



T.C.
MALTEPE ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

ARAŞTIRMACI AĞI ARAMA MOTORU

YASİN YENER
Yüksek Lisans Tezi

Tez Danışmanı
Doç. Dr. Turgay Tugay Bilgin

İSTANBUL – 2016

T.C.
MALTEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

ARAŞTIRMACI AĞI ARAMA MOTORU

YÜKSEK LİSANS TEZİ

Yasin Yener

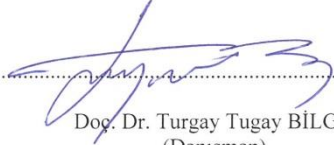
Tez Danışmanı


Doç. Dr. Turgay Tugay Bilgin

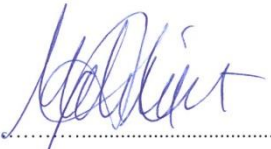
İSTANBUL – 2016

T.C. Maltepe Üniversitesi
Fen Bilimleri Enstitüsü Müdürlüğü'ne,

29.06.2016 tarihinde tezinin savunmasını yapan Yasin YENER'e ait "Araştırmacı Ağı Arama Motoru" başlıklı çalışma, Jürimiz tarafından Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Bilgisayar Mühendisliği Tezli Yüksek Lisans Programında Yüksek Lisans Tezi Olarak **Oy Birliği/Oy Çokluğuyla** Kabul Edilmiştir.


Doç. Dr. Turgay Tugay BİLGİN
(Danışman)


Doç. Dr. Serhat ÖZEKES
(Üye)


Yrd. Doç. Dr. Mehmet Ali Aksoy TÜYSÜZ
(Üye)

ÖZET

Yüksek Lisans Tezi, Arařtırmacı Ađı Arama Motoru, T.C. Maltepe Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliđi Anabilim Dalı.

Bu tez alıřması, TÜBİTAK tarafından geliřtirilen ARBİS sistemine kayıtlı arařtırmacıların, veri madenciliđi ve sosyal ađ analizi yöntemleri kullanılarak birbirlerine olan benzerliklerinin tespit edilmesini amalamaktadır.

ARBİS sistemine kayıtlı bir arařtırmacı, ARBİS üzerinden sorgulama servisini kullanarak bir faaliyet alanındaki diđer arařtırmacıların bilgilerine ulaşabilmektedir. Ancak sorgulama, faaliyet alanlarında yapılabildiđi için, sadece sorgulanan alanda yüzlerce kiřinin bilgisi verilmektedir. Verilen bu bilgilerden, bir arařtırmacının diđer bir arařtırmacı ile olan yakınlıđı tespit edilememektedir.

Tez kapsamında geliřtirilen bu alıřma ile, bir arařtırmacının ismi sisteme girilip arama yapıldıđında, ilgili kiřiye yakın otuz arařtırmacının sonucu gösterilmektedir. Bu sayede arařtırmacı kimlerin kendisine daha yakın olduđunu yakınlık derecesine göre görsel olarak görerek kendisine benzer kiřiler ile bađlantı sađlayabilmektedir. Bu sayede ARBİS sistemine kayıtlı arařtırmacıların sosyal ađı ıkartılmıřtır.

Anahtar Kelimeler: Veri Madenciliđi, Komřuluk Arama, Graf Yapılar, Sosyal Ađ Analizi, Arařtırmacı Ađı, ARBİS.

ABSTRACT

Master's Thesis, Researcher network search engine, T.C Maltepe University, Institute of Sciences, Department of Computer Engineering.

This graduate thesis aims to determine the similarity of researchers to each other who registered to ARBIS system developed by TUBITAK using data mining and social network analysis.

Researcher registered on ARBIS system, is able to query information of the other researchers working on any activity by using ARBIS system. However, as the query can be made only for areas of activity, the information of the hundreds people is given. The relation of researchers in the system can not be determined with this provided information.

As a part of this study, once the name of a researcher is searched in the system, thirty researchers who are relevant with related person are shown. Researcher visually sees the relatives who are close to him/her and could provide connections with those people. Social network of registered researchers to ARBIS has been created by that way.

Keywords: Data Mining, Neighborhood Search, Graph Structures, Social Network Analysis, Researcher Network, ARBIS.

TEŞEKKÜR

Tez konumu seçmeme yardımcı olan, çalışmaya teşvik eden, bu süreçte yol göstericiliğini ve bilgisini benden esirgemeyen değerli danışman hocam Doç. Dr. Turgay Tugay Bilgin'e, teşekkürlerimi sunarım.

Yüksek lisans dönemi boyunca eğitimime verdikleri destek ve hassasiyetleri için arkadaşlarıma, sevgili oğlum Yağız Arda Yener'e, sevgili eşim Gülçin Yener'e ve aileme saygı, sevgi ve sonsuz teşekkürlerimi sunarım.

Ayrıca eğitimim konusunda tarafıma desteğini esirgemeyen, İnsan Kaynakları Müdürümüz Ertan Bilek'e, firma yöneticilerimiz, Han Schut ve Brett Argirakis'e, çalışma arkadaşlarıma teşekkürlerimi sunarım.

Mayıs – 2016

Yasin Yener

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER	iv
ŞEKİLLER DİZİNİ.....	vi
TABLolar DİZİNİ	vii
DENKLEMLER.....	viii
SİMGELER DİZİNİ VE KISALTMALAR	ix
1. GİRİŞ.....	1
1.1 Tez Çalışmasının Amacı	1
1.2 Kapsam.....	1
1.3 Benzer Çalışmalar	2
2. PROBLEMİN TANIMI.....	4
2.1 Araştırmacı Ağlarının Önemi	4
2.2 ARBİS Portalının Yapısı	5
2.3 ARBİS Portalının Eksiklikleri.....	8
2.4 YÖK Akademik Arama Uygulaması Yapısı	9
2.5 Önerilen Modelin Avantajları.....	10
3. LİTERATÜR İNCELEMESİ	11
3.1 Sosyal Ağ Kavramı	11
3.2 Sosyal Ağ Arama Motorları	12
3.3 Sosyal Ağ Analizi Yaklaşımları	12
3.2.1 Bağlantı Tabanlı ve Yapısal Analizler	13
3.2.2 Hareketli ve Sabit Analiz	13

3.4	Sosyal Ağ Analizi Sınıflama	13
4.	ARAŞTIRMACI AĞI ARAMA MOTORU UYGULAMASI.....	14
4.1	Veri Toplama ve Ön İşleme	14
4.1.1	ARBİS’de Teknik Bilimler Faaliyet Alanının Alt Kategorilerinin İncelenmesi	16
4.1.2	Otomatik Sorgulama İçin Excel Makro ve Data Load Programının Kullanımı	18
4.1.3	Teknik Bilimler Altındaki Kategorilerin Sorgulanması ve Sonuçların html Olarak Kaydedilmesi	21
4.1.4	HTML Sonuçların Ayrıştırılarak Veri Tabanına Kaydı.....	21
4.1.5	Veritabanından Matris yapıda CSV Dosyasının Hazırlanması.....	22
4.2	Uygulamanın Hazırlanması	24
4.2.1	R Dilinde Kodlanan K En Yakın Komşu Algoritmasının PHP ile Entegrasyonu	24
4.2.2	KNN Algoritmasının Yapısı	30
4.2.3	R Uygulamasında Dönen Sonuca Göre Çizge Çizilmesi.....	30
5.	SONUÇLAR VE DEĞERLENDİRME.....	34
5.1	Elde Edilen Sonuçlar	34
5.2	Değerlendirme	34
5.3	Öneriler.....	35
	KAYNAKLAR	36
	ÖZGEÇMİŞ	38

ŞEKİLLER DİZİNİ

Şekil 1.Microsoft Akademik Arama Aracı	5
Şekil 2.ARBİS Sorgulama Ekranı.....	6
Şekil 3.ARBİS Sorgulama Ekranı.....	7
Şekil 4.ARBİS Sorgulama Sonucu	8
Şekil 5.YÖK Akademik Ana Sayfa	9
Şekil 6.YÖK Akademik Arama Sonucu	10
Şekil 7.YÖK Akademik Çizge Grafik Çizimi	10
Şekil 8.Sosyal Ağ Gösterimi.....	12
Şekil 9.Sosyal Ağ Sınıflaması [24]	14
Şekil 10.Ana Koordinat Hesaplaması	18
Şekil 11.Data Load Programı Örnek Komutları	19
Şekil 12.Data Load Programı Ekran Görüntüsü.....	20
Şekil 13.ARBİS den bilgilerin alınması.....	21
Şekil 14.HTML Dosyaların Ayrıştırılması ve DB'ye aktarılması.....	22
Şekil 15.Data Load Programı Ekran Görüntüsü.....	22
Şekil 16.Veritabanından CSV Dosyasının oluşturulması.....	23
Şekil 17.CSV Dosyası Ekran Görüntüsü.....	23
Şekil 18.PHP ile R'nin Haberleşmesi ve Çizge Grafik Çizimi.....	24
Şekil 19.Sorgulama Ekran Görüntüsü.....	25
Şekil 20.Query değişkeni içeriği.....	27
Şekil 21.knndata verisi.....	28
Şekil 22.Nodes Değişkeni	31
Şekil 23.Links Değişkeni	32
Şekil 24.Çizge Grafik Görünümü	33

TABLolar DİZİNİ

Tablo 1.Excel Tablosu Ekran Görüntüsü.....	15
Tablo 2.X ve Y Koordinatı Hesaplaması	17
Tablo 3.Data Load komutları tablosu.....	19
Tablo 4.CSV DosyaTablo Görünümü.....	26
Tablo 5.Satır ve Sütun Adları tanımlaması sonucu.....	26
Tablo 6. Şekil 21'in detay tablosu.	29

DENKLEMLER

Denklem 1.Öklid Uzaklığı Denklemi	28
---	----

SİMGELER DİZİNİ VE KISALTMALAR

SNA	: Social Network Analysis (Sosyal Ağ Analizi)
KNN	: K- Nearest Neighbors Algorithm (K En yakın komşu Algoritması)
PHP	: Hypertext Preprocessor (Üstünyazı Önışlemcisi)
SQL	: Structured Query Language (Yapılandırılmış Sorgu Dili)
ARBİS	: Arařtırmacı Bilgi Sistemi
TÜBİTAK	: Türkiye Bilimsel ve Teknolojik Arařtırma Kurumu
TC	: Türkiye Cumhuriyeti
YÖK	: Yüksek Öğretim Kurumu
CSV	: Comma Separated Values (Virgöl ile Ayrılmış Değerler)
DB	: Database (Veri Tabanı)
ID	: Defination of id (Tanımlama numarası)
Oracle	: Merkezi Amerika'da olan bir yazılım firmasının yazılımının adıdır.
SAP	: System Analysis and Program Development (Sistem analizi ve Program Geliştirme) Avrupanın en büyük kurumsal yazılım geliřtiren yazılım řirketidir ve aynı zamanda geliřtirdikleri yazılımın adıdır.
ERP	: Enterprise Resources Planning (Kurumsal Kaynak Planlaması)
Excel	: Microsoft tarafından Microsoft Windows ve Apple Macintosh iřletim sistemleri tabanında çalışmak üzere yazılan ve dağıtımı yapılan bir hesap çizelgesi programıdır.

1. GİRİŞ

1.1 Tez Çalışmasının Amacı

Günümüzde sosyal ağ kavramı sıklıkla kullanılmakta ve internetin yaygınlaşması ile her bir grup kendisine özgü bir ağ kurmaktadır. Bu ağlardan çeşitli analizler yapılarak reklam, satış, müşteri ve insan davranışları gibi konularda bir bilgi oluşturulmaya çalışılmaktadır.

Türkiye'deki araştırmacıların bilgilerinin tutulabilmesi ve araştırmacıların birbirleri ile bağlantılarının sağlanması amaçlı TÜBİTAK tarafından ARBİS sistemi geliştirilmiştir. Ancak bu sistem veri tabanı gibi bilgilerin tutulduğu bir yapı olarak çalışmaktadır.

Bu tez çalışması ile araştırmacıların birbirleri ile olan bağları yani sosyal ağları görsel olarak sunulmaktadır. Geliştirilen uygulama kullanılarak bağlantıların tespiti kolaylaştırılmıştır.

1.2 Kapsam

Bu tez çalışması ile Türkiye'de TÜBİTAK'ın oluşturduğu ARBİS portalına kayıtlı araştırmacıların ilgili faaliyet alanlarına göre birbirleri ile olan benzerliklerinin tespit edilerek çizge grafikte gösterilmesi amaçlanmıştır. Bu sayede günümüzde oldukça yaygınlaşan sosyal ağ kavramını araştırmacılar için kullanılarak Araştırmacı Ağı adlı bir uygulama geliştirilmiştir.

ARBİS'de, Teknik Bilimler faaliyet alanının altında bulunan her bir faaliyet alanı tek tek sorgulanarak elde edilen araştırmacı bilgileri kayıt edilmiştir. Bu sonuçlar ayrıştırıldıktan sonra veri madenciliği teknikleri uygulanıp çizge grafik ile sorgulama yapan kişiye gösterilmektedir. Çizge grafik ile, bir araştırmacının, kendisine benzer diğer araştırmacıları tespit etmesi amaçlanmıştır.

ARBİS üzerinde var olan arařtırmacı ađının, grselleřtirilmesi sađlanmıřtır. Yani Arbis Sosyal Ađının verileri iřlenerek gnmz Sosyal Ađları gibi grsellik ile bađlantıları gsteririr bir izge grafiđe dnřtrlmřtir. Veriler anlamlı bilgiye dnřtrlerek arařtırmacıların zaman kaybını nleyecek bir sistem ortaya ıkartılmıřtır.

1.3 Benzer alıřmalar

İnternet kullanımının yaygınlařması ile insanların birbirleri ile iletiřimi kolaylařmıřtır. İnsanlar dođaları geređi, gerek hayatta olduđu gibi sanal ortamda da topluluklar oluřtırmaya bařlamıřlardır. İnsanların internet kullanımlarının artması ile byk veriler oluřmaktadır. Bu verilerin anlamlı bilgilere dnřtrlmesi de veri madenciliđi yntemlerinin kullanılması ile sađlanmaktadır. Bu uygulamalar ile eřitli analizler yapılması ihtiyaı dođmuřtur. Bu ihtiyaa ynelik yapılmıř olan sosyal ađ analizleri, komřuluk iliřkilerine dair bazı alıřmalar ařađıda aıklanmaktadır.

Necla Tektař [1] niversite đrencilerinin sosyal ađları kullanımlarına ynelik bir arařtırması řu řekildedir; Her geen gn geliřen bilgi ve iletiřim teknolojisindeki geliřmeler interneti gnlk yařamın vazgeilmezleri arasına almıřtır. İnternet ortamında, kullanıcıların kendilerini tanımlayarak her kltrden farklı kullanıcı ile iletiřime getiđi, bunun yanında normal sosyal yařamda kullanılan jest ve mimik hareketlerini simgeleyen sembollerle duygu ve dřncelerini sanal olarak ifade ettiđi, bu řekilde sosyal iletiřim kurduđu ortamlar “sosyal ađ” olarak tanımlanmaktadır. Sosyal ađların yaygınlařması, gnmzde artık her řey gibi sosyalliđin de llebilir bir hal almasını sađlamıřtır. Sosyalliđin lt de byk oranda sosyal paylařım siteleri olmuřtur. zellikle đrenciler arasında oka kullanılan sosyal ađların hayatın vazgeilmezi haline gelmesi dikkat ekmektedir. alıřmasını; Marmara niversitesi Teknik Bilimler Meslek Yksekokulu’nda okuyan đrencilerin sosyal ađlara bakıř aılarının analizini tespit etmek zere hazırlanmıřtır. Bu bađlamda đrencilere anket uygulanmıř, elde edilen veriler analiz edilerek sosyal ađları hangi amala kullandıkları deđerlendirilmiřtir.

Doç Dr.Ali Karıcı, Onur Boy [2] Sosyal ağların web madenciliği teknikleri ile analizi ve ortak atıf analizi ile benzerlik tahmini konulu araştırması şu şekildedir; Sosyal ağlar, barındırdıkları sosyal aktörlerin ilişkilerine ait oldukça faydalı bilgiler ihtiva etmektedirler. Bu yapı ve ilişkilerin analiz edilmesi aracılığıyla yakınlıklar, benzeşmeler, eğilimler ve etkileşimler gibi verilere ulaşılarak ağdaki ilişkiler hakkında çeşitli yorumlara veya tahminlere varılabilir. Özellikle internet teknolojisinde yaşanan teknik ve kültürel gelişmeler ile sosyal ağlara ilişkin veriler somutlaşmış ve ölçülebilir hale gelmiştir. İnternet üzerinde gerçekleşen sosyal iletişim sonucu oluşan veri yığını çok büyük boyutlara ulaşmaktadır. Söz konusu devasa boyuttaki veri yığını içerisinde anlamlı bilginin çıkarılabilmesi için “Web Madenciliği” teknikleri kullanılmaktadır. Dijital sosyal ağların web madenciliği teknikleri kullanılarak analiz edilmesi ve bu sayede akademik, ticari, sosyolojik vb. alanlarda kıymetli verilere ulaşılması önemli bir çalışma konusu haline gelmiştir. Çalışmasında, genel hatları ile sosyal ağların web madenciliği teknikleri ile analizine ilişkin yöntemler incelenmiştir. Anlatılan konulara ilişkin somut bir örnek uygulama sunmak amacıyla web madenciliği tekniklerinden biri olan “Web Yapı Madenciliği” seçilmiş olup “Ortak Atıf Analizi ile Benzerlik Tahmini” uygulaması geliştirilmiştir.

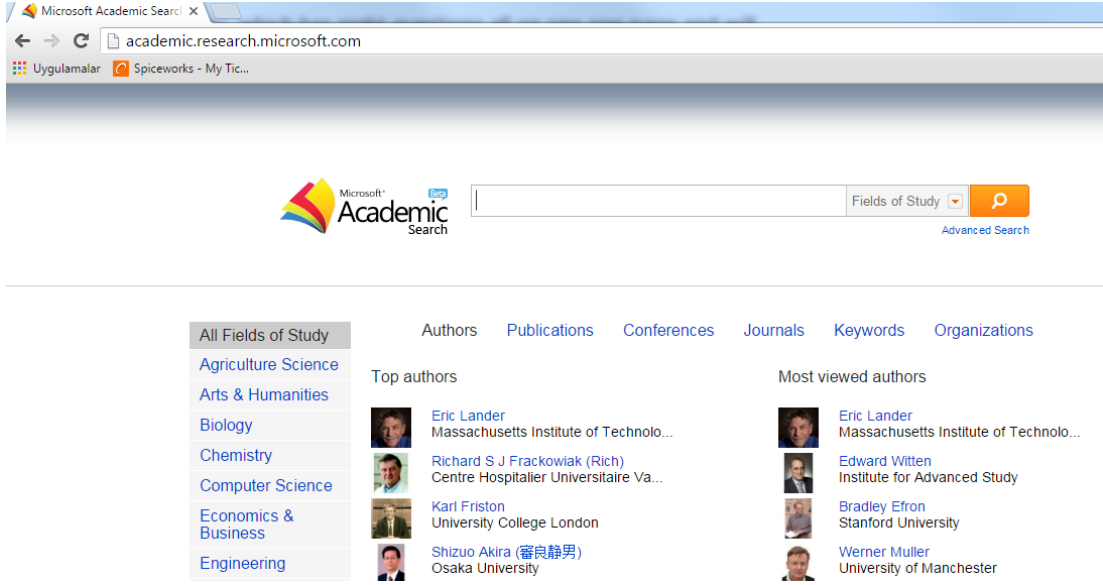
Mehmet Ulvi Şimşek [3] Bu tez çalışmasında sosyal ağlarda veri madenciliği ile yapılan uygulamalar, alandaki sorunlar ve açık çalışma alanları incelenmiştir. Çalışmasında, twitter sosyal ağı verileri toplanmış ve duygu belirten Türkçe kelimeler belirlenmiştir. Twitter verileri her gün için ayrı olarak duygu belirten kelimelerin kullanılmasına göre analiz edilmiştir. Türk Twitter kullanıcılarının duygusal durum değişiminin İMKB indeksi ile ilişkili olup olmadığı araştırılmıştır. Twitter mesajlarını analiz etmek için frekans yöntemi ve ortalama mutluluk analizi yöntemi kullanılmış ve bu iki metot karşılaştırılmıştır. Tweet veri seti içerisinde mesajları olan kullanıcıların yaptıkları yorumlara göre benzerliği incelenmiştir. Bu tez çalışması, alanda Türkçe tweetler ve borsa indeksi üzerinde yapılmış ilk çalışma olması açısından önem taşımaktadır.

2. PROBLEMİN TANIMI

2.1 Araştırmacı Ağlarının Önemi

Akademik alanda faaliyet gösteren insanların birbirleri ile bağlantı sağlamaları amaçlı bazı çalışmalar yapılmıştır. Örnek göstermek gerekirse, Microsoft firmasının <http://academic.research.microsoft.com> adlı web uygulamasını gösterilebiliriz. Bu uygulamada, arama yapılan akademik personele benzer diğer akademik personeller görsel bir çizge grafik ile gösterilmektedir. Yine burada çıkan sonuca göre her bir akademik personelin bağlantılarına gidilebilmektedir. Ancak bu uygulama Microsoft firması tarafından sonlandırılmış olup çizge grafik yerine metinsel sonuçlar verilmektedir.

Zamanın önem kazandığı günümüzde, bilgisayar teknolojisinin ve yazılımın ilerlemesi ile artık insanların büyük veriler ile uğraşarak zaman kaybetmesi hiç kabul edilebilir değildir. Bu zaman kaybını zamanları son derece önemli bilim insanları için düşündüğümüzde telafisi olmayacaktır. Facebook, twitter ve benzeri uygulamalarda insanların birbirleri ile olan bağlantıları çıkartılırken, araştırmacılar için yapılmış çalışmalar diğer sosyal ağlar kadar başarılı değil ya da bağlantıyı bulmak kolay değildir. Araştırmacılar da kendi bağlantılarını kolaylıkla bulabilmelidirler. Bağlantılarına erişimin kolay olması sayesinde, yaptıkları ya da yapacak oldukları çalışmalarda etkileşim halinde olacaklardır. Hem kendilerine hem de ülkemize katma değer sağlanacaktır. Bu kapsamda araştırmacı ağları önem kazanmaktadır.



Şekil 1. Microsoft Akademik Arama Aracı.

2.2 ARBİS Portalının Yapısı

Araştırmacı Bilgi Sistemi (ARBİS), Türkiye'nin araştırmacı veritabanını oluşturmak ve sürekli olarak güncellemek amacıyla, TÜBİTAK tarafından tasarlanan ve geliştirilen web tabanlı bir uygulamadır.

ARBİS ile, Türkiye'de görev yapan tüm araştırmacılar ve yurtdışında çalışan Türk araştırmacıların, veritabanına tek bir noktadan kayıt olarak bilgilerini girebilmeleri ve güncelleyebilmelerinin sağlanması amaçlanmıştır.

ARBİS'e kayıt yaptıran ve bilgilerini girerek onay alan araştırmacılar, AB Yedinci Çerçeve Programı gibi çeşitli iletişim platformlarından yararlanabilmektedirler. TÜBİTAK projelerine, başvurma ve hakem-izleyici hizmeti verme gibi etkinlikler için bir ön koşulu yerine getirmiş olmaktadır[26].

Araştırmacı öncelikle, TC vatandaşları için TC Kimlik numarasını, Yabancı Uyruklu araştırmacı için Pasaport Numarasını kullanarak ARBİS'e kaydını yapar. TÜBİTAK tarafından kayıt onaylandıktan sonra sistem kullanılmaya başlanır. Şekil 2'de

görüldüğü gibi Sorgulama Servisi menüsüne gidilerek, ekranda bulunan kriterlere göre giriş yapılarak sonuçlara ulaşılır.

SORGULAMA Kriterleri

Sorgulama sonunda bilgilerinin görüntülenmesine izin veren arařtırmacılar görüntülenir. Ayrıca bu arařtırmacıların sadece izin verdikleri bilgileri ekrana gelecektir. Lütfen sorgulama sonucunu deęerlendirirken bu özellikleri dikkate alınız.

Arařtırmacının uzmanlařtığı bilimsel ve teknolojik faaliyet alanı
 Alt alanları ile birlikte

Arařtırmacının AR-GE çalışmalarının yöneldiđi iktisadi faaliyet alanı
 Alt alanları ile birlikte

Arařtırmacının AR-GE faaliyet alanını tanımlayan anahtar kelimeler
(Aranan kelimeleri yazarken aralarına bir boşluk bırakınız)

Arařtırmacının çalıştığı işyerine ait bilgiler

İşyeri Türü

Arařtırmacının İşyerindeki Görevi

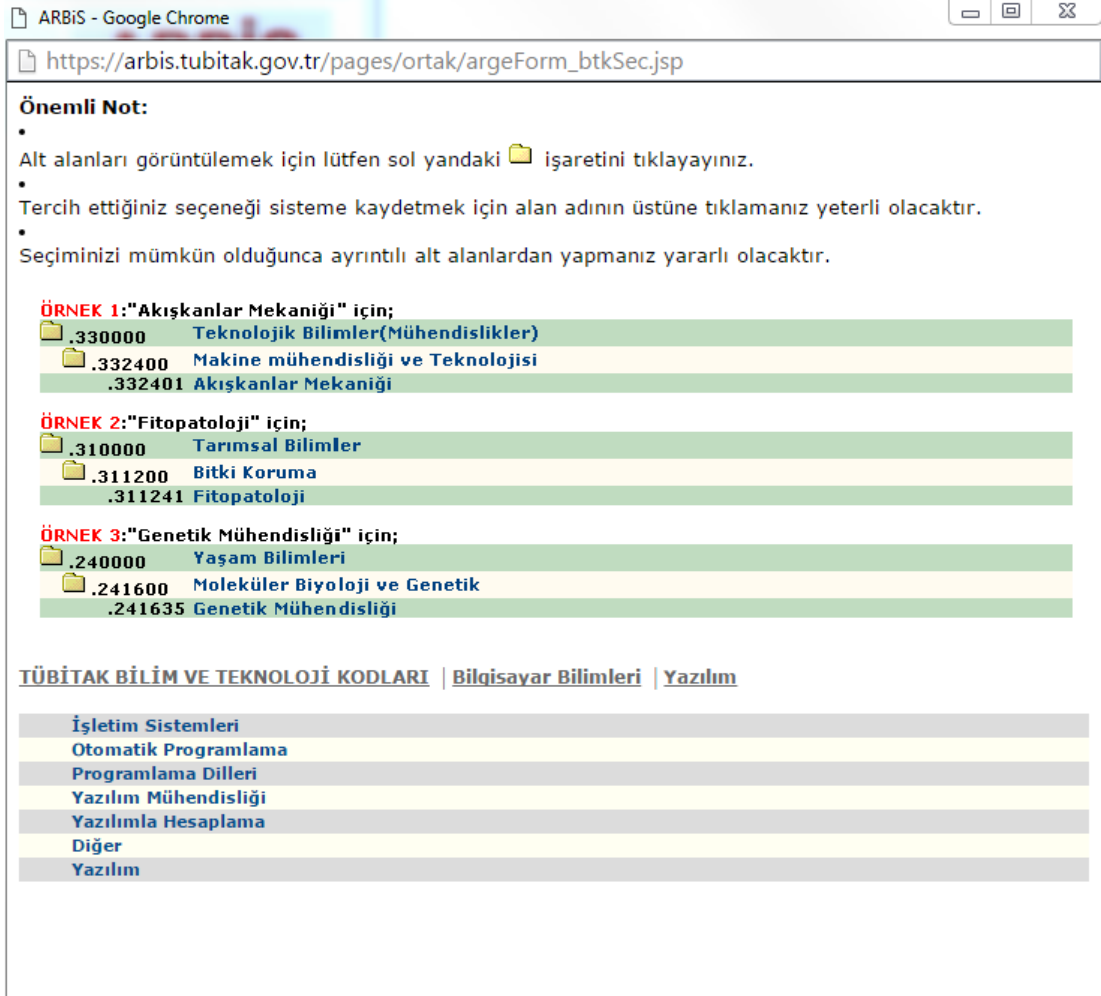
İşyerinin Bulunduđu Ülke

İşyerinin Bulunduđu Şehir

İşyeri Faaliyet Alanı
 Alt alanları ile birlikte

Şekil 2.ARBİS Sorgulama Ekranı.

Teknik Bilimler menüsü altında sırayla Bilgisayar Bilimleri > Yazılım > Yazılım Mühendisliđi faaliyet alanına Şekil 3’de gösterildiđi şekilde ulaşarak sorgulama yapıldığında Şekil 4’de görüldüğü gibi sonuçlar elde edilebilmektedir.



Şekil 3.ARBİS Sorgulama Ekranı.

Sorgulama yapıldıktan sonra Şekil 4’te sonuçlar gösterilmektedir. Bu sonuçlar, ilgili faaliyet alanına kendisini dahil etmiş bir araştırmacının, Ad, Soyad, E-ileti ve telefon bilgisini vermektedir. İletişim bilgileri ile araştırmacıların bağlantı kurmaları amaçlanmıştır.

Arama Sonucu: 3489 adet belirlenen kriterlere uygun kayıt bulundu...

<< Arama sayfasına Geri Dön

	Adı	Soyadı	E-ileti	Telefon
1.	ABDİ BORA	AYDEMİR	bora.aydemir@gmail.com	--
2.	ABDİL BURAK	BAŞLI	aburakbasli@gmail.com	--
3.	ABDİL KADİR	AVCI	kadir.avci@dokutek.com.tr	0216 456 2140
4.	ABDULKADİR	GENÇ	abdulkadir.genc@bizitek.com	212-2869966
5.	ABDULKADİR	YILMAZ	abdulkadiryilmaz@gmail.com	312-5087243
6.	ABDULKADİR	KAYIKLI	abdulkadir.kayikli@gmail.com	02122762772
7.	ABDULKADİR	ÖLMEZ	abdulkadir.olmez@eb.com.tr	0312 463 88 88
8.	ABDULKADİR CAN	ÇAKMAK	ccakmak@tai.com.tr	312-81118007653
9.	ABDULKADİR SELÇUK	GÜCEĞLİOĞLU	aselcuk@yahoo.com	--
10.	ABDULKERİM	ÖNCÜ	kerim.oncu@tubitak.gov.tr	--
11.	ABDULLAH	ÇAVUŞOĞLU	abdullah.cavusoglu@gmail.com	312 3241502
12.	ABDULLAH	TURHAN	info@abdullahturhan.com	--
13.	ABDULLAH	ATMACA	abdullah.atmaca@tcmb.gov.tr	--
14.	ABDULLAH	ÖNDEN	abdullah.onden@fabrikod.com	--
15.	ABDULLAH	KUZHAN	abdullahkuzhan@hotmail.com	216-5561755/1755
16.	ABDULLAH	APAYDIN	abdullah.apaydin@edata.com.tr	212-274 28 72
17.	ABDULLAH	SÖNMEZ	asonmez@amvg.com	212-2862910
18.	ABDULLAH	ÇİTİL	a.citil@gmail.com	--
19.	ABDULLAH	ASLAN	a.aslana@gmail.com	--
20.	ABDULLAH	SÖNMEZ	abdullah.sonmez@gmail.com	312 5087243

Copyright 2007-2010 Türkiye Bilimsel ve Teknolojik Araştırma Kurumu. Her hakkı saklıdır.

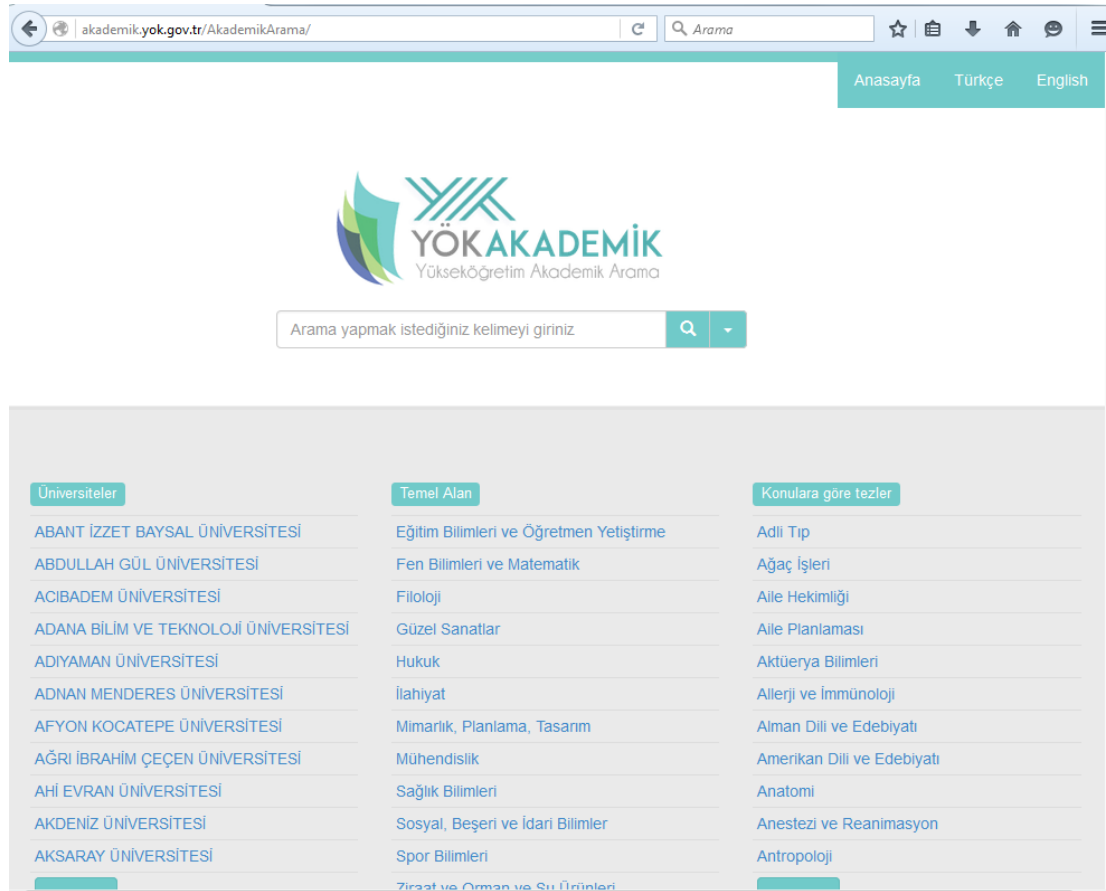
Şekil 4.ARBİS Sorgulama Sonucu

2.3 ARBİS Portalının Eksiklikleri

Şekil 4’de bulunan sorgulama sonucunda 3489 adet kayıt bulunmaktadır (25 Şubat 2016). Araştırmacının kendisine benzer diğer araştırmacıyı bu sonuçlardan tespit etmesi zordur. Bu sonuçlardan benzerliğin anlaşılabilmesi için tek tek araştırmacılar ile iletişime geçilmesi gerekmektedir. En basit şu iki soruya yanıt alınamamaktadır. Araştırmacı ile hangi alanlarda çalışılmış? Aslında Yazılım Mühendisliği altında gösterilen araştırmacı Jeoloji Mühendisi ama hobi olarak Yazılım Mühendisliği faaliyet alanı ile mi ilgileniyor? ARBİS, veritabanı olarak bir alanda büyük veri verebilmektedir. Ancak bu verilerden, sosyal ağ şeklinde bilgi almak mümkün değildir.

2.4 YÖK Akademik Arama Uygulaması Yapısı

Yüksek Öğretim Kurumu tarafından hazırlanan bu uygulamaya da, kurumun kendi veri tabanında kayıtlı akademisyenlerin tüm detaylı bilgileri yayınlanır. Akademisyen Adı, Üniversite adı, Temel Alan ve Konulara Göre Tezler şeklinde arama yapılabilmektedir.



Şekil 5. YÖK Akademik Ana Sayfa

Yapılan arama sonucunda, ilgili akademisyenin daha önceden bir çalışmada ortaklıkları bulunan akademisyenler ile olan bağlantısı “Birlikte Çalıştığı Kişiler” linki kullanılarak gösterilir.

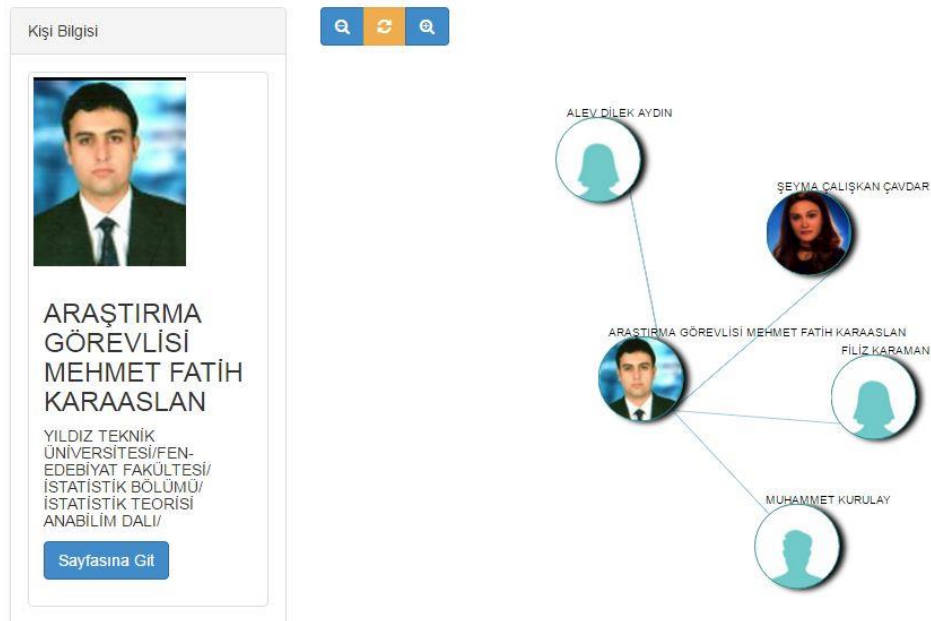
Bu sistemde bağlantıyı oluşturan unsur beraber yayınlanmış makale ve bildirimlerdir. Veri madenciliği çalışması olmadan tanımlı bilgiler üzerinden bağlantılar

gösterilmektedir. Şekil 6 ve 7’de görüldüğü üzere Mehmet Fatih Karaaslanın, yayınladığı makalede ismi geçen kişiler çizge grafik üzerinde gösterilmektedir.



The screenshot shows a profile page for Mehmet Fatih Karaaslan, an Assistant Professor at Yıldız Teknik Üniversitesi. The page includes a search bar, a list of publications, and a network graph. The first publication listed is "A Geometric Interpretation and Comparison of the Methods Of Ordinary Least Square (OLS) and Bivariate Lagrange Interpolation" by Şeyma Çalışkan Çavdar, Alev Dilek Aydın, Mehmet Fatih Karaaslan, published in 2015. The network graph shows connections between Mehmet Fatih Karaaslan and four other individuals: Alev Dilek Aydın, Şeyma Çalışkan Çavdar, Filiz Karaman, and Muhammet Kurulay.

Şekil 6.YÖK Akademik Arama Sonucu



The screenshot displays a network graph for Mehmet Fatih Karaaslan. On the left, there is a detailed profile card for Mehmet Fatih Karaaslan, including his name, title (ARAŞTIRMA GÖREVLİSİ), and affiliation (YILDIZ TEKNİK ÜNİVERSİTESİ/FEN-EDEBİYAT FAKÜLTESİ/İSTATİSTİK BÖLÜMÜ/İSTATİSTİK TEORİSİ ANABİLİM DALI). To the right, a network graph shows Mehmet Fatih Karaaslan at the center, connected to four other individuals: Alev Dilek Aydın, Şeyma Çalışkan Çavdar, Filiz Karaman, and Muhammet Kurulay. Each individual is represented by a circular profile picture.

Şekil 7.YÖK Akademik Çizge Grafik Çizimi

2.5 Önerilen Modelin Avantajları

Bu tez çalışması kapsamında gerçekleştirilen uygulamada, veri madenciliği teknikleri kullanılarak, doğrudan bağlantısı olmayan akademisyenlerin benzerlikleri tespit edilmektedir. Tespit edilen bu benzerlik bilgisi kullanılarak çizge grafik çizilmektedir.

ARBİS'den alınmış olan yaklaşık 40.000 arařtırmacının, 900 faaliyet alanındaki verileri en yakın komřu yöntemi ile anlam kazanarak, sorgulama yapan arařtırmacı ya büyük veri yerine çizge grafik řeklinde bilgi vermektedir.

Arařtırmacı; kendisinin veya hakkında bilgi sahibi olmak istediđi bir arařtırmacının ad ve soyadını sisteme girerek arama yaptırır. Arama sonucu olarak, aranan kiřiye en yakın 30 arařtırmacının sonucu ekrana getirilir. Benzerliđi en çok olan merkeze en yakın olacak řekilde çizge grafik çizimi yapılır.

Bu sayede sorgulama yapan kiři, karřısında satırlar dolusu bilgi yerine görsel bir sonuç elde eder. Arařtırmacının sosyal ađı çizge grafikde gösterilir.

3. LİTERATÜR İNCELEMESİ

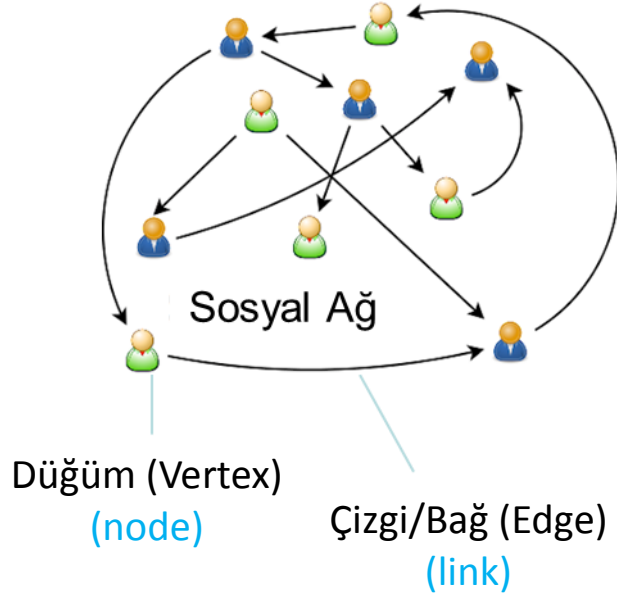
3.1 Sosyal Ađ Kavramı

Sosyal ađlar insanlıđın ilk var oluřundan itibaren var olmuř bir kavramdır. İnsanlar arasındaki resmi, gayri-resmi, politik, ailevi, bölgesel ya da herhangi başka bir řekildeki iliřkiler sosyal ađların oluřmasını sađlar.

Sosyal ađ tanımı, insanların birbirine bađlı oldukları bađlantının tanımıdır. Sosyal ađ analizi sosyoloji, antropoloji, sosyal psikoloji, iletiřim, ekonomi, matematik gibi birçok alanda sıklıkla kullanılır.

Sosyal iliři kurma amacıyla kurulmuř internet sitelerinde bireylerin diđerleriyle kurdukları bađlantıların bilimsel dilde anlařılabilmesi için Sosyal Ađ Analizleri yoğun bir řekilde kullanılmaktadır.

Sosyal ađların gösteriminde; düđümler (nodes) kiřileri ve/veya organizasyonları, çizgiler (links/edges) ise düđümler arasındaki iliřkileri göstermektedir. İliřkiler tek taraflı ve/veya karřılıklı zayıf veya güçlü olabilmektedir (řekil 5).



Şekil 8.Sosyal Ağ Gösterimi.

3.2 Sosyal Ağ Arama Motorları

Yine bu amaçla, bu sitelerde (mesela Facebook, MySpace, Linked-in vs.) sosyal ağ grupları kurulmakta ve bilginin bireyler arasında ne yönde taşındığı konusunda araştırmalar ortaya konmaktadır.

Akademik alandaki insanlar için Microsoft firmasının oluşturmuş olduğu <http://academic.research.microsoft.com/default.aspx> internet sayfası uygulaması mevcuttur. (Şekil 1). Bu uygulamaya bilgilerini giren araştırmacılar yine birbirleri ile bağlantılarını tespit ederek iletişime geçebilmektedirler.

3.3 Sosyal Ağ Analizi Yaklaşımları

Sosyal ağ analizi, elde bulunan veriye bazı yaklaşım yöntemlerini kullanarak veriden anlamlı bilgi üretilmesini sağlar.

3.2.1 Bağlantı Tabanlı ve Yapısal Analizler

Sosyal ağda var olan bağlantılar kullanılarak bir topluluk veya alandaki ilgili düğümlerin, bağlantıların, alt çizgelerin veya ilgi alanlarının çıkarılması hedeflenir.

3.2.2 Hareketli ve Sabit Analiz

Ağdaki hareketliliğe bağlı olarak kullanılan sabit (static) analiz veya hareketli (dynamic) analiz yöntemi seçilebilir. Örneğin kitapların atıflarının takip edildiği bir ağda sabit analiz yapılabilirken, verinin hızla aktığı bir ağda hareketli analiz yöntemlerine ihtiyaç duyulmaktadır. Temel olarak sabit analiz yöntemi, bir sosyal ağın değişmemesi veya değişimin sabit boyutlarda olması durumunda kullanılırken, dinamik analiz yöntemi Facebook veya Youtube gibi değişimin ön görülemeyeceği veya üstel olarak değiştiği durumlarda kullanılabilir (Khaled Al-Naami, 2014)

3.4 Sosyal Ağ Analizi Sınıflama

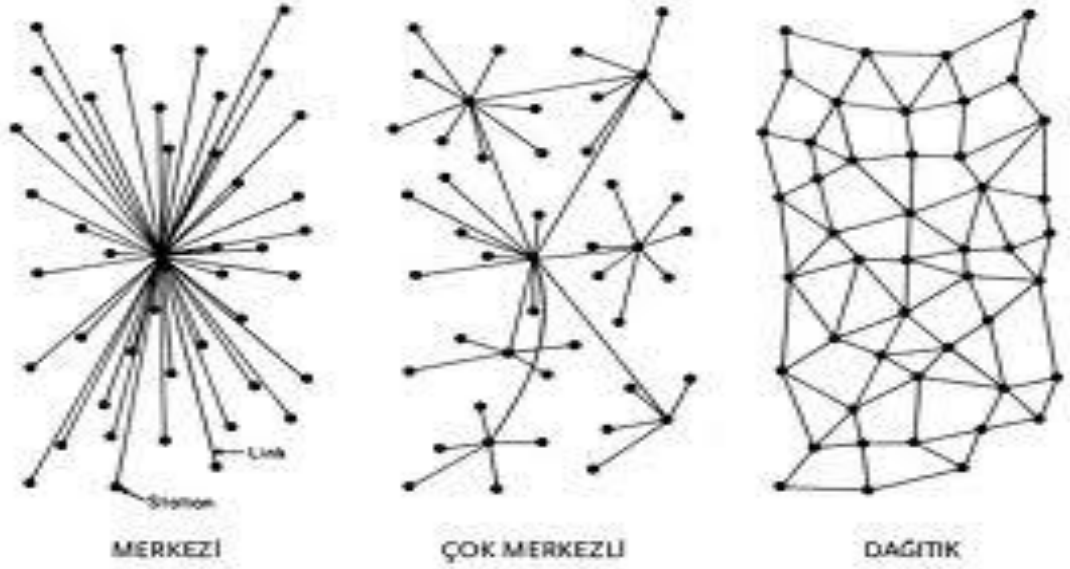
Ağların topolojisini incelediğimizde, sosyal ağ analizinden çıkan bulgular, üç ayrı sınıfta ele alınır. Bunlar; merkezi, çok merkezli ve dağıtık ağ topolojileridir (Şekil 9).

Merkezi Ağlar (Yıldız Topoloji): Merkezi ağ yapısında tüm düğümler tek bir düğüme bağlıdır. Dallar arasında bağlantı yoktur. Hiyerarşiktir. Tek bir otorite vardır. Merkezi ağlara örnek olarak Turkcell'in gözetleme merkezi verilebilir.

Çok Merkezli Ağlar (Ağaç Topoloji): Birçok merkezi ağın merkezlerinden birbiriyle bağlanmış halidir. Çok merkezli ağa örnek olarak hava yolu sistemi verilebilir. Hava yolu sisteminde aktarma yapılan her bir nokta ayrı bir ağın merkezini oluşturur.

Dağıtık Ağlar (Mesh Topoloji): Bir dağıtık ağın merkezi yoktur. Her düğüm bağımsızdır. Bir düğümden diğer düğüme pek çok yoldan gidilebilir. Bu ağ sistemine örnek olarak da kara yolu ve torrent paylaşım ağı verilebilir.

Bu topolojileri dışında, topolojiler yoğun (dense), seyrek (sparse), merkez-çevre (core-periphery), yüksek bağlantılı ve küçük dünya (small world) şeklinde sıralanabilir [24].



Şekil 9. Sosyal Ağ Sınıflaması [24]

4. ARAŞTIRMACI AĞI ARAMA MOTORU UYGULAMASI

4.1 Veri Toplama ve Ön İşleme

ARBİS sistemi üzerinde, bir faaliyet alanındaki araştırmacıya yada araştırmacılara ulaşabilmek için, ilgili faaliyet alanının sorgulanması yolu ile o faaliyet alanına, tanımlı araştırmacıya ya da araştırmacıların bilgilerine ulaşılabilir (Şekil 2-3-4).

Tez çalışması kapsamında, Teknik Bilimler altında yapılmıştır. Tüm faaliyet alanlarının ağaç yapısı çıkartılmıştır. 1095 Adet alan tespit edildi ve hangi faaliyet alanının altında hangi alt alanın olduğu bir excel tablosuna işlendi (Tablo 1). Excel

tablosunda B kolonunda bulunan ID sütununda her bir alana ayrı bir numara verilmiştir.

ARBİS’de alt faaliyet alanları veya buna bağlı üst faaliyet alanı ayrı ayrı sorgulanabilir. Örneğin Teknik Bilimler altında Bilgisayar Mühendisliği sorgulanabildiği gibi Bilgisayar Mühendisliği altında Yazılım Mühendisliği sorgulanabilir. Bu ayırım excel dosyasında G kolonunda bulunan Evet/Hayır ayırımı ile uygulanmıştır.

Verilerin toplanmasında, 1600x900 çözünürlüğü olan bir bilgisayar kullanılmıştır. Linklerin X ve Y koordinat hesaplamaları yapılarak her bir faaliyet alanı tıklanmıştır. Her bir faaliyet alanının otomatik sorgulanmasıyla dönen sonuçlarının, Tablo 1’de B kolonunda bulunan ID kullanılarak <<id>>.html dosyası olarak kaydedilmesi sağlanmıştır. Daha sonrasında bu sorgulama sonuçları hazırlanan başka bir uygulama ile ayrıştırılarak veri tabanına kaydedilmiştir.

A	B	C	D	E	F	G	H	I
ID	Kat No	Ust ID		Main Menu	Alt Menu	Alt Kırılım	Durum	Sno
1	5	0	0		0 Teknik Bilimler	Evet		
1	6	1	5	Teknik Bilimler	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Evet	Ok	3
1	7	1	5	Teknik Bilimler	Çevre Mühendisliği	Evet	Ok	4
1	8	1	5	Teknik Bilimler	Deniz Bilimleri ve Teknolojisi	Evet	Ok	5
1	9	1	5	Teknik Bilimler	Harita Mühendisliği-Geomatik	Evet	Ok	6
1	10	1	5	Teknik Bilimler	Jeofizik Mühendisliği	Evet	Ok	7
1	11	1	5	Teknik Bilimler	Jeoloji Mühendisliği	Evet	Ok	8
1	12	1	5	Teknik Bilimler	Bilgi Sistemleri, Haberleşme ve Kontrol Mühendisliği	Evet	Ok	9
1	13	1	5	Teknik Bilimler	Bilgisayar Bilimleri	Evet	Ok	10
1	14	1	5	Teknik Bilimler	Biyomedikal Mühendisliği	Evet	Ok	11
1	15	1	5	Teknik Bilimler	Elektrik-Elektronik Mühendisliği	Evet	Ok	12
1	16	1	5	Teknik Bilimler	Kimya Mühendisliği ve Teknolojisi	Evet	Ok	13
1	17	1	5	Teknik Bilimler	İnşaat Mühendisliği	Evet	Ok	14
1	18	1	5	Teknik Bilimler	Endüstri Mühendisliği	Evet	Ok	15
1	19	1	5	Teknik Bilimler	Makina Mühendisliği	Evet	Ok	16
1	20	1	5	Teknik Bilimler	Maden Mühendisliği ve Teknolojisi	Evet	Ok	17
1	21	1	5	Teknik Bilimler	Petrol ve Doğal Gaz Mühendisliği	Evet	Ok	18
1	22	1	5	Teknik Bilimler	Tekstil Mühendisliği ve Teknolojisi	Evet	Ok	19
1	23	1	5	Teknik Bilimler	Mimarlık	Evet	Ok	20
1	24	1	5	Teknik Bilimler	Gemi İnşaatı ve Gemi Makinaları Mühendisliği	Evet	Ok	21
1	25	1	5	Teknik Bilimler	Havacılık ve Uzay Mühendisliği	Evet	Ok	22
1	26	1	5	Teknik Bilimler	Metaller ve Malzeme Mühendisliği	Evet	Ok	23
1	27	1	5	Teknik Bilimler	Gıda Mühendisliği	Evet	Ok	24
1	28	1	5	Teknik Bilimler	Teknik Bilimler	Hayır	No Need	25
1	29	2	6	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Atmosfer Fiziği	Evet	Ok	2
1	30	2	6	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Atmosfer-Okyanus Etkileşimi	Hayır	No Need	3
1	31	2	6	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Hidroloji-Hidrometeoroloji	Hayır	No Need	4
1	32	2	6	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	İklim	Evet	Ok	5
1	33	2	6	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Meteorolojik Afetler	Evet	Ok	6
1	34	2	6	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Diğer	Hayır	No Need	7
1	35	2	6	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Atmosfer Bilimleri ve Meteoroloji Mühendisliği	Hayır	No Need	8
1	36	3	29	Atmosfer Fiziği	Atmosfer Dinamiği ve Termodinamiği	Hayır	No Need	2
1	37	3	29	Atmosfer Fiziği	Atmosferik Elektrite	Hayır	No Need	3

Tablo 1.Excel Tablosu Ekran Görüntüsü.

4.1.1 ARBİS’de Teknik Bilimler Faaliyet Alanının Alt Katogorilerinin İncelenmesi

ARBİS’den verilerin alınabilmesi için excel dosyasına çıkartılan menü ağacının, sistemde sanki bir insan tarafından tıklama yapılmış gibi çalışabilmesi için tüm faaliyet alanları incelenip buna göre bir algoritma oluşturuldu.

Bu algoritma, faaliyet alanlarının uzunluklarına göre satırda kayma olması durumunu da hesaba katarak tıklanacak linkin Y ekseninin hesaplamasını yapar. Linkin X ve Y koordinatlarını hesaplar. Aşağıda Tablo 2’de görüldüğü üzere X koordinatı sabit tutularak Y ekseninde tıklamanın seviyesine göre link için seçilmesi gereken adres hesaplanmıştır.

	A	B	C	D	E	F	G	H	I
15	1	2	3	4	5	6	7	8	
16		85	Tıksız	1 Tık	2 Tık	3 Tık	4 Tık	5 Tık	
17	Artış Miktarı	X1	Y1	Y2	Y3	Y4	Y5	Y6	
18		90	512	440	375	314	202	202	1. Kayıt
19	16	90	528	456	391	330	218	218	2. Kayıt
20	16	90	544	472	407	346	234	234	3. Kayıt
21	16	90	560	488	423	362	250	250	4. Kayıt
22	16	90	576	504	439	378	266	266	5. Kayıt
23	16	90	592	520	455	394	282	282	6. Kayıt
24	16	90	608	536	471	410	298	298	7. Kayıt
25	16	90	624	552	487	426	314	314	8. Kayıt
26	16	90	640	568	503	442	330	330	9. Kayıt
27	16	90	656	584	519	458	346	346	10. Kayıt
28	16	90	672	600	535	474	362	362	11. Kayıt
29	16	90	688	616	551	490	378	378	12. Kayıt
30	16	90	704	632	567	506	394	394	13. Kayıt
31	16	90	720	648	583	522	410	410	14. Kayıt
32	16	90	736	664	599	538	426	426	15. Kayıt
33	16	90	752	680	615	554	442	442	16. Kayıt
34	16	90	768	696	631	570	458	458	17. Kayıt
35	16	90	784	712	647	586	474	474	18. Kayıt
36	16	90	800	728	663	602	490	490	19. Kayıt
37	16	90	816	744	679	618	506	506	20. Kayıt
38	16	90	832	760	695	634	522	522	21. Kayıt
39	16	90	848	776	711	650	538	538	22. Kayıt
40	16	90	864	792	727	666	554	554	23. Kayıt
41	16	90	880	808	743	682	570	570	24. Kayıt
42	16	90	896	824	759	698	586	586	25. Kayıt
43	16	90	912	840	775	714	602	602	26. Kayıt
44	16	90	928	856	791	730	618	618	27. Kayıt
45	16	90	944	872	807	746	634	634	28. Kayıt
46	16	90	960	888	823	762	650	650	29. Kayıt
47	16	90	976	904	839	778	666	666	30. Kayıt
48	16	90	992	920	855	794	682	682	31. Kayıt
49	16	90	1008	936	871	810	698	698	32. Kayıt

Tablo 2.X ve Y Koordinatı Hesaplaması

ARBİS uygulamasının yapısında sabit olarak bulunan “Liste”, “Ara” veya açılır penceredeki sağ kaydırma çubuğunun tıklama adresleri Şekil 10’da hesaplanmıştır. Örneğin, Şekil 2’de görünen “Liste” düğmesinin koordinatı her zaman 1125,370 (x,y) olduğu gibi.

X	Y	
1125	370	Uzmanlaşılın Faaliyet Alanı "Liste" tıkla
90	560	Teknik Bilimler Tıkla
750	715	Sağ ScroolBarı Aşağı kaydırma
357	699	Ara Click
17	70	Dosya Menuşu Tıkla
55	247	Farklı Kaydet Tıkla
		Faaliyet alanının ID si ile dosya adı girilecek ve ENTER yapılacak
771	320	Arama Sayfasına geri dön
1295	501	Ara Demeden önce bir kere clicklenecek (420)

Şekil 10.Ana Koordinat Hesaplaması

4.1.2 Otomatik Sorgulama İçin Excel Makro ve Data Load Programının Kullanımı

Faaliyet alanlarının ARBİS portalından tek tek sorgulanabilmesi için portal uygulamasında linklerin yerlerinin koordinat hesaplamaları ve tabloları oluşturulduktan sonra, otomatik tıklama işleminin ARBİS'e uygulanabilmesi için Data Load programı kullanılmıştır. [25]

Data Load programı, Oracle, SAP gibi çeşitli ERP uygulamalarına otomatik veri girilmesi için hazırlanmış bir uygulama olmasına rağmen farklı amaçlar için de kullanılabilir. Bir uygulamada, kordinatları belirlenmiş bir yerin tıklanması, bu alana veri girişi, Enter tuşunun işlevi gibi bir çok işlem yapılabilir (Şekil 11). Genellikle bir uygulamaya otomatik veri girişi yapmak amacı ile kullanılır ve Classic versiyonu ücretsiz temin edilebilir.

Command	EBS - R12	EBS - 11i	EBS - 10.7/	EBS - 10.7S Siebel	SAP	Peoplesoft	JDE	Dynamics	Other	Description
TAB	\{TAB}	\{TAB}	\{TAB}	\{TAB}	\{TAB}	\{TAB}	\{TAB}	\{TAB}	\{TAB}	TAB key
ENT	\{ENTER}	\{ENTER}	\{ENTER}	\{ENTER}	\{ENTER}	\{ENTER}	\{ENTER}	\{ENTER}	\{ENTER}	Enter key
*UP	\{UP}	\{UP}	\{UP}	\{UP}	\{UP}	\{UP}	\{UP}	\{UP}	\{UP}	Up arrow key
*DN	\{DOWN}	\{DOWN}	\{DOWN}	\{DOWN}	\{DOWN}	\{DOWN}	\{DOWN}	\{DOWN}	\{DOWN}	Down arrow key
*LT	\{LEFT}	\{LEFT}	\{LEFT}	\{LEFT}	\{LEFT}	\{LEFT}	\{LEFT}	\{LEFT}	\{LEFT}	Left arrow key
*RT	\{RIGHT}	\{RIGHT}	\{RIGHT}	\{RIGHT}	\{RIGHT}	\{RIGHT}	\{RIGHT}	\{RIGHT}	\{RIGHT}	Right arrow key
*SP	\%F%V	\%F%V	\%A%\{DOW	\%A%A						Save & process
*SAVE	\^S	\^S	\^S	\{F10}						Save
*NB	\+{PGDN}	\+{PGDN}	\+{PGDN}	\^PGDN}						Next Block
*PB	\+{PGUP}	\+{PGUP}	\+{PGUP}	\^PGUP}						Previous Block
*NF	\{TAB}	\{TAB}	\%G%\{DOW	\{TAB}						Next Field
*PF	\+{TAB}	\+{TAB}	\%G%\{DOW	\+{TAB}						Previous Field
*NR	\{DOWN}	\{DOWN}	\%G%\{DOW	\+{DOWN}						Next record
*PR	\{UP}	\{UP}	\%G%\{DOW	\+{UP}						Previous record
*FR	\%V%\{DF	\%V%\{DF	\%G%\{DOW	\%G%\{F						First record
*LR	\%V%\{DL	\%V%\{DL	\%G%\{DOW	\%G%\{L						Last record
*FR	\{F6}	\{F6}	\{F6}	\+{F4}						Refresh record

Şekil 11.Data Load Programı Örnek Komutları

Data Load programının bu işlevselliği, ARBİS sisteminin otomatik sorgulamasında kullanılmıştır. Bu amaçla, link koordinatlarının otomatik tıklanması için aşağıda bulunan Tablo'3 deki komutlar hazırlanmıştır.

*SL(5)	*ML(1125,370)	*SL(5)	*ML(90,560)							
*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)					
*SL(3)	*ML(90,346)									
*SL(3)	*ML(90,528)									
*SL(3)	*ML(90,528)									
*SL(5)	*ML(357,699)	*SL(5)	*ML(17,70)	*SL(1)	*ML(55,247)	*SL(1)	36	*SL(1)	ENT	*SL(5) *ML(771,320)
*SL(5)	*ML(1125,370)	*SL(5)	*ML(90,560)							
*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)					
*SL(3)	*ML(90,346)									
*SL(3)	*ML(90,528)									
*SL(3)	*ML(90,544)									
*SL(5)	*ML(357,699)	*SL(5)	*ML(17,70)	*SL(1)	*ML(55,247)	*SL(1)	37	*SL(1)	ENT	*SL(5) *ML(771,320)
*SL(5)	*ML(1125,370)	*SL(5)	*ML(90,560)							
*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)					
*SL(3)	*ML(90,346)									
*SL(3)	*ML(90,528)									
*SL(3)	*ML(90,560)									
*SL(5)	*ML(357,699)	*SL(5)	*ML(17,70)	*SL(1)	*ML(55,247)	*SL(1)	38	*SL(1)	ENT	*SL(5) *ML(771,320)
*SL(5)	*ML(1125,370)	*SL(5)	*ML(90,560)							
*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)	*SL(1)	*ML(750,715)					
*SL(3)	*ML(90,346)									
*SL(3)	*ML(90,528)									
*SL(3)	*ML(90,576)									
*SL(5)	*ML(357,699)	*SL(5)	*ML(17,70)	*SL(1)	*ML(55,247)	*SL(1)	39	*SL(1)	ENT	*SL(5) *ML(771,320)

Tablo 3.Data Load komutları tablosu.

Excel programının içerisinde bulunan, Makro yöntemi kullanılarak hazırlanan Visual Basic program kodları ile Tablo 3’de görünen 6100 satırlık Data Load komutları hazırlanmıştır. Burada *SL(n) komutu içerisinde yazan n kadar uygulama bekleme yapar. *ML(x,y) komutu içerisinde bulunan x ve y koordinatı tıklanır. ENT komutu klavye de bulunan enter tuşunun işlemini yapar. Hücre de yazan tek değer ise o an aktif olan alana girilecek olan değerdir.

Data Load programı oldukça basit olduğu gibi komutlarının dizilimi oldukça önemlidir. İlgili uygulamada 4 farklı komut kullanılarak 6100 satırlık komut bloğu oluşturulmuştur. Verilerin, ARBİS sisteminden her bir faaliyet alanının sorgulanarak alınması sağlanmıştır.

The screenshot shows the DataLoad Classic application window. The title bar reads 'DataLoad Classic C:\Users\yyener\Documents\EskiDocuments\Tez\DataLoad.dld'. The menu bar includes 'File', 'Edit', 'Tools', and 'Help'. The toolbar contains various icons for file operations and execution. The 'Window' menu is open, showing 'SORGULAMA SERVİSİ - ARBİS - Windows Internet Explo' and 'Command Group' with a dropdown arrow. The 'Notes' area contains 'R1 *SL(1)' and 'C1'. The 'Data' tab is active, displaying a grid with columns for 'Title' and various data points. The grid contains 17 rows of commands, including *SL(1), *ML(1269,673), *UP, *DN, *ML(940,325), *SL(5), *ML(198,530), *ML(904,680), *ML(198,500), *ML(243,500), *ML(290,655), *ML(17,70), *ML(59,208), *SL(1), 34, *ML(635,320), *ML(940,325), *ML(198,530), *ML(904,680), *ML(904,680), *ML(198,330), *ML(198,500), *ML(243,516), *ML(290,655), *ML(17,70), *ML(59,208), *SL(1), 35, *ML(635,320), *UP, *DN, *ML(940,325), *ML(198,530), *ML(904,680), *ML(904,680), *ML(198,330), *ML(198,500), *ML(243,532), *ML(290,655), *ML(17,70), *ML(59,208), *SL(1), 36, *ML(635,320), *UP, *DN, *ML(940,325), *ML(198,530).

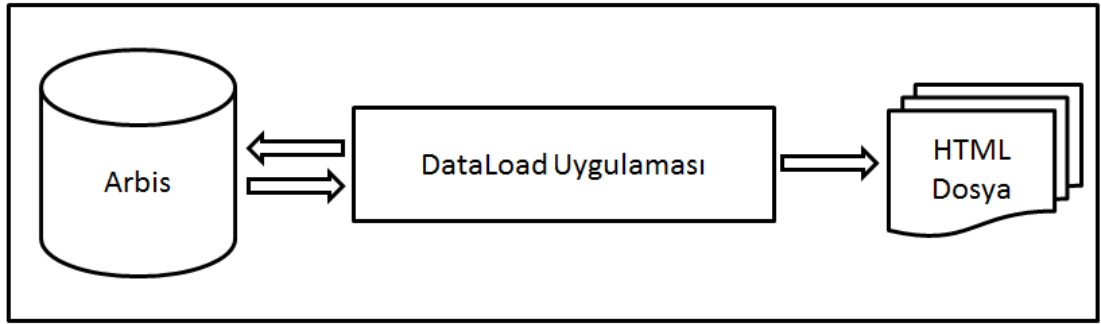
Title	*SL(1)	*ML(1269,673)	Data	Data	Data	Data	Data	Data	Data	Data	Data	Data
1	*SL(1)	*ML(1269,673)										
2	*SL(5)	*UP	*DN	*ML(940,325)	*SL(5)	*ML(198,530)						
3	*SL(5)	*ML(904,680)	*ML(904,680)	*ML(904,680)	*SL(1)	*ML(198,330)						
4	*SL(5)	*ML(198,500)										
5	*SL(5)	*ML(243,500)										
6	*SL(5)	*ML(290,655)	*SL(5)	*ML(17,70)	*SL(1)	*ML(59,208)	*SL(1)	34	*SL(1)	ENT	*SL(5)	*ML(635,320)
7	*SL(5)	*UP	*DN	*ML(940,325)	*SL(5)	*ML(198,530)						
8	*SL(5)	*ML(904,680)	*ML(904,680)	*ML(904,680)	*SL(1)	*ML(198,330)						
9	*SL(5)	*ML(198,500)										
10	*SL(5)	*ML(243,516)										
11	*SL(5)	*ML(290,655)	*SL(5)	*ML(17,70)	*SL(1)	*ML(59,208)	*SL(1)	35	*SL(1)	ENT	*SL(5)	*ML(635,320)
12	*SL(5)	*UP	*DN	*ML(940,325)	*SL(5)	*ML(198,530)						
13	*SL(5)	*ML(904,680)	*ML(904,680)	*ML(904,680)	*SL(1)	*ML(198,330)						
14	*SL(5)	*ML(198,500)										
15	*SL(5)	*ML(243,532)										
16	*SL(5)	*ML(290,655)	*SL(5)	*ML(17,70)	*SL(1)	*ML(59,208)	*SL(1)	36	*SL(1)	ENT	*SL(5)	*ML(635,320)
17	*SL(5)	*UP	*DN	*ML(940,325)	*SL(5)	*ML(198,530)						

Şekil 12.Data Load Programı Ekran Görüntüsü.

4.1.3 Teknik Bilimler Altındaki Kategorilerin Sorgulanması ve Sonuçların html Olarak Kaydedilmesi

Teknik Bilimler altında bulunan 1095 adet faaliyet alanı, oluşturulan Data Load programı komutları ile sanki bir insan tarafından sorgulama yapılmış gibi hareket ederek sorgulanmıştır. Sorgulama sonuçları, <<id>>.mht olarak sorgulama yapan bilgisayara kayıt edilmiştir. Daha sonra bu dosyalar üzerinde ayrıştırma işlemi yapılarak bilgiler alınmıştır.

Bu işlem için tahsis edilen bilgisayar yaklaşık 1 hafta boyunca aralıklı olarak çalışarak, tüm alanların sorgulamasını sağlayıp araştırmacı bilgilerinin ARBİS sisteminden alınmasını sağlamıştır.



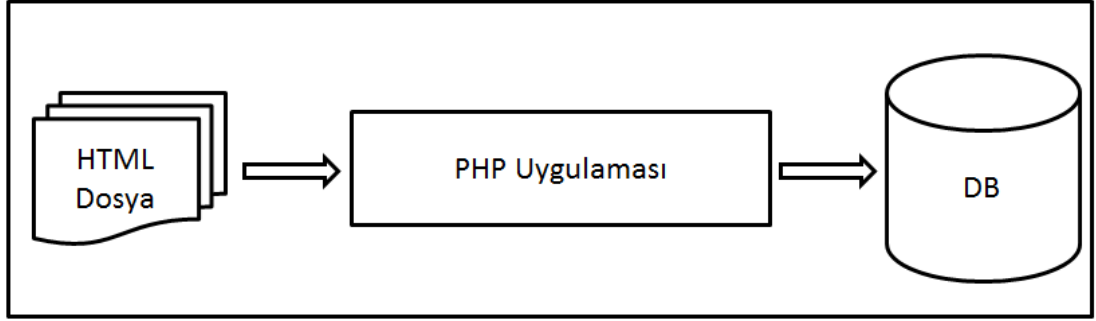
Şekil 13.ARBİS den bilgilerin alınması.

4.1.4 HTML Sonuçların Ayrıştırılarak Veri Tabanına Kaydı

Sorgulamalar sonucunda <<id>>.mht olarak her bir faaliyet alanı için kayıt edilen dosyalar, PHP programlama dilinde hazırlanan uygulama ile tek tek ayrıştırılarak veri tabanına kayıt edilmiştir.

PHP uygulamasında, dosya okuma fonksiyonu ile dosyalar açılmış ve içerisinden araştırmacıların bilgileri ayrıştırılmıştır. Araştırmacıların; ad, soyad ve email bilgileri alınarak veritabanında ilgili tabloya kayıt edilmiştir (Şekil 15). Her bir

arařtırmacının hangi faaliyet alanında bulunduđunun bilgisi, iliřki tablosuna iřlenmiřtir.



řekil 14.HTML Dosyaların Ayırırırılması ve DB'ye aktarılması.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	kategorilerdb				arastirmacldb				iliskidb					
2	id	MEDIUMINT	AI	PK	id	INT	AI	PK	id	BIGINT	AI	PK		
3	katno	MEDIUMINT			ad	VARCHAR	100		katid	MEDIUMINT				
4	ustid	MEDIUMINT	50		soyad	VARCHAR	100		arastirmaciid	INT				
5	mainmenuad	VARCHAR	200		email	VARCHAR	200							
6	menuad	VARCHAR	200											
7	altmenu	VARCHAR	5	Evet, Hayır										
8	sno	MEDIUMINT												

řekil 15.Data Load Programı Ekran Görüntüsü.

4.1.5 Veritabanından Matris yapıda CSV Dosyasının Hazırlanması

Veri madenciliđinde, verilerin temin edildikten sonra, iřlenecek forma getirilmesi gerekmektedir. Uygulamada bunun için CSV formatı kullanılmıřtır. Arařtırmacı ad-soyad ve email bilgilerinin satırlara, faaliyet alanlarının ID numaraları ile sütünlara gelecek řekilde CSV dosyası hazırlandı. Bu dosyada eđer bir arařtırmacı ilgili sütündeki faaliyet alanında var ise 1, yok ise 0 ile iřaretlenmiřtir.

4.2 Uygulamanın Hazırlanması

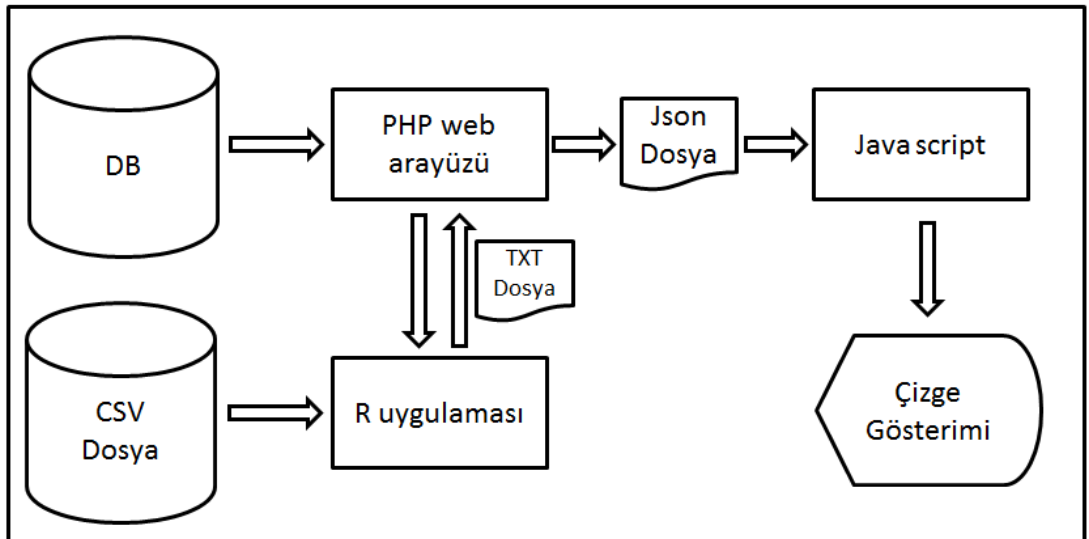
4.2.1 R Dilinde Kodlanan K En Yakın Komşu Algoritmasının PHP ile Entegrasyonu

Araştırmacıların birbirleri ile olan bağlantılarını (komşuluklarını) tespit etmek amaçlı, R programlama dili kullanılmıştır. R kendine özgü fonksiyonları ile karmaşık problemleri dahi saniyeler içerisinde çözümlenip sonuç verebilen bir analiz yapma aracıdır.

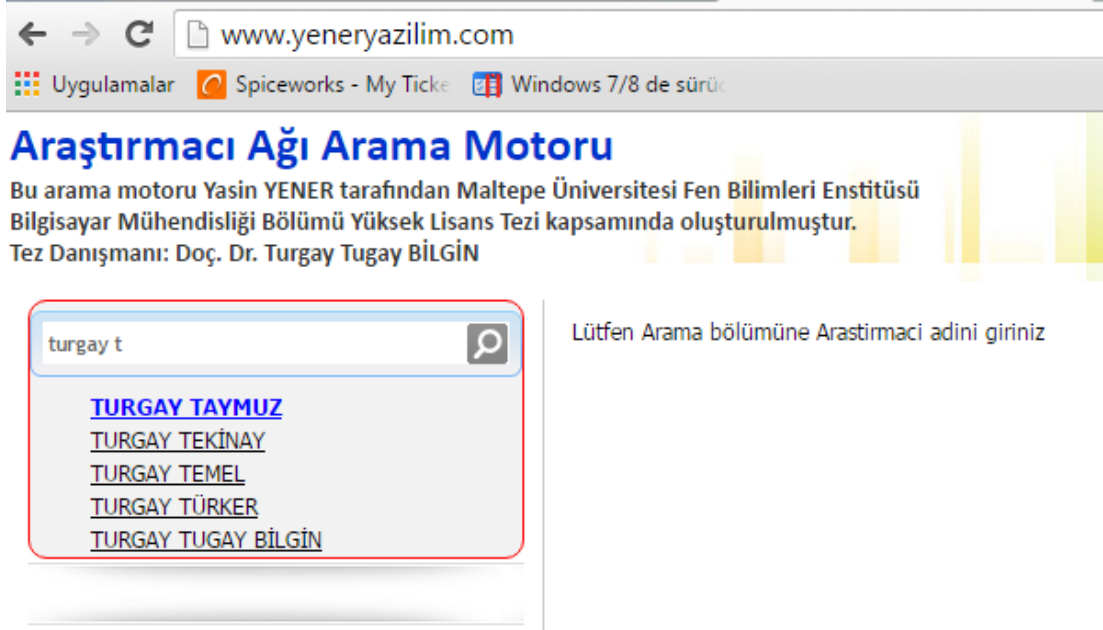
PHP programı ile R programlama dili, PHP programında çalıştırılan exec komutu ile çağırılırken parametre olarak geçilen ve içerisinde R komutlarını barındıran R uzantılı dosyayı çalıştırır. "Rscript" komutu ile R çağırılır. Bu sayede iki program birbiri ile haberleşir.

```
<?php exec("Rscript *.R", $return); ?>
```

R programında kendisine parametre geçilen dosyadaki komutları çalıştırıp sonucunu dizi halinde txt dosyaya yazar. PHP programı txt dosyadaki sonuçları alıp json uzantılı bir dosyaya yazar ve bunları kullanarak çizge grafiğinin çizimini sağlar.



Şekil 18.PHP ile R'nin Haberleşmesi ve Çizge Grafik Çizimi



Şekil 19.Sorgulama Ekran Görüntüsü.

Uygulamada, Şekil 19’da görüldüğü gibi arama alanına yazılan her bir karakterden sonra arka tarafta çalışan ajax komutu ile kullanıcının girdiği metin ile uyuşan arařtırmacıların listesi veritabanından ekrana getirilir. O an ENTER’a basılması halinde, koyu mavi olarak en başta gösterilen arařtırmacının bilgileri formdan gönderilir ve bu arařtırmacının komşuluk hesaplaması yapılır. Arařtırmacıların listelendiđi listeden herhangi bir arařtırmacı seçilerek de komşuluk hesaplaması yaptırılabilir.

AJAX yöntemi sayesinde, bir web uygulamasında aktif form, sayfa yenilenmeden formdaki bilgiler arka tarafta sunucuya gönderilerek, sunucuda işlenmesini ve sunucudan dönen sonucun yine ekran yenilenmeden alınabilmesi sağlanır.

Veritabanındaki kişinin ID numarası ile oluşturulan CSV dosyadaki satır numarası aynı olacak şekilde CSV dosyası hazırlanmıştır. Bu sayede arama sonucunda elde edilen kişinin ID numarası kullanılarak CSV dosyasının hangi satırında çalışma yapılacağı tespit edilir.

R dili komutları ve açıklamaları aşağıdaki şekildedir:

CSV dosyadan bilgilerin okunması işlemi:

```
csvdosya=read.table("csvdosya.csv", sep=";" , quote= "\\")
```

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V
1	arastirmaci	28	30	31	34	35	36	38	39	40	4
2	ABDİ-KÜKNER-kukner@itu.edu.tr	1	0	0	0	0	0	0	0	0	0
3	ABDİ-ATILGAN-dashing0343@gmail.com	1	0	0	0	0	0	0	0	0	0
4	ABDUL HAFEZ-ABDUL HAFEZ-abdul.hafez@hku.edu.tr	1	0	0	0	0	0	0	0	0	0
5	ABDULCELİL-BUĞÜTEKİN-abugutekin@adiyaman.edu.tr	1	0	0	0	0	0	0	0	0	0
6	ABDULKADİR-BİLİŞİK-kadirbilisik@gmail.com	1	0	0	0	0	0	0	0	0	0
7	ABDULKADİR-GÜLLÜ-agullu@gazi.edu.tr	1	0	0	0	0	0	0	0	0	0
8	ABDULKADİR-ÖZCAN-akadirzcn@gmail.com	1	0	0	0	0	0	0	0	0	0
9	ABDULKADİR-TURAN-abdulkadir.turan@yapikredi.com.tr	1	0	0	0	0	0	0	0	0	0

Tablo 4.CSV DosyaTablo Görünümü

Bilgilerin matris yapıya çevrilmesi, satır ve sütun adlarının tanımlanması

```
csvm=as.matrix(csvdosya[2:nrow(csvdosya),2:ncol(csvdosya)
])
rownames(csvm)= csvdosya[2:nrow(csvdosya),1:1]
colnames(csvm)= csvdosya[1:1, 2:ncol(csvdosya)]
```

	row.names	28	30	31	34	35	36	38	39	40	41	42	43
1	ABDİ-KÜKNER-kukner@itu.edu.tr	1	0	0	0	0	0	0	0	0	0	0	0
2	ABDİ-ATILGAN-dashing0343@gmail.com	1	0	0	0	0	0	0	0	0	0	0	0
3	ABDUL HAFEZ-ABDUL HAFEZ-abdul.hafez@hku.edu.tr	1	0	0	0	0	0	0	0	0	0	0	0
4	ABDULCELİL-BUĞÜTEKİN-abugutekin@adiyaman.edu.tr	1	0	0	0	0	0	0	0	0	0	0	0
5	ABDULKADİR-BİLİŞİK-kadirbilisik@gmail.com	1	0	0	0	0	0	0	0	0	0	0	0
6	ABDULKADİR-GÜLLÜ-agullu@gazi.edu.tr	1	0	0	0	0	0	0	0	0	0	0	0
7	ABDULKADİR-ÖZCAN-akadirzcn@gmail.com	1	0	0	0	0	0	0	0	0	0	0	0
8	ABDULKADİR-TURAN-abdulkadir.turan@yapikredi.com.tr	1	0	0	0	0	0	0	0	0	0	0	0
9	ABDULKERİM-BARAN-abdulkerimbaran@tulomsas.com.tr	1	0	0	0	0	0	0	0	0	0	0	0

Tablo 5.Satır ve Sütun Adları tanımlaması sonucu.

R programında, knnx fonksiyonunun çalıştırılabilmesi için ihtiyaç duyulan ve daha önceden indirilen FNN kütüphanesi library komutu ile çağrılarak kullanıma hazır hale getirilmiştir.

```
library(FNN)
```

Matris yapıya dönüştürülen dosyadan sögulanacak olan satırın bilgileri matris olarak query değişkenine atanır. Burada \$id olarak tanımlı değişken araştırmacının veritabanındaki ID numarasıdır. Bu numara aynı zamanda csv dosyadaki satır numarasıdır.

```
query=t(as.matrix(csvm[$id:$id , ]))
```

Şekil 20’de gösterilen sonuçta Turgay Tugay Bilgin incelenmiş olup veritabanındaki ID numarası 10402 dir. R programının konsol ekranında işlemler yapıldığında şekil 20’deki gibi sonuç alınır.

```
> query=t(as.matrix(csvm[10402:10402,]))
> row.names(csvm)[10402]
[1] "TURGAY TUGAY-BÄ° LGÄ° N-turgaybilgin@maltepe.edu.tr"
> |
```

Şekil 20.Query değişkeni içeriği

En yakın komşunun bulunması için knnx fonksiyonu brute algoritması ile kullanılmıştır. Parametre olarak, matris yapıdaki csvm değişkeni, sögulanacak olan satır, en yakın kaç komşunun bulunacağını bilgisi ve yöntem parametre olarak geçilir.

```
knndata=get.knnx(csvm,query , k=30, algorithm=c(brute))
```

get.knnx fonksiyonu çalışmasını tamamladığında, satır indis numarası ve hesaplanan uzaklık bilgisi ile beraber knndata değişkenine bulunan komşulukları matris olarak döndürür.

```
> knndata=get.knnx(csvm,query , k=30, algorithm=c("brute"))
> knndata
$nn.index
  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15] [,16]
[1,] 10402 10395 7155 12033 7403 10226 11856 14840 14896 14949 537 10169 10319 10987 12025 12085
  [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24] [,25] [,26] [,27] [,28] [,29] [,30]
[1,] 12900 13345 13483 9464 10179 10192 10306 11442 11464 12006 12086 12875 12904 12939

$nn.dist
  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11]
[1,] 0 3.741657 4.123106 4.123106 4.242641 4.242641 4.242641 4.242641 4.242641 4.242641 4.358899
  [,12] [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22]
[1,] 4.358899 4.358899 4.358899 4.358899 4.358899 4.358899 4.358899 4.358899 4.472136 4.472136 4.472136
  [,23] [,24] [,25] [,26] [,27] [,28] [,29] [,30]
[1,] 4.472136 4.472136 4.472136 4.472136 4.472136 4.472136 4.472136 4.472136
```

Şekil 21.knndata verisi

Dizide bulunan 1 numaralı kayıt kişinin kendisidir ve değeri 0 dönmüştür. Sonuçlar 0'a yaklaştıkça komşuluğun daha yakın olduğunu gösterir. Uzaklaştıkça ise komşuluğun yani benzerliğin azaldığını gösterir. Dönen bu değerler txt uzantılı dosyaya çıktı alınarak sonuçların PHP kodlarına aktarımı sağlanır.

$$d(i, j) = \sqrt{\sum_{k=1}^P (x_{ik} - y_{jk})^2}$$

Denklem 1.Öklid Uzaklığı Denklemi

KNN (K en yakın komşu) basit bir algoritmadır. Eldeki tüm durumları ve sınıfları yeni duruma göre ölçümler. KNN 1970 li yıllarda tanındı ve günümüze kadar istatistiksel amaçlı kullanılmaktadır. KNN algoritması; Öklid bağlantısı uzaklığı formülünü kullanır (Denklem 1). Öklid bağlantısı iki boyutlu uzayda Pisagor teoreminin uygulamasıdır.

4.2.2 KNN Algoritmasının Yapısı

K En Yakın Komşu yöntemi, sınıflandırma problemini çözen denetimli öğrenme yöntemleri arasında yer alır. Yöntemde; sınıflandırma yapılacak verilerin öğrenme kümesindeki normal davranış verilerine benzerlikleri hesaplanarak; en yakın olduğu düşünülen k verinin ortalamasıyla, belirlenen eşik değere göre sınıflara atamaları yapılır. Önemli olan, her bir sınıfın özelliklerinin önceden net bir şekilde belirlenmiş olmasıdır. Yöntemin performansını k en yakın komşu sayısı, eşik değeri, benzerlik ölçümü ve öğrenme kümesindeki normal davranışların yeterli sayıda olması kriterleri etkilemektedir[28].

Her ne kadar KNN algoritması k-means algoritmasındaki benzer özellikler taşısa da büyük farklılıklar da içermektedir. KNN algoritması bir eğitim verisi içerirken k-means algoritması bir eğitim verisi içermez. Yeni bir değer geldiğinde K değerine mesafeler hesaplanır ve yeni değer bir kümeye ilave edilir. Mesafe hesaplama işleminde ise k-means ve hiyerarşik kümelemede kullanılan öklid uzaklığı, manhattan uzaklığı gibi mesafe hesaplama yöntemleri kullanılabilir. Bu algoritma beş adımdan oluşur [29].

1. Öncelikle K değeri belirlenir.
2. Diğer nesnelere hedef nesneye olan öklid uzaklıkları hesaplanır.
3. Uzaklıklar sıralanır ve en minimum uzaklığa bağlı olarak en yakın komşular bulunur.
4. En yakın komşu kategorileri toplanır.
5. En uygun komşu kategorisi seçilir.

4.2.3 R Uygulamasında Dönen Sonuca Göre Çizge Çizilmesi

R uygulamasından txt dosyasına aktarılan sonuçlar tekrar PHP de işlenerek çizge grafik oluşturulması için Json uzantılı veri dosyası hazırlanır. Graph çizimi için

kullanılan javascript uygulaması json uzantılı dosyayı bir veritabanı tablosu gibi kullanarak buradan ilgili alanlardan bilgileri alır ve çizge grafik çizer.

Json uzantılı dosya hazırlanırken, mesafe değerlerinde minimum ve maximum değerleri alınarak buna göre uzaklık değerlerine 4 aralık hesaplanarak 4 ayrı grup oluşturulur. En yakın grup merkeze en yakın, en uzak grup da merkeze en uzak olacak şekilde çizge grafik çizimi yapılır.

Json dosyası, araştırmacıların, Ad, Soyad, email ve grup numarası bilgisi sütunlarını, satırlar halinde tutar. Javascript uygulaması bu bilgileri alarak çizge grafikte uygun yerlere yerleştirir ve çizge grafiği oluşturur.

```
1 { "nodes": [
2   { "name": "ADİL", "surname": "ALPKOÇAK", "group": 5, "email": "alpkocak@cs.deu.edu.tr" }
3   , { "name": "AHMET ERCAN", "surname": "TOPCU", "group": 5, "email": "atopcu@gmail.com" }
4   , { "name": "AHMET FATİH", "surname": "MUSTAÇOĞLU", "group": 5, "email": "afatih.mustacoglu@tubitak.gov.tr" }
5   , { "name": "ALİ", "surname": "UĞUR", "group": 5, "email": "aliugur@gmail.com" }
6   , { "name": "ALİ KEMAL", "surname": "TAŞCI", "group": 5, "email": "ali_kemal_tasci@yahoo.com" }
7   , { "name": "ATAKAN", "surname": "KURT", "group": 5, "email": "atakanhoca@gmail.com" }
8   , { "name": "ÇAĞATAY", "surname": "ÇATAL", "group": 5, "email": "c.catal@iku.edu.tr" }
9   , { "name": "ÇAĞDAŞ EVREN", "surname": "GEREDE", "group": 5, "email": "cagdas.gerede@gmail.com" }
10  , { "name": "CANAN", "surname": "GİRGİN", "group": 4, "email": "canan.girgin@tubitak.gov.tr" }
11  , { "name": "CEM", "surname": "TUTYOL", "group": 5, "email": "tutyolc@yahoo.com" }
12  , { "name": "CENGİZ", "surname": "ÇELİK", "group": 4, "email": "ccelik@gmail.com" }
13  , { "name": "CENK", "surname": "DERİNÖZLÜ", "group": 5, "email": "cenk.derinozlu@htr.com.tr" }
14  , { "name": "FERHAT ÖZGÜR", "surname": "ÇATAK", "group": 4, "email": "ozgur.catak@tubitak.gov.tr" }
15  , { "name": "GALİP", "surname": "AYDIN", "group": 5, "email": "gaydin@firat.edu.tr" }
16  , { "name": "GEYLANİ", "surname": "KARDAŞ", "group": 5, "email": "geylani.kardas@ege.edu.tr" }
17  , { "name": "GÜRHAN", "surname": "GÜNDÜZ", "group": 5, "email": "ggunduz@pamukkale.edu.tr" }
18  , { "name": "HİDAYET", "surname": "TAKCI", "group": 4, "email": "htakci@cumhuriyet.edu.tr" }
19  , { "name": "HÜSEYİN", "surname": "PEHLİVAN", "group": 4, "email": "pehlivan@ktu.edu.tr" }
20  , { "name": "İLYAS", "surname": "ÇİÇEKLİ", "group": 5, "email": "ilyas@cs.hacettepe.edu.tr" }
21  , { "name": "KEMAL", "surname": "YÜKSEK", "group": 5, "email": "k.yukseki@iku.edu.tr" }
22  , { "name": "KERİM", "surname": "ÖZTEMEL", "group": 5, "email": "kerim.oztemel@agmlab.com" }
23  , { "name": "KORAY", "surname": "PAKYÜREK", "group": 5, "email": "korayp@gmail.com" }
24  , { "name": "MEHMET OZAN", "surname": "KÜSMEN", "group": 5, "email": "okusmen@aselsan.com.tr" }
25  , { "name": "MEHMET SİDDİK", "surname": "AKTAŞ", "group": 3, "email": "msaktas@gmail.com" }
26  , { "name": "MURAT", "surname": "NOHUTCU", "group": 4, "email": "murat@tuana.com.tr" }
27  , { "name": "TANSEL", "surname": "ÖZYER", "group": 4, "email": "ozyer@etu.edu.tr" }
28  , { "name": "TOLGA", "surname": "BERBER", "group": 2, "email": "tolga.berber@fen.ktu.edu.tr" }
29  , { "name": "TURGAY TUGAY", "surname": "BİLGİN", "group": 0, "email": "turgaybilgin@maltepe.edu.tr" }
30  , { "name": "ZEYNEL ABİDİN", "surname": "TOKUOĞLU", "group": 5, "email": "zeynel@devco.com.tr" }
31  , { "name": "ZİYA", "surname": "KARAKAYA", "group": 5, "email": "ziya@atilim.edu.tr" }
32 ] }
```

Şekil 22.Nodes Değişkeni

Json dosyası XML dosya formatı gibi bilgileri tutar. Uygulamada Nodes ve Links adı ile 2 ayrı ağaç yapısı kullanılmıştır. Nodes kısmında R uygulamasından dönen sonuca göre 30 tane araştırmacının bilgilerinin yanında bu kişilerin yakınlık değerine göre oluşturulmuş grup numarası tutulmuştur. 0 numaralı grup aranan araştırmacının

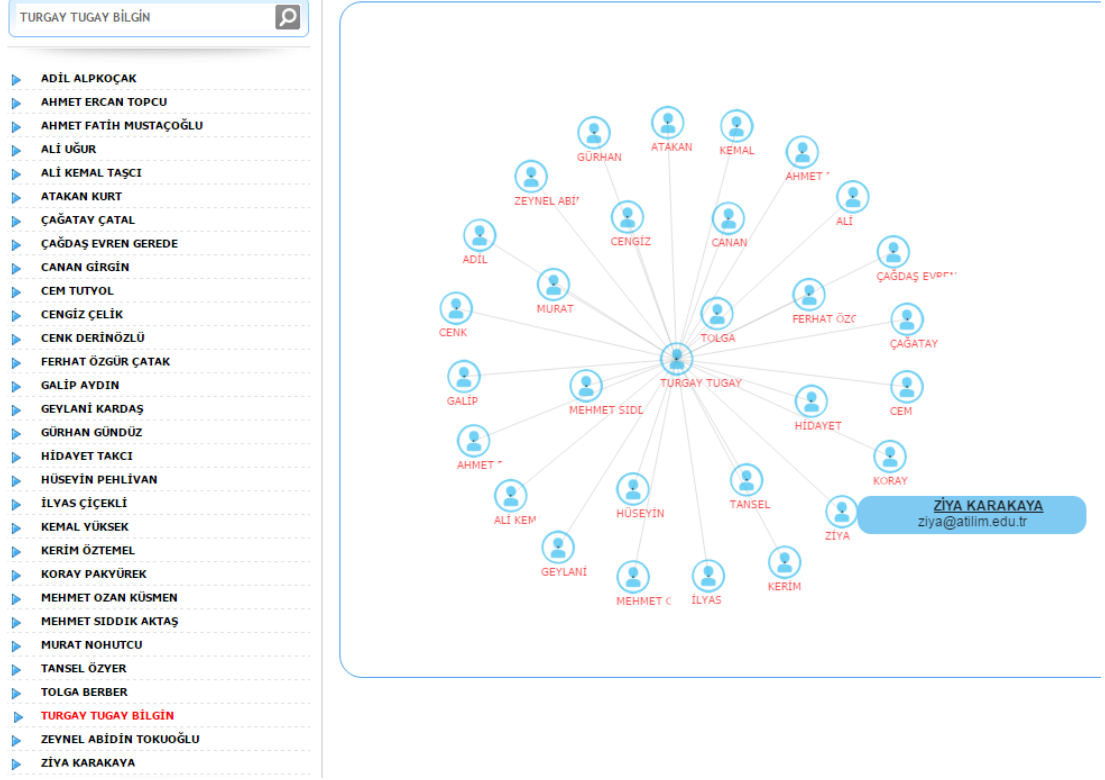
kendisidir. Diğer grup numaraları ise benzerliklerine göre sırasıyla 2,3,4 ve 5 olarak devam etmiştir.

```
, "links": [  
  {"source":27,"target":0,"value":120}  
  , {"source":27,"target":1,"value":120}  
  , {"source":27,"target":2,"value":120}  
  , {"source":27,"target":3,"value":120}  
  , {"source":27,"target":4,"value":120}  
  , {"source":27,"target":5,"value":120}  
  , {"source":27,"target":6,"value":120}  
  , {"source":27,"target":7,"value":120}  
  , {"source":27,"target":8,"value":90}  
  , {"source":27,"target":9,"value":120}  
  , {"source":27,"target":10,"value":90}  
  , {"source":27,"target":11,"value":120}  
  , {"source":27,"target":12,"value":90}  
  , {"source":27,"target":13,"value":120}  
  , {"source":27,"target":14,"value":120}  
  , {"source":27,"target":15,"value":120}  
  , {"source":27,"target":16,"value":90}  
  , {"source":27,"target":17,"value":90}  
  , {"source":27,"target":18,"value":120}  
  , {"source":27,"target":19,"value":120}  
  , {"source":27,"target":20,"value":120}  
  , {"source":27,"target":21,"value":120}  
  , {"source":27,"target":22,"value":120}  
  , {"source":27,"target":23,"value":60}  
  , {"source":27,"target":24,"value":90}  
  , {"source":27,"target":25,"value":90}  
  , {"source":27,"target":26,"value":30}  
  , {"source":27,"target":28,"value":120}  
  , {"source":27,"target":29,"value":120}  
]
```

Şekil 23.Links Değişkeni

Links değişkeninde ise her bir düğümün birbiri ile bağlantısının bilgisi bulunur. Source (Kaynak) ve Target (Hedef) değişkeni bu bilgiyi tutar. Merkezi çizge grafik yöntemi kullanılmasından dolayı merkezde (source) 27 nolu düğüm var, hedefte ise diğer düğümler vardır. Value (Değer) değişkeninde ise düğümlerin merkez düğüme

olan uzaklığı tutulur. Bu sayede değere göre düğümlerin çizge grafiğe yerleşimi sağlanır.



Şekil 24.Çizge Grafik Görünümü

Arama alanına, araştırmacı olarak Turgay Tugay Bilgin'in adı girilmiştir ve bu araştırmacı için yapılan hesaplama neticesinde Şekil 24'de bulunan çizge grafik hazırlanarak ekrana verilmiştir.

Çizge grafikte, araştırmacı Tolga Berber, Turgay Tugay Bilgin'e diğer araştırmacılara göre daha benzerdir. Bu nedenle Turgay Tugay Bilgin'in yakınına gelecek şekilde çizge grafiğe yerleştirilmiştir.

5. SONUÇLAR VE DEĞERLENDİRME

5.1 Elde Edilen Sonuçlar

Bu tez çalışması kapsamında geliştirilen uygulamanın; bir araştırmacının, diğer araştırmacıların ARBİS uygulamasındaki faaliyet alanları bilgisini kullanarak birbirleri ile olan yakınlıklarını (benzerliklerini) doğru bir şekilde hesaplama kabiliyetinde olduğu görülmüştür.

Bu sonuçlar kullanılarak, araştırmacıların sosyal ağı oluşturulmaktadır. Araştırmacı, kendine benzer diğer araştırmacıları benzerlik derecesine göre görsel olarak görebilmekte ve iletişim sağlayabilmektedir.

5.2 Değerlendirme

Yapılan çalışmada, oluşturulan çizge grafik kişilerin kendileri için girmiş olduğu faaliyet alanına göre çizilmektedir. Burada, sisteme doğru bilgilerin girildiğini düşündüğümüzde uygulama doğru sonuç vermektedir. Bir araştırmacı, geliştirilen uygulama ile sorgulama yaparak kendisine benzer araştırmacıları bulup iletişime geçebilmektedir.

Günümüzde sosyal ağları ve hergün üretilen verinin büyüklüğü göz önüne alındığında, insanların bu kadar veriden kendilerine kısa zamanda anlamlı bilgi verebilecek sistemlere, uygulamalara ihtiyaçları doğmuştur. Bu nedenle geliştirilen uygulama sayesinde, yaklaşık 40000 araştırmacının kaydının bulunduğu bir sistemde, araştırmacı kim benimle benzer alanlarda çalışıyor sorusunun cevabını hızlı bir şekilde elde edebilmektedir. Dolayısıyla araştırmacı; bir konu üzerinde yardıma yada beraber çalışmaya ihtiyaç duyduğunda, kendisine benzer kişileri arayıp onlara ulaşmak için harcayacak olduğu zamanı, çalışmakta olduğu projesinde kullanarak zamandan tasarruf edebilmektedir. ARBİS sistemine tanımlı

arařtırmacılar kendileri hakkında, hatalı ya da eksik faaliyet alanı beyan etmeleri durumunda hatalı sonuçlar döndürülebilmektedir.

5.3 Öneriler

Bu tez çalışması sayesinde arařtırmacıların birbirleri ile bağlantı kurmalarının kolaylaştırılması amaçlanmıştır. Arařtırmacıların ARBİS'e bilgilerini doğru eklediğini kabul ettiğimizde geliştirilen bu uygulama da doğru sonuç verilmektedir.

Uygulama, hesaplama işlemini sorgulama esnasında yapmakta ve bu esnada CSV dosyası R programlama dilinde hazırlanan uygulama ile işlenmektedir. Bu işlem ortalama 10-12 saniye zaman almaktadır. Dolayısıyla bir arařtırmacının sonucunun getirilmesi yaklaşık olarak 13-15 saniye kadar sürmektedir. Bu zamanın kısaltılması için çalışma yapılarak sistemden daha hızlı sonuç alınması sağlanabilir.

Geliştirilen yazılım, ARBİS sisteminden güncel veri alamamaktadır. ARBİS ile canlı bir bağlantı sağlanarak ve anlık veriler kullanılarak daha güncel benzerlik sonucu elde edilebilir.

Bu çalışmada, Teknik Bilimler altındaki faaliyet alanlarına tanımlı arařtırmacıların bilgileri kullanılmıştır. Bu işlem, Sağlık Bilimleri, Sosyal ve Beşeri Bilimler, Tarımsal Bilimler ve Temel Bilimler gibi diğere tüm alanlara da uygulanabilir.

KAYNAKLAR

1. Necla Tektaş. Üniversite Öğrencilerinin Sosyal Ağları Kullanımlarına Yönelik Bir Araştırma, Tarih Okulu Dergisi, Sayfa 851-870, Mart 2014.
2. Doç.Dr. Ali Karcı, Onur Boy, Sosyal Ağların Web Madenciliği Teknikleri ile Analizi ve Ortak Atıf Analizi ile Benzerlik Tahmini, Elektrik-Elektronik ve Bilgisayar Sempozyumu 2011.
3. Mehmet Ulvi Şimşek, Sosyal Ağlarda Veri Madenciliği Üzerine Bir Uygulama, Gazi Üniversitesi Fen Bilimleri Enstitüsü; ANKARA; yüksek lisans tezi, Eylül 2012
4. Kuan, J., Fast k nearest neighbour search for R-tree family, Dept. of Electron. & Comput. Sci., Southampton Univ., UK ; Lewis, P, 924-928, IEEE, 1997
5. Yanchang Zhao, An Example of Social Network Analysis with R using Package igraph, blog.RdataMining.com; <https://rdatamining.wordpress.com/2012/05/17/an-example-of-social-network-analysis-with-r-using-package-igraph/>, 2012
6. G.Xu, Y. Zhang, L. Li, “Web Mining and Social Networking” 2011.
7. Baykal, A., Coşkun, C. “Web Madenciliği Teknikleri”, Akademik Bilişim ’09 – XI. Akademik Bilişim Konferansı Bildirileri s.797-800, Harran Üniversitesi Şanlıurfa, Şubat 2009.
8. Sadi Evren Seker, Sosyal Ağlarda Veri Madenciliği (Data Mining on Social Networks), YBS Ansiklopedi www.YBSAnsiklopedi.com Cilt 2, Sayı 2, Haziran 2015
9. P. Hall; B. U. Park; R. J. ,Choice of neighbor order in nearest-neighbor classification, Samworth (2008)
10. Sibel Kırmızıgül Çalışkan, İbrahim Soğukpınar, KxKNN: K-means ve K en yakın komşu yöntemleri ile ağlarda nüfus tespiti. EMO Yayınları, 2008
11. Chelms, C., & V.K., P. (2011). Social networking analysis: A stat of the art and the effect of semantics. Third International Conference on Social Computing, IEEE. IEEE.
12. Borgatti, S., & Everett, M. (2006). A Graph-Theoretic prespective on centrality. Social Networks , 28 (4), 466-484.

13. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research* , 993-1022.
14. Kaschesky, M., Sobkowicz, P., & Bouchard, G. (2011). Opinion Mining in Social network: Modelling, Simulating, and Visualizing Political Opinion Formation in the Web . *The Proceedings of 12th Annual International Conference on Digital Government Research* .
15. Beygelzimer A., Kakade S. and Langford J. (2006), “Cover trees for nearest neighbor,” *ACM Proc. 23rd international conference on Machine learning*, 148, 97-104.
16. Arya S., Mount D.M., Netanyahu N.S., Silverman R. and Wu A.Y. (1998), “An optimal algorithm for approximate nearest neighbor searching,” *Journal of the ACM*, 45, 891-923.
17. Arya S. and Mount D.M. (1993), “Approximate nearest neighbor searching,” *Proc. 4th Ann. ACM-SIAM Symposium on Discrete Algorithms (SODA'93)*, 271-280.
18. Bentley J.L. (1975), “Multidimensional binary search trees used for associative search,” *Communication ACM*, 18, 309-517.
19. Khaled Al-Naami, S. E. (2014). GISQF: An Efficient Spatial Query Processing System. *Cloud Computing, 2014 IEEE 7th International Conference on* (s. 681-688). IEEE.
20. <http://aliunlu.com/2009/01/28/sosyal-ag-analizi-nedir/> (03.03.2016)
21. <http://akademik.yok.gov.tr/AkademikArama/view/viewAuthor.jsp> (03.03.2016)
22. <http://ab.org.tr/ab08/bildiri/71.pdf> (08.05.2016)
23. <https://pinargul.wordpress.com/tag/sosyal-ag-analizi/> (08.05.2016)
24. <https://pinargul.wordpress.com/tag/ag-analizi-ornekleri/> (27.05.2016)
25. <http://www.dataload.com/>
26. <https://arbis.tubitak.gov.tr>
27. <http://akademik.yok.gov.tr/AkademikArama/>
28. Sibel Kırmızıgül Çalışkan, İbrahim Soğukpınar “KxKNN: K-MEANS VE K EN YAKIN KOMŞU YÖNTEMLERİ İLE AĞLARDA NÜFUZ TESPİTİ”, 2.Ağ ve Bilgi Güvenliği Ulusal Sempozyumu Bildirileri, 2008.
29. <http://bilgmuh.nku.edu.tr/erdincuzun/post/k-nn-algoritmasi> (27.05.2016)

ÖZGEÇMİŞ

Yasin Yener, 1983 yılı Amasya doğumludur. İlköğrenim, ortaokulu ve Liseyi Amasya'da tamamladı. 2000 yılında girdiği Cumhuriyet Üniversitesi Bilgisayar Programcılığı bölümünden 2002 yılında mezun oldu. 2005 yılında girdiği Anadolu Üniversitesi İşletme Fakültesi'nden 2008 yılında mezun oldu. 2003 yılından itibaren Asmaş A.Ş. firmasında Bilgi İşlem ve Raporlama Koordinatörü olarak çalışmaktadır. 2013 yılında başladığı Maltepe Üniversitesi Bilgisayar Mühendisliği Yüksek Lisans programına tez aşamasında devam etmektedir.