., "Missing and spurious interactions and the reconstruction of complex networks", *Proceedings Of The National Academy Of Sciences*, 106 (52): 22073–22078 (2009).
22. Getoor, L. and Diehl, C. P., "Link mining", *ACM SIGKDD Explorations Newsletter*, 7 (2): 3–12 (2005).
23. Huang, Z., "Link Prediction Based on Graph Topology: The Predictive Value of Generalized Clustering Coefficient", *SSRN Electronic Journal*, 1: 289–297 (2010).
24. Newman, M. E. J., "Clustering and preferential attachment in growing networks", *Physical Review E*, 64 (2): 025102 (2001).

25. Carminati, B., Ferrari, E., and Perego, A., "Rule-Based Access Control for Social Networks", *Lecture Notes in Computer Science*, 1734–1744 (2006).
26. Tan, F., Xia, Y., and Zhu, B., "Link Prediction in Complex Networks: A Mutual Information Perspective", *PLoS ONE*, 9 (9): e107056 (2014).
27. Murata, T. and Moriyasu, S., "Link Prediction based on Structural Properties of Online Social Networks", *New Generation Computing*, 26 (3): 245–257 (2008).
28. Adamic, L. A. and Adar, E., "Friends and neighbors on the Web", *Social Networks*, 25 (3): 211–230 (2003).
29. Zhou, T., Lü, L., and Zhang, Y.-C., "Predicting missing links via local information", *The European Physical Journal B*, 71 (4): 623–630 (2009).
30. Jaccard, P., "Etude de la distribution florale dans une portion des Alpes et du Jura", *Bulletin De La Societe Vaudoise Des Sciences Naturelles*, 37: 547–579 (1901).
31. Barabási, A.-L. and Albert, R., "Emergence of Scaling in Random Networks", *Science*, 286 (5439): 509–512 (1999).
32. SORENSEN, T. A., "A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons", *Biol. Skar.*, 5: 1–34 (1948).
33. Salton, G. and McGill, M., "Introduction to Modern Information Retrieval", *Mc Graw Hill*, New York, 448 (1983).
34. Allali, O., Magnien, C., and Latapy, M., "Link prediction in bipartite graphs using internal links and weighted projection", *2011 IEEE Conference On Computer Communications Workshops*, Shangai, 936–941 (2011).
35. Caiyan, D., Chen, L., and Li, B., "Link prediction in complex network based on modularity", *Soft Computing*, 21 (15): 4197–4214 (2017).
36. Park, K., Lai, Y.-C., and Ye, N., "Characterization of weighted complex networks", *Physical Review E*, 70 (2): 026109 (2004).
37. Murata, T. and Moriyasu, S., "Link Prediction of Social Networks Based on Weighted Proximity Measures", *IEEE/WIC/ACM International Conference on Web Intelligence (WI'07)*, Fremont, 85–88 (2007).
38. Lü, L. and Zhou, T., "Link prediction in weighted networks: The role of weak ties", *EPL (Europhysics Letters)*, 89 (1): 18001 (2010).
39. de Sa, H. R. and Prudencio, R. B. C., "Supervised link prediction in weighted networks", *The 2011 International Joint Conference on Neural Networks*,

San Jose, 2281–2288 (2011).

40. Bütün, E., Kaya, M., and Alhajj, R., "Extension of neighbor-based link prediction methods for directed, weighted and temporal social networks", *Information Sciences*, 463–464: 152–165 (2018).
41. Liu, J. and Deng, G., "Link prediction in a user–object network based on time-weighted resource allocation", *Physica A: Statistical Mechanics And Its Applications*, 388 (17): 3643–3650 (2009).
42. Moradabadi, B. and Meybodi, M. R., "Link prediction in weighted social networks using learning automata", *Engineering Applications Of Artificial Intelligence*, 70 (November 2017): 16–24 (2018).
43. Aslan, S. and Kaya, B., "Time-aware link prediction based on strengthened projection in bipartite networks", *Information Sciences*, 506: 217–233 (2020).
44. Zhu, B. and Xia, Y., "Link Prediction in Weighted Networks: A Weighted Mutual Information Model", *PLOS ONE*, 11 (2): e0148265 (2016).
45. Barrat, A., Barthelemy, M., and Vespignani, A., "Dynamical Processes on Complex Networks", *Cambridge University Press*, Cambridge, 369–378 (2008).
46. van der Hofstad, R., "Random Graphs and Complex Networks", *Cambridge University Press*, Cambridge, 255–263 (2017).
47. Estrada, E., "Graph and Network Theory, Mathematical Tools for Physicist", 2nd ed., Grinfeld, M., *Wiley-Vch*, Weinheim, 134–139 (2015).
48. AKAL, Ş., "Gerçek ve Model Ağların Karakteristik Özelliklerinin Karşılaştırılması", *Öneri Dergisi*, 11 (41): 251 (2014).
49. Yılmaz, H., Sulak, E. E. and Özkaynak, E., "Complex Network Analysis of UEFA Europe League Competitions", *International Conference On Advanced Technologies, Computer Engineering And Science (ICATCES 2018) Proceeding Book*, Karabük, 389–393 (2018).
50. Çavuşoğlu, A. and Türker, İ., "Scientific collaboration network of Turkey", *Chaos, Solitons & Fractals*, 57: 9–18 (2013).
51. Albert, R. and Barabási, A.-L., "Statistical mechanics of complex networks", *Reviews Of Modern Physics*, 74 (1): 47–97 (2002).
52. Barabási, A. L., Albert, R., and Jeong, H., "Scale-free characteristics of random networks: The topology of the world-wide web", *Physica A: Statistical Mechanics And Its Applications*, 281 (1): 69–77 (2000).
53. Barabási, A. ., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., and Vicsek, T.,

- "Evolution of the social network of scientific collaborations", *Physica A: Statistical Mechanics And Its Applications*, 311 (3–4): 590–614 (2002).
54. Newman, M. E. J., "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality", *Physical Review E*, 64 (1): 016132 (2001).
 55. Newman, M. E. J., "Spread of epidemic disease on networks", *Physical Review E*, 66 (1): 016128 (2002).
 56. Türker, İ., "Evaluation of the Turkish Highway Network Analysis With Traffic Data", *Balkan Journal Of Electrical And Computer Engineering*, (2018).
 57. Türker, İ. and Sulak, E. E., "A multilayer network analysis of hashtags in twitter via co-occurrence and semantic links", *International Journal Of Modern Physics B*, 32 (04): 1850029 (2018).
 58. Strogatz, S. H., "Exploring complex networks", *Nature*, 410 (6825): 268–276 (2001).
 59. Wasserman, S. and Faust, K., "Social Network Analysis", *Cambridge University Press*, (1994).
 60. Aslan, S. and Kaya, M., "Topic recommendation for authors as a link prediction problem", *Future Generation Computer Systems*, 89: 249–264 (2018).
 61. Harary, F., R.Z. Norman, "Graph Theory as a Mathematical Model in Social Science", *Bulletin De L'Institut De Recherches Économiques Et Sociales*, 26 (8): 737 (1960).
 62. Mason, O. and Verwoerd, M., "Graph theory and networks in Biology", *IET Systems Biology*, 1 (2): 89–119 (2007).
 63. Likaj, R., Shala, A., Mehmetaj, M., Hyseni, P., and Bajrami, X., "Application of graph theory to find optimal paths for the transportation problem", *IFAC Proceedings Volumes*, 46 (8): 235–240 (2013).
 64. Boffey, T. B., "Graph Theory in Operations Research", *Macmillan Education UK*, London, (1982).
 65. Adams, C., "Leonhard Euler and the Seven Bridges of Königsberg", *The Mathematical Intelligencer*, 33 (4): 18–20 (2011).
 66. Erdős, P. and Rényi, A., "On Random Graphs I", *Publicationes Mathematicae Debrecen*, 6: 290 (1959).
 67. Erdos, P. and Renyi, A., "On the evolution of random graphs", *Publ. Math. Inst. Hungary. Acad. Sci.*, 5: 17–61 (1960).
 68. Bollobás, B., "Random Graphs", *Cambridge University Press*, Cambridge, 75–

- 88 (2001).
69. Gilbert, E. N., "Random Graphs", *The Annals Of Mathematical Statistics*, 30 (4): 1141–1144 (1959).
 70. Aiello, W., Chung, F., and Lu, L., "A random graph model for massive graphs", *STOC '00: Proceedings of the thirty-second annual ACM symposium on Theory of computing*, New York, 171–180 (2000).
 71. Karoński, M., "A review of random graphs", *Journal Of Graph Theory*, 6 (4): 349–389 (1982).
 72. Breiman, L., "Random Forests", *Machine Learning*, 45 (1): 5–32 (2001).
 73. Allali, O., Magnien, C., and Latapy, M., "Internal link prediction: A new approach for predicting links in bipartite graphs", *Intelligent Data Analysis*, 17 (1): 5–25 (2013).
 74. Ciaccio, A. Di and Giorgi, G. M., "Statistical Analysis of Social Networks", *RIEDS - Rivista Italiana Di Economia, Demografia E Statistica - Italian Review Of Economics, Demography And Statistics*, 67 (3–4): 103–110 (2013).
 75. Reichardt, J. and Bornholdt, S., "Statistical mechanics of community detection", *Physical Review E*, 74 (1): 16110 (2006).
 76. Lancichinetti, A., Fortunato, S., and Radicchi, F., "Benchmark graphs for testing community detection algorithms", *Physical Review E*, 78 (4): 046110 (2008).
 77. Fortunato, S., "Community detection in graphs", *Physics Reports*, 486 (3–5): 75–174 (2010).
 78. Kumar, R., Novak, J., and Tomkins, A., "Structure and evolution of online social networks", *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining*, Philadelphia , 611–617 (2006).
 79. Weitz, J. S., Benfey, P. N., and Wingreen, N. S., "Evolution, Interactions, and Biological Networks", *PLoS Biology*, 5 (1): e11 (2007).
 80. Dorogovtsev, S. N. and Mendes, J. F. F., "Evolution of networks", *Advances In Physics*, 51 (4): 1079–1187 (2002).
 81. Barabási, A.-L., "Evolution of Networks: From Biological Nets to the Internet and WWW Evolution of Networks: From Biological Nets to the Internet and WWW", S. N. Dorogovtsev and J. F. F. Mendes Oxford U. Press, New York, 2003. \$95.00 (264 pp.). ISBN 0-19-851590-1", *Physics Today*, 57 (10): 81–82 (2004).
 82. Bakshy, E., Rosenn, I., Marlow, C., and Adamic, L., "The role of social

- networks in information diffusion", *WWW '12: Proceedings of the 21st international conference on World Wide Web*, Lyon, 519–528 (2012).
83. Yu, K., Chu, W., Yu, S., Tresp, V., and Xu, Z., "Stochastic Relational Models for Discriminative Link Prediction", *Advances in Neural Information Processing Systems 19*, Vancouver, 1553-1560 (2006).
 84. Stieglitz, S. and Dang-Xuan, L., "Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior", *Journal Of Management Information Systems*, 29 (4): 217–248 (2013).
 85. Gomez-Rodriguez, M., Leskovec, J., and Krause, A., "Inferring Networks of Diffusion and Influence", *ACM Transactions On Knowledge Discovery From Data*, 5 (4): 1–37 (2012).
 86. Srivastava, J., "Data mining for social network analysis", *IEEE International Conference on Intelligence and Security Informatics*, Taipei, 33-34 (2008).
 87. Domingos, P., "Mining social networks for viral marketing", *Journal Of Retailing And Consumer Services*, 20 (1); (2005).
 88. Thelwall, M., Wilkinson, D., and Uppal, S., "Data mining emotion in social network communication: Gender differences in MySpace", *Journal Of The American Society For Information Science And Technology*, 61 (1): 190–199 (2010).
 89. Xia, M., Wang, J., and He, Y., "BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics", *PLoS ONE*, 8 (7): e68910 (2013).
 90. Breitkreutz, B.-J., Stark, C., and Tyers, M., "Osprey: a network visualization system", *Genome Biology*, 4 (3): R22 (2003).
 91. Moody, J., McFarland, D., and Bender-deMoll, S., "Dynamic Network Visualization", *American Journal Of Sociology*, 110 (4): 1206–1241 (2005).
 92. Christakis, N. A. and Fowler, J. H., "Social Network Visualization in Epidemiology.", *Norsk Epidemiologi = Norwegian Journal Of Epidemiology*, 19 (1): 5–16 (2009).
 93. Bastian, M., Heymann, S., and Jacomy, M., Gephi: An Open Source Software for Exploring and Manipulating Networks, *Proceedings of the Third International Conference on Weblogs and Social Media, ICWSM 2009*, California, 361–362 (2009).
 94. Findık, O. and Özkaynak, E., "Link prediction based on node weighting in complex networks", *Soft Computing*, <https://doi.org/10.1007/s00500-020-05314-8> (2020).
 95. Caldarelli, G. and Vespignani, A., "Preliminaries and Basic Definitions in

- Network Theory", *Complex Systems and Interdisciplinary Science*, 5–16 (2007).
96. Newman, M. E. J. and Girvan, M., "Finding and evaluating community structure in networks", *Phys. Rev. E*, 69 (2): 26113 (2004).
 97. Freeman, L. C., "Centrality in social networks conceptual clarification", *Social Networks*, 1 (3): 215–239 (1978).
 98. Sabidussi, G., "The centrality index of a graph", *Psychometrika*, 31 (4): 581–603 (1966).
 99. Bonacich, P., "Factoring and weighting approaches to status scores and clique identification", *The Journal Of Mathematical Sociology*, 2 (1): 113–120 (1972).
 100. Martínez, V., Berzal, F., and Cubero, J.-C., "A Survey of Link Prediction in Complex Networks", *ACM Computing Surveys*, 49 (4): 1–33 (2017).
 101. Song, H. H., Cho, T. W., Dave, V., Zhang, Y., and Qiu, L., "Scalable proximity estimation and link prediction in online social networks", *IMC '09: Internet Measurement Conference*, Chicago, 332–335 (2009).
 102. Hristova, D., Noulas, A., Brown, C., Musolesi, M., and Mascolo, C., "A multilayer approach to multiplexity and link prediction in online geo-social networks", *EPJ Data Science*, 5 (1): 24 (2016).
 103. Jalili, M., Orouskhani, Y., Asgari, M., Alipourfard, N., and Perc, M., "Link prediction in multiplex online social networks", *Royal Society Open Science*, 4 (2): 160863 (2017).
 104. Lei, C. and Ruan, J., "A novel link prediction algorithm for reconstructing protein–protein interaction networks by topological similarity", *Bioinformatics*, 29 (3): 355–364 (2013).
 105. Iakovidou, N., Symeonidis, P., and Manolopoulos, Y., "Multiway spectral clustering link prediction in protein-protein interaction networks", *Proceedings of the 10th IEEE International Conference on Information Technology and Applications in Biomedicine*, Corfu, 1–4 (2010).
 106. Airoldi, E., Blei, D., Fienberg, S., and Xing, E., "Mixed Membership Stochastic Blockmodels", *Mach Learn Res*, 1981–2014 (2008).
 107. Zhang, M., Cui, Z., Jiang, S., and Chen, Y., "Beyond Link Prediction: Predicting Hyperlinks in Adjacency Space", *AAAI*, 142–149 (2018).
 108. Ma, Y., Liang, X., Huang, J., and Cheng, G., "Intercity Transportation Construction Based on Link Prediction", *29th International Conference on Tools with Artificial Intelligence (ICTAI)*, 1135–1138 (2017).

109. Desmarais, B. A. and Cranmer, S. J., "Forecasting the locational dynamics of transnational terrorism: a network analytic approach", *Security Informatics*, 2 (1): 8 (2013).
110. Mutlu, E. C. and Oghaz, T. A., "Review on Graph Feature Learning and Feature Extraction Techniques for Link Prediction", *ArXiv*, abs/1901.0: (2019).
111. Nigam, A. and Chawla, N. V., "Link Prediction in a Semi-bipartite Network for Recommendation", *Lecture Notes in Computer Science* , 9622:127–135 (2016).
112. Sarukkai, R. R., "Link prediction and path analysis using Markov chains", *Computer Networks*, 33 (1–6): 377–386 (2000).
113. Guo, F., Yang, Z., and Zhou, T., "Predicting link directions via a recursive subgraph-based ranking", *Physica A: Statistical Mechanics And Its Applications*, 392 (16): 3402–3408 (2013).
114. Grabowski, A., Kruszewska, N., and Kosiński, R. A., "Dynamic phenomena and human activity in an artificial society", *Phys. Rev. E*, 78 (6): 66110 (2008).
115. Newman, M. E. J., "Clustering and preferential attachment in growing networks", *Physical Review E - Statistical Physics, Plasmas, Fluids, And Related Interdisciplinary Topics*, 64 (2): 4 (2001).
116. Ravasz, E. and Barabási, A.-L., "Hierarchical organization in complex networks", *Physical Review E*, 67 (2): 026112 (2003).
117. Leicht, E. A., Holme, P., and Newman, M. E. J., "Vertex similarity in networks", *Phys. Rev. E*, 73 (2): 26120 (2006).
118. Katz, L., "A new status index derived from sociometric analysis", *Psychometrika*, 18 (1): 39–43 (1953).
119. Neville, J. and Jensen, D., "Relational Dependency Networks", *J. Mach. Learn. Res.*, 8: 653–692 (2007).
120. Liben-Nowell, D., "An Algorithmic Approach to Social Networks", *Massachusetts Institute of Technology*, USA, (2005).
121. Lü, L., Jin, C.-H., and Zhou, T., "Similarity index based on local paths for link prediction of complex networks.", *Physical Review. E, Statistical, Nonlinear, And Soft Matter Physics*, 80 4 Pt 2: 46122 (2009).
122. Papadimitriou, A., Symeonidis, P., and Manolopoulos, Y., "Fast and accurate link prediction in social networking systems", *Journal Of Systems And Software*, 85 (9): 2119–2132 (2012).

123. Lichtenwalter, R. N., Lussier, J. T., and Chawla, N. V., "New Perspectives and Methods in Link Prediction", *The 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, 243-252 (2010).
124. Clauset, A., Moore, C., and Newman, M. E. J., "Hierarchical structure and the prediction of missing links in networks", *Nature*, 453 (7191): 98–101 (2008).
125. Goldenberg, A., "A Survey of Statistical Network Models", *Foundations And Trends® In Machine Learning*, 2 (2): 129–233 (2009).
126. Guimerà, R. and Sales-Pardo, M., "Missing and spurious interactions and the reconstruction of complex networks", *Proceedings Of The National Academy Of Sciences*, 106 (52): 22073–22078 (2009).
127. Peixoto, T. P., "Hierarchical Block Structures and High-Resolution Model Selection in Large Networks", *Phys. Rev. X*, 4 (1): 11047 (2014).
128. Vallès-Català, T., Peixoto, T. P., Sales-Pardo, M., and Guimerà, R., "Consistencies and inconsistencies between model selection and link prediction in networks", *Phys. Rev. E*, 97 (6): 62316 (2018).
129. Guimerà, R. and Sales-Pardo, M., "A Network Inference Method for Large-Scale Unsupervised Identification of Novel Drug-Drug Interactions", *PLoS Computational Biology*, 9 (12): e1003374 (2013).
130. Rovira-Asenjo, N., Gumí, T., Sales-Pardo, M., and Guimerà, R., "Predicting future conflict between team-members with parameter-free models of social networks", *Scientific Reports*, 3 (1): 1999 (2013).
131. Kashima, H. and Abe, N., "A Parameterized Probabilistic Model of Network Evolution for Supervised Link Prediction", **Sixth International Conference on Data Mining (ICDM'06)**, Hong Kong, 340–349 (2006).
132. Wang, C., Satuluri, V., and Parthasarathy, S., "Local Probabilistic Models for Link Prediction", *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, Omaha, 322–331 (2007).
133. Bilgic, M., Namata, G. M., and Getoor, L., "Combining Collective Classification and Link Prediction", *Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007)*, Omaha, 381–386(2007).
134. Taskar, B., Abbeel, P., and Koller, D., "Discriminative Probabilistic Models for Relational Data", *Appears in Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI2002)*, Alberta, 485–492 (2002).
135. Taskar, B., Wong, M., Abbeel, P., and Koller, D., "Link Prediction in Relational Data.", *NIPS'03: Proceedings of the 16th International Conference on Neural Information Processing Systems*, Cambridge, 243–252 (2003).

136. Heckerman, D., Meek, C., and Koller, D., "Probabilistic Models for Relational Data", *Technical report, Microsoft Research, Redmond*, 47–62(2004).
137. Heckerman, D., Chickering, D., Meek, C., Rounthwaite, R., and Kadie, C., "Dependency Networks for Collaborative Filtering and Data Visualization.", *Appears in Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI2000)*, Stanford, 264–273(2000).
138. Xu, Z., Tresp, V., Yu, K., Yu, S., and Kriegel, H.-P., "Dirichlet Enhanced Relational Learning", *Proc. 22nd Int. Conf. on Machine Learning*, Bonn, 1004–1011 (2005).
139. Duan, L., Ma, S., Aggarwal, C., Ma, T., and Huai, J., "An Ensemble Approach to Link Prediction", *IEEE Transactions On Knowledge And Data Engineering*, 29 (11): 2402–2416 (2017).
140. Hamilton, W. L., Ying, R., and Leskovec, J., "Representation Learning on Graphs: Methods and Applications", *ArXiv*, abs/1709.0: (2017).
141. Grover, A. and Leskovec, J., "Node2vec: Scalable Feature Learning for Networks", *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Stanford, 855–864 (2016).
142. Goyal, P. and Ferrara, E., "Graph embedding techniques, applications, and performance: A survey", *Knowledge-Based Systems*, 151: 78–94 (2018).
143. Goyal, P., Chhetri, S. R., and Canedo, A., "Dyngraph2vec: Capturing network dynamics using dynamic graph representation learning", *Knowledge-Based Systems*, 187: 104816 (2020).
144. Shultz, T. R., Fahlman, S. E., Craw, S., Andritsos, P., Tsaparas, P., Silva, R., Drummond, C., Ling, C. X., Sheng, V. S., Drummond, C., Lanzi, P. L., Gama, J., Wiegand, R. P., Sen, P., Namata, G., Bilgic, M., Getoor, L., He, J., Jain, S., Stephan, F., Jain, S., Stephan, F., Sammut, C., Harries, M., Sammut, C., Ting, K. M., Pfahringer, B., Case, J., Jain, S., Wagstaff, K. L., Nijssen, S., Wirth, A., Ling, C. X., Sheng, V. S., Zhang, X., Sammut, C., Cancedda, N., Renders, J.-M., Michelucci, P., Oblinger, D., Keogh, E., and Mueen, A., "Confusion Matrix", *Encyclopedia of Machine Learning, Springer US*, Boston, MA, 209–209 (2011).
145. Hanley, J. A. and McNeil, B. J., "The meaning and use of the area under a receiver operating characteristic (ROC) curve.", *Radiology*, 143 (1): 29–36 (1982).
146. Zheng, J., Egger, C., and Lienert, J., "A scenario-based MCDA framework for wastewater infrastructure planning under uncertainty", *Journal Of Environmental Management*, 183: 895–908 (2016).

147. Lundie, A., "Review of multi-criteria decision aid for integrated sustainability assessment of urban water systems", *Urban Water Journal*, 5: 315–327 (2008).
148. Roy, B., "Paradigms and Challenges", Multiple Criteria Decision Analysis: State of the Art Surveys, *Springer-Verlag*, New York, 3–24 (2005) .
149. Amine, M. El, Pailhes, J., and Perry, N., "Critical Review of Multi-criteria Decision Aid Methods in Conceptual Design Phases: Application to the Development of a Solar Collector Structure", *Procedia CIRP*, 21: 497–502 (2014).
150. Cinelli, M., Coles, S. R., and Kirwan, K., "Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment", *Ecological Indicators*, 46: 138–148 (2014).
151. Zavadskas, E. K., Turskis, Z., and Kildienė, S., "State Of Art Surveys Of Overviews On Mcdm/Madm Methods", *Technological And Economic Development Of Economy*, 20 (1): 165–179 (2014).
152. Bulut, T., "A New Multi Criteria Decision Making Method: Approach of Logarithmic Concept (APLOCO)", *International Journal Of Artificial Intelligence & Applications*, 9 (1): 15–33 (2018).
153. Mardani, A., Jusoh, A., MD Nor, K., Khalifah, Z., Zakwan, N., and Valipour, A., "Multiple criteria decision-making techniques and their applications – a review of the literature from 2000 to 2014", *Economic Research-Ekonomska Istraživanja*, 28 (1): 516–571 (2015).
154. Linkov, I., Varghese, A., Jamil, S., Seager, T. P., Kiker, G., and Bridges, T., "Multi-Criteria Decision Analysis: A Framework for Structuring Remedial Decisions at Contaminated Sites", Comparative Risk Assessment and Environmental Decision Making, *Kluwer Academic Publishers*, Dordrecht, 15–54 .
155. Findık, O., Özkaynak, E., "Node Weighting Method in Centrality Measure of Complex Networks", *Icatces 2019 Proceeding Book*, Antalya, 342–348 (2019).
156. Athanassopoulos, A. D. and Giokas, D., "The Use of Data Envelopment Analysis in Banking Institutions: Evidence from the Commercial Bank of Greece", *Interfaces*, 30 (2): 81–95 (2000).
157. Dursun, M., Karsak, E. E., and Karadayi, M. A., "Assessment of health-care waste treatment alternatives using fuzzy multi-criteria decision making approaches", *Resources, Conservation And Recycling*, 57: 98–107 (2011).
158. Peng, Y., Zhang, Y., Tang, Y., and Li, S., "An incident information management framework based on data integration, data mining, and multi-criteria decision making", *Decision Support Systems*, 51 (2): 316–327 (2011).

159. Rezaian, S. and Jozi, S. A., "Health- Safety and Environmental Risk Assessment of Refineries Using of Multi Criteria Decision Making Method", *APCBEE Procedia*, 3: 235–238 (2012).
160. Aldlaigan, A. H. and Buttle, F. A., "SYSTRA-SQ: a new measure of bank service quality", *International Journal Of Service Industry Management*, 13 (4): 362–381 (2002).
161. Tansel İç, Y., "Development of a credit limit allocation model for banks using an integrated Fuzzy TOPSIS and linear programming", *Expert Systems With Applications*, 39 (5): 5309–5316 (2012).
162. Özcan, T., Çelebi, N., and Esnaf, Ş., "Comparative analysis of multi-criteria decision making methodologies and implementation of a warehouse location selection problem", *Expert Systems With Applications*, 38 (8): 9773–9779 (2011).
163. Watts, D. J. and Strogatz, S. H., "Collective dynamics of “small-world” networks", *Nature*, 393 (6684): 440–442 (1998).
164. İnternet: Tennis Results and Tennis Betting Data, “2000 - 2017 ATP Men's Tour”,
<http://www.tuik.gov.tr/PreHaberBultenleri.do?id>
165. İnternet: IEEE Xplore, “2000-2017 Arası Yayınlanmış Makaleler”,
<http://www.tennis-data.co.uk/alldata.php>

ÖZGEÇMİŞ

Emrah ÖZKAYNAK 1979 yılında Konya'nın Ereğli ilçesinde doğdu; ilk ve orta öğrenimini aynı şehirde tamamladı. Ereğli Anadolu Teknik Lisesi Bilgisayar Yazılımı Bölümü'nden mezun oldu. 2001 yılında Zonguldak Karaelmas Üniversitesi Karabük Teknik Eğitim Fakültesi Elektronik ve Bilgisayar Eğitimi Bölümü'nde öğrenime başlayıp 2005 yılında mezun oldu. 2006 yılında Niğde Ulukışla Endüstri Meslek Lisesi'nde öğretmen olarak göreve başladı. 2011 yılında Karabük Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda başladığı yüksek lisans programını 2013 yılında tamamladı. 2010 yılında Karabük Üniversitesi'nde Öğretim Görevlisi olarak göreve başladı ve halen aynı yerde çalışmaya devam etmektedir.

ADRES BİLGİLERİ

Adres : Karabük Üniversitesi
Uzaktan Eğitim Uygulama ve Araştırma Merkezi
Balıklarkayası Mevkii / KARABÜK
Tel : (507) 796 37 21
E-posta : eozkaynak@karabuk.edu.tr / emrahozkaynak@gmail.com