

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

SOSYAL AĞLARDA BAĞLANTI TAHMİNİ

YÜKSEK LİSANS TEZİ

Melike Merve TEMEL

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Abdullah Erhan AKKAYA

TEMMUZ-2020

**İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

SOSYAL AĞLARDA BAĞLANTI TAHMİNİ



YÜKSEK LİSANS TEZİ

**Melike Merve TEMEL
(36183619023)**

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Abdullah Erhan AKKAYA

TEMMUZ-2020

TEŐEKKÜR VE ÖNSÖZ

Bu tez alıőmasının her aőamasında yardım, öneri, bilgi, tecrübe ve desteklerini esirgemededen beni her konuda yönlendiren danışman hocam Sayın Dr. Öğr. Üyesi Abdullah Erhan AKKAYA'ya,

Bu tez alıőmasında kullandığımız veri setini bizimle paylaşan, tecrübe ve desteklerini hiçbir zaman esirgemeyen kıymetli hocam Sayın Dr. Öğr. Üyesi Kenan İNCE'ye,

Manevi desteęiyle her zaman yanımda olan deęerli dostum Esra KARADENİZ'e,

Hayatımın her alanında maddi ve manevi olarak her zaman yanımda olan ve beni bugünlere getiren deęerli anneme, babama ve kardeşlerime;

Tez alıőmam süresince bana karşı sabırla ve anlayıőla yaklaşarak her zaman yanımda olan sevgili eőime;

teőekkür ederim.

ONUR SÖZÜ

Yüksek Lisans Tezi olarak sunduđum ‘‘Sosyal Ağlarda Bağlantı Tahmini’’ başlıklı bu çalışmanın bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığını ve yararlandığım bütün kaynakların hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuđunu belirtir, bunu onurumla doğrularım.

Melike Merve TEMEL



İÇİNDEKİLER

TEŞEKKÜR VE ÖNSÖZ	i
ONUR SÖZÜ.....	ii
İÇİNDEKİLER.....	iii
ÇİZELGELER DİZİNİ.....	iv
ŞEKİLLER DİZİNİ.....	v
SEMBOLLER VE KISALTMALAR.....	vi
ÖZET	vii
ABSTRACT	viii
1. GİRİŞ	1
2. ÇİZGE TEORİSİ	2
2.1.Çizge Çeşitleri	4
2.2.Matematiksel Gösterimi	5
3. AĞ BİLİMİ	7
3.1.Sosyal Ağlar	7
3.2. Sosyal Ağ Analizi	9
3.3. Sosyal Ağ Analizindeki Adımlar.....	9
3.4. Sosyal Ağ Analizindeki Ölçütler.....	10
Sosyal ağ analizinde kullanılan başlıca ölçütler anlatılmıştır.....	10
3.4.1. Merkezilik (Centrality).....	10
3.4.2. Diğer ölçütler.....	13
4. BAĞLANTI TAHMİNİ	14
4.1. Denetimli Yöntemler	15
4.2. Denetimsiz Yöntemler.....	16
4.2.1. Düğüm tabanlı ölçütler (Yerel benzerlik indeksleri).....	16
4.2.2. Yol tabanlı ölçütler (Genel benzerlik indeksleri)	18
5. UYGULAMA	19
5.1. Çalışılan Veri Seti.....	19
5.2. Uygulama Sonuçları	19
5.2.1. Marmara Bölgesi	20
5.2.2. İç Anadolu Bölgesi	21
5.2.3. Karadeniz Bölgesi	23
5.2.4. Doğu Anadolu Bölgesi	29
5.2.5. Ege Bölgesi.....	35
5.2.6. Akdeniz Bölgesi	42
5.2.7. Güney Doğu Anadolu Bölgesi.....	46
6. SONUÇLAR	51
KAYNAKLAR	52
ÖZGEÇMİŞ	54

ÇİZELGELER DİZİNİ

Çizelge 5.1 : Türkiye’de bölgeler bazında üniversite sayıları.	19
Çizelge 5.2 : Karadeniz Bölgesi adamic adar bağlantı tahmin skorları.	26
Çizelge 5.3 : Karadeniz Bölgesi ortak komşu bağlantı tahmin skorları.	27
Çizelge 5.4 : Karadeniz Bölgesi jaccard bağlantı tahmin skorları.	28
Çizelge 5.5 : Karadeniz Bölgesi kaynak tahsisli bağlantı tahmin skorları	29
Çizelge 5.6 : Doğu Anadolu Bölgesi adamic adar bağlantı tahmin skorları.	32
Çizelge 5.7 : Doğu Anadolu Bölgesi ortak komşu bağlantı tahmin skorları.	33
Çizelge 5.8 : Doğu Anadolu Bölgesi jaccard bağlantı tahmin skorları.	34
Çizelge 5.9 : Doğu Anadolu Bölgesi kaynak tahsisli bağlantı tahmin skorları.	35
Çizelge 5.10 : Ege Bölgesi adamic adar bağlantı tahmin skorları.	38
Çizelge 5.11 : Ege Bölgesi ortak komşu bağlantı tahmin skorları.	39
Çizelge 5.12 : Ege Bölgesi jaccard bağlantı tahmin skorları.	40
Çizelge 5.13 : Ege Bölgesi kaynak tahsisli bağlantı tahmin skorları.	41
Çizelge 5.14 : Akdeniz Bölgesi adamic adar bağlantı tahmin skorları.	44
Çizelge 5.15 : Akdeniz Bölgesi ortak komşu bağlantı tahmin skorları.	44
Çizelge 5.16 : Akdeniz Bölgesi jaccard bağlantı tahmin skorları.	45
Çizelge 5.17 : Akdeniz Bölgesi kaynak tahsisli bağlantı tahmin skorları.	45
Çizelge 5.18 : Güney Doğu Anadolu Bölgesi adamic adar bağlantı tahmin skorları.	48
Çizelge 5.19 : Güney Doğu Anadolu Bölgesi ortak komşu bağlantı tahmin skorları.	49
Çizelge 5.20 : Güney Doğu Anadolu Bölgesi jaccard bağlantı tahmin skorları.	49
Çizelge 5.21 : Güney Doğu Anadolu Bölgesi kaynak tahsisli bağlantı tahmin skorları.	50

ŞEKİLLER DİZİNİ

Şekil 2.1 : 18. yüzyılda Königsberg ve 7 köprü görseli [4].	2
Şekil 2.2 : Königsberg grafi [3].	3
Şekil 2.3 : Çizge çeşitleri [5].	4
Şekil 2.4 : Yönsüz ve yönlü çizgeler.	5
Şekil 2.5 : İlişki listesi (edge list).	6
Şekil 2.6 : İlişki matrisi (adjacency matrix).	6
Şekil 2.8 : Ağ modelinin çizge ile ifade edilmesi.	6
Şekil 3.1 : Sosyal ağ örnekleri.	8
Şekil 3.2 : Derece merkeziliği hesaplaması için örnek bir ağ.	11
Şekil 3.3 : Yakınlık merkeziliği hesaplaması için örnek bir ağ.	12
Şekil 4.1 : Farklı zamanlarda ağın durumu [23].	14
Şekil 5.1 : Marmara Bölgesi Üniversiteleri Arası İş Birliği.	20
Şekil 5.2 : Marmara Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini.	21
Şekil 5.3 : İç Anadolu Bölgesi Üniversiteleri Arası İş Birliği.	22
Şekil 5.4 : İç Anadolu Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini.	23
Şekil 5.5 : Karadeniz Bölgesi Üniversiteleri Arası İş Birliği.	24
Şekil 5.6 : Karadeniz Bölgesi Üniversiteleri Arası İş Birliği ve Gelecekteki Bağlantı Tahmini.	25
Şekil 5.7 : Karadeniz Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini.	25
Şekil 5.8 : Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birliği.	30
Şekil 5.9 : Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birliği ve Gelecekteki Bağlantı Tahmini.	31
Şekil 5.10 : Doğu Anadolu Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini.	31
Şekil 5.11 : Ege Bölgesi Üniversiteleri Arası İş Birliği.	36
Şekil 5.12 : Ege Bölgesi Üniversiteleri Arası İş Birliği ve Gelecekteki Bağlantı Tahmini.	37
Şekil 5.13 : Ege Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini.	37
Şekil 5.14 : Akdeniz Bölgesi Üniversiteleri Arası İş Birlikleri.	42
Şekil 5.15 : Akdeniz Bölgesi Üniversiteleri Arası İş Birlikleri ve Gelecekteki Bağlantı Tahmini.	43
Şekil 5.16 : Akdeniz Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini.	43
Şekil 5.17 : Güney Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birlikleri.	46
Şekil 5.18 : Güney Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birlikleri ve Gelecekteki Bağlantı Tahmini.	47
Şekil 5.19 : Güney Doğu Anadolu Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini.	47

SEMBOLLER VE KISALTMALAR

$\Gamma(x)$: Ağ içerisindeki x düğümünün komşularının kümesi
$ \Gamma(x) $: Ağ içerisindeki x düğümünün derecesi (komşularının sayısı)
$w(x, y)$: Ağ içerisindeki x ve y düğümü arasındaki ağırlık
V	: Düğümler kümesi
$G = \langle V, E \rangle$: V düğüm E ayrıttan oluşan bir graf
t	: Zaman
$C_D(x)$: Ağ içerisindeki x düğümünün derece merkeziliği
$C_C(x)$: Ağ içerisindeki x düğümünün yakınlık merkeziliği
WOS	: Web of Science İnternet Sayfası

ÖZET

Yüksek Lisans Tezi

SOSYAL AĞLARDA BAĞLANTI TAHMİNİ

Melike Merve TEMEL

İnönü Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

54+VIII sayfa

2020

Danışman: Dr. Öğr. Üyesi Abdullah Erhan AKKAYA

Günümüzde internetin gelişmesi sonucu sosyal medyanın yaygın ve yoğun kullanımı sosyal ağ analizi ve bağlantı tahmini alanlarındaki çalışmaların artmasına zemin hazırlamıştır. Bağlantı tahmini problemi, çizge yapısı ile ağ içerisindeki düğümler ve bu düğümler arasındaki mevcut ilişkiler referans alınarak, ağın ve ağ içindeki ilişkilerin zamanla nasıl değişeceği sorusunun cevabını aramaktadır.

Bu çalışmada sosyal ağlarda bağlantı tahmini konusuna giriş yapıldıktan sonra, bu alanda yapılan çalışmalar üzerinde durulmuştur. Bağlantı tahmini için önerilen yöntemler incelenmiştir. Türkiye Üniversiteleri Arasındaki Bağlantıların bulunduğu veri setine bağlantı tahmini yöntemleri uygulanarak elde edilen sonuçlar değerlendirilmiş ve analiz edilmiştir.

Anahtar Kelimeler: Sosyal Ağ Analizi, Bağlantı Tahmini, Veri Madenciliği, Çizge Teorisi, Sosyal Ağlar

ABSTRACT

Master Thesis

LINK PREDICTION IN SOCIAL NETWORKS

Melike Merve TEMEL

Inonu University
Graduate School of Nature and Applied Sciences
Department of Computer Engineering

54+VIII pages

2020

Supervisor: Asst. Prof. Abdullah Erhan AKKAYA

Today, as a result of the development of the internet, the widespread and intensive use of social media has led to increased studies in the areas of social network analysis and connection prediction.

The connection prediction problem, takes the graph structure, the nodes in the network and the current relationships between these nodes as a reference and searches for the answer of the question of how the network and relationships within the network will change over time.

In this study, after the introduction of connection prediction in social networks, the studies in this area are emphasized. Suggested methods for connection prediction are examined. Connection prediction methods have been applied to the dataset of the connections between Turkish universities. The results obtained were evaluated and analyzed.

Keywords: Social Network Analysis, Link Prediction, Data Mining, Graph Theory, Social Network

1. GİRİŞ

İnsanlar sosyal varlıklardır ve birbirleriyle çeşitli şekillerde etkileşirler, bu ağlara doğada her yerde rastlarız. Bu ağlar akrabalık, arkadaşlık gibi ilişkilerden doğan bağlar olabileceği gibi çevrimiçi sosyal ağların uygulama alanları çok daha geniştir. Sosyal ağlar, çizgilerle temsil edilir ve sosyal ağlar ile ilgili kapsamlı bir literatür vardır [1].

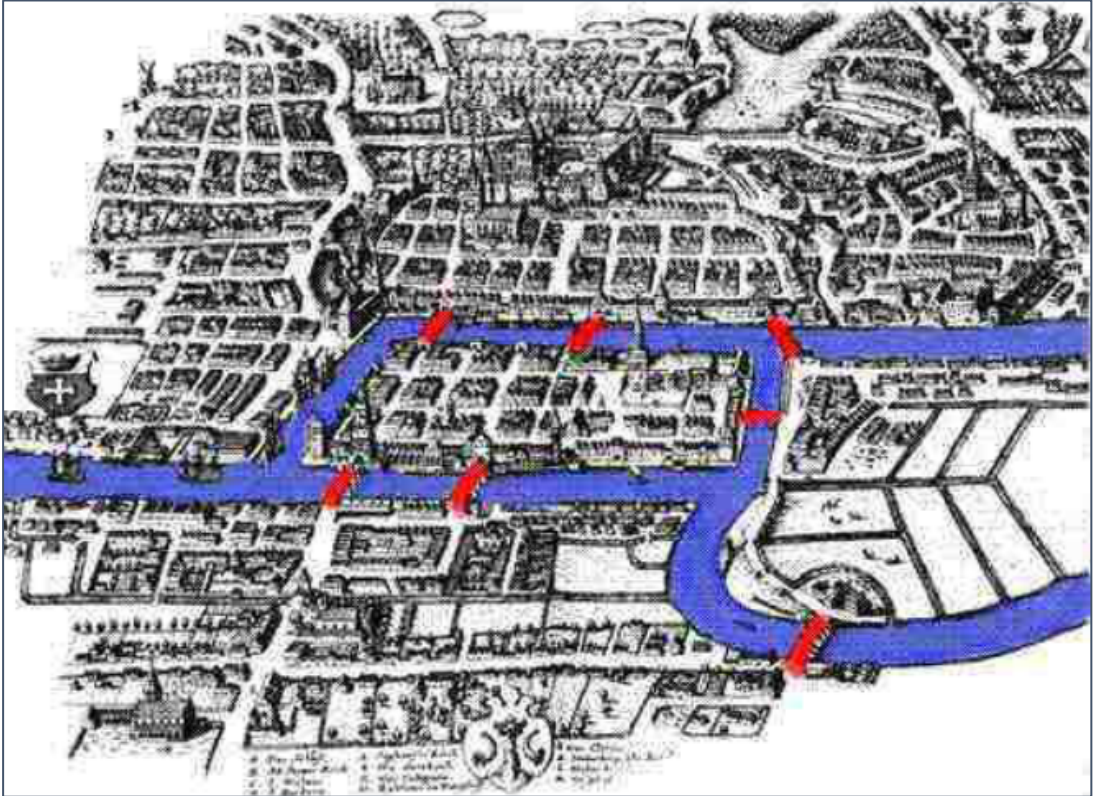
Sosyal ağların ortaya çıkışı ve gelişimi son yıllardaki en heyecan verici olaylardan biridir. Sosyal ağların sürekli gelişmesi, birçok araştırmacının dikkatini çeken farklı alanlarda büyük bir bilgi kaynağı oluşturan bir alan oldu. Sosyal ağlar sağladıkları verinin çeşitliliği ve boyutu bakımından oldukça zengin sistemlerdir. Bu yönüyle eş bulunmaz kaynaklardır. Bu durum, büyük veriyi nasıl işleyeceğimiz ile ilgili alanların öne çıkmasına zemin hazırlamıştır. Sosyal ağ analizi, sosyal ağ yapısı, sosyal ağlarda topluluk analizi, bağlantı tahmini gibi alanlara yönelimi hızlandırmıştır. Sosyal ağlarda bağlantı tahmini yöntemi kişilerin birbirleriyle kurdukları ilişkiler hakkında çeşitli değerlendirmeler yaparak, gelecekteki durumlarının nasıl olacağını öngörme temeline dayanır. Sürekli olarak değişen yapılarda bağlantı tahmininde bulunmak oldukça karmaşık bir konudur. Ağ içerisinde bulunan verilerden başka veriler çıkarma ve bu verilerin ilerideki durumları hakkında tahminler geliştirmek hedeflenmektedir. Ağ içerisinde bulunan verilerden başka veriler çıkarma ve bu verilerin ilerideki durumları hakkında tahminler geliştirmek hedeflenmektedir.

Bu tezde, ağ bilimi, sosyal ağlar, sosyal ağ analizi, sosyal ağlarda bağlantı tahmini konuları açıklanarak sosyal ağlarda bağlantı tahmini alanında yapılmış çalışmalar incelenmiştir. Türkiye’de bulunan üniversiteler arası bağlantıların bulunduğu veri setine bağlantı tahmini algoritmaları uygulanarak sonuçlar karşılaştırılmıştır.

2. ÇİZGE TEORİSİ

Graf teorisi, çizge teorisi veya çizit teorisi grafları inceleyen matematik dalıdır. Graf, düğümler ve bu düğümleri birbirine bağlayan kenarlardan oluşan bir ağ yapısı türüdür. Bir graf, çizge veya çizit, düğümlerden (köşeler) ve bu düğümleri birbirine bağlayan kenarlardan (yaylardan, bağıntılardan) oluşur [2]. Çizge teorisinin temelleri 18. yüzyıla dayanmaktadır. Königsberg kentinde bir problemin çözümünü İsviçreli matematikçi Leonhard Euler'in bulmasıyla çizge teorisi ortaya çıkmıştır.

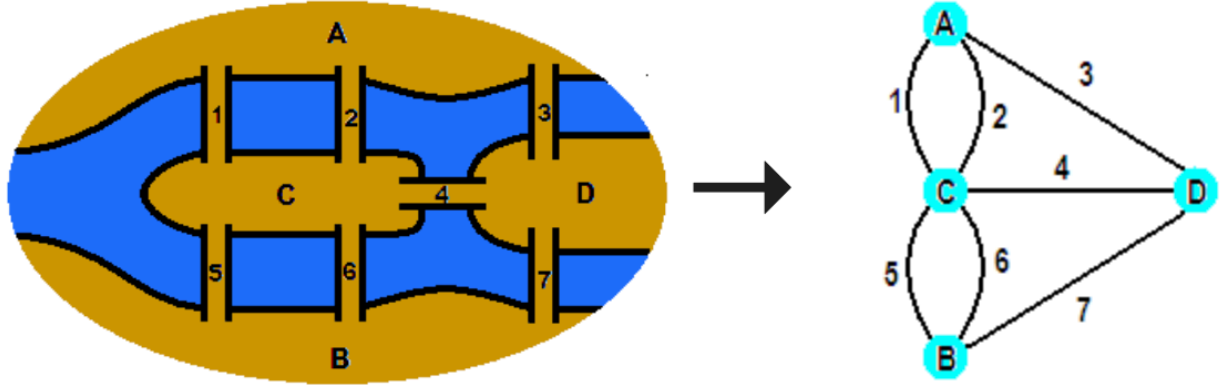
Königsberg şehrinde Eski ve Yeni Pregel nehirlerinin birleşiminden oluşan Pregel nehri kenti tam dört anakaraya bölmüştü. Nehirler üzerindeki yedi köprü aracılığıyla bu dört parça arasında ulaşım sağlanıyordu. 'Şehrin herhangi bir noktasından başlayıp her köprüden yalnızca bir defa geçmek şartıyla şehir turu yapılabilir mi?' probleminin çözümünü arayan Euler çizge teorisinin temelini atmıştır. Euler problemi basitleştirmek ve daha iyi analiz etmek amacıyla Şekil 2.2. de gösterilen grafi çizmiştir.



Şekil 2.1 : 18. yüzyılda Königsberg ve 7 köprü görseli [4].

Königsberg şehrinde Eski ve Yeni Pregel nehirlerinin birleşiminden oluşan Pregel nehri kenti tam dört anakaraya bölmüştü. Nehirler üzerindeki yedi köprü aracılığıyla bu dört parça arasında ulaşım sağlanıyordu. 'Şehrin herhangi bir noktasından başlayıp her

köprüden yalnızca bir defa geçmek şartıyla şehir turu yapılabilir mi?’ probleminin çözümünü arayan Euler çizge teorisinin temelini atmıştır. Euler problemi basitleştirmek ve daha iyi analiz etmek amacıyla Şekil 2.1.2. de gösterilen grafi çizmiştir



Şekil 2.2 : Königsberg grafi [3].

Euler, şehrin bölündüğü dört parçanın her biri için bir nokta, köprüler için ise kenar adı verilen bağlantı çizgilerini kullanmıştır. Euler problemi "Herhangi bir noktadan harekete başlayıp, bütün kenarlardan ancak ve ancak bir defa geçerek, bütün noktaları ziyaret ettikten sonra başlangıç noktasına varabilir miyiz?" şeklinde değiştirmiştir.

Euler, çalışmalarının sonucunda bunun mümkün olabilmesi için tüm noktaların çift dereceli olması gerektiğini ortaya koymuştur. Şekil 2.2’ de gösterilen çizgede tüm düğümlerin dereceleri tek olduğundan her köprüden bir kere geçmenin imkânsız olduğunu göstermiştir.

Euler, bütün düğümlerin derecelerinin çift olması gerekliliğine şöyle bir düşünce ile karar vermiştir. Böylesi problemlerde başlangıç düğümünden diğer düğümlere gidildiğinde o düğümün bir giren bir de çıkan kenarı bulunmalıdır, yani düğüm çift dereceli olmalıdır. Başlangıç düğümü ve bitiş düğümü hariç graftaki tüm düğümler bu kurala uymalıdır. Bu iki düğümün derecesi tek olabilir. Çünkü bu düğümlerden başlangıç düğümünün giriş kenarı bitiş düğümünün ise çıkış kenarı olmasa da olur. Fakat Königsberg köprüsü problemi için bu durum böyle değildir. Çünkü problemin tanımında “başlanan noktaya tekrar dönmek” ifadesi yer almaktadır. Yani başlangıç düğümünün giriş kenarına bitiş düğümünün çıkış kenarına ihtiyacı vardır. Sonuç olarak Königsberg köprüsü probleminin çözülebilmesi için koşul tüm noktaların derecelerinin çift olmasıdır.

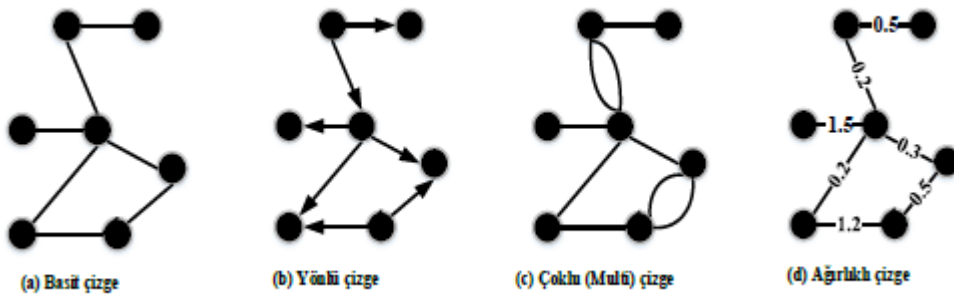
Çizge teorisi zamanla geliştirilip birçok disiplinde kullanım alanı bulmuştur. Ağlardaki karmaşıklık basitleştirilerek ağların görsel biçimi olan çizgeler ile ifade edilebilir. Konuma bağlı uygulamalarda, elektrik devrelerinde, şebeke yapılarında, yazılımlarda graf gösterimleri kullanılmaktadır. Yoğun miktarda etkileşimin olduğu karmaşık sistemleri çözümlmek için ağları anlayıp, çizgeler ile analiz etmek gerekmektedir.

Canlı sistemler içerisindeki etkileşimler, süreçler bu şekilde çözümlenebilir. Ağ canlı veya cansız bazı bileşenler ve bu bileşenler arasındaki etkileşimlerden oluşur. Ağ içerisindeki bu bileşenler “düğüm” olarak adlandırılır. Düğüm bazı ağlar için insan olabileceği gibi bazı ağlar için protein, arı veya ülke de olabilir. Ağlardaki düğümler bağlantılar aracılığıyla birbirlerine bağlıdır. İnsan beyni, bedeni, sinir sistemi ağlar için muhteşem örneklerdir.

2.1.Çizge Çeşitleri

Düğümlerin birbirleriyle ilişkileri yönlü veya yönsüz olarak ifade edilmektedir. Ayrıca kenarlar ağırlıklı ya da ağırlıksız olabilir. Kenar ağırlığı iki düğüm arasındaki bağlantının kuvvetini temsil eden özelliktir.

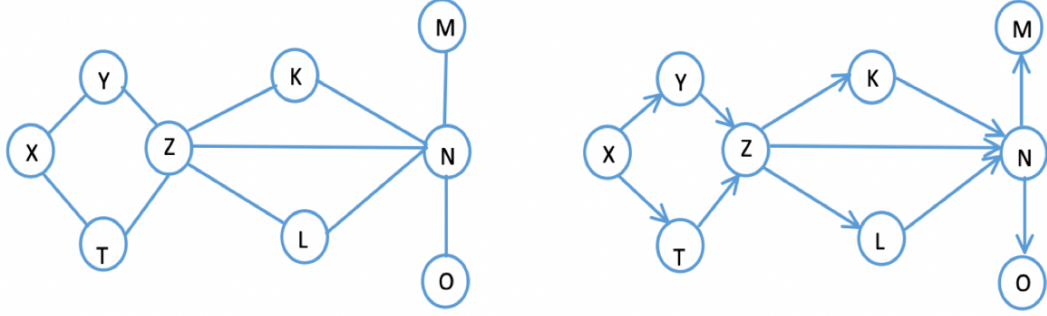
Çizgeleri kategorilere ayırmanın farklı yolları vardır. Çizgeler, kenarların içeriği, ağırlıkların varlığı, yönü ve kenar sayısı gibi özelliklere dayalı gruplara ayrılır. Farklı çizge modelleri Şekil 2.3’de gösterilmiştir.



Şekil 2.3 : Çizge çeşitleri [5].

a. Basit çizgeler, herhangi bir ağırlık olmaksızın oluşturulan yönsüz, paralel kenar olmayan ve döngü içermeyen (bir düğümden kendisine olan bağlantı) çizge türüdür. Kenarların herhangi bir yönü yoktur ve simetriktir. Düğümler arasındaki ikili ilişkileri temsil ederler.

b. Yönlü çizgeler, düğümler arasındaki kenarların yöne sahip olduğu çizge türüdür. Yönlü ağ veya yönlü çizge kısaca graf olarak adlandırılır. Yönlü ağlara örnek olarak; bir sayfadan diğerine yönlendirilerek çalıştırılan internet ağı, avdan avcıya gelen enerji ile gıda ağı, bir yayından diğerine alıntı yapılan atıf ağı gösterilebilir. Her bir kenarı sıralı bir düğüm çifti ile ilişkilendirilmiştir.



Şekil 2.4 : Yönsüz ve yönlü çizgeler.

c. Çoklu (Multi) çizgeler, basit çizgelerin yeterli olmadığı durumlarda kullanılır. Yönsüz, paralel kenarı olan ve döngü içermeyen çizgelerdir. Basit çizgeler, çoklu çizgedir fakat çoklu çizgeler basit çizge değildir.

d. Ağırlıklı çizgeler her bir kenarına nümerik bir değer, ağırlık verilmiş çizge türüdür. Kenar ağırlığı düğümler arasındaki bağlantının gücünü temsil ettiğinden dolayı güçlü ve zayıf olmak üzere ikiye ayrılır. Birçok durumda çoklu grafikler, iki düğümü birbirine bağlayan kenarların sayısı kenar ağırlığına yansıtılacak şekilde ağırlıklı çizgelere dönüştürülür.

2.2. Matematiksel Gösterimi

Sayısal verilerin oluşturduğu tablolardan çizge diyagramları oluşturulabilir ve görselleştirilebilir. Böylelikle çizgeler yardımıyla ağlar çok daha iyi analiz edilebilir. Örneğin Elif, Bora, Cenk ve Ayşe adlı kişilerin birbirlerine gönderdikleri mesaj sayılarının listesi aşağıdaki gibi olsun:

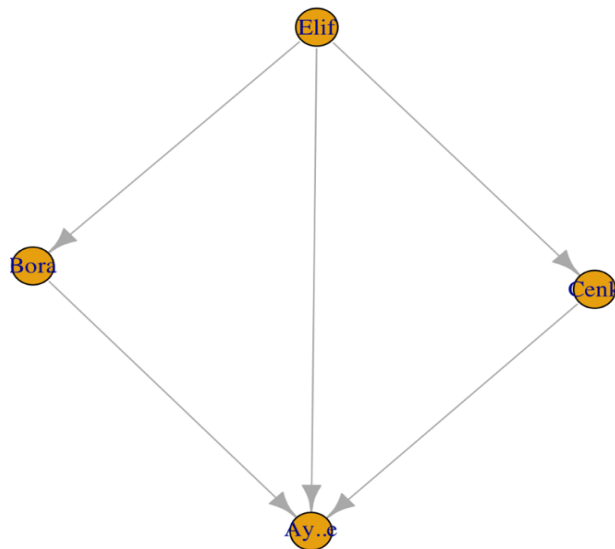
Mesaj Gönderen	Mesaj Alan	Sayı
Elif	Bora	10
Elif	Cenk	8
Elif	Ayşe	2
Bora	Ayşe	7
Cenk	Ayşe	3

Şekil 2.5 : İlişki listesi (edge list).

*	Elif	Bora	Cenk	Ayşe
Elif	0	10	8	2
Bora	0	0	0	7
Cenk	0	0	0	3
Ayşe	0	0	0	0

Şekil 2.6 : İlişki matrisi (adjacency matrix).

Bu ağ modeli görselleştirildiğinde daha somut ve anlaşılır olabilir. Bu ağın görsel hali aşağıda gösterilmiştir.



Şekil 2.7 : Ağ modelinin çizge ile ifade edilmesi.

3. AĞ BİLİMİ

İki veya ikiden fazla kişi, şirket, yazar, hücre veya düğümden oluşan varlıklar arasında kurulan ilişki “ağ” kavramını tanımlamaktadır. Bu ilişkiler ise ağ içerisindeki varlıklar arasında mesaj, kredi, atıf, sinyal veya başka şekillerde olabilir. Bahsedilen ağ tanımı ile hali hazırdaki ağların yapısını anlamak mümkün değildir. Çünkü anlatılan ağ tanımı oldukça sade ve basittir. Mevcut ağların ortaya çıkardığı ağ yapıları karmaşık yapılardır. “Karmaşıklık” ise bir modelinin formüle edilmesinde karşılaşılan zorluktur. Ağlarda bileşen çeşitliliği, miktarı, yani büyüklük ve çok boyutluluk karmaşıklığın en temel iki sebebidir. Ağlar karmaşıklaştıkça ağ biliminin sunduğu imkânlar ve ağ olarak kabul edilebilecek yapıların her birinin üyelerinin bireysel özelliklerinin ötesindeki ağ doğası ortaya çıkmaktadır. Bu doğa birey psikolojisi ile kitle psikolojisi arasındaki ayrım kadar geniştir.

3.1.Sosyal Ağlar

"Sosyal Ağ" kelimesi insanların birbirleriyle olan ilişkilerini açıklamak için ilk olarak 1954 yılında John Arundel Barnes tarafından kullanılmıştır [1]. Sosyal ağlar, kişilerin birbirleriyle ilişki kurdukları, fotoğraf, fikir, ses, konum gibi verileri aktardıkları platformlardır. Bu ağların en popüler olanları Facebook, WhatsApp, Youtube, Instagram, LinkedIn, Twitter, Tumblr vb. olarak sıralanabilir. Sosyal ağlar aracılığıyla devasa boyutta veriler elde edilmektedir. Sosyal ağlar, her bir düğümü bir kişi veya kurumu temsil eden ve birbirleriyle bağlantılı birçok farklı düğümden oluşmaktadır. Sosyal bir ağ kenar ve düğümlerden oluşan bir yapıdır. Kenarlar düğümler arasındaki ilişkiyi göstermek adına düğümleri birbirine bağlar. Sosyal ağ yapısı bir defa doğru olarak oluşturulduğunda, etrafımızda olup biten problemleri çözmek için gerekli bilgileri bu ağ içerisinde bulabiliriz.



Şekil 3.1 : Sosyal ağ örnekleri.

Sosyal ağlar sürekli olarak güncellenme özelliğine sahip olduğundan, kenar ve düğümler zaman içerisinde ağa eklenmektedir. Bu özelliğinden dolayı sosyal ağları dinamik yapılar olarak tanımlayabiliriz. Giderek artan kenar ve düğüm sayısı, geleceğe yönelik tahminlerde bulunmamızı gerektirebilir. Bu ihtiyaçtan dolayı bağlantı tahmini (link prediction) metotları ortaya çıkmıştır. Sosyal ağlarda bağlantı tahmini, kişilerin birbirleriyle kurdukları ilişkiler hakkında çeşitli değerlendirmeler yaparak, gelecekteki durumlarının nasıl olacağını öngörme temeline dayanır. Bağlantı tahmini, bilim adamları veya üniversiteler arasında potansiyel arkadaşlıklar veya iş birlikleri bulmasına yardımcı olan birçok uygulamada kullanılabilir. Bu nedenle bağlantı tahmini sosyal ağ analizinde ve diğer birçok alanda önemli bir role sahiptir [2,3].

3.2. Sosyal Ağ Analizi

Sosyal ağlar, sahip oldukları sosyal aktörlerin ilişkilerine ait yararlı bilgiler içerirler. Bu yapı ve ilişkilerin analiz edilmesi sonucu yönelimler, benzerlikler ve etkileşimler gibi sonuçlara ulaşılarak ağ hakkında bazı yorumlar veya tahminler yapılabilir. Özellikle internet teknolojisinde günden güne artan teknik ve kültürel gelişmeler, sosyal ağların yaygın kullanımı, buradan elde edilen verilerin ölçülebilmesi ve anlamlandırılması için sosyal ağ analizinin önemini ortaya çıkarmıştır. Sosyal ağ analizi, sosyal ağ içerisindeki aktörlerin yani bilgisayarlar, insanlar ya da nesnelerin arasındaki ilişkilerin sayısal hale getirilip sonuçlar çıkarılmasını sağlayan bir yöntemdir. Aynı zamanda sosyal ağın yapı karakteristiğinin ortaya çıkarılmasında da önemli bir rol oynar. Sosyal ağ analizi, ağ içerisinde var olan ancak anlaşılması çok kolay olmayan ilişkileri ağları modelleyerek, ilişkileri görselleştirerek ortaya koyabilir.

Disiplinler arası bir çalışma alanı olan sosyal ağ analizi konusu psikoloji, ekonomi, bilişim, sosyoloji, biyoloji gibi birçok alanda etkin bir biçimde kullanılabilir. Sosyal ağ analizi temel anlamda ağ içindeki aktör, bağlantı ve ilişkilere yoğunlaşır. Ağ içerisindeki temel elemanlar aktörlerdir. Her bir aktör bir düğüm ile ifade edilir. Düğümler arasında bağlantılar vardır bu bağlantılar ilişkilerin görselleştirilmiş şekilleridir. İlişkilerde içerik, yön, yoğunluk, aktiflik gibi bazı parametreler öne çıkar. İçerik iki düğüm arasındaki ilişkiyi anlatır. İki düğüm arasındaki ilişki yönlü veya yönsüz olabilmektedir. Mesela Facebook platformunda arkadaşlık yönsüz bir graftır. Çünkü Facebook arkadaşlığında ilişki çift taraflıdır. Twitter platformu ise yönlü graflara örnek olarak verilebilir. Çünkü Twitter için aynı durum söz konusu değildir. Bir kişi diğerini takip edebilirken takip edilen kişinin de takip isteğine karşılık verme zorunluluğu yoktur [6].

3.3. Sosyal Ağ Analizindeki Adımlar

Analizi yapılacak olan sosyal ağ verisi oldukça geniş bir alandan toplanmaktadır. Bu veriler, kullanıcı bilgilerinden, oturum hareketlerinden, paylaşım etkileşimlerinden elde edilebildiği gibi bazen devasa boyutlara sahip günlük kayıtların tutulduğu dosyalardan da oluşabilir. Bahsedilen büyüklükteki ve karmaşıklıkta verilerin iyi analiz edilebilmesi için izlenecek adımlar dört ana başlıkta toplanmıştır [7].

Veri Seçimi: Analiz edilmek istenen verinin çevrimiçi veya çevrimdışı olarak elde edilmesi aşamasıdır. Elde edilen veri bir veri tabanında tutulur. Verinin elde edilme biçimleri, yöntemleri değişiklik gösterebilir. Bu konu ile ilgili herhangi bir standart yoktur.

Bilgi Çıkarımı ve Ön İşleme: Elde edilen veri ham haldedir. Bu veri ancak analiz edilerek anlamlı, işe yarar hale getirilebilir. Verinin analiz edilip anlamlı hale getirilmesi ile bilgi elde edilir. Bilgi elde edildikten sonra anlaşılabilir olabilmesi için modelleme yapılır. Gerekli filtre işlemleri uygulanarak etkin bir analize hazır hale getirilir.

Genelleştirme: Bu safhada artık elde edilen veri analiz edilerek bilgi çıkarılmış, modellenmiş ve filtrelenmiştir. Bu işlemler sonucundan bir model ortaya çıkmıştır. Ancak bu model spesifik bir modeldir. Bu modelin genelleştirilmesi ve genel-geçer kurallar üretebilmesi hedeflenmektedir.

Analiz: Genelleştirilen model aracılığı ile çıkarılan kurallar veriden anlamlı sonuçlar elde etmemize yardımcı eder. Gelecekte olabilecek durum ve olayları tahmin etmemize ve değerlendirmemize olanak sağlar [8].

3.4. Sosyal Ağ Analizindeki Ölçütler

Sosyal ağ analizinde kullanılan başlıca ölçütler anlatılmıştır.

3.4.1. Merkezilik (Centrality)

Merkezilik, ağdaki bir düğümün öteki düğümlere nazaran daha merkezde bulunmasıdır. Merkezilik kavramı en çok sosyal ağın mikro katmanı olan düğüm katmanında kullanılan analitik araçtır. Bir düğümün merkeziliği ile ağ ilişkilerine katılma derecesi doğru orantılıdır. Merkezi olan düğüm ağda sosyal derece ve önem açısından öne çıkar. Yani merkezilik aslında ağın en önemli ve merkezi düğümlerinin aranmasıdır.

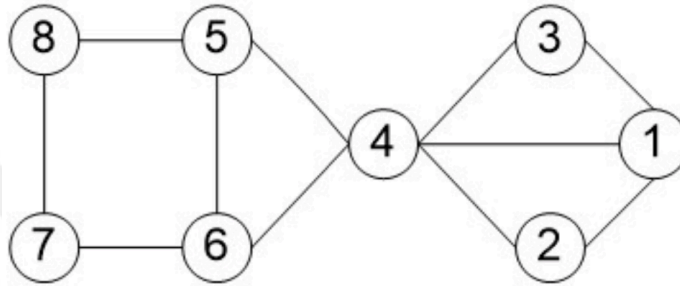
Derece Merkeziliği (Degree Centrality)

Derece merkeziliği, bir düğümde meydana gelen bağlantı sayısı, düğümün sahip olduğu bağların sayısı olarak tanımlanabilir. Ağdaki düğümler arasındaki bağlantıların yönleri vardır. Bu yönler iki şekilde olabilir. Birincisi düğüme yönlendirilen bağlantılar,

ikincisi ise düğümün diğer düğümlere yönlendirdiği bağlantılardır. [9, 10].

$G = \langle V, E \rangle$ ile ifade edilen n adet düğüme sahip bir ağda tepe düğüm olan v_i derecesi $CD(v_i) = \frac{d_i}{n-1}$ (3.1) şeklinde hesaplanmaktadır.

Şekil 3.2'de, düğüm 4 en yüksek dereceli merkeziyete sahipken, 7 ve 8 en düşük merkeziğe sahiptir.



Şekil 3.2 : Derece merkeziliği hesaplaması için örnek bir ağ.

$$CD(4) = \frac{5}{7} = 0.71 \quad CD(8) = \frac{2}{7} = 0.28$$

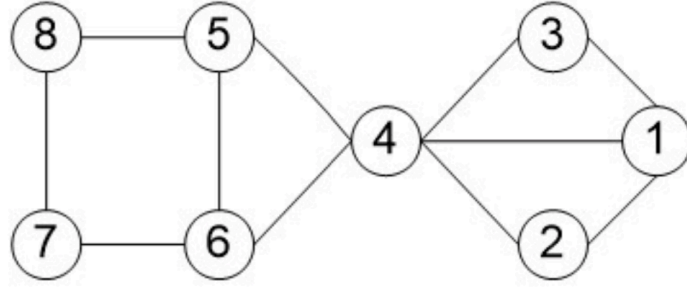
Yakınlık Merkeziliği (Closeness Centrality)

Yakınlık merkeziliği, bir düğümün ağdaki diğer tüm düğümlere ne kadar yakın olduğunu gösterir. Ağ teorisinde, yakınlık karmaşık bir merkeziyet ölçüsüdür. Bir v noktası ile ondan ulaşılabilen diğer tüm köşeler arasındaki en kısa mesafe ortalaması olarak tanımlanır. Düğümden ağdaki diğer düğümlere kadar olan en kısa yol uzunluğunun ortalaması olarak hesaplanır [9, 11].

$G = \langle V, E \rangle$ ile ifade edilen n adet düğüme sahip bir ağda tepe düğüm olan v_i olan bir ağın yakınlık merkeziliği;

$$CC(v_i) = \frac{n-1}{\sum_{j \neq i} g(v_i, v_j)} \quad (3.2)$$

şeklinde hesaplanmaktadır.



Şekil 3.3 : Yakınlık merkeziliği hesaplaması için örnek bir ağ.

$$CC(4) = \frac{8 - 1}{1 + 1 + 1 + 1 + 1 + 2 + 2} = \frac{7}{9} = 0.77$$

$$CC(5) = \frac{8 - 1}{1 + 1 + 1 + 2 + 2 + 2 + 3} = \frac{7}{12} = 0.58$$

Şekil 3.3'de verilen örnek ağda, düğüm 4'ün düğüm 5'ten daha merkezi olduğu hesaplanmıştır.

Arasındalık Merkeziliği (Betweenness Centrality)

Arasındalık merkeziliği, bir düğümün ağdaki en kısa yollarda ne kadar önemli olduğunu ölçer. Bu hesaplamayı şu şekilde yapmaktadır. Öncelikle düğüm seçilir. Ardından ağdaki bütün düğüm çiftleri arasındaki en kısa yollar hesaplanır. Bulunan bu yolların kaç tanesinde seçilen düğümün var olduğunun oranı şeklinde hesaplanır.

Ancak bu hesaplama bazı büyük ağlar için oldukça maliyetli olabilmektedir. Bu gibi durumlarda belirli seviye komşuluklara inilerek hesaplama yapılabilir. Arasındalık merkeziliği büyük olan düğümler için diğer düğümlerde önemli oldukları söylenebilir [9, 10].

Öz vektör Merkeziliği (Eigenvector Centrality)

Öz vektör merkeziliği merkeziyetin daha karmaşık bir alanıdır: az sayıda bağlantıya sahip bir kişi, eğer bu birkaç bağlantı başkalarına çok iyi bağlıysa, çok yüksek öz vektör merkeziyetine sahip olabilir. Öz vektör merkeziliği, bağlantıların değişken bir değere sahip olmasına izin verir, böylece bazı düğümlere bağlanmak diğerlerine bağlanmaktan daha fazla fayda sağlar. Google'ın arama motoru tarafından kullanılan PageRank algoritması, esas olarak yönlendirilmiş ağlar için kullanılan öz vektör merkeziliğinin bir çeşididir.

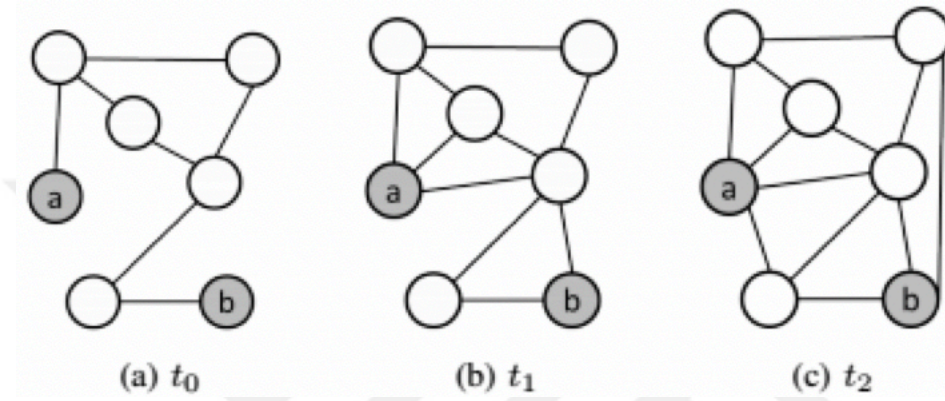
3.4.2. Diğer ölçütler

- **Merkezileştirme:** Merkezileştirme ağın daha düzenli, sağlıklı sonuç üreten hale getirilmesidir. Merkezi bir ağ, bağlantılarının çoğunu bir veya birkaç düğüm etrafında dağıtırken, merkezi olmayan bir ağ, her düğümün sahip olduğu bağlantı sayısı arasında çok az değişiklik olduğu bir ağdır.
- **Kümelenme Katsayısı:** Kümelenme katsayısı bir düğümün, bağlı olduğu düğümler ile arasındaki bağlantının ne kadar güçlü olduğunu söylemektedir. Bir düğümün bağlantılı olduğu düğümlerin kendi aralarındaki, gerçek bağlantıların olası tüm bağlantılara oranıyla hesaplanmaktadır [12].
- **Bütünlük:** Düğümlerin birbiriyle kurdukları bağları inceler. En güçlü bağları kuran düğümlerin benzerliğinin fazla olduğunu söyler.
- **Derece:** Düğümün bir ağ içerisinde sahip olduğu bağlantıların toplamıdır. Yönlü bir ağda çalışılıyor ise düğüme yönlendirilen ve düğümün yönlendiği tüm bağlantı sayılarının toplamına denir. Basit bir hesaplamadır ancak düğümün önemini gösterir. Birçok ağ için düğüm derecesi önem ve gücü direkt ifade edebilir. [9, 10].
- **Yoğunluk:** Ağ içerisinde bulunan bütün bağlantıların, olası bağlantıların tümüne oranına denir. Bir ağ içerisinde bütün düğümler birbirine bağlı ise yoğunluk değeri 1 olacaktır. Ağın yoğunluk değeri yükseldikçe ağ içerisinde güçlü bir bağlantının olduğu ve düğümlerin birbirinden ayrı olmadığı anlaşılmaktadır.
- **Köprü:** Ağ içerisinde var olan bir bağlantıya köprü diyebilmemiz için bu bağlantı kaldırıldığında oluşan sonuçlara bakmalıyız. Bağlantı kaldırıldığında eğer diğer uç düğümlere erişim sağlanamıyorsa bu bağlantı köprüdür.
- **Lokal Köprü:** Bitiş noktasını öteki komşular kullanmıyorsa bu kenar lokal köprüdür.
- **Yol Uzunluğu:** Ağ içerisindeki iki düğüm arasındaki mesafedir. Bu mesafe kısaldıkça aradaki düğüm ve bağlantı sayısı da azalır. Aynı zamanda yol uzunluğu azaldıkça bağlantı kurma ihtimali artar.
- **Prestij:** Yönlü bir ağda düğümün merkeziliğini tanımlamak için kullanılan terimdir.

Bu anlatılan ölçütlerin her biri ağın yapısı hakkında bilgiler verir. Ağ analizinde kullanılan ölçütler bütün bir ağa da uygulanabilir, gruplara ya da bireylere özelleştirilerek de kullanılabilir.

4. BAĞLANTI TAHMİNİ

Sosyal ağlar, bir grup veya topluluktaki insanlar arasındaki etkileşimleri modellemenin popüler bir yoludur. Bir düğümün gruptaki bir kişiye karşılık geldiği ve kenarların ise kişiler arasındaki ilişkileri temsil ettiği çizgeler ile görselleştirilebilirler. Zaman içinde çizgeye yeni kenarlar ve düğümler eklendiğinden, sosyal ağlar da çok dinamiktir. Bir sosyal ağın evrimini yönlendiren dinamikleri anlamak, çok sayıda değişken parametresi nedeniyle karmaşık bir sorundur.



Şekil 4.1 : Farklı zamanlarda ağın durumu [23].

İki spesifik düğüm arasındaki ilişkiyi anlamak nispeten daha kolay bir problemdir. Sorulabilecek bazı sorular şunlardır: İlişkilendirme modelleri zaman içinde nasıl değişir? Bağlantıları yönlendiren faktörler nelerdir? İki düğüm arasındaki ilişki diğer düğümlerden nasıl etkilenir? Burada ele almak istediğimiz sorun, çizgenin mevcut durumundaki düğümler arasında bir ilişki olmadığını bilerek, iki düğüm arasında gelecekteki bir ilişki olasılığını tahmin etmektir. Bu sorun yaygın olarak Bağlantı Tahmin problemi olarak bilinmektedir [13].

Bağlantı tahmin problemi bir sosyal ağın evrimini, ağ topolojisinin kendisine özgü özelliklerini kullanarak ne ölçüde modellenebileceğini, ağın mevcut durumunun gelecekteki bağlantıları tahmin etmek için kullanılıp kullanılmayacağını araştırır.

Bağlantı tahmin problemi ayrıca, gözlemlenen bir ağdan eksik bağlantıların bulunması sorunuyla da ilgilenir. Bir dizi alanda, gözlemlenebilir verilere dayanan bir ağ oluşturulur ve daha sonra doğrudan görünür olmayan ancak var olabilecek bağlantılar ortaya çıkarılmaya çalışılır. Bir sosyal ağın arkadaşlık önerisi gerçek hayatta da sosyal ağlarda da henüz başlamamış bir arkadaşlığı başlatabilir.

Sosyal ağların ötesinde, bağlantı tahmininin birçok başka uygulaması vardır. Biyoinformatik alanı örnek olarak verilebilir. Proteinler arasındaki etkileşimleri bulmak için bağlantı tahmini kullanılabilir. E-ticarete öneri sistemlerinin oluşturulmasına yardımcı olabilir. Amazon'da "bunu satın alanlar bunu da satın aldı" özelliği ve güvenlik alanı bağlantı tahmininde gizli terörist veya suçlu gruplarının belirlenmesine yardımcı olabilir. Ayrıca, ortak yazarlık ağları (örneğin bilimsel dergilerde, ortak yazarlı makaleleri olan kenarları birleştiren çiftler) üzerinde birçok çalışma yapılmıştır [14].

Ağda "yakın" olan iki bilim adamının ortak meslektaşları olacak ve benzer çevrelerde seyahat edecekler; bu sosyal yakınlık, kendilerinin yakın gelecekte iş birliği yapma olasılıklarının daha yüksek olduğunu düşündürmektedir. Bu nedenle, bu uygulamadaki bağlantı tahmini, tesadüfen oluşması daha uzun sürecek karşılıklı olarak yararlı profesyonel veya akademik bağlantı / iş birliklerini hızlandırmak için kullanılabilir.

Bağlantı tahmin probleminin klasik bir tanımı şu şekilde ifade edilir: "t zamanında bir sosyal ağın anlık görüntüsü verildiğinde, t'den belirli bir gelecekteki zamana kadar ağa eklenecek kenarları doğru bir şekilde tahmin etmeye çalışıyoruz.". Bu sorunla başa çıkmak için çeşitli yaklaşımlar önerilmiştir.

Bağlantı tahmin problemiyle ilgilenen birçok yöntem vardır; iki düğüm arasındaki yakınlık ölçütlerini kullanan yöntem, bunlar arasında en iyi bilinen yöntemdir [14]. Genellikle, benzerlik puanları hesaplandıktan sonra tahminle ilgilenen iki tür yaklaşım vardır. Bu yaklaşımlar makine öğrenmesinde de kullanılan denetimli yöntemler ve denetimsiz yöntemlerdir.

4.1. Denetimli Yöntemler

Bağlantı tahmin problemi, denetimli yaklaşımda bir sınıflandırma problemi olarak ele alınır, burada pratik olarak bağlı olan düğüm çiftleri, pozitif olarak atanır, ancak bağlantısız düğüm çiftleri, negatif olarak atanır. Denetimsiz yöntemin aksine, denetimli yaklaşım, kullanılacak sınıflandırıcıyı eğitmek için bir eğitim setine ihtiyaç duyar. Denetimli yöntemde yapısal özelliklere göre sınıflandırma ve karşılaştırmalar yapılır.

4.2. Denetimsiz Yöntemler

Bağılı olmayan düğüm çiftlerini bir liste olarak tanımlar, ardından listenin en üstüne en yüksek puanlara sahip düğüm çiftlerini koyar. Sonra listelenen üst düğüm çiftlerinin bağlantılı olması beklenir. Bu yöntemin uygulanması çok kolaydır ve tahmini gerçekleştirmek için bir eğitim seti gerektirmez. Bununla birlikte, zorlukları da vardır. Eşik değer tanımlama ihtiyacı ve birden fazla ölçüt tarafından sağlanan bilgileri birleştirmede zorluk gibi bazı sınırlamalar gösterir.

Denetimli yöntemde karar ağaçları, destek vektör makineleri gibi bazı denetimli öğrenme algoritmalarından da yararlanır.

4.2.1. Düğüm tabanlı ölçütler (Yerel benzerlik indeksleri)

Düğüm çiftleri arasındaki benzerliği gösteren ölçümlerin değerlerini çıkarmak, düğüm bazlı yaklaşımın temel noktasıdır. Düğümler, yakınlıklarını bulmak için her bir düğüm çiftine uygulanan benzerlik ölçütlerini bir özellik vektörü olarak görür. Bundan sonra puanlar denetimsiz yöntemle kullanılabilir [14, 15, 16]. Denetimsiz stratejide, bir yakınlık metriği seçilir ve ağdaki düğüm çiftlerine uygulanır. En üst sıradaki düğüm çiftinin gelecekte bir bağlantı yapacağı tahmin edilmektedir.

- $\Gamma(x)$: Ağ içerisindeki x düğümünün komşularının kümesini,
- $|\Gamma(x)|$: Ağ içerisindeki x düğümünün derecesini (komşularının sayısı),
- $w(x, y)$: Ağ içerisindeki x ve y düğümü arasındaki ağırlığı ifade etmektedir.

- **Ortak Komşu:** Ortak komşu sayısı, bağlantı tahmin problemine uyarlanmış en yaygın yöntemlerden biridir ve oldukça basittir. Bu yöntem, iki düğümün gelecekteki bağlantısının, bu düğümler arasındaki çok sayıda ortak komşuyla ilişkili olduğunu söyler [14]. t zamanında x ve y 'nin ortak komşularının sayısı ile t 'den sonraki bir zamanda x ve y 'nin işbirliği olasılığı arasında pozitif bir ilişki olduğunu doğrulamıştır. $\Gamma(x)$ x 'in komşularının ve $\Gamma(y)$ y 'nin komşularının kümesi ise,

$$\text{skor}(x, y) = |\Gamma(x) \cap \Gamma(y)| \quad (4.1)$$

- **Jaccard İndeksleme:** Bilgi alımında yaygın olarak kullanılan bir benzerlik ölçüsüdür. X veya y'nin rastgele seçilmiş bir özelliği f için hem x hem de y'nin f özelliğine sahip olma olasılığını ölçer. Ortak komşuluklar toplam komşuluk sayısına bölünerek hesaplanır.

$$skor(x, y) = \frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x) \cup \Gamma(y)|} \quad (4.2)$$

- **Tercihli Bağlantı İndeksleme:** Sosyal ağlarda iyi bilinen bir kavram, birçok arkadaşı olan kullanıcıların gelecekte daha fazla bağlantı oluşturma eğiliminde olmasıdır. Bunun nedeni, sosyal ağlarda zenginlerin zenginleşmesidir. Her düğümün sahip olduğu bağlantı sayıları çarpımını hesaplayarak iki düğümümüzün ne kadar “zengin” olduğunu tahmin ediyoruz.

$$skor(x, y) = |\Gamma(x)| \cdot |\Gamma(y)| \quad (4.3)$$

- **Adamic Adar İndeksleme:** Bu yöntem, az bağlantılı komşulara daha fazla değer vererek ortak komşulukları hesaplar. Bir örnekle açıklanmak istenirse, içerisinde ‘örneğin’ kelimesi geçen yazıların benzerlik ihtimali içerisinde ‘kümeleme katsayısı’ geçen yazıların benzerlik ihtimalinden azdır.

$$skor(x, y) = \sum_{z: x \text{ ve } y \text{ nin özelliği}} \frac{1}{\log(\text{frekans}(z))} \quad (4.4)$$

- **Kaynak Tahsisli İndeksleme:** Bu yöntem kompleks ağlarda tercih edilir. Birbiriyle doğrudan bağlantılı olmayan düğüm çiftleri arasındaki bağlantıları ölçer. Birbirleri arasında bağlantı olmadığı halde düğümler ortak komşuları vasıtası ile iletim sağlarlar. İletim sağlayan her düğümün kaynak ünitesi vardır ve bu kaynağı komşularına eşit dağıtır. Bu düğümler arasındaki benzerlik birbirlerinden aldıkları kaynaklara göre hesaplanır.

$$skor(x, y) = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{k_z} \quad (4.5)$$

4.2.2. Yol tabanlı ölçütler (Genel benzerlik indeksleri)

- **Katz İndeksleme:** Leo Katz'ın [17] önerdiği bu yöntem, x ve y düğümleri arasındaki en kısa yolların sayısının toplamını baz almaktadır. Benzerlik hesaplamasında uzun yolların hesaplama üzerindeki dezavantajını azaltmak için yolun uzunluğu bir parametreye (β^l) üs olarak eklenmiştir. (l : yolun uzunluğunu ifade etmektedir.) Bu sayede uzun yolların hesaplamaya etkisi daha az olmaktadır. Bu denklem (2.6) aşağıdaki eşitlikle hesaplanır.

$$skor(x, y) = \sum_{l=1}^{\infty} \beta^l \cdot |yollar_{x,y}^{(l)}| \quad (4.6)$$

- **Köklü PageRank:** Ulaşma zamanında x ve y düğümleri arasındaki mesafe çok kısa olsa da ağın geri kalanındaki yolların çok uzun olması, yani x düğümünden y düğümüne gidebilmek için geçilmesi gereken düğüm sayısının fazla oluşu bu indeks için bir dezavantaj oluşturmaktadır. x ve y arasındaki rasgele yürüyüş her bir adımda α parametresindeki olasılık değeri ile yeniden başa döndürülebilir. Böylece ağ içinde rasgele yürüyüşler olabilecek en kısa yollardan yapılmış olur. Rasgele yürüyüşün belli bir olasılıkla yeniden başlatılması web sayfalarındaki PageRank ölçütünün temelini oluşturur. Rasgele yürüyüş $\alpha[0,1]$ olasılığı ile başa döner, $1 - \alpha$ olasılığı ile o an bulunan düğümün komşularından rasgele birine gider. Bu işlem her adımda uygulanır [18].

i düğümünün tüm koşulları için D diagonal derece matrisinde

$$D[i, i] = \sum_j A[i, j] \text{ 'dir.}$$

$N = D^{-1}A$, komşuluk matrisinin satırlarının 1'e normleştirilmesidir. Formülü denklem (4.7)'de verilmiştir.

$$RPR = (1 - \alpha)(I - \alpha N)^{-1} \quad (4.7)$$

- **SimRank:** Bu yol tabanlı benzerlik indeksinin dayandığı esas şudur; iki düğüm benzer düğümlerle ilişkili ise bu düğümler benzerdir. Öz yinelemeli olarak formülü aşağıdaki gibidir:

$$similarity(x, y) = \gamma \cdot \frac{\sum_{a \in \Gamma(x)} \sum_{b \in \Gamma(y)} similarity(a, b)}{|\Gamma(x)| \cdot |\Gamma(y)|} \quad (4.8)$$

5. UYGULAMA

5.1. Çalışılan Veri Seti

Türkiye' deki 131 üniversite ve bunların 1980'den 2015'e kadar tüm işbirlikli çalışmalarını içeren iş birliği grafi geliştirilen bir uygulama aracılığıyla Web of science (WOS) [19] dan elde edilmiştir [20-21]. Veri seti 131 özel veya devlet üniversitesi, 270.000' in üzerinde yazar, 151 araştırma alanı ve 40 belge türü içermektedir. Bir iş birliği grafiği oluşturabilmek için, farklı kurumlardan yazarlar içeren yayınlar kullanılmış ve bu ilişkiler kurumlar arası bağlantılar olarak tanımlanmıştır [20].

5.2. Uygulama Sonuçları

Verinin büyük olmasından dolayı, analizlerin bölgesel bazda yapılması, görselleştirme açısından uygun görülmüştür. Ayrıca, yapılan geçmiş çalışmalarda da, üniversiteler arası iş birliklerini etkileyen faktörler arasında mekânsal yakınlığın etkili olduğu görülmüştür. [21, 22]. Bu sebeple, bölgesel bazda analiz etmenin aslında çok fazla veri kaybına sebep olmadığı görülmüş oldu.

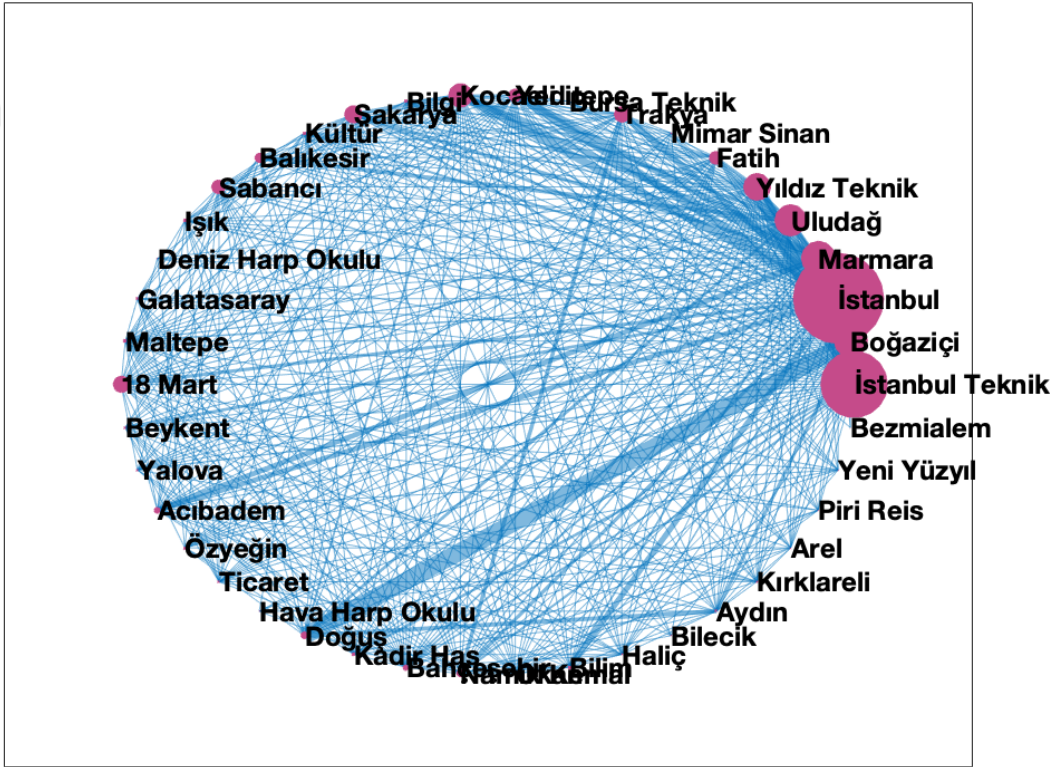
Akdeniz, Doğu Anadolu, Ege, Güney Doğu Anadolu, İç Anadolu, Karadeniz ve Marmara bölgeleri ele alınmıştır. Veri setinde bulunan 131 üniversitenin bölgelere göre dağılımı Çizelge 5.1' de verilmiştir.

Çizelge 5.1 : Türkiye' de bölgeler bazında üniversite sayıları.

Bölge	Üniversite Sayısı
Marmara Bölgesi	42
İç Anadolu Bölgesi	26
Karadeniz Bölgesi	17
Doğu Anadolu Bölgesi	15
Ege Bölgesi	13
Akdeniz Bölgesi	9
Güney Doğu Anadolu Bölgesi	9

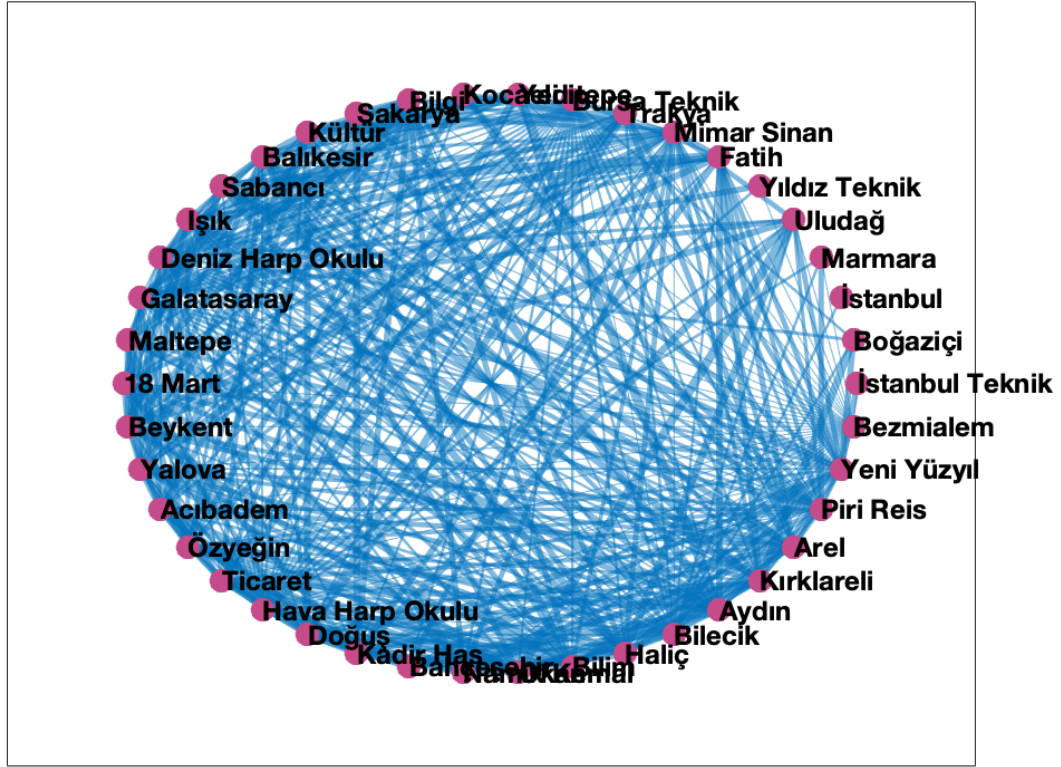
5.2.1. Marmara Bölgesi

Veri seti içerisinde Akdeniz bölgesinde 42 üniversite yer almaktadır. Bu üniversitelerin mevcut durumdaki (veri seti içerisindeki) iş birlikleri Şekil 5.1’de gösterilmektedir. Ancak üniversite sayısının ve iş birliklerinin fazla oluşu sebebi ile görsel olarak iyi ifade edilememiştir. Düğümlerin büyüklüğü ağ içerisindeki bağlantı sayısına göre değişiklik göstermektedir. Ayrıtların kalınlığı ise birbirine bağladığı düğümlerin ne kadar etkileşimde olduğuna, yani üniversite iş birliklerinin sayısına bağlı olarak değişiklik göstermektedir.



Şekil 5.1 : Marmara Bölgesi Üniversiteleri Arası İş Birliği

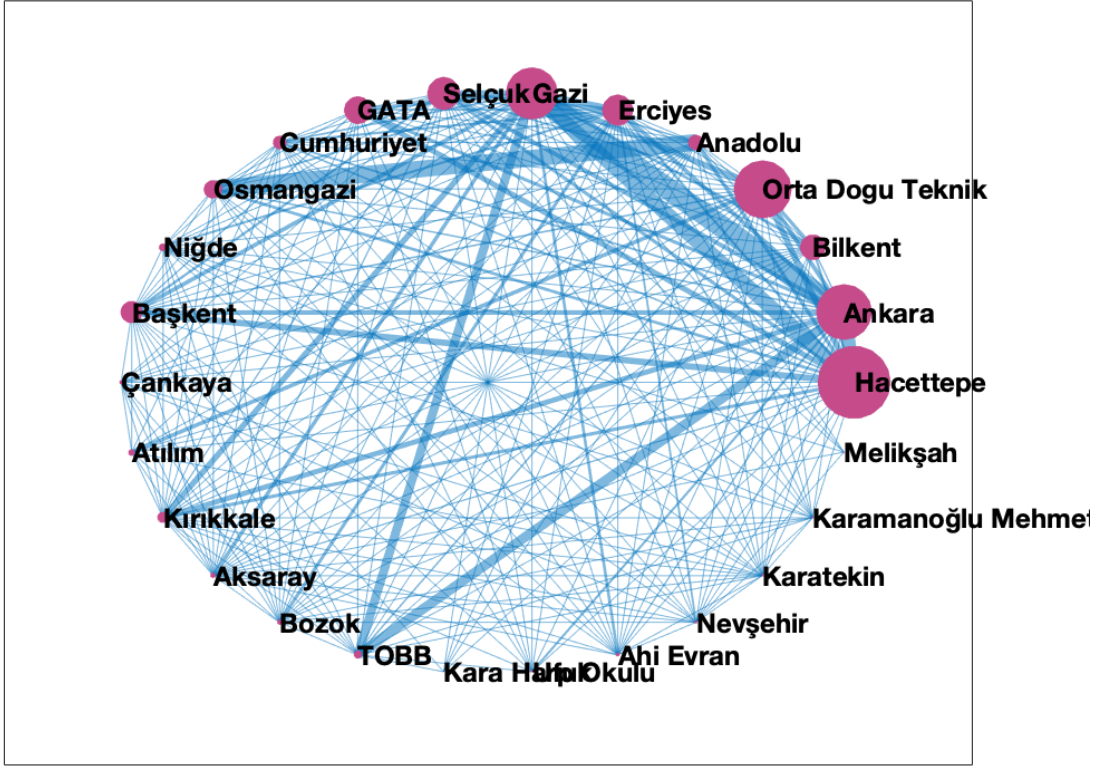
Şekil 5.2. Marmara Bölgesi üniversiteleri arasındaki bağlantı tahmininin görüntüsüdür. Ancak 42 üniversite bulunduğundan bu görsel net ve anlaşılır değildir. Bu tahmin, veri setine bağlantı tahmin algoritmalarından adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme uygulanarak elde edilmiştir. Bütün bu algoritmalar aynı sonucu farklı skorlar ile vermiştir.



Şekil 5.2 : Marmara Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini

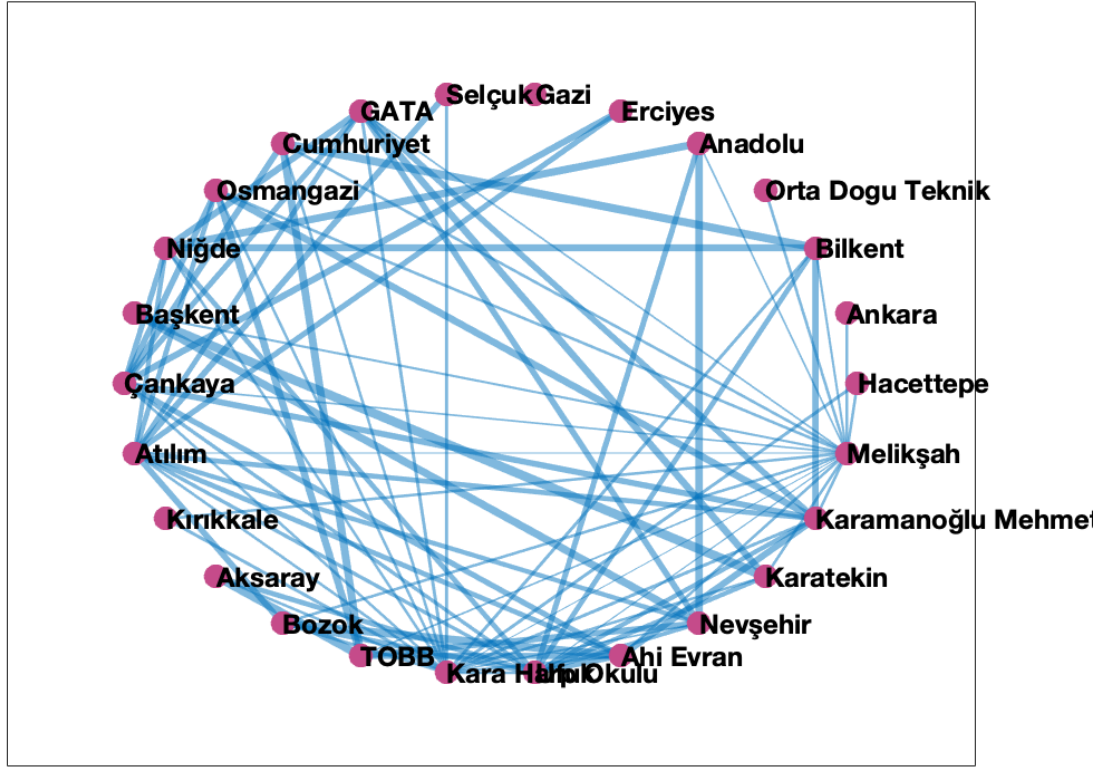
5.2.2. İç Anadolu Bölgesi

Veri seti içerisinde İç Anadolu bölgesinde 26 üniversite yer almaktadır. Bu üniversitelerin mevcut durumdaki etkileşimleri Şekil 5.3 de gösterilmektedir. Ancak üniversite sayısının ve iş birliklerinin fazla oluşu sebebi ile görsel olarak iyi ifade edilememiştir. Düğümlerin büyüklüğü ağ içerisindeki bağlantı sayısına göre değişiklik göstermektedir. Ayrıtların kalınlığı ise birbirine bağladığı düğümlerin ne kadar etkileşimde olduğuna, yani üniversite iş birliklerinin sayısına bağlı olarak değişkenlik göstermektedir.



Şekil 5.3 : İç Anadolu Bölgesi Üniversiteleri Arası İş Birliği

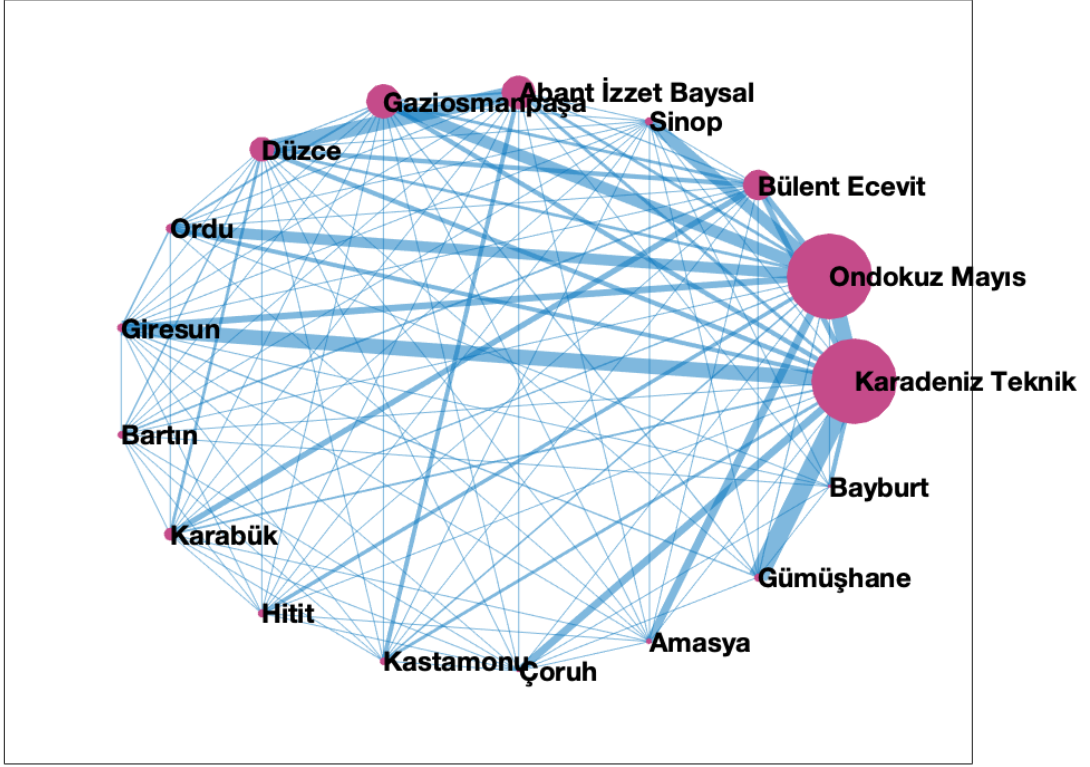
Şekil 5.4. İç Anadolu Bölgesi üniversiteleri arasındaki bağlantı tahmininin görüntüsüdür. Ancak 26 üniversite bulunduğundan bu görsel net ve anlaşılır değildir. Bu tahmin, veri setine bağlantı tahmin algoritmalarından adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme uygulanarak elde edilmiştir. Bütün bu algoritmalar aynı sonucu farklı skorlar ile vermiştir.



Şekil 5.4 : İç Anadolu Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini

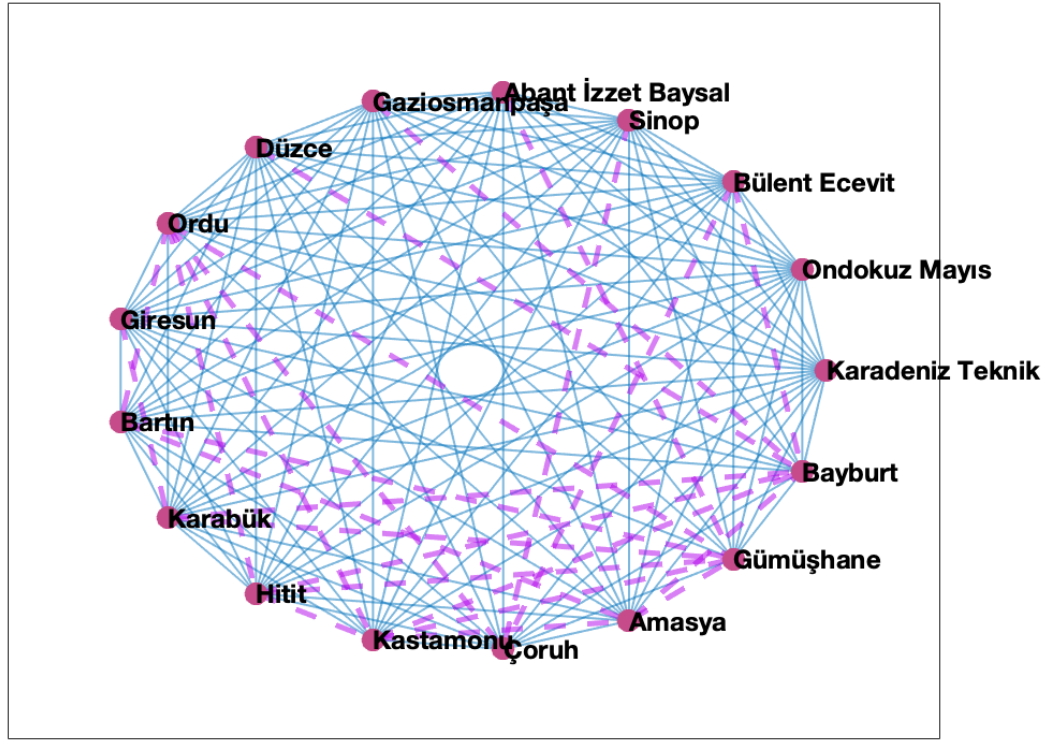
5.2.3. Karadeniz Bölgesi

Veri seti içerisinde Karadeniz bölgesinde 17 üniversite yer almaktadır. Bu üniversitelerin mevcut durumdaki etkileşimleri Şekil 5.5. de gösterilmektedir. Düğümlerin büyüklüğü ağ içerisindeki bağlantı sayısına göre değişiklik göstermektedir. Ayrıkların kalınlığı ise birbirine bağladığı düğümlerin ne kadar etkileşimde olduğuna, yani üniversite iş birliklerinin sayısına bağlı olarak değişkenlik göstermektedir.

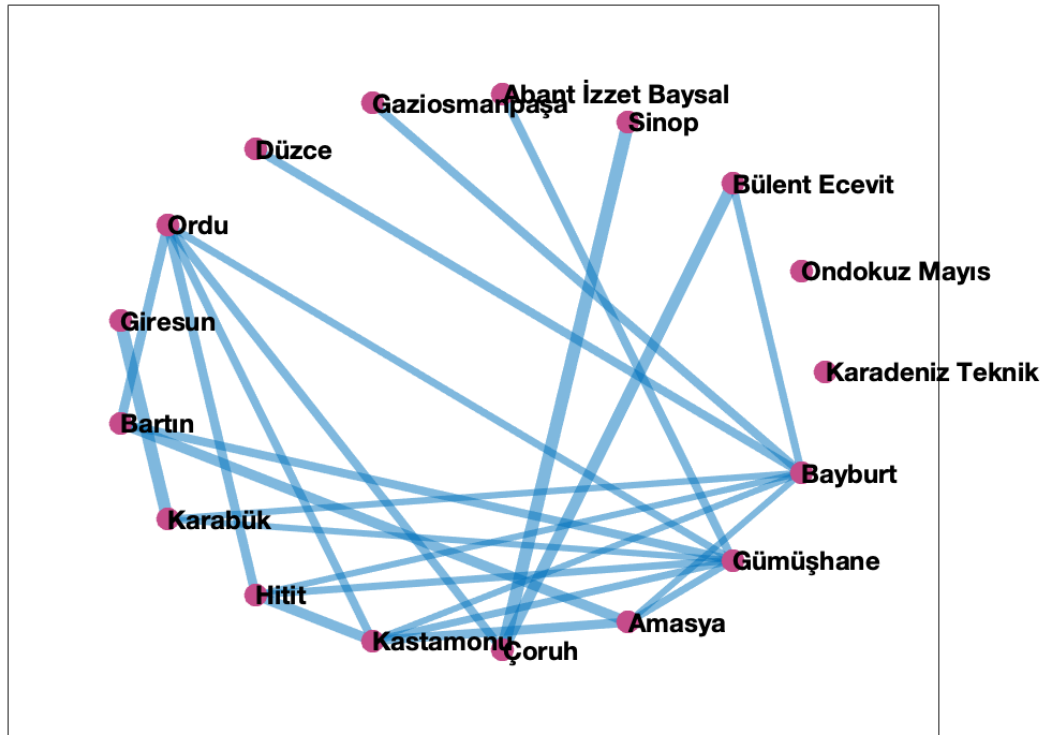


Şekil 5.5 : Karadeniz Bölgesi Üniversiteleri Arası İş Birliği

Şekil 5.6. ve Şekil 5.7. de gösterilen gelecekte Karadeniz Bölgesindeki üniversiteler arası bağlantı tahminleridir. Bu tahmin, veri setine bağlantı tahmin algoritmalarından adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme uygulanarak elde edilmiştir. Bütün bu algoritmalar aynı sonucu farklı skorlar ile vermiştir.



Şekil 5.6 : Karadeniz Bölgesi Üniversiteleri Arası İş Birliği ve Gelecekteki Bağlantı Tahmini



Şekil 5.7 : Karadeniz Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini

Adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme algoritmalarına göre bulunan sonuçlar aynıdır. Fakat algoritmaların referans aldıkları benzerlik ölçütleri farklı olduğundan farklı skorlar elde edilmiştir. Aşağıda gelecekte oluşması muhtemel bağlantı tahminleri ve skorları detaylı olarak gösterilmiştir.

Çizelge 5.2 : Karadeniz Bölgesi adamic adar bağlantı tahmin skorları.

	KATÜ	O. Mayıs	B. Ecevit	Sinop	A. İ. Baysal	Gaziosmanpaşa	Düzce	Ordu	Giresun	Bartın	Karabük	Hitit	Kastamonu	Çoruh	Amasya	Gümüşhane	Bayburt
KATÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O. Mayıs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B. Ecevit	0	0	0	0	0	0	0	0	0	0	0	0	0	4,6	0	0	3,0
Sinop	0	0	0	0	0	0	0	0	0	0	0	0	0	5,1	0	0	0
A.İ. Baysal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3,4	0
Gaziosmanpaşa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3,4
Düzce	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3,4
Ordu	0	0	0	0	0	0	0	0	0	3,7	0	3,7	3,3	3,4	0	3,0	0
Giresun	0	0	0	0	0	0	0	0	0	0	4,9	0	0	0	0	0	0
Bartın	0	0	0	0	0	0	0	3,7	0	0	0	0	0	0	4,1	3,4	0
Karabük	0	0	0	0	0	0	0	0	4,9	0	0	0	0	0	0	2,5	2,6
Hitit	0	0	0	0	0	0	0	3,7	0	0	0	0	4,1	0	0	2,9	2,6
Kastamonu	0	0	0	0	0	0	0	3,3	0	0	0	4,1	0	0	3,7	2,9	2,6
Çoruh	0	0	4,6	5,1	0	0	0	3,4	0	0	0	0	0	0	0	0	0
Amasya	0	0	0	0	0	0	0	0	0	4,1	0	0	3,7	0	0	2,9	2,6
Gümüşhane	0	0	0	0	3,4	0	0	3,0	0	3,4	2,5	2,9	2,9	0	2,9	0	0
Bayburt	0	0	3,0	0	0	3,4	3,4	0	0	0	2,6	2,6	2,6	0	2,6	0	0

Çizelge 5.3 : Karadeniz Bölgesi ortak komşu bağlantı tahmin skorları.

	KATÜ	O. Mayıs	B. Ecevit	Sinop	A. İ. Baysal	Gaziosmanpaşa	Düzce	Ordu	Giresun	Bartın	Karabük	Hitit	Kastamonu	Çoruh	Amasya	Gümüşhane	Bayburt
KATÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O. Mayıs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B. Ecevit	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	8
Sinop	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0
A.İ. Baysal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0
Gaziosmanpaşa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
Düzce	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
Ordu	0	0	0	0	0	0	0	0	0	10	0	10	9	9	0	8	0
Giresun	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0
Bartın	0	0	0	0	0	0	0	10	0	0	0	0	0	0	11	9	0
Karabük	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	7	7
Hitit	0	0	0	0	0	0	0	10	0	0	0	0	11	0	0	8	7
Kastamonu	0	0	0	0	0	0	0	9	0	0	0	11	0	0	10	8	7
Çoruh	0	0	12	13	0	0	0	9	0	0	0	0	0	0	0	0	0
Amasya	0	0	0	0	0	0	0	0	0	11	0	0	10	0	0	8	7
Gümüşhane	0	0	0	0	9	0	0	8	0	9	7	8	8	0	8	0	0
Bayburt	0	0	8	0	0	9	9	0	0	0	7	7	7	0	7	0	0

Çizelge 5.4 : Karadeniz Bölgesi jaccard bağlantı tahmin skorları.

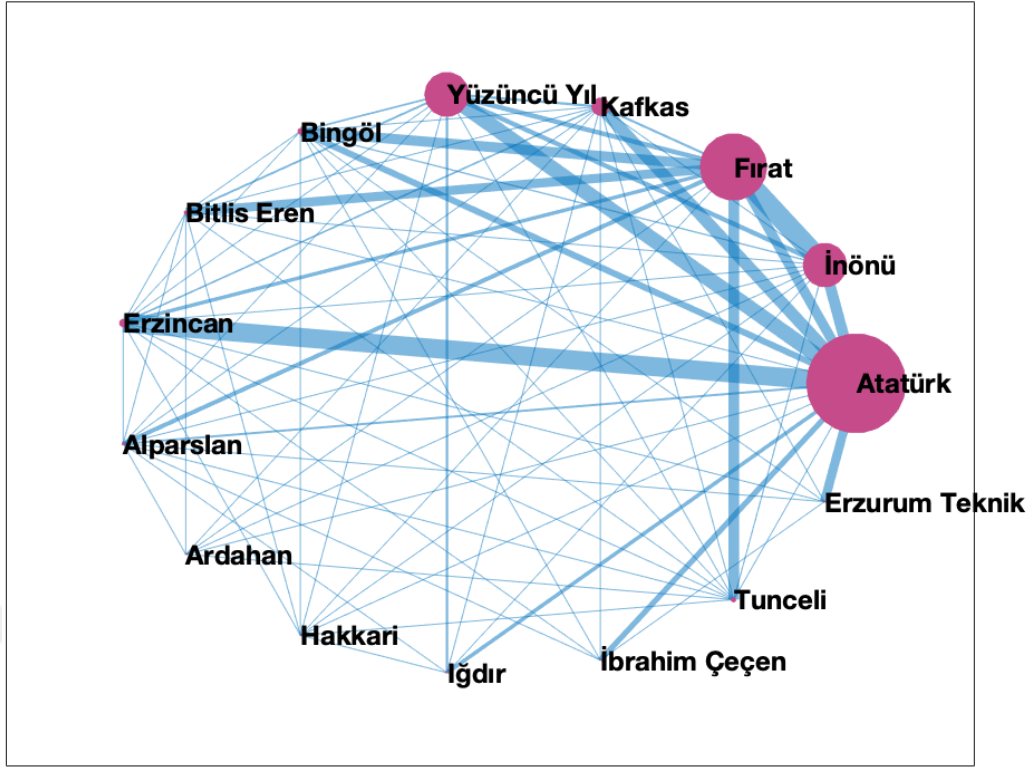
	KATÜ	O. Mayıs	B. Ecevit	Sinop	A. İ. Baysal	Gaziosmanpaşa	Düzce	Ordu	Giresun	Bartın	Karabük	Hitit	Kastamonu	Çoruh	Amasya	Gümüşhane	Bayburt
KATÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O. Mayıs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B. Ecevit	0	0	0	0	0	0	0	0	0	0	0	0	0	0,8	0	0	0,5
Sinop	0	0	0	0	0	0	0	0	0	0	0	0	0	0,8	0	0	0
A. İ. Baysal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,6	0
Gaziosmanpaşa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,6
Düzce	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,6
Ordu	0	0	0	0	0	0	0	0	0	0,7	0	0,7	0,6	0,6	0	0,6	0
Giresun	0	0	0	0	0	0	0	0	0	0	0,8	0	0	0	0	0	0
Bartın	0	0	0	0	0	0	0	0,7	0	0	0	0	0	0	0,7	0,6	0
Karabük	0	0	0	0	0	0	0	0	0,8	0	0	0	0	0	0	0,4	0,4
Hitit	0	0	0	0	0	0	0	0,7	0	0	0	0	0,9	0	0	0,6	0,5
Kastamonu	0	0	0	0	0	0	0	0,6	0	0	0	0,9	0	0	0,7	0,6	0,5
Çoruh	0	0	0,8	0,8	0	0	0	0,6	0	0	0	0	0	0	0	0	0
Amasya	0	0	0	0	0	0	0	0	0	0,7	0	0	0,7	0	0	0,6	0,5
Gümüşhane	0	0	0	0	0,6	0	0	0,6	0	0,6	0,4	0,6	0,6	0	0,6	0	0
Bayburt	0	0	0,5	0	0	0,6	0,6	0	0	0	0,4	0,5	0,5	0	0,5	0	0

Çizelge 5.5 : Karadeniz Bölgesi kaynak tahsisli bağlantı tahmin skorları

	KATÜ	O. Mayıs	B. Ecevit	Sinop	A. İ. Baysal	Gaziosmanpaşa	Düzce	Ordu	Giresun	Bartın	Karabük	Hitit	Kastamonu	Çoruh	Amasya	Gümüşhane	Bayburt
KATÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O. Mayıs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B. Ecevit	0	0	0	0	0	0	0	0	0	0	0	0	0	0,9	0	0	0,6
Sinop	0	0	0	0	0	0	0	0	0	0	0	0	0	1,0	0	0	0
A. İ. Baysal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,6	0
Gaziosmanpaşa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,6
Düzce	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,6
Ordu	0	0	0	0	0	0	0	0	0	0,7	0	0,6	0,6	0,6	0	0,5	0
Giresun	0	0	0	0	0	0	0	0	0	0	0,9	0	0	0	0	0	0
Bartın	0	0	0	0	0	0	0	0,7	0	0	0	0	0	0	0,7	0,6	0
Karabük	0	0	0	0	0	0	0	0	0,9	0	0	0	0	0	0	0,4	0,5
Hitit	0	0	0	0	0	0	0	0,6	0	0	0	0	0,7	0	0	0,5	0,4
Kastamonu	0	0	0	0	0	0	0	0,6	0	0	0	0,7	0	0	0,6	0,5	0,4
Çoruh	0	0	0,9	1,0	0	0	0	0,6	0	0	0	0	0	0	0	0	0
Amasya	0	0	0	0	0	0	0	0	0	0,7	0	0	0,6	0	0	0,5	0,4
Gümüşhane	0	0	0	0	0,6	0	0	0,5	0	0,6	0,4	0,5	0,5	0	0,5	0	0
Bayburt	0	0	0,6	0	0	0,6	0,6	0	0	0	0,5	0,4	0,4	0	0,4	0	0

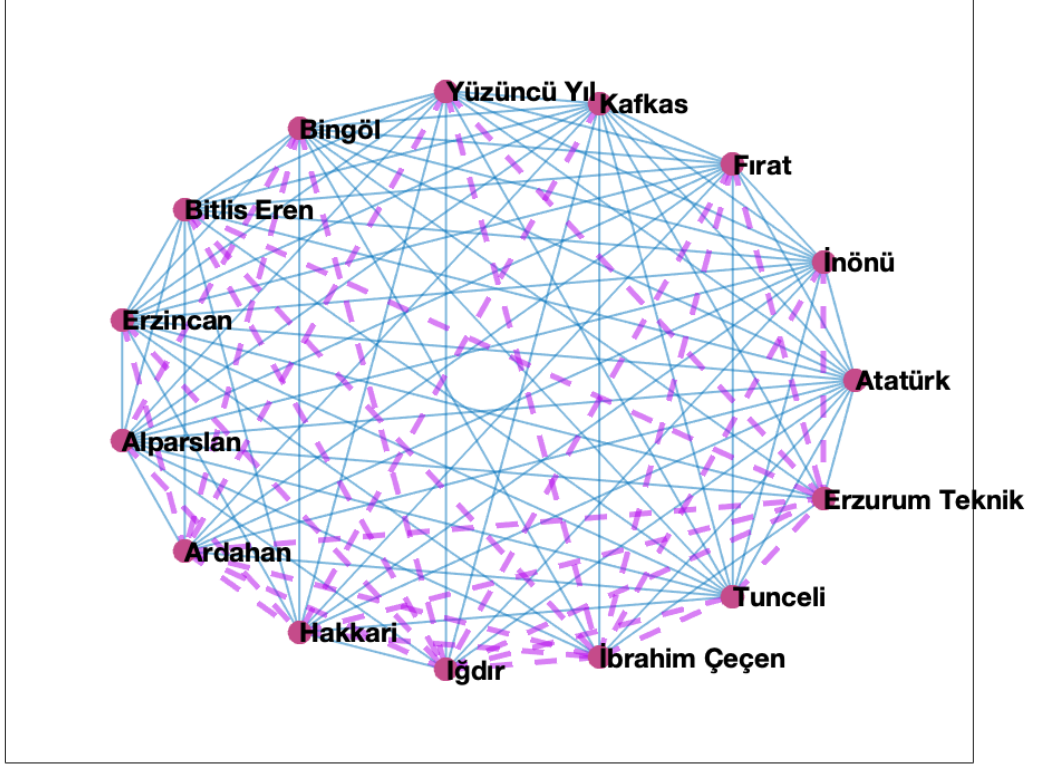
5.2.4. Doğu Anadolu Bölgesi

Veri seti içerisinde Doğu Anadolu Bölgesinde 15 üniversite yer almaktadır. Bu üniversitelerin mevcut durumdaki etkileşimleri Şekil 5.8. de gösterilmektedir. Düğümlerin büyüklüğü ağ içerisindeki bağlantı sayısına göre değişiklik göstermektedir. Ayrıtların kalınlığı ise birbirine bağladığı düğümlerin ne kadar etkileşimde olduğuna, yani üniversite iş birliklerinin sayısına bağlı olarak değişkenlik göstermektedir.

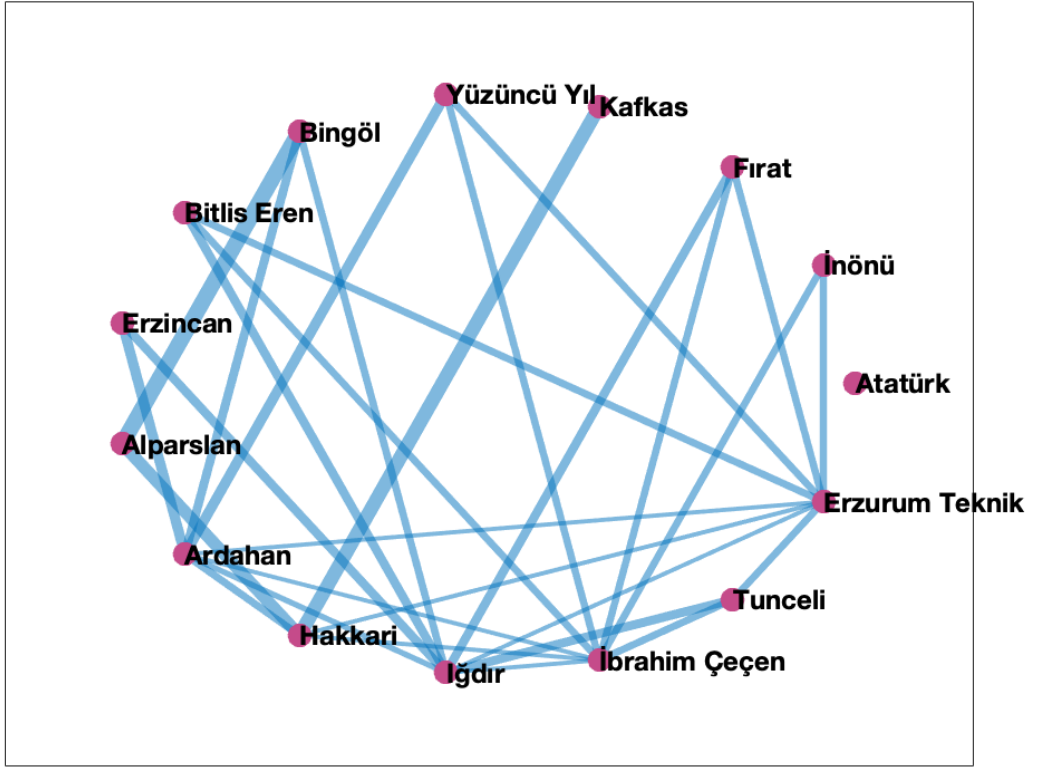


Şekil 5.8 : Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birliği

Şekil 5.9. ve Şekil 5.10. da gösterilen gelecekte Doğu Anadolu Bölgesindeki üniversiteler arası bağlantı tahminleridir. Bu tahmin, veri setine bağlantı tahmin algoritmalarından adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme uygulanarak elde edilmiştir. Bütün bu algoritmalar aynı sonucu farklı skorlar ile vermiştir.



Şekil 5.9 : Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birliği ve Gelecekteki Bağlantı Tahmini



Şekil 5.10 : Doğu Anadolu Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini

Adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme algoritmalarına göre bulunan sonuçlar aynıdır. Fakat algoritmaların referans aldıkları benzerlik ölçütleri farklı olduğundan farklı skorlar elde edilmiştir. Aşağıda gelecekte oluşması muhtemel bağlantı tahminleri ve skorları detaylı olarak gösterilmiştir.

Çizelge 5.6 : Doğu Anadolu Bölgesi adamic adar bağlantı tahmin skorları.

	Atatürk	İnönü	Fırat	Kafkas	YüzüncüYıl	Bingöl	Bitlis Eren	Erzincan	Alparslan	Ardahan	Hakkâri	Iğdır	İ.Çeçen	Tunceli	E.Teknik
Atatürk	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
İnönü	0	0	0	0	0	0	0	0	0	0	0	0	1,9	0	1,9
Fırat	0	0	0	0	0	0	0	0	0	0	0	2,4	1,9	0	1,9
Kafkas	0	0	0	0	0	0	0	0	0	0	3,8	0	0	0	0
YüzüncüYıl	0	0	0	0	0	0	0	0	0	2,8	0	0	1,9	0	1,9
Bingöl	0	0	0	0	0	0	0	0	4,3	2,4	0	2,0	0	0	0
Bitlis Eren	0	0	0	0	0	0	0	0	0	0	0	2,4	1,9	0	1,9
Erzincan	0	0	0	0	0	0	0	0	0	2,8	0	2,4	0	0	0
Alparslan	0	0	0	0	0	4,3	0	0	0	0	3,4	0	0	0	0
Ardahan	0	0	0	0	2,8	2,4	0	2,8	0	0	2,0	1,5	1,1	0	1,1
Hakkâri	0	0	0	3,8	0	0	0	0	3,4	2,0	0	0	1,1	0	1,1
Iğdır	0	0	2,4	0	0	2,0	2,4	2,4	0	1,5	0	0	1,1	2,4	1,1
İ.Çeçen	0	1,9	1,9	0	1,9	0	1,9	0	0	1,1	1,1	1,1	0	1,9	0
Tunceli	0	0	0	0	0	0	0	0	0	0	0	2,4	1,9	0	1,9
E.Teknik	0	1,9	1,9	0	1,9	0	1,9	0	0	1,1	1,1	1,1	0	1,9	0

Çizelge 5.7 : Doğu Anadolu Bölgesi ortak komşu bağlantı tahmin skorları.

	Atatürk	İnönü	Fırat	Kafkas	YüzüncüYıl	Bingöl	Bitlis Eren	Erzincan	Alparslan	Ardahan	Hakkâri	Iğdır	İ.Çeçen	Tunceli	E.Teknik
Atatürk	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
İnönü	0	0	0	0	0	0	0	0	0	0	0	0	5	0	5
Fırat	0	0	0	0	0	0	0	0	0	0	0	6	5	0	5
Kafkas	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0
YüzüncüYıl	0	0	0	0	0	0	0	0	0	7	0	0	5	0	5
Bingöl	0	0	0	0	0	0	0	0	10	6	0	5	0	0	0
Bitlis Eren	0	0	0	0	0	0	0	0	0	0	0	6	5	0	5
Erzincan	0	0	0	0	0	0	0	0	0	7	0	6	0	0	0
Alparslan	0	0	0	0	0	10	0	0	0	0	8	0	0	0	0
Ardahan	0	0	0	0	7	6	0	7	0	0	5	4	3	0	3
Hakkâri	0	0	0	9	0	0	0	0	8	5	0	0	3	0	3
Iğdır	0	0	6	0	0	5	6	6	0	4	0	0	3	6	3
İ.Çeçen	0	5	5	0	5	0	5	0	0	3	3	3	0	5	0
Tunceli	0	0	0	0	0	0	0	0	0	0	0	6	5	0	5
E.Teknik	0	5	5	0	5	0	5	0	0	3	3	3	0	3	0

Çizelge 5.8 : Doğu Anadolu Bölgesi jaccard bağlantı tahmin skorları.

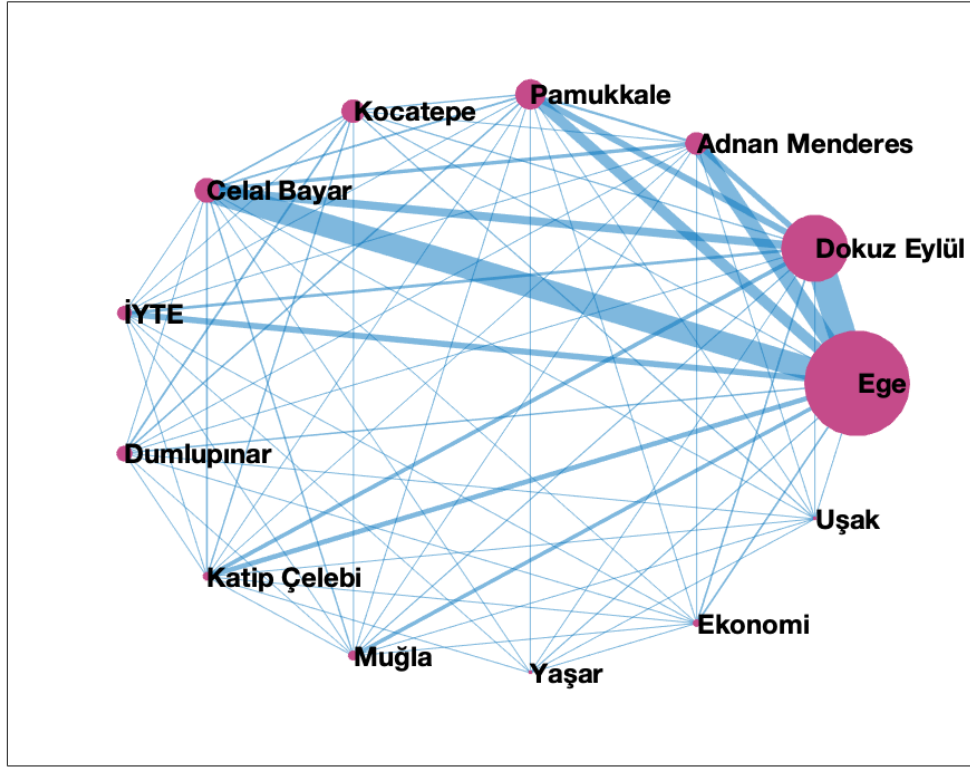
	Atatürk	İnönü	Fırat	Kafkas	Yüzüncü Yıl	Bingöl	Bitlis Eren	Erzincan	Alparslan	Ardahan	Hakkâri	İğdır	İ.Çeçen	Tunceli	E.Teknik
Atatürk	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
İnönü	0	0	0	0	0	0	0	0	0	0	0	0	0,3	0	0,3
Fırat	0	0	0	0	0	0	0	0	0	0	0	0,5	0,4	0	0,4
Kafkas	0	0	0	0	0	0	0	0	0	0	0,6	0	0	0	0
Yüzüncü Yıl	0	0	0	0	0	0	0	0	0	0,6	0	0	0,4	0	0,4
Bingöl	0	0	0	0	0	0	0	0	0,7	0,5	0	0,4	0	0	0
Bitlis Eren	0	0	0	0	0	0	0	0	0	0	0	0,5	0,4	0	0,4
Erzincan	0	0	0	0	0	0	0	0	0	0,5	0	0,5	0	0	0
Alparslan	0	0	0	0	0	0,7	0	0	0	0	0,6	0	0	0	0
Ardahan	0	0	0	0	0,6	0,5	0	0,5	0	0	0,4	0,4	0,3	0	0,3
Hakkâri	0	0	0	0,6	0	0	0	0	0,6	0,4	0	0	0,2	0	0,2
İğdır	0	0	0,5	0	0	0,4	0,5	0,5	0	0,4	0	0	0,3	0,5	0,3
İ.Çeçen	0	3,8	0,4	0	0,4	0	0,4	0	0	0,3	0,2	0,3	0	0,4	0
Tunceli	0	0	0	0	0	0	0	0	0	0	0	0,5	0,4	0	0,4
E.Teknik	0	3,8	0,4	0	0,4	0	0,4	0	0	0,3	0,2	0,3	0	0,4	0

Çizelge 5.9 : Doğu Anadolu Bölgesi kaynak tahsisli bağlantı tahmin skorları.

	Atatürk	İnönü	Fırat	Kafkas	YüzüncüYıl	Bingöl	Bitlis Eren	Erzincan	Alparslan	Ardahan	Hakkâri	Iğdır	İ.Çeçen	Tunceli	E.Teknik
Atatürk	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
İnönü	0	0	0	0	0	0	0	0	0	0	0	0	0,4	0	0,2
Fırat	0	0	0	0	0	0	0	0	0	0	0	0,5	0,4	0	0,4
Kafkas	0	0	0	0	0	0	0	0	0	0	0,8	0	0	0	0
YüzüncüYıl	0	0	0	0	0	0	0	0	0	0,5	0	0	0,4	0	0,4
Bingöl	0	0	0	0	0	0	0	0	1,0	0,5	0	0,4	0	0	0
Bitlis Eren	0	0	0	0	0	0	0	0	0	0	0	0,5	0,4	0	0,4
Erzincan	0	0	0	0	0	0	0	0	0	0,5	0	0,5	0	0	0
Alparslan	0	0	0	0	0	1,0	0	0	0	0	0,7	0	0	0	0
Ardahan	0	0	0	0	0,5	0,5	0	0,5	0	0	0,4	0,3	0,3	0	0,2
Hakkâri	0	0	0	0,8	0	0	0	0	0,7	0,4	0	0	0,2	0	0,2
Iğdır	0	0	0,5	0	0	0,4	0,5	0,5	0	0,3	0	0	0,2	0,5	0,2
İ.Çeçen	0	0,4	0,4	0	0,4	0	0,4	0	0	0,2	0,2	0,2	0	0,4	0
Tunceli	0	0	0	0	0	0	0	0	0	0	0	0,5	0,4	0	0,4
E.Teknik	0	0,4	0,4	0	0,4	0	0,4	0	0	0,2	0,2	0,2	0	0,4	0

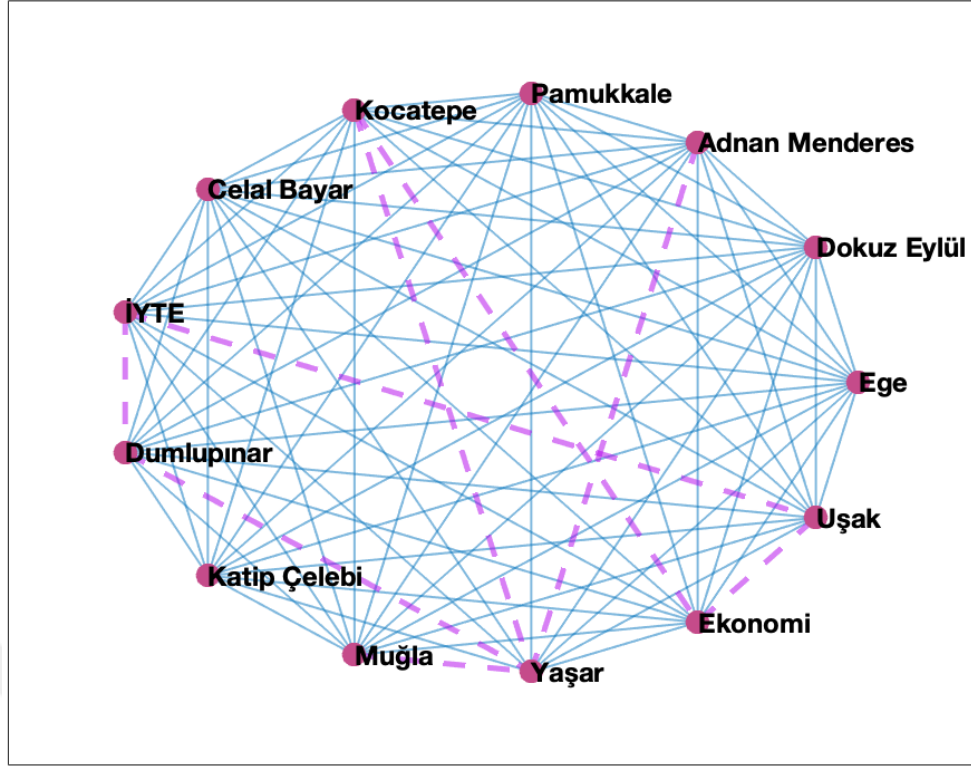
5.2.5. Ege Bölgesi

Veri seti içerisinde Ege Bölgesinde 13 üniversite yer almaktadır. Bu üniversitelerin mevcut durumdaki etkileşimleri Şekil 5.11. de gösterilmektedir. Düzgümlerin büyüklüğü ağ içerisindeki bağlantı sayısına göre değişiklik göstermektedir. Ayrıtların kalınlığı ise birbirine bağladığı düğümlerin ne kadar etkileşimde olduğuna, yani üniversite iş birliklerinin sayısına bağlı olarak değişkenlik göstermektedir.

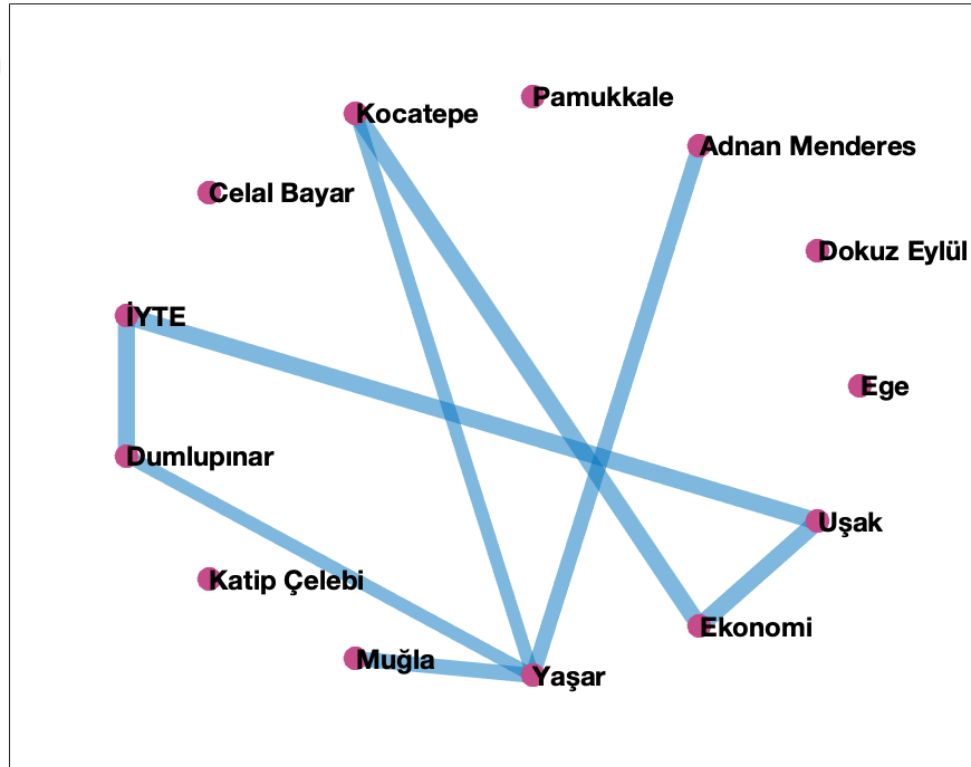


Şekil 5.11 : Ege Bölgesi Üniversiteleri Arası İş Birliği

Şekil 5.12. ve Şekil 5.13. da gösterilen gelecekte Ege Bölgesindeki üniversiteler arası bağlantı tahminleridir. Bu tahmin, veri setine bağlantı tahmin algoritmalarından adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme uygulanarak elde edilmiştir. Bütün bu algoritmalar aynı sonucu farklı skorlar ile vermiştir.



Şekil 5.12 : Ege Bölgesi Üniversiteleri Arası İş Birliği ve Gelecekteki Bağlantı Tahmini



Şekil 5.13 : Ege Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini

Adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme algoritmalarına göre bulunan sonuçlar aynıdır. Fakat algoritmaların referans aldıkları benzerlik ölçütleri farklı olduğundan farklı skorlar elde edilmiştir. Aşağıda gelecekte oluşması muhtemel bağlantı tahminleri ve skorları detaylı olarak gösterilmiştir.

Çizelge 5.10 : Ege Bölgesi adamic adar bağlantı tahmin skorları.

	Ege	Dokuz Eylül	A.Menderes	Pamukkale	Kocatepe	C.Bayar	IYTE	Dumlupınar	Kâtip Çelebi	Muğla	Yaşar	Ekonomi	Uşak
Ege	0	0	0	0	0	0	0	0	0	0	0	0	0
D.Eylül	0	0	0	0	0	0	0	0	0	0	0	0	0
A.Menderea	0	0	0	0	0	0	0	0	0	0	3,2	0	0
Pamukkale	0	0	0	0	0	0	0	0	0	0	0	0	0
Kocatepe	0	0	0	0	0	0	0	0	0	0	2,8	3,7	0
C.Bayar	0	0	0	0	0	0	0	0	0	0	0	0	0
IYTE	0	0	0	0	0	0	0	3,7	0	0	0	0	3,7
Dumlupınar	0	0	0	0	0	0	3,7	0	0	0	2,8	0	0
Kâtip Çelebi	0	0	0	0	0	0	0	0	0	0	0	0	0
Muğla	0	0	0	0	0	0	0	0	0	0	3,3	0	0
Yaşar	0	0	3,3	0	2,8	0	0	2,8	0	3,3	0	0	0
Ekonomi	0	0	0	0	3,7	0	0	0	0	0	0	0	3,7
Uşak	0	0	0	0	0	0	3,7	0	0	0	0	3,7	0

Çizelge 5.11 : Ege Bölgesi ortak komşu bağlantı tahmin skorları.

	Ege	Dokuz Eylül	A.Menderes	Pamukkale	Kocatepe	C.Bayar	IYTE	Dumlupınar	Kâtip Çelebi	Muğla	Yaşar	Ekonomi	Uşak
Ege	0	0	0	0	0	0	0	0	0	0	0	0	0
D.Eylül	0	0	0	0	0	0	0	0	0	0	0	0	0
A.Menderes	0	0	0	0	0	0	0	0	0	0	8	0	0
Pamukkale	0	0	0	0	0	0	0	0	0	0	0	0	0
Kocatepe	0	0	0	0	0	0	0	0	0	0	7	9	0
C.Bayar	0	0	0	0	0	0	0	0	0	0	0	0	0
IYTE	0	0	0	0	0	0	0	9	0	0	0	0	9
Dumlupınar	0	0	0	0	0	0	9	0	0	0	7	0	0
Kâtip Çelebi	0	0	0	0	0	0	0	0	0	0	0	0	0
Muğla	0	0	0	0	0	0	0	0	0	0	8	0	0
Yaşar	0	0	8	0	7	0	0	7	0	8	0	0	0
Ekonomi	0	0	0	0	9	0	0	0	0	0	0	0	9
Uşak	0	0	0	0	0	0	9	0	0	0	0	9	0

Çizelge 5.12 : Ege Bölgesi jaccard bağlantı tahmin skorları.

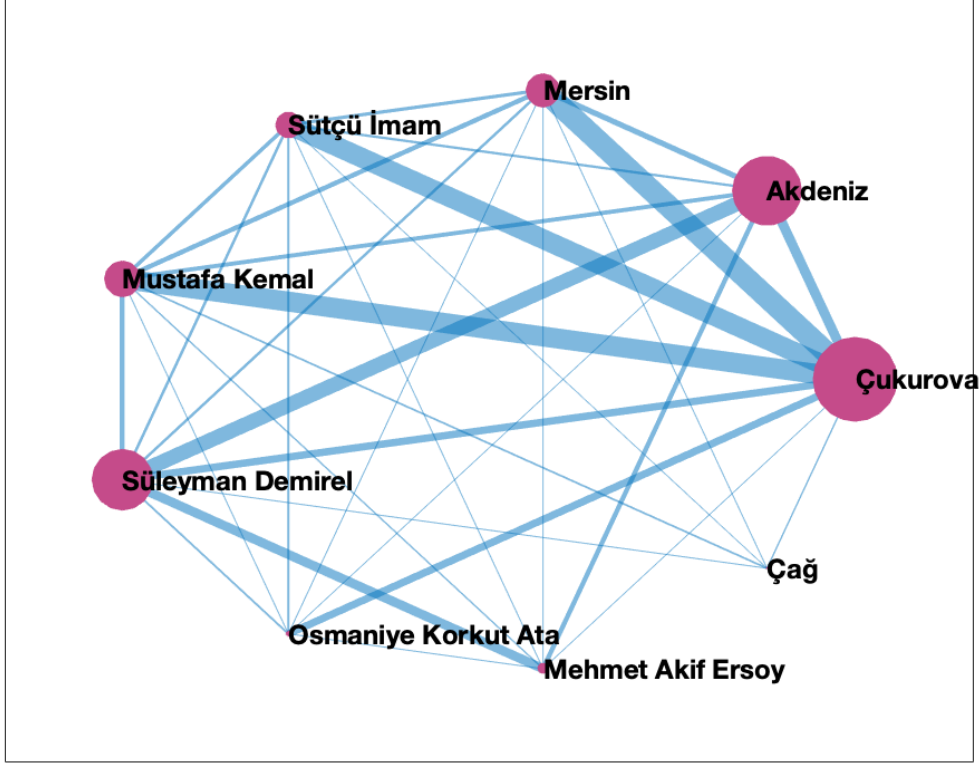
	Ege	Dokuz Eylül	A.Menderes	Pamukkale	Kocatepe	C.Bayar	IYTE	Dumlupınar	Kâtip Çelebi	Muğla	Yaşar	Ekonomi	Uşak
Ege	0	0	0	0	0	0	0	0	0	0	0	0	0
D.Eylül	0	0	0	0	0	0	0	0	0	0	0	0	0
A.Menderea	0	0	0	0	0	0	0	0	0	0	0,7	0	0
Pamukkale	0	0	0	0	0	0	0	0	0	0	0	0	0
Kocatepe	0	0	0	0	0	0	0	0	0	0	0,6	0,8	0
C.Bayar	0	0	0	0	0	0	0	0	0	0	0	0	0,8
IYTE	0	0	0	0	0	0	0	0,8	0	0	0	0	0
Dumlupınar	0	0	0	0	0	0	0,8	0	0	0	0,6	0	0
Kâtip Çelebi	0	0	0	0	0	0	0	0	0	0	0	0	0
Muğla	0	0	0	0	0	0	0	0	0	0	0,7	0	0
Yaşar	0	0	0,7	0	0,6	0	0	0,6	0	0,7	0	0	0
Ekonomi	0	0	0	0	0,8	0	0	0	0	0	0	0	0,8
Uşak	0	0	0	0	0	0	0,8	0	0	0	0	0,8	0

Çizelge 5.13 : Ege Bölgesi kaynak tahsisli bağlantı tahmin skorları.

	Ege	Dokuz Eylül	A.Menderes	Pamukkale	Kocatepe	C.Bayar	IYTE	Dumlupınar	Kâtip Çelebi	Muğla	Yaşar	Ekonomi	Uşak
Ege	0	0	0	0	0	0	0	0	0	0	0	0	0
D.Eylül	0	0	0	0	0	0	0	0	0	0	0	0	0
A.Menderea	0	0	0	0	0	0	0	0	0	0	0,7	0	0
Pamukkale	0	0	0	0	0	0	0	0	0	0	0	0	0
Kocatepe	0	0	0	0	0	0	0	0	0	0	0,6	0,7	0
C.Bayar	0	0	0	0	0	0	0	0	0	0	0	0	0,8
IYTE	0	0	0	0	0	0	0	0,7	0	0	0	0	0
Dumlupınar	0	0	0	0	0	0	0,7	0	0	0	0,6	0	0
Kâtip Çelebi	0	0	0	0	0	0	0	0	0	0	0	0	0
Muğla	0	0	0	0	0	0	0	0	0	0	0,7	0	0
Yaşar	0	0	0,7	0	0,6	0	0	0,6	0	0,7	0	0	0
Ekonomi	0	0	0	0	0,7	0	0	0	0	0	0	0	0,8
Uşak	0	0	0	0	0	0	0,8	0	0	0	0	0,8	0

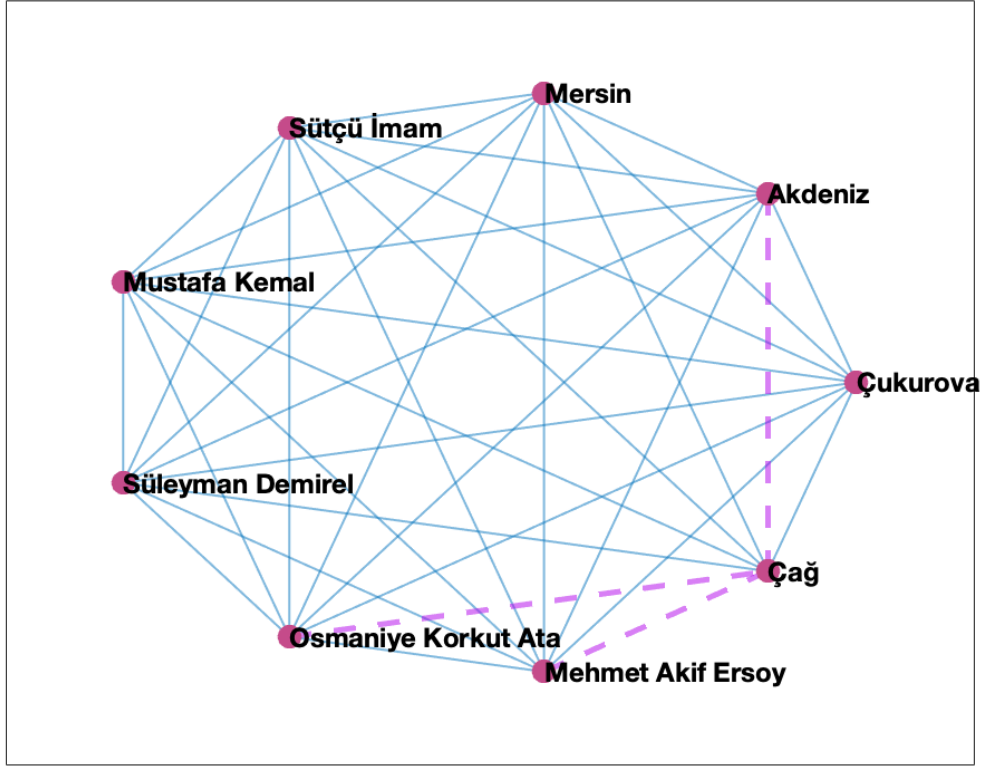
5.2.6. Akdeniz Bölgesi

Veri seti içerisinde Akdeniz Bölgesinde 9 üniversite yer almaktadır. Bu üniversitelerin mevcut durumdaki etkileşimleri Şekil 5.14. de gösterilmektedir. Düğümlerin büyüklüğü ağ içerisindeki bağlantı sayısına göre değişiklik göstermektedir. Ayrıtların kalınlığı ise birbirine bağladığı düğümlerin ne kadar etkileşimde olduğuna, yani üniversite iş birliklerinin sayısına bağlı olarak değişkenlik göstermektedir.

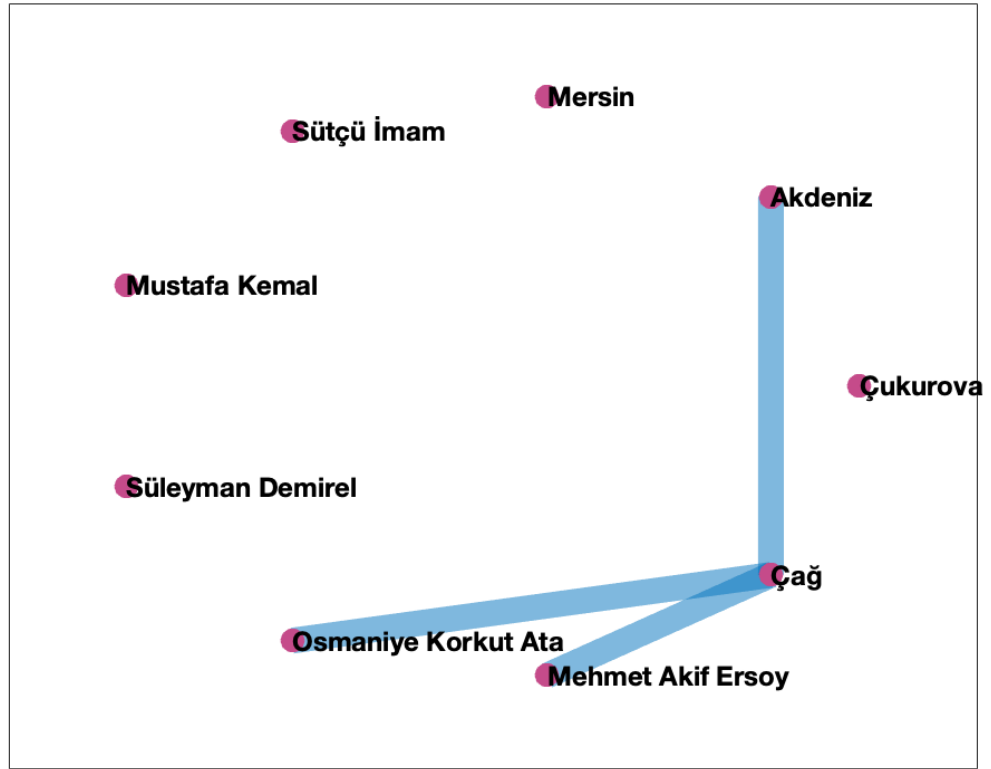


Şekil 5.14 : Akdeniz Bölgesi Üniversiteleri Arası İş Birlikleri

Şekil 5.15. ve Şekil 5.16. da gösterilen gelecekte Akdeniz Bölgesindeki üniversiteler arası bağlantı tahminleridir. Bu tahmin, veri setine bağlantı tahmin algoritmalarından adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme uygulanarak elde edilmiştir. Bütün bu algoritmalar aynı sonucu farklı skorlar ile vermiştir.



Şekil 5.15 : Akdeniz Bölgesi Üniversiteleri Arası İş Birlikleri ve Gelecekteki Bağlantı Tahmini



Şekil 5.16 : Akdeniz Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini

Adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme algoritmalarına göre bulunan sonuçlar aynıdır. Fakat algoritmaların referans aldıkları benzerlik ölçütleri farklı olduğundan farklı skorlar elde edilmiştir. Aşağıda gelecekte oluşması muhtemel bağlantı tahminleri ve skorları detaylı olarak gösterilmiştir.

Çizelge 5.14 : Akdeniz Bölgesi adamic adar bağlantı tahmin skorları.

	Çukurova	Akdeniz	Mersin	Sütçü İmam	M. Kemal	S. Demirel	O. Korkut Ata	M. Akif Ersoy	Çağ
Çukurova	0	0	0	0	0	0	0	0	0
Akdeniz	0	0	0	0	0	0	0	0	2,4
Mersin	0	0	0	0	0	0	0	0	0
Sütçü İmam	0	0	0	0	0	0	0	0	0
M. Kemal	0	0	0	0	0	0	0	0	0
S. Demirel	0	0	0	0	0	0	0	0	0
O. Korkut Ata	0	0	0	0	0	0	0	0	2,4
M. Akif Ersoy	0	0	0	0	0	0	0	0	2,4
Çağ	0	2,4	0	0	0	0	2,4	2,4	0

Çizelge 5.15 : Akdeniz Bölgesi ortak komşu bağlantı tahmin skorları.

	Çukurova	Akdeniz	Mersin	Sütçü İmam	M. Kemal	S. Demirel	O. Korkut Ata	M. Akif Ersoy	Çağ
Çukurova	0	0	0	0	0	0	0	0	0
Akdeniz	0	0	0	0	0	0	0	0	5
Mersin	0	0	0	0	0	0	0	0	0
Sütçü İmam	0	0	0	0	0	0	0	0	0
M. Kemal	0	0	0	0	0	0	0	0	0
S. Demirel	0	0	0	0	0	0	0	0	0
O. Korkut Ata	0	0	0	0	0	0	0	0	5
M. Akif Ersoy	0	0	0	0	0	0	0	0	5
Çağ	0	5	0	0	0	0	5	5	0

Çizelge 5.16 : Akdeniz Bölgesi jaccard bağlantı tahmin skorları.

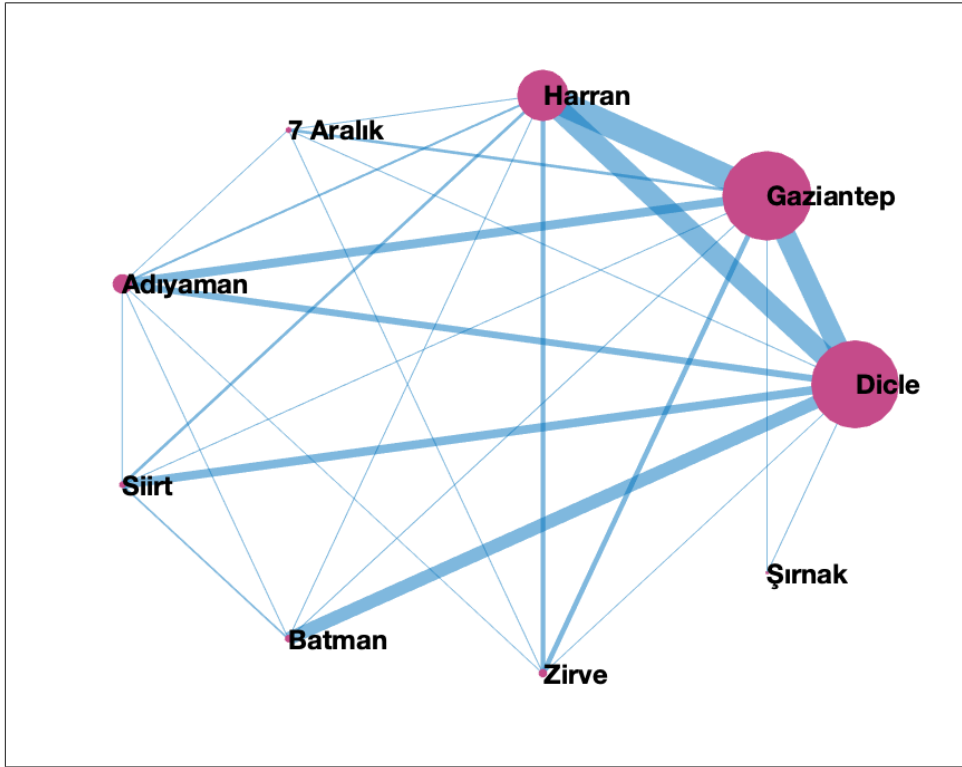
	Çukurova	Akdeniz	Mersin	Sütçü İmam	M. Kemal	S. Demirel	O. Korkut Ata	M. Akif Ersoy	Çağ
Çukurova	0	0	0	0	0	0	0	0	0
Akdeniz	0	0	0	0	0	0	0	0	0,7
Mersin	0	0	0	0	0	0	0	0	0
Sütçü İmam	0	0	0	0	0	0	0	0	0
M. Kemal	0	0	0	0	0	0	0	0	0
S. Demirel	0	0	0	0	0	0	0	0	0
O. Korkut Ata	0	0	0	0	0	0	0	0	0,7
M. Akif Ersoy	0	0	0	0	0	0	0	0	0,7
Çağ	0	0,7	0	0	0	0	0,7	0,7	0

Çizelge 5.17 : Akdeniz Bölgesi kaynak tahsisli bağlantı tahmin skorları.

	Çukurova	Akdeniz	Mersin	Sütçü İmam	M. Kemal	S. Demirel	O. Korkut Ata	M. Akif Ersoy	Çağ
Çukurova	0	0	0	0	0	0	0	0	0
Akdeniz	0	0	0	0	0	0	0	0	0,6
Mersin	0	0	0	0	0	0	0	0	0
Sütçü İmam	0	0	0	0	0	0	0	0	0
M. Kemal	0	0	0	0	0	0	0	0	0
S. Demirel	0	0	0	0	0	0	0	0	0
O. Korkut Ata	0	0	0	0	0	0	0	0	0,6
M. Akif Ersoy	0	0	0	0	0	0	0	0	0,6
Çağ	0	0,6	0	0	0	0	0,6	0,6	0

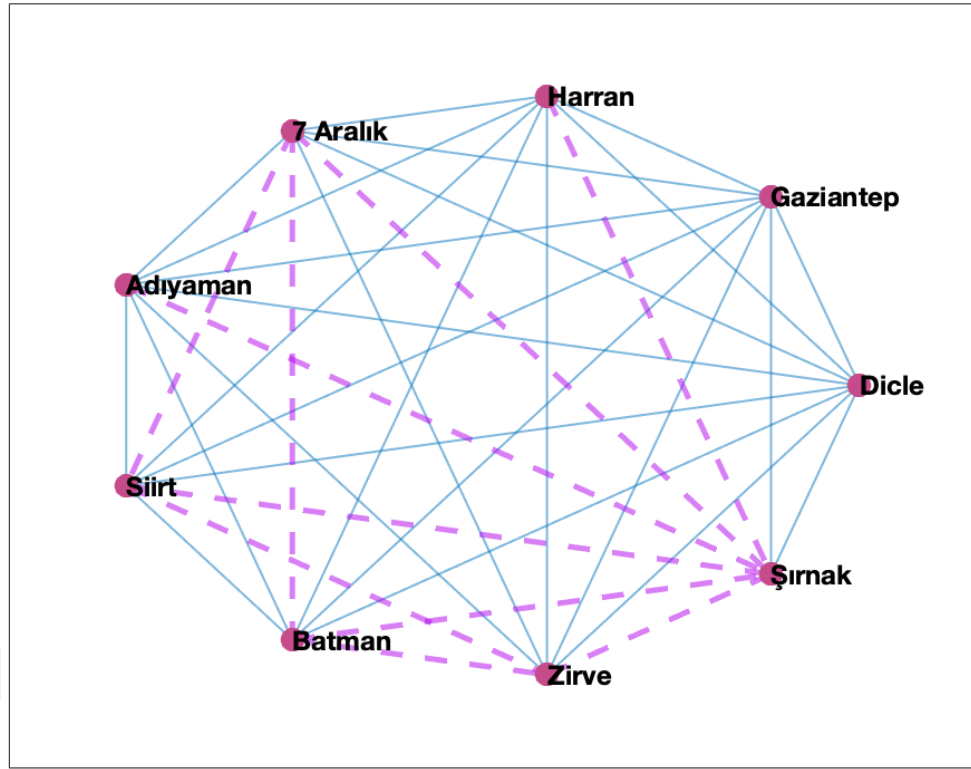
5.2.7. Güney Doğu Anadolu Bölgesi

Veri seti içerisinde Güney Doğu Anadolu Bölgesinde 9 üniversite yer almaktadır. Bu üniversitelerin mevcut durumdaki etkileşimleri Şekil 5.17. de gösterilmektedir. Düğümlerin büyüklüğü ağ içerisindeki bağlantı sayısına göre değişiklik göstermektedir. Ayrıkların kalınlığı ise birbirine bağladığı düğümlerin ne kadar etkileşimde olduğuna, yani üniversite iş birliklerinin sayısına bağlı olarak değişiklik göstermektedir.

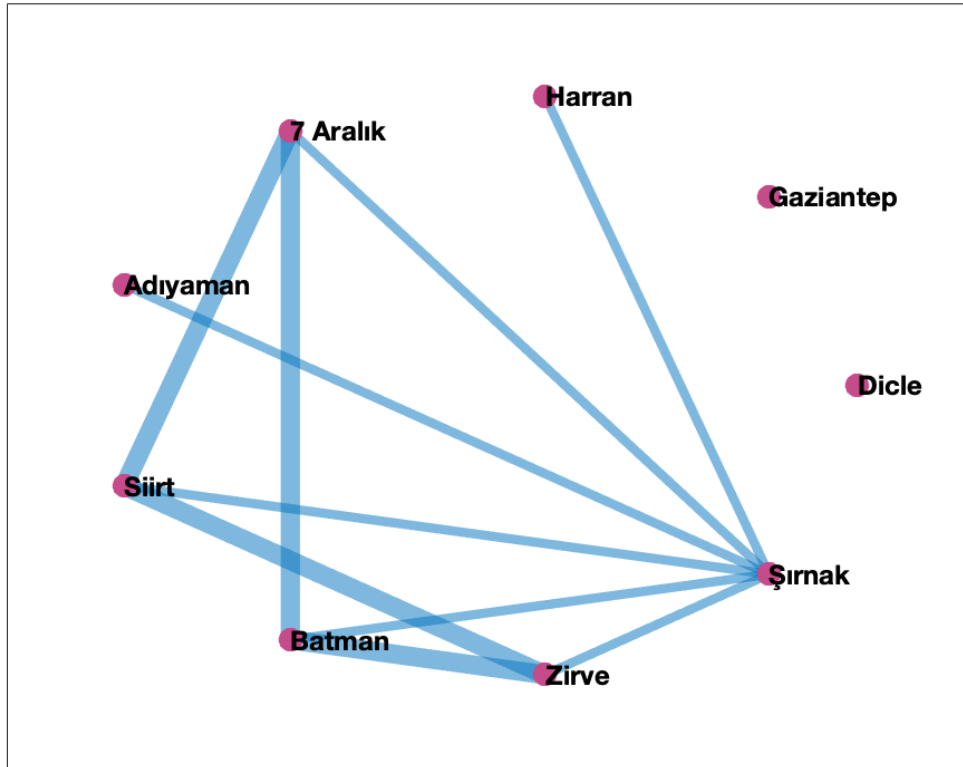


Şekil 5.17 : Güney Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birlikleri

Şekil 5.18. ve Şekil 5.19. da gösterilen gelecekte Güney Doğu Anadolu Bölgesindeki üniversiteler arası bağlantı tahminleridir. Bu tahmin, veri setine bağlantı tahmin algoritmalarından adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme uygulanarak elde edilmiştir. Bütün bu algoritmalar aynı sonucu farklı skorlar ile vermiştir.



Şekil 5.18 : Güney Doğu Anadolu Bölgesi Üniversiteleri Arası İş Birlikleri ve Gelecekteki Bağlantı Tahmini



Şekil 5.19 : Güney Doğu Anadolu Bölgesi Üniversiteleri Arası Gelecekteki Bağlantı Tahmini

Adamic adar, ortak komşu, jaccard, kaynak tahsisli indeksleme algoritmalarına göre bulunan sonuçlar aynıdır. Fakat algoritmaların referans aldıkları benzerlik ölçütleri farklı olduğundan farklı skorlar elde edilmiştir. Aşağıda gelecekte oluşması muhtemel bağlantı tahminleri ve skorları detaylı olarak gösterilmiştir.

Çizelge 5.18 : Güney Doğu Anadolu Bölgesi adamic adar bağlantı tahmin skorları.

	Dicle	Gaziantep	Harran	7 Aralık	Adıyaman	Siirt	Batman	Zirve	Şırnak
Dicle	0	0	0	0	0	0	0	0	0
Gaziantep	0	0	0	0	0	0	0	0	0
Harran	0	0	0	0	0	0	0	0	0,9
7 Aralık	0	0	0	0	0	1,9	1,9	0	0,9
Adıyaman	0	0	0	0	0	0	0	0	0,9
Siirt	0	0	0	1,9	0	0	0	1,9	0,9
Batman	0	0	0	1,9	0	0	0	1,9	0,9
Zirve	0	0	0	0	0	1,9	1,9	0	0,9
Şırnak	0	0	0,9	0,9	0,9	0,9	0,9	0,9	0

Çizelge 5.19 : Güney Doğu Anadolu Bölgesi ortak komşu bağlantı tahmin skorları.

	Dicle	Gaziantep	Harran	7 Aralık	Adıyaman	Siirt	Batman	Zirve	Şırnak
Dicle	0	0	0	0	0	0	0	0	0
Gaziantep	0	0	0	0	0	0	0	0	0
Harran	0	0	0	0	0	0	0	0	2
7 Aralık	0	0	0	0	0	4	4	0	2
Adıyaman	0	0	0	0	0	0	0	0	2
Siirt	0	0	0	4	0	0	0	4	2
Batman	0	0	0	4	0	0	0	4	2
Zirve	0	0	0	0	0	4	4	0	2
Şırnak	0	0	2	2	2	2	2	2	0

Çizelge 5.20 : Güney Doğu Anadolu Bölgesi jaccard bağlantı tahmin skorları.

	Dicle	Gaziantep	Harran	7 Aralık	Adıyaman	Siirt	Batman	Zirve	Şırnak
Dicle	0	0	0	0	0	0	0	0	0
Gaziantep	0	0	0	0	0	0	0	0	0
Harran	0	0	0	0	0	0	0	0	0,2
7 Aralık	0	0	0	0	0	0,6	0,6	0	0,4
Adıyaman	0	0	0	0	0	0	0	0	0,2
Siirt	0	0	0	0,6	0	0	0	0,6	0,4
Batman	0	0	0	0,6	0	0	0	0,6	0,4
Zirve	0	0	0	0	0	0,6	0,6	0	0,4
Şırnak	0	0	0,2	0,4	0,2	0,4	0,4	0,4	0

Çizelge 5.21 : Güney Doğu Anadolu Bölgesi kaynak tahsisli bağlantı tahmin skorları.

	Dicle	Gaziantep	Harran	7 Aralık	Adıyaman	Siirt	Batman	Zirve	Şırnak
Dicle	0	0	0	0	0	0	0	0	0
Gaziantep	0	0	0	0	0	0	0	0	0
Harran	0	0	0	0	0	0	0	0	0,2
7 Aralık	0	0	0	0	0	0,5	0,5	0	0,2
Adıyaman	0	0	0	0	0	0	0	0	0,2
Siirt	0	0	0	0,5	0	0	0	0,5	0,2
Batman	0	0	0	0,5	0	0	0	0,5	0,2
Zirve	0	0	0	0	0	0,5	0,5	0	0,2
Şırnak	0	0	0,2	0,2	0,2	0,2	0,2	0,2	0

6. SONUÇLAR

Günümüzde siber ortamda üretilen anlık veri miktarının devasa boyutlara ulaştığı aşıkardır. Verilerin ölçümlenebilir olmasından dolayı birçok bilim alanında veri yığınları üzerinde gerekli analizler yapılarak, bu veriler faydalı bilgilere dönüştürülmektedir. Bu tez çalışmasının konusu olan üniversiteler arası iş birliği alanında yapılan tahminlerle de akademik çalışmaların hızlandırılması ve iş birliklerinin güçlendirilmesi hedeflenmiştir.

Bütün analizlerin gerçekleştirilmesi için sosyal ağ analizinde de sıklıkla kullanılan metotlardan biri olan bağlantı tahmini yöntemleri kullanılmıştır. Bağlantı tahmini veri madenciliği yöntemlerinin sosyal ağlar üzerine uygulanması ile edilmiştir. Bazı yerel indeksleme yöntemleri kullanılarak iş birlikleri hesaplanmış ve Türkiye'deki akademik iş birlikleri baz alınarak gelecekte oluşması tahmin edilen yeni bağlantılar tahmin edilmiştir. Önceki bölümlerde açıklanan Ortak Komşu, Adamic Adar, Jaccard, Kaynak Tahsisli yerel benzerlik algoritmaları kullanılarak bağlantı tahmini yapılmıştır. Bu yöntemlerin seçilmesinin sebebi Liben-Nowell vd. [24] ve Zhou vd. [25] yerel benzerlik indislerini kullanarak gerçek ağlar üzerinde kıyaslamalar yapmaları ve kullanılan yöntemler içerisinde en iyi sonuçları bu yöntemlerin vermesidir. [26]

Çalışmada komşuluk ilişkilerini kullanan yöntemler üzerinden hesaplamalar yapılmıştır. Kullanılan bütün algoritmalar hemen hemen aynı sonucu vermiştir. Hesaplamaların doğruluğu irdelenmiştir. Uygulamada elde edilen bilgiler incelendiğinde yapılan tahminlerin gerçekçi sonuçlar verdiği görülmüştür. [26]

Türkiye'nin bilimsel ağı hakkında öngörülerde bulunulup, ilerde ilişki kurması muhtemel üniversiteler belirlenmiştir. Özellikle gelişmekte olan veya geliştirilmesi hedeflenen üniversitelerle ilişki kurması muhtemel üniversitelerin ortak projeler geliştirmesi veya bu ve bunun gibi aktivitelerde kullanılabilecek referans bir çalışma oluşturulmaya çalışılmıştır.

KAYNAKLAR

- [1] Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.
- [2] Wikipedia: *Çizge Teorisi*. FL: Wikipedia Foundation, Inc. Retrieved May, 2020, from https://tr.wikipedia.org/wiki/Çizge_teorisi
- [3] Wikipedia: *Königsberg'in yedi köprüsü*. FL: Wikipedia Foundation, Inc. Retrieved May, 2020, from https://tr.wikipedia.org/wiki/Königsberg%27in_yedi_köprüsü
- [4] Euler, L. (1741). Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, 128-140.
- [5] Aslan, S., & Kaya, M. (2018, September). Predicting Links in Complex Bipartite Networks Based on Strengthening Projections. In 2018 *International Conference on Artificial Intelligence and Data Processing (IDAP)* (pp. 1-6). IEEE.
- [6] Chang, W. L., & Lin, T. H. (2010, August). A cluster-based approach for automatic social network construction. In 2010 *IEEE Second International Conference on Social Computing* (pp. 601-606). IEEE.
- [7] Ting, I. H. (2008, June). Web mining techniques for on-line social networks analysis. In 2008 *International Conference on Service Systems and Service Management* (pp. 1-5). IEEE.
- [8] Karıcı, A., & Boy, O. (2011). Sosyal ağların web madenciliği teknikleri ile analizi ve ortak atıf analizi ile benzerlik tahmini. *Elektrik-Elektronik ve Bilgisayar Sempozyumu*, 154-161.
- [9] Scott, J. (2000). *Social network analysis: A handbook*. Sage London. 2nd edition.
- [10] Jaina, R. K., & Sheth, A. (Eds.). (2007). *Social Networks and the Semantic Web. Semantic Web and Beyond: Computing for Human Experience*. Springer.
- [11] Newman, M. E. (2008). The mathematics of networks. *The new palgrave encyclopedia of economics*, 2(2008), 1-12.
- [12] Schank, T., & Wagner, D. (2004). *Approximating clustering-coefficient and transitivity*. Universität Karlsruhe, Fakultät für Informatik.
- [13] Al Hasan, M., & Zaki, M. J. (2010). Link Prediction in Social Networks. *Department of Computer Science, Rensselaer Polytechnic Institute, Troy, USA*.
- [14] Zhou, T., Lü, L., & Zhang, Y. C. (2009). Predicting missing links via local information. *The European Physical Journal B*, 71(4), 623-630.
- [15] Murata, T., & Moriyasu, S. (2008). Link prediction based on structural properties of online social networks. *New Generation Computing*, 26(3), 245-257.
- [16] Radicchi, F., Fortunato, S., Markines, B., & Vespignani, A. (2009). Diffusion of scientific credits and the ranking of scientists. *Physical Review E*, 80(5), 056103.
- [17] Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18(1), 39-43.
- [18] Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine.
- [19] Reuters, T. (2012). Web of Science. [Online]. Available: <https://apps.webofknowledge.com>. [Accessed: 01-Jan-2015].

- [20] İnce, K., & Karci, A. (2017, September). Collaboration network analysis of Turkey in regional basis. In *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1-6). IEEE.
- [21] İnce, K. (2016). Analyzing Tendency of Academic Growth in Turkey. *International Artificial Intelligence and Data Processing Symposium*, 400-403.
- [22] İnce, K., & Karci, A. (2016). Analysing and Modelling of Academic Collaboration of Turkey. In *International Conference on Natural Science and Engineering* (pp. 2772-2776).
- [23] da Silva Soares, P. R., & Prudêncio, R. B. C. (2012, June). Time series based link prediction. In *The 2012 international joint conference on neural networks (IJCNN)* (pp. 1-7). IEEE.
- [24] Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. *Journal of the American society for information science and technology*, 58(7), 1019-1031.
- [25] Murata, T., & Moriyasu, S. (2008). Link prediction based on structural properties of online social networks. *New Generation Computing*, 26(3), 245-257.
- [26] Tuğal, İ., Kaya, M., & Tuncer, T. (2013). Link prediction in disease and drug networks. In *6-th International Conference of Advanced Computer Systems and Networks: Design and Application (ACSN 2013)* (pp. 46-50).

ÖZGEÇMİŞ

Ad Soyad : Melike Merve TEMEL
Doğum Yeri ve Tarihi : Malatya / 10.07.1991
E-Posta : melikemervetemel@gmail.com

ÖĞRENİM DURUMU:

- **Lisans** : 2013, İnönü Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği

MESLEKİ DENEYİM:

- 2013-2014 Büyük Doğu Yazılım şirketinde yazılım uzmanı olarak çalışmıştır.
- 2014-halen Çalık Denim şirketinde yazılım uzmanı olarak çalışmaktadır.

YÜKSEK LİSANS TEZİNDEN TÜRETİLEN ÇALIŞMALAR

- **KARADENİZ, E., TEMEL, M. M., & KARCI, A. (2018, September)**. Prediction of Collaboration Between Universities of Turkey. In 2018 International Conference on Artificial Intelligence and Data Processing (IDAP) (pp. 1-4). IEEE.