

T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

**KARMAŞIK AĞLARDA GÜVENİLİR REFERANS
SİSTEMİ OLUŞTURACAK DİNAMİK PUANLAMA
ALGORİTMASI**

Yüksek Lisans Tezi

Tezi Hazırlayan :

Samet ESEN

İSTANBUL, 2017

T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

**KARMAŞIK AĞLARDA GÜVENİLİR REFERANS
SİSTEMİ OLUŞTURACAK DİNAMİK PUANLAMA
ALGORİTMASI**

Yüksek Lisans Tezi

Tezi Hazırlayan :

Samet ESEN

Öğrenci No:

130820020

Danışman:

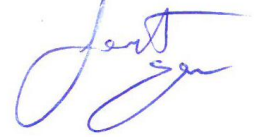
Yrd. Doç. Dr. Ediz ŞAYKOL

İSTANBUL, 2017

YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum “**Karmaşık Ağlarda Güvenilir Referans Sistemi Oluşturacak Dinamik Puanlama Algoritması**” başlıklı bu çalışmanın, bilimsel ahlak ve geleneklere uygun şekilde tarafımdan yazıldığını, yararlandığım eserlerin tamamının kaynaklarda gösterildiğini ve çalışmamın içinde kullanıldıkları her yerde bunlara atıf yapıldığını belirtir ve bunu onurumla doğrularım. 03/08/2017

Samet ESEN



T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZ SAVUNMA SINAVI SONUÇ TUTANAĞI

Beykent Üniversitesi
Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Aşağıda tez adı belirtilen yüksek lisans öğrencisi 130820020 SAMET ESEN no'luin .../.../... tarihinde yapılan tez savunma sınavı¹ sonucunda 45 dakika süreyle sunduğu ve savunduğu tezi hakkında² oybirliğiyle, KABUL kararı verilmiştir.

Bilgilerinize saygılarımızla arz ederiz.

Anabilim Dalı : Bilgisayar Mühendisliği
Programı : Bilgisayar Mühendisliği
Tez Başlığı³ : Karmaşık Ağlarda Güvenilir Referans Sistemi Oluşturacak Dinamik Puanlama Algoritması

Tez Sınav Jürisi

Öğretim Üyesi

Danışman : Yrd. Doç. Dr. Ediz ŞAYKOL
Üye : Yrd. Doç. Dr. Atılgan YILMAZ
Üye : Doç. Dr. Gökhan SİLİHTANLIOĞLU

İmza



¹ Jüri üyeleri söz konusu tezin kendilerine teslim edildiği tarihten itibaren en geç bir ay içinde toplanarak öğrenciyi tez savunma sınavına alır. Belirlenen günde yapılamayan jüri toplantısı, katılanların hazırladığı bir tutanakla enstitü yönetimine bildirilir. Bu durumda jüri en geç onbeş gün içinde toplanarak aday tez savunma sınavına alır. Tez savunma sınav süresi en az 45 dakikadır. Yüksek lisans tez savunma sınavı, tez çalışmasının sunulması ve bunu izleyen soru-yanıt bölümlerinden oluşur ve dinleyiciye açıktır. (Beykent Lisansüstü eğitim ve Öğretim Yönetmeliği-Madde30-3)

² Tez sınavının tamamlanmasından sonra jüri, tez hakkında "kabul", "düzeltme" veya "red" kararı verir. Jüri başkanı, jüri üyelerince imzalanmış sınav tutanağını, tez sınavını izleyen üç gün içinde ilgili enstitü yönetimine teslim eder. Tezi başarısız bulunan öğrencinin Enstitü ile ilişkisi kesilir. Tezi hakkında düzeltme kararı verilen öğrenci en geç üç ay içinde gerekli düzeltmeleri yaparak ve yönetmelikte belirtilen usullere uygun olarak tezini aynı jüri önünde yeniden savunur. Bu savunma sınavında da tezi kabul edilmeyen öğrencinin enstitü ile ilişkisi kesilir. (Beykent Lisansüstü eğitim ve Öğretim Yönetmeliği-Madde30-4)

³ İlerde doğabilecek aksaklıkların engellenmesi için tezin başlığının yazılması gerekmektedir.

Adı ve Soyadı : Samet ESEN
Danışmanı : Yrd. Doç. Dr. Ediz ŞAYKOL
Türü ve Tarihi : Yüksek Lisans, 2017
Alanı : Bilgisayar Mühendisliği
Anahtar Kelimeler : Sosyal Ağ, Karmaşık Ağlar, Kariyer Ağları, Profesyonel Grup Etkileşimi

ÖZ

KARMAŞIK AĞLARDA GÜVENİLİR REFERANS SİSTEMİ OLUŞTURACAK DİNAMİK PUANLAMA ALGORİTMASI

Günümüz internet kullanıcılarının en sık ve yoğun olarak giriş sağladığı internet siteleri, sosyal ağ olarak tabir edilen, Facebook, Twitter, LinkedIn gibi sitelerdir. Sadece sosyalleşmek için değil, bilgi edinme veya kariyer gelişimi için de bu siteler tercih edilir. Bu sosyal ağlar da, kullanıcıların birbirleri ile etkileşimleri karmaşık ağlar üzerine kurulmuştur. Kullanıcılar, takip ettikleri kişi ya da kurumlar ile istedikleri gibi iletişime geçmekte veya çeşitli paylaşımlarda bulunabilmektedirler. Ayrıca kullanıcılar, kariyerlerine de bu muazzam uzayda yön verebilir duruma gelmişlerdir. Doğal olarak kurumlar da ihtiyaçları dahilindeki kişilere bu ağlar üzerinden kolayca ulaşabilir duruma gelmişlerdir. Ancak kurumlar hala, yeni bir çalışanı kabul etmeden önce güvenilir olduklarını düşündükleri kişiler üzerinden referans kontrolü yapmak mecburiyetindedirler. İnsanları sanal ortamda ifade edebilecek, diğer tüm kullanıcılar tarafından da kabul görmüş bir öz değer kavramının oluşması, bu zorunluluğu ortadan kaldıracaktır.

Bu tez çalışması kapsamında geliştirilecek olan nihai sistem ile; karmaşık ağlar da, güvenilir bir referans sistemi oluşturabilmek için dinamik puanlama algoritması tasarlanmıştır.

Name and Surname : Samet ESEN
Supervisor : Assist. Prof. Dr. Ediz ŞAYKOL
Degree and Date : Master, 2017
Major : Computer Engineering
Key Words : Social Networks, Complex Networks, Career Networks,
Professional Group Interaction

ABSTRACT

DYNAMIC SCORING ALGORITHMS TO CREATE RELIABLE REFERENCE SYSTEMS IN COMPLEX NETWORKS

The most frequently and intensely accessed internet sites of today's internet users are social networking sites such as Facebook, Twitter, LinkedIn. These sites are preferred not only for socialization, but also for information or career development. In these social networks, the interaction of users with each other is based on complex networks. Users are able to communicate with each other or with the institutions they follow, or they can be found in various communities. In addition, users have been able to direct their careers to this enormous space. Naturally, they have become easily accessible to the people involved in their needs through these networks. However, institutions are still obliged to perform reference checks on those whom they think are reliable before accepting a new employee. The creation of a concept of self-worth which can be expressed in the virtual environment by people and accepted by all other users, can eliminate this necessity.

With the final system to be developed under this thesis study; in complex networks, a dynamic scoring algorithm is designed to be able to create a reliable reference system.

İÇİNDEKİLER

ÖZ	i
ABSTRACT	ii
TABLolar LİSTESİ	iv
ŞEKİLLER LİSTESİ	v
KISALTMALAR	vi
1.GİRİŞ	1
1.1 Literatür Taraması.....	1
1.1.1 Mevcut Örneklerin İncelenmesi	2
1.2 Tezin Amacı.....	3
1.3 Tez Çalışmasının Yapısı	4
2. SOSYAL AĞLARA GENEL BAKIŞ	5
2.1 Sosyal Ağ Analizi	5
2.2 Geçmişten Günümüze Sosyal Ağlar	6
3. KULLANILAN YÖNTEMLER	9
3.1 Google Pagerank Algoritması	9
3.2 Sosyal Ağlarda Merkezîyet (Centrality)	12
3.3 Tarjan Algoritması	13
4. KARMAŞIK AĞ YAPISI	16
4.1 Karmaşık Ağ Yapısının Kurulması	16
4.1.1 Örnek Veri Kümesinin Oluşturulması.....	17
4.1.2 Karmaşık Ağdaki Referans Değerlerinin Hesaplanması	21
4.2 Power Iteration Yöntemi	23
4.3 Puanlama Sistemindeki Anormallikler	24
4.4 Puanlama Sisteminin Matematiksel Modellemesi	27
SONUÇ	30
KAYNAKÇA	31

TABLÖLAR LİSTESİ

Tablo 1. Oyların veri tabanı tablosundaki kayıt satırları.....	18
Tablo 2. Data Set Puan Hesaplama Sonucu-1.....	22
Tablo 3. Data Set Puan Hesaplama Sonucu-2.....	24



ŞEKİLLER LİSTESİ

Şekil 1: Küçük bir sosyal ağ örneği.....	2
Şekil 2: Genel iş aramak için kullanılan sosyal ağlar.....	12
Şekil 3: Tarjan algoritması pseudo(sözde) kodu	14
Şekil 4: MVC yetkinliği için temas noktaları (vertex) ve bağlantılar	19
Şekil 5: Grafikte bir kişinin seçilmesi ile elde edilen bağlantılar ve kişinin temel oy bilgileri.....	20
Şekil 6: Örnek Veri Kümesi Parçası	21
Şekil 7: Kişiler ve oylar	24
Şekil 8: Kuvvetli bağlı elemanlar	25

KISALTMALAR

BBS	Bulletin Board Service
PR	Pagerank
DÖA	Derin Öncelikli Arama



1.GİRİŞ

İnternet kullanımının da artmasıyla özellikle son 10 yıldır sosyal ağlar insan hayatının vazgeçilmez bir parçası haline gelmiştir. Kullanıcılar sosyal platformlar üzerinden kendileri ile ilgili istedikleri her türlü bilgiyi paylaşır hale gelmişlerdir. Bu bilgiler arasında kişilerin demografik bilgileri, eğitim bilgileri, iş bilgileri vb. türde bilgileri düşünebiliriz. Ancak hala, paylaşılan bu bilgiler doğrultusunda, kişinin sahip olduğu yetkinliklerin düzeyini belirleyebilecek bir sistem mevcut değildir.

Günümüzde LinkedIn üzerinde yetkinlik paylaşma ve puanlama sistemi bulunmaktadır. Ancak bu sistemde de, kişinin ilgili yetkinlikte ne derece uzman olduğunu belirleyebilecek bir yapı mevcut değildir. Bu sistemlerde kişiye oy veren diğer kullanıcıların ilgili yetkinlik ile ilgili bilgi düzeyi hesaba katılmaksızın puanlama yapılmaktadır, doğal olarak ortaya çıkan sonuç gerçeği yansıtmayabilmektedir.

Y yetkinliğe sahip A profesörü ve B öğrencisi için; A kişinin 5 oyu ve B öğrencisinin 50 oyu varsa bu öğrencinin Y yetkinliğinde çok daha uzman ya da bilgili olduğu anlamı çıkarılabilmektedir. Fakat bu sonuç oldukça yanıltıcıdır; A profesörüne verilen 5 oy aslında alanında uzman 5 farklı kişiden alınmış olabilir. Bu 5 oyun herhangi bir tanesi, B öğrencisi için verilmiş 50 oydan çok daha değerli olabilir. Sistemin bu farklılığı anlayabilecek seviyede bir algoritma yapısına sahip olması gerekmektedir.

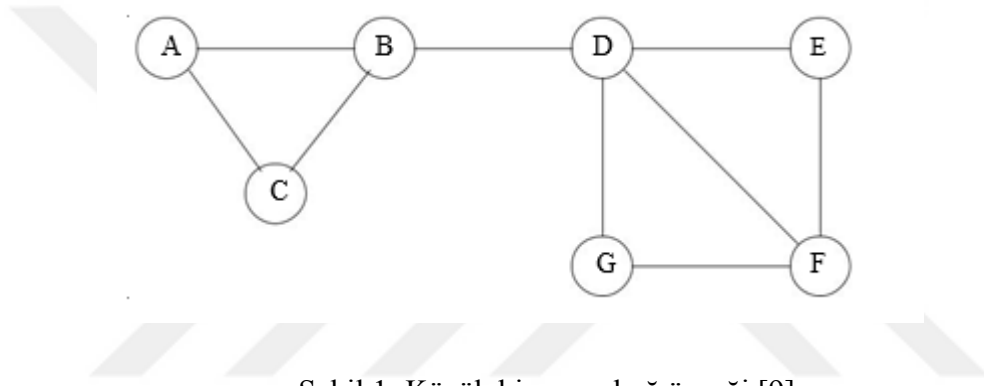
1.1 Literatür Taraması

Ağ kavramı kelime anlamı olarak “birçok nokta ile bunlar arasındaki bağlantılarla gösterilebilen bir dizgeye ilişkin yapı” olarak tanımlanmaktadır [7]. Sosyal bir ağ ise, düğümler olarak adlandırılan bireylerden veya elemanlardan oluşan ve bu düğümleri kendi aralarındaki çeşitli ilişki ve etkileşimlere bağlı olarak birbirine bağlayan bir yapı olarak tanımlanabilir [8].

Sosyal ağlar, yapısındaki dinamizm gereği zaman içinde değişen sosyal kavramlar arasındaki sınırlı ikili ilişki kümelerini içeren sosyal yapılardır [5]. Sosyal ağlar; bireyler, gruplar, organizasyonlar ve toplumlar olmak üzere kavram

kümelerinden oluşmakta ve bu birimler düğüm (node) olarak isimlendirilmektedir. Bu kavramlar arasında olan ilişkiler ise bağ (tie) ya da kenarlar (edge) ile ifade edilmektedir [6].

Sosyal ağlar doğal olarak grafikler olarak modellenir ve biz bunu sosyal grafik olarak isimlendiririz [9]. Sosyal bir ağ hakkında, "topluluklar" 1, yani düğümlerin alt gruplarını (ağı oluşturan insanlar veya diğer varlıklar) olağandışı güçlü bağlantılarla nasıl tanımlanacağı önemli bir sorundur [9]. Bu tip güçlü bağların anlaşılabilmesi ve hareketlerinin analizi için 3. Bölümde değineceğimiz çeşitli algoritmalar mevcuttur.



Şekil 1: Küçük bir sosyal ağ örneği [9].

Bir sosyal ağın en temel özellikleri şunlardır:

- Ağı oluşturan kavramlar olmalı. Örneğin insanlar, kurumlar vb.
- Bu ağdaki kavramlar arasında en az bir bağ olmalı.
- Rassallık ya da yeryüzünün rastlantısal olmadığı varsayımı olmalı. Formülize edilmesi en zor kavram budur. Örneğin bir A işletmesi hem B hem de C kişisiyle ilgileniyorsa, B ve C kişilerinin birbirleriyle olan ilişkileri ortalamaya göre daha güçlü olabilir [9].

1.1.1 Mevcut Örneklerin İncelenmesi

Yetkinlik odaklı sosyal ağ paylaşımının popülaritesi gün geçtikçe artmaktadır. Birçok sosyal ağ özelliklerine kullanıcı yetkinliklerini de ekleyerek, yetenek bazlı arama ve yetkinlik haritası/havuzu oluşturmaya başlamışlardır.

Yetenek odaklı sosyal ağlara örnek olarak;

- LinkedIn: Kişilerin üye profil sayfalarında iş geçmişleri, eğitim bilgileri ve özelleştirilebilir modüller bulunmaktadır. Bu ağdaki üyeler bağlantı olarak tanımlanmaktadır. LinkedIn ağında ücretli veya ücretsiz olmak üzere iki hesap alınabilmektedir. Ücretsiz olarak alınan hesaplarda aynı iş ve okul ortamındaki kişilere erişilebilme, kendi sayfalarında çeşitli bağlantılar yayınlama ve kişilerin profil sayfalarını görüntüleyerek yeteneklerine onaylayarak onlara oy verebilme gibi çeşitli işlemler gerçekleştirilebilmektedir. Ücretli üyeliklerde bu işlemlere ek olarak daha geniş aramalar yapılabilmektedir, yani LinkedIn veri tabanındaki kişilere daha kolay erişim sağlanabilmektedir.
- Ryze.com: Bu sosyal ağda kişiler ücretsiz olarak kendi sayfalarını oluşturarak işleri ve ilgi alanları ile ilgili çeşitli ağlara katılabilmektedirler. Diğer üyelere mesaj yollayabilir ve üyelerin kendi profil sayfalarını görüntüleyebilmektedirler. Belirli üyelik ücretleri ile gelişmiş arama özelliği sağlayan bir hizmet sunmaktadır.
- Xing.com: LinkedIn ile benzerlik göstermektedir. Üyelerinin kişisel profil sayfalarını sunmaktadır ve çeşitli tartışma panoları içermektedir. Basit bir hesap söz konusu ise bu hesap ücretsiz olarak alınabilmektedir; fakat çeşitli aramalar yapmak, mesaj göndermek gibi işlemler ise ücretli üyelik gerektirmektedir.

1.2 Tezin Amacı

Sosyal bir ağın karakteristik özelliklerinde de bahsettiğimiz gibi; raslantısal hiçbir şeyin olmadığı varsayımıyla, birbirleriyle ilişkisi olan kavram ya da kişilerin, karmaşık bir ağ üzerinde birbirlerine dokundukları noktaların değerinin, diğer kavram ya da kişiler ile olan değerinden farklı olması gerekliliği aşikardır. Ayrıca sosyal ağ kavramının gün geçtikçe arttığı günümüzde, kullanım arttıkça, insanların istekleri de şekillenecektir. Kullanıcılar merak ettikleri bir konu veya bir yetkinlik üzerinde, hem kendilerinin hem de diğer kullanıcıların bilgi seviyelerini görmek isteyeceklerdir.

Kariyer amaçlı sosyal ağ kavramı ile bilgi amaçlı sosyal ağ kavramının birleştirilmesi ve bu sürecin dengeli ölçüm yöntemleri ile yürütülmesi bu tez çalışmasını oluşturan temel kavramdır. Güvenilir bir referans sistemi oluşması, kariyer maksatlı kullanımını arttıracaktır.

Bu tez çalışması ile sosyal ağ kavramını oluşturan karmaşık ağlar da, dinamik bir puanlama algoritması tasarlanmıştır. Bu algoritma ile güvenilir bir referans sistemi oluşturmak hedeflenmiştir. Oluşan referans sistemi, kişilerin uzmanlık alanlarında birbirleri ile ilişkiler kurmasına ve konu merkezli etkileşimde bulunmalarına da araç olacaktır.

1.3 Tez Çalışmasının Yapısı

Bu tez çalışmasının I. bölümü olan Giriş bölümünde, tezin amacı anlatılmış ve günümüzdeki örnekler incelenmiştir. II. Bölümde, tez çalışması esnasında kullanılan yöntemler detaylı bir şekilde irdelenmiştir. III. Bölümde, belirlenen yöntem anlatılmış ve örnek veri kümesi üzerinde uygulanan test sonuçlarına değinilmiştir. IV. Bölümde, ağdaki olası usulsüz ve kötüye kullanım ile haksız referans elde etme konusuna getirilen öneriden bahsedilmiştir. Tezin sonunda konu ile ilgili kaynaklar ve referanslar sunulmuştur.

2. SOSYAL AĞLARA GENEL BAKIŞ

2.1 Sosyal Ağ Analizi

Sosyal ağların iş bulma, iş sorunlarına çözüm getirme, iş konusunda profesyonel arkadaşlıklar kurma amacı ile kullanımı, internet öncesi dönemlerden bu yana süre gelmektedir. Sosyal ağ, (social network), içinde bir grup sosyal bireyin ve bu bireyleri birbirine bağlayan ilişkilerin tanımlandığı kümeye verilen genel addır. Bireyleri birbirine bağlayan özellikler içeren, bu bireylerin profesyonel, amatör ya da tamamen sosyal amaçlı etkileşimde bulunmasını sağlayan sosyal ağlar, araştırma açısından sosyal yapıların ya da bu yapılarda görülen özelliklerin analizi amacıyla da kullanılabilirler. Bu bağlamda sosyal ağ analizi ile sosyal yapılardaki yerel ve global desenlerin bulunması, etkili olan bireylerin tesbit edilmesi, ve ağ dinamiklerinin keşfi gibi göreceli olarak kompleks işlemler gerçekleştirilebilir.

Sosyal ağlar ve bu ağların analizi disiplinlerarası bir araştırma konusu olup sosyal psikoloji, sosyoloji, istatistik, çizge teorisi, veri madenciliği gibi birçok ayrı alanda çalışmayı gerektirmektedir. Sosyoloji alanında Simmel [10] grup ilişki ağları üzerinde ilk çalışmaları yapmış ve teori geliştirmiştir. Benzer biçimde Moreno [11] kişilerarası ilişkilerin tanımlanması için “sociogram” adı verilen yapıyı gerçekleştirmiştir. Bu çalışmalar 1950 li yıllarda matematiksel olarak formalize edilmiştir. Günümüzde modern sosyolojide sosyal ağ analizi en önemli paradigmalardan biri durumunda olup diğer birtakım bilim dallarında da uygulama alanı bulmaktadır. Diğer kompleks ağ problemleri ile birlikte sosyal ağ problemleri “ağ bilimi” bilim dalı altında incelenmektedir. Burada “sosyal ağ” kavramının çok geniş ve bilişim biliminden çok daha eskiye uzanan bir spektrumu kapsadığını tekrar etmek ve bu anlamda kullanıldığı ifade etmek gereklidir.

Sosyal Ağlar, internet tabanlı sosyal ağ sistemlerinin yaygınlaşması ile birlikte yaygın olarak popüler kavrama dönüşmüş ve internet tabanlı sosyal ağ sistemlerini doğrudan tanımlamak amacıyla da kullanılmaya başlamıştır. Günümüzde sosyal ağlar profesyonel ve eğlence amaçları ile kullanılmaktadır.

2.2 Geçmişten Günümüze Sosyal Ağlar

Sosyal ağlar kapsamında ilk akla gelen teknolojik yapı “facebook” olmakla birlikte, ağ tabanlı sosyal ilişkilerin 1970 lerden başlayarak BBS, AOL, USENET ve CompuServe hizmetleri ile başladığını söylemek mümkündür. Özellikle kısa mesafe yerel telefon hattı ile yerel BBS (Bulletin Board Service) üzerinden birbirleri ile haberleşen, fikir alışverişi yapan, problemleri tartışan profesyoneller, günümüz sosyal ağları anlamına en yakın etkileşimi 1980’li yıllarda yaygın olarak gerçekleştirmişlerdir. Fidonet ve USENET ile sosyal ağlar daha da dünya çapına yaygınlaşarak aynı konu ile ilgili kişilerin bir arada etkileşimde bulunmalarına ve haberleşmelerine olanak sağlamıştır. 1990 ların sonuna doğru classmates.com ya da dünyada 6 kişilik uzaklıkta herkesin bir tanıdığına aynı olması teorisine dayanılarak oluşturulan sixdegrees.com örnekleri internet üzerinde oluşturulmuştur. Hızlı, gerçek zamanlı ya da kısmi gerçek zamanlı, mesajlaşma ve sohbet, yine sosyal ağ kavramına paralel kavram olarak IRC (Internet Relay Chat Network) ve BBS’ler üzerinde ortaya çıkmakla birlikte daha sonraları ICQ, AOL, MSN, GoogleHangout gibi yeni varyasyonlarla devam etmektedir. Bu sistemlerde birbirlerini önceden tanıyan kişilere özel gruplar oluşturulabildiği gibi, oluşmuş gruplara tanınmamış kişilerin de gruplara katılabilmesi olanağı bulunmaktadır.

Sonraki dönemde ise Friendster, LinkedIn, MySpace, Facebook, Google+ gibi yapıların oluştuğu, bu yapıların bir kısmının profesyonel, bir kısmının eğlence, bir kısmının ise her ikisini de içerecek türde hazırlandığı gözlenmektedir. Bunlara ek olarak hızlı mesajlaşma ve paylaşım üzerine kurgulanmış twitter yapısının çok yüksek sayıda kullanıcı kitlesine sahip olduğu gözlenmektedir.

Sosyal medyanın profesyonel kullanıcılar tarafından kullanımının ise dramatik bir biçimde arttığı, kurum içi ve dışı ile etkileşimin profesyonel olarak geliştiği gözlenmektedir[12].

Başlangıçta okul öğrencileri ve mezunları için tasarlandığı iddia edilen Facebook, günümüzde iş başvurularında ve kişisel referansların oluşturulmasında önemli rol oynayan bir konuma gelmiştir[12]. Bu paralelde sadece profesyonel kullanımı hedefleyen LinkedIn sistemi ardından kurulmuş ve operasyona geçmiştir.

LinkedIn sisteminde birbirini profesyonel olarak tanıyan kişiler kendilerini ilişkilendirebilir, birbirlerinin hangi alanda uzman olduklarını etiketleyebilmektedirler. Ancak LinkedIn günümüzde çoğunlukla özgeçmiş tanımlaması ve yeni iş aramada sıçrama noktası olarak kullanımı daha yaygındır.

Sosyal ağlar bilgi paylaşımında da ciddi olarak kullanılmaya başlamıştır. Yukarıda bahsedildiği gibi BBS, Fidonet ve USENET dönemlerinde yalnızca sosyal etkileşim yanında profesyonellerin ve amatörlerin belirli bir konu üzerinde fikir alışverişi ve soru-yanıt etkileşimi gösterdiği bilinmektedir. Günümüzde binlerce konuda hazırlanmış olan forumlarda kullanıcılar belirli bir konuda soru-cevap ilişkisi içine girmektedirler. Bazı forumlar yalnızca üyelerin erişebileceği içeriklere, bazıları herkesin erişebileceği içerikleri bazıları da kısmen üyelerin kısmen herkesin erişebileceği içeriklere sahiptir. İçeriği açık olan bölümlere arama motorlarının da erişebilmesi sayesinde belirli konularda sorulan soru ve cevaplar arama motorlarından doğrudan bulunmaktadır. Bu da sözkonusu forum sistemlerinin yeni üyelere erişiminde bir model olarak gözlenmektedir.

Başlangıçta bilgi ve belirli konulara hedeflenmiş sosyal forum sistemlerinde kullanıcılar arası ilişki yalnızca moderatör ve normal kullanıcılar seviyesi ile sınırlı iken, zamanla “rütbe” ya da “rank” olarak adlandırılacak olan ve forum üyesinin kredibilitelerini belirleyen yapılar ortaya çıkmıştır. Bu model ile uzun süre sistemde üye olan, faydalı paylaşımlarda bulunan ve sözüne güvenilir olduğu anlaşılan bireylerin rütbeleri yukarı doğru geliştirilir. Değişken rütbe yapıları forum sistemlerinde verilen yanıtların, yapılan katkıların güvenilirliklerini de doğrudan belirleyen etken olmuştur.

Katılımcıların birbirine oy verdiği model, forum gibi yapılardan e-ticaret yapılarına da ulaşmıştır. E-ticaret sistemlerinde satıcılar ve alıcılar birbirlerine puanlama yardımı ile derecelendirme yaparak güvenilirlikleri hakkında sonraki müşteri ya da satıcıları bilgilendirmektedirler. Bu sayede katılımcıların sistemdeki göreceli pozisyonu, sisteme dahil olan diğer katılımcılar tarafından bilinmekte, böylelikle gerçek fiziksel ortamdaki etkileşimlere daha yakın bir sosyal etkileşim ile e-ticaret desteklenmektedir.

Sosyal etkileşim ile belirli soruların yanıtlarının aranması modelini Usenet döneminden gelen eski forum tabanlı yapıdan yeni modern yapıya dönüştüren başarılı sistemlerden biri StackExchange sistemidir. Bu sistemde sorular ve yanıtlar kişiler üzerinden eşleştirilir ve belirli konularda uzmanlar zaman içinde oluşur. Verilen yanıtlar ve yanıtların doğru kabul edilmesi ile belirli bir zaman içinde katılımcıların ustalık seviyesi kayıt altına alınır ve paylaşılır. Bu sayede belirli bir soruya verilen yanıtın doğruluğu üzerinde kullanıcıların fikir sahibi olması sağlanmış olmaktadır.

Profesyonel iş ilişkisi ve “tanıdık” yapısını doğrudan iş modeli olarak ele alan LinkedIn benzeri bir yapının, içerik geliştirme ve profesyonel bilgi kaynağı olma modeline dayanan yapıdaki unsurları içermediği görülmektedir. Her ne kadar mevcut durumda üyeler hakkında bilgi alınması mümkün olsa da hem bilgi, hem de referans ve içerik bileşeninin aynı sistemde entegre edildiği yaygın bir çözüm piyasada mevcut değildir. LinkedIn halihazırda kullanıcıların daha sık uğraması amacıyla konu bazlı Facebook tarzında profesyonel tartışma grupları oluşturmuş olsa da tam olarak istenilen neticeyi alamamıştır. Buna ek olarak referans verenlerin ağırlıklarının etken olarak kabul edildiği bir yapının halihazırda mevcut olmadığı görülmektedir.

3. KULLANILAN YÖNTEMLER

Yukarıda anlatılan gerekçelere istinaden sosyal ağ özelliği taşıyan, profesyonel kariyer bilgilerine dayalı ve aynı zamanda bilgi paylaşımı ve konu/uzmanlık merkezli düğüm noktası özellikleri içeren bir sistemin gerekliliği aşikardır.

Bu çalışma ile yukarıda ifade edilen ihtiyaca yönelik geliştirmeler yapılmıştır. Bu sistem kişilerin kariyer gelişimlerini takip etme ve profesyonel sosyal ağlarını genişletme yanında, başlangıç noktası bulma amacıyla diğer kullanıcılar ile etkileşimini sağlayacaktır. İçerik miktarını geliştiren kullanıcıların kredibilitesi artacak, bu kişilerin “önerdiği” diğer kullanıcıların tavsiye değeri de otomatik olarak yükselecektir. Kullanıcıların diğer kullanıcılar tarafından puanlaması, onay toplam puanına ve kademesine sahip olması, bu puan ve kademenin gücü ile paralel bir biçimde diğer uzmanları etiketleme olanağına sahip olmaları sağlanmıştır.

Halihazırdaki sosyal ağlarda ve diğer sosyal etkileşim ortamlarında bir başkasına ya da bir ürüne verilen puanlar gözönüne alındığında tüm kullanıcılar eşit etki sahibi olarak değerlendirilmektedir. Oysa rank değeri zaman içinde sistem içinden yükselmiş olan kullanıcıların yaptıkları etiketlemeler ve puanlamaların değerinin çok daha yüksek olması gerekmektedir.

Halihazırdaki mevcut yaygın sistemlerde, verilen puanların veya yapılan etiketlemelerin doğruluğunun test olanağı yoktur. Oysa ki zincir ilişkilerin yapay etkisini kullanarak kendisini belirli bir konuda uzman olarak gösterebilmek bu sistemlerde mümkündür.

3.1 Google Pagerank Algoritması

Pagerank algoritması, birbirine referans veren web sayfalarının önem ve kalite derecelerini belirlemek amacıyla yaygın kullanılan algoritmadır. Arama motorları internetteki sayfaların ne kadar ziyaretçi aldığını, içeriklerinin ne kadar kaliteli olduğunu ve bizim o sayfaya hangi olasılıkla ulaşabileceğimizi ölçmek için her sayfaya bir değer atar. Sayfa değeri olarak Türkçe’ye çevirilen pagerank, bu değere verilen addır. Sayfa değeri hesaplaması belli bir algoritmaya göre

yapılmaktadır. Bu algoritma internet sayfalarının birbirleriyle olan ilişkilerini ölçebildiği gibi, aslında birbirleriyle bir şekilde ilişkili olan herhangi bir veri yığına da uygulanabilir. Sayfa değeri bir sayfanın değerini hesaplar, o sayfaya diğer sayfalardan verilen bağlantıları göz önüne alır. Bağlantı veren sayfaların çok ziyaret edilmesi onların verdikleri bağlantıların değerini ancak artırır; çünkü bu sayfalardan yola çıkarak diğer bir sayfaya ulaşma olasılığımız daha yüksektir. Sayfa değeri birçok internet tarayıcısının sayfaları sıralandırmak için başvurduğu yöntemdir. Bu değer ile arama sonuçlarında sayfa ilk sıralarda yer alır.[13]

Aşağıdaki formül, sayfanın değerinin hesaplanabilmesi için kullanılan pagerank algoritmasının formülüdür.

$$PR(A) = (1 - d) + d (PR(t1) / C(t1) + \dots + PR(tn) / C(tn))$$

- **t1...tn:** Değerini bulmak istediğimiz sayfaya bağlantı veren sayfalar
- **PR(tn):** Bize bağlantı veren sayfaların kendi değerleri
- **C(tn):** Her sayfanın diğer sayfalara verdiği bağlantı sayısı
- **d:** Tüm oranların toplamının 1' i geçmesini engellemek için çarpılan katsayı (damping factor)
- **1-d:** (Tüm sayfaların sayısı) / (Sayfa değeri) oranının 1 olduğu varsayılarak d(PR(t1)/C(t1)...) değerine, d'den geri kalan değer burada eklenmiştir.

Basit bir örnek ile formülü inceleyelim; örneğimizde en az bir bağlantısı olan sayfalar kullanacağız. A,B,C,D isimli 4 sayfamız olsun. Sayfa değeri'nin bu sayfalar arasında eşit olarak dağıldığını varsayarak her sayfaya 0.25 değerini verelim. Eğer bu sayfalardan B,C ve D sayfaları, A sayfasına bağlantı verirse her biri A sayfasına 0.25'lik bir katkı sağlamış olur. O zaman bütün sayfa değeri A'da toplanmış olur çünkü tüm bağlantılar A'ya gitmektedir. Yani A'nın değeri toplam=0.75 olur. [13]

O zaman başka bağlantıların da var olduğunu varsayarak devam edelim. B'nin C'ye , D'nin diğer tüm sayfalara bağlantı vermesi durumunda bağlantı değerleri dışı verilen bağlantılar arasında paylaşılır. Yani B, A'ya 0.125 değerinde bir oy vermiş olur. D ise yalnızca 0.083. Yani ilk değerinin 1/3 ü kadar. [13]

$$PR(A) = \frac{PR(B)}{2} + \frac{PR(C)}{1} + \frac{PR(D)}{3}.$$

Kısacası, dışarıdan verilen bağlantı değerlerinin toplamını, bu değerlerin normalleştirilmiş haline bölerek toplam sayfa değerini bulabiliriz. [13]

PageRank teorisindeki “d” değerini ise internette gezinen ve sayfalara rastgele tıklayan ve en sonunda tıklamaktan vazgeçen birinin her adımda bir sonraki sayfaya da tıklama olasılığını sabittir. Uzun hesaplamalar ve araştırmalar sonunda 0.85 değeri kabul edilmiştir. Azaltan katsayı (damping factor) 1 den çıkartılarak sayfa değeri değerine eklenir. [13]

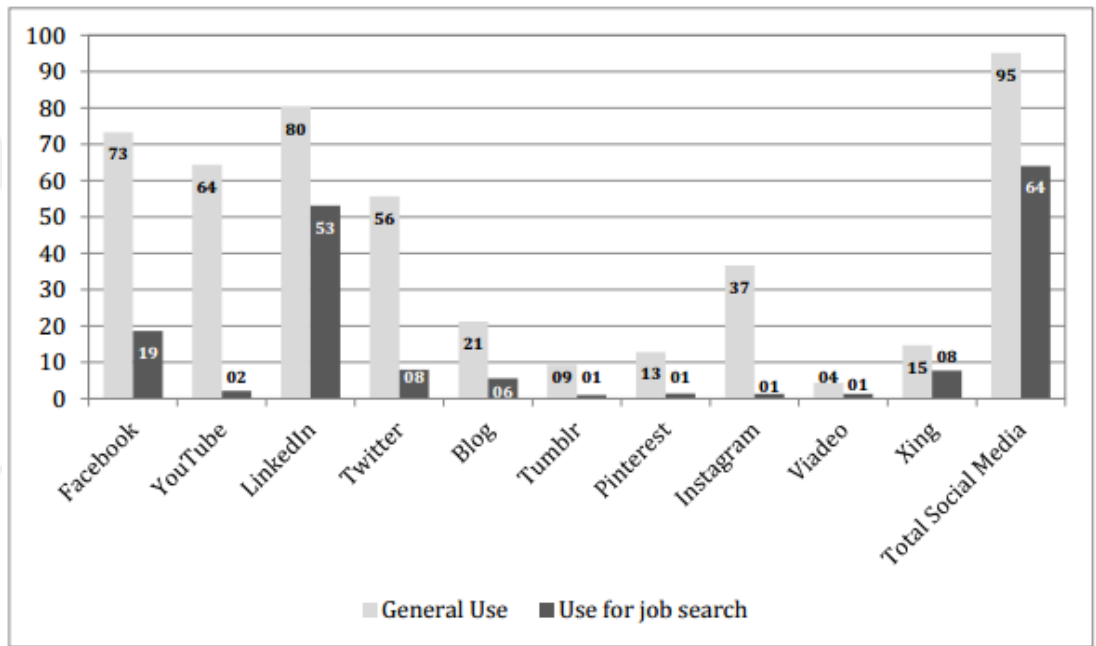
$$PR(A) = \frac{1-d}{N} + d \left(\frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)} + \dots \right).$$

Yukarıdaki örnekte, en az bir bağlantısı olan sayfaları baz almıştık. Bu 4 sayfanın da başlangıç değerini 0.25 almıştık. Her sayfanın değerinin ona bağlantı veren sayfaların değeriyle hesaplandığı düşünüldüğünde, henüz hiç sayfa değeri hesaplanmamış bir sayfanın diğer bir sayfaya verdiği değeri bulmak için farklı işlemler yapılmalıdır. Google’ın yöntemine göre; sayfa değeri, tekrar tekrar uygulanan bir algoritma ile hesaplanır ve internetteki sayfaların normalleştirilmiş link matrisinin öz vektörüne (eigenvector) karşılık gelir. Yani, sayfaların son değerleri bilinmeden de hesaplama yapılabilir. Bunun için aynı formül hesaplanan geçici değerler üzerinden tekrar tekrar (iterative) uygulanınca, her seferinde olması gereken değere bir adım daha yaklaşılr. Bu hesaplanan değerler saklanarak bir sonraki adımda kullanılır ve yeni değerleri bulunur. Bu işlemi değerler artık belli bir sayıya yaklaşına ve adımlar arası değişme miktarları yeteri kadar azalana dek sürdürmemiz durumunda henüz hiç sayfa değeri hesaplanmamış bir sayfanın da sayfa değeri hesaplanır. [13]

Pagerank algoritmalarının ihtiyaca göre modifikasyonları, alana özel sıralama ve sosyal ağlara yönelik çalışmalarda yaygın görülmektedir. Genellikle standart algoritmanın çok fazla link verenlerin dezavantajlı duruma düşürmesi gibi bazı sorunlarının uygulama alanında mevcut olmaması nedeniyle, konu duyarlı pagerank algoritmalarında bazı modifikasyonlar yapılmaktadır.

3.2 Sosyal Ağlarda Merkezîyet (Centrality)

İş bulma süreçlerinde sosyal ağların kullanımı konusunda çalışmalar yapılmıştır.[14]



Şekil 2: Genel iş aramak için kullanılan sosyal ağlar [14]*

Bunlar iş arama ve sosyal ağlar, kariyer planlamada sosyal ağlar gibi alt problemlere odaklanmış hep problemin sosyal boyutunu hem de matematiksel boyutunu incelemişlerdir. Sosyal ağlar konusunda yapılan çalışmaların bir bölümü Çizge (Graph) algoritmaları içerir. Bu algoritmaların birçoğu temel çizge algoritmalarını uyarlanması ya da doğrudan uygulanmasıdır. Çizge üzerinde dolaşma algoritmaları, bu algoritmaların zaman karmaşıklığı konularının incelendiği görülmektedir. Çizge algoritmalarında döngülerin bulunması, bu döngülerin etiketlenmesi, iki düğüm arasındaki bağın kuvvetinin tespiti, etki analizi, geleneksel klik problemi gibi problemlerin sosyal ağlar kapsamında ve sıralama problemleri kapsamında literatürde incelendiği görülmektedir. [15]

Çizgelerden yola çıkarak sosyal ağlarda “centrality” ölçümleri, etki bırakan kişilerin tespiti çalışmaları incelenmiştir. Bu çalışmalarda çizge bağlarının ağırlıkları belirli noktalarda düğüm ve çekim (gravity) oluşturabilmektedir. Bu durumda o noktadaki kişilerin etkileyici ya da usta pozisyonda olduğu sonucu çıkarılmaktadır. Bunun akabinde sosyal ağların formal analizi çalışmaları göze çarpmaktadır. Bu kapsamda temel işlemler arasında uzaklık, döngü, çizgede yürüme, çizgede düğüme erişilebilirlik, düğüm kümeleme, düğüm grup yapıları, yoğunluk hesapları ele alınmaktadır. Kategorik değerler kullanıldığında özdeğer ve yakınlık algoritmaları çizgeler üzerinde yine işlenebilmektedir.

Bu sistemde diğer sosyal ağlarda olmayan referans veren kişiye göre ve onun yetkinliklerine göre ağırlıklandırma ve yetkinlik sıralandırmaları yapılabilecektir. Buradaki puanlama etkisi, puanlayan kişinin yetkinliğine ve sistem içindeki aktivitesinin miktarı ve türüne göre değişim gösterecektir. Kullanıcının işveren, tecrübeli bir personel, konu ile ilgili çalışmış bir akademisyen olması durumlarında puanlama gücü farklı olacaktır. Buna proje kapsamında "kullanıcı özdeğeri" adı verilmiştir ve projeye özgün bir hesaplama yöntemi ile elde edilmesi planlanmaktadır. Yetkinlikler hakkında bilgi ve deneyim sahibi kişilerin puan etkisi fazla olacağından kişilerin profillerinde belirttikleri onay almış yetkinlik değerleri daha tutarlı ve güvenilir olacaktır.

3.3 Tarjan Algoritması

Özellikle usulsüz kullanım ya da şişirme olarak adlandırılan yapay referanslama hareketlerinin önüne geçilmesi sistemin önemli gereksinimlerinden biridir. Birbirlerine puan vererek kapalı bir devre oluşturan 2,3,4,...,n kullanıcının puanlarının zincirleme reaksiyon ile yükselmesine izin vermeyen bir sistem planlanmıştır. Bu amaçla çizge algoritmalarında bilinen gelişmiş tekniklerden uygun olanları kullanılacaktır.

Proje kapsamında birbiri ile çok fazla yetkinlik oyu verenlerin gruplarının bulunabilmesi için sıkı bağlı çizge bileşenleri (Strongly Connected Components) kullanılmasının uygun olduğu değerlendirilmektedir [16]. Bu amaçla kullanılan

algoritmalar arasında en popüler, hızlı olan ve anlaşılır olan Tarjan algoritması yaygın kullanılmaktadır.

Tarjan'ın algoritması, grafiğin teorisinde, grafiğin kuvvetle bağlı bileşenlerini bulmak için kullanılan bir algoritmadır. Kosaraju'nun algoritması ve yol tabanlı güçlü bileşen algoritması da dahil olmak üzere alternatif yöntemler için zamanı eşleştiren doğrusal zaman içinde çalışır. Tarjan Algoritması ismi, keşfedicisi Robert Tarjan tarafından seçilmiştir. [1]

```
algorithm tarjan is
input: graph  $G = (V, E)$ 
output: set of strongly connected components (sets of vertices)

index := 0
S := empty array
for each  $v$  in  $V$  do
  if ( $v.index$  is undefined) then
    strongconnect( $v$ )
  end if
end for

function strongconnect( $v$ )
  // Set the depth index for  $v$  to the smallest unused index
   $v.index := index$ 
   $v.lowlink := index$ 
   $index := index + 1$ 
  S.push( $v$ )
   $v.onStack := true$ 

  // Consider successors of  $v$ 
  for each ( $v, w$ ) in  $E$  do
    if ( $w.index$  is undefined) then
      // Successor  $w$  has not yet been visited; recurse on it
      strongconnect( $w$ )
       $v.lowlink := \min(v.lowlink, w.lowlink)$ 
    else if ( $w.onStack$ ) then
      // Successor  $w$  is in stack  $S$  and hence in the current SCC
      // Note: The next line may look odd - but is correct.
      // It says  $w.index$  not  $w.lowlink$ ; that is deliberate and from the original paper
       $v.lowlink := \min(v.lowlink, w.index)$ 
    end if
  end for

  // If  $v$  is a root node, pop the stack and generate an SCC
  if ( $v.lowlink = v.index$ ) then
    start a new strongly connected component
    repeat
       $w := S.pop()$ 
       $w.onStack := false$ 
      add  $w$  to current strongly connected component
    while ( $w != v$ )
    output the current strongly connected component
  end if
end function
```

Şekil 3. Tarjan algoritması pseudo(sözde) kodu

Donald Knuth Tarjan'ın algoritmasını, The Stanford GraphBase adlı kitabında Knuth'un en sevdiği uygulamalardan biri olarak tanımladı. [3]

Ayrıca şunları yazdı:

Bu problem için tasarladığı veri yapıları inanılmaz derecede güzel bir şekilde bir araya geldi, böylece yönlendirilmiş bir grafiği araştırırken bakmanız gereken nicelikler her zaman sihirli bir şekilde parmaklarınızın ucunda. Ve aynı zamanda Tarjan'ın algoritması topolojik sıralama da yapar.

Tarjan algoritması yönlendirilmiş çizelgede kuvvetli bağlı grupları bulmak için kullanılan bir yöntemdir. Kuvvetli bağlı grup herhangi iki köşe arasında en az bir yönlü yol bulunan maksimum sayıda köşeye sahip gruptur. Tarjan algoritmasının tanımını Derin Öncelikli Arama (DÖA)'yı temel almaktadır. Köşeler DÖA yöntemine göre gezilir ve sıralanır. L köşesi , V köşesinden ulaşılabilen en düşük sıra numaralı köşedir. Aynı temsilciye sahip köşeler aynı kuvvetli bağlı gruba yerleştirilir.

Temsili olarak işaretlenen aynı ağlar, aynı kuvvetli bağlı gruplar olarak yerleştirilir.[4] Tarjan algoritması değiştirilmiş bir Derin Öncelikli Arama algoritmasıdır. Bu yüzden karmaşıklığı $O(|V| + |E|)$ 'dir.

4. KARMAŞIK AĞ YAPISI

Bu çalışma ile, çok kullanıcı, kompleks profesyonel ağ oluşturma özellikli, bilgi paylaşma ortamı destekli, puanlama kabiliyetine sahip profesyonel bir sosyal ağ ve kariyer etkileşim sistemi oluşturulması hedeflenmiştir. Sosyal etkileşim sistemindeki kullanıcılar, özgeçmiş bilgilerini ve uzmanlık alanlarını girebilecektir. Uzmanlık alanı ile ilişkilendirilmiş link bilgilerini ve referansları (URL, DOI vb), diğer kullanıcıların özümsemesi ve kendisi için kullanabilmesi amacıyla sisteme ekleyeceklerdir. Bu referanslar editörlük gereksiniminin olmaması sebebiyle link biçiminde olacaktır. Kullanıcılar, diğer kullanıcılardan bidirim alabilecekler, belirli bir konuda yetkinlik düğümü açabilecekler ve yetkinlik konusu üzerinde gerek bağlantıları paylaşabileceklerdir.

Kompleks sosyal ağ etkileşimi ile sistemde bulunan diğer kullanıcılar ile mevcut kullanıcı arasındaki farklı ağırlıkta profesyonel uzaklık ve bağ gücü hesaplanacaktır. Bu sayede kişiler arası bağlar ve etki güçleri hesaba katılarak belirli yetkinliklerde uzmanlar sıralanabileceklerdir.

4.1 Karmaşık Ağ Yapısının Kurulması

Ağ yapıları doğrudan düğümlerden elde edilmesi zor olan bilgileri ortaya çıkarmak için kullanılmaktadır. Örneğin, protein etkileşim ağı içindeki hastalık ile ilişkili proteinlerin keşfi [17], sosyal ağ kullanarak reklamlarda potansiyel müşteri hedefleme [18], ve kimyasal molekül yapılarında benzer ağ örüntüleri araştırma gibi konular ağ tabanlı analitik metodolojileri kullanarak gerçekleştirilmektedir.

Düğüm etiketli ağ (node-annotated network) şunlardan oluşur:

- Düğüm listesi
- Düğümler arasında ilişkiyi temsil eden bağlantılar
- Düğümlerle ilişkili özellikler

4.1.1 Örnek Veri Kümesinin Oluşturulması

Örnek veri setini oluşturmak için 1000 kullanıcı oluşturulmuş ve 1000*1000'lik bir matris kullanılmıştır. 48 adet yetkinlik sisteme eklenmiş ve 11 adet yetkinlik grubu oluşturulmuştur. 1000 kişiyi %10'luk hata payı bırakarak 100'er setler halinde böldük. Her bir 100'lük set için 1 yetkinlik grubuna dahil olan yetkinliklerden çeşitli yetkinlikler atanmıştır.

Örneğin;

Yetkinlik: Futbol, Tenis, Basketbol, Yüzme

Yetkinlik Grup: Spor

Bu kullanıcıların %70'i 10-20 kişiye referans, %30'u 20-30 kişiye referans olmuşlardır. Sistemde test için minimum 10.000 maksimum 30.000 referansın verilmiş olması planlanmıştır.

Kişilerin birbirine olan bağlantı tipleri de; İş Arkadaşı, Yönetici, Akademisyen ve Arkadaş olarak tanımlanmıştır.

Herhangi bir referans, bu yıl mı yoksa 5 yıl önce mi verilmiş olması matematiksel puan hesaplamada önemli bir rol oynayacaktır. Bu yüzden kullanıcıların aldığı referansların tarihleri de sistem tarafından tutulmaktadır.

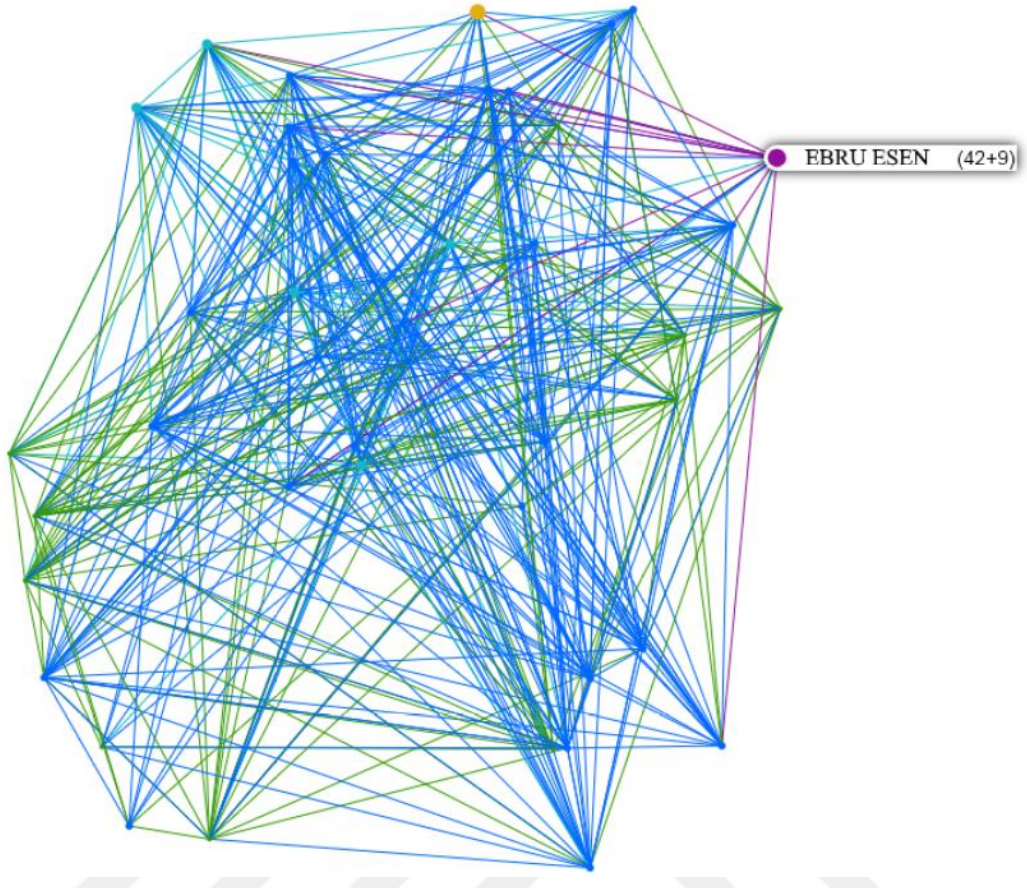
Oylama işleminde oluşturulan 100 kişilik kümeler arasında rasgele oylama yapılmıştır. Bu sayede hatalı kayıt girilmemesi hedeflenmiştir.

Oylama ile ilgili tüm veriler veri tabanında oluşturulan tablolara kayıt edilmektedir. Bu veriler tabloda kodlar ve değerler şeklinde tutulurlar. Tablo 1 de bu verilen örneği görülebilir.

	UserSkillVoteID	VoterID	RatedID	SkillID	VoteDate	ConnectionTypeID	Vote
1	114543	902	962	19	2014-09-17 14:45:38.507	1	1
2	114609	902	962	20	2014-09-17 14:46:12.170	1	1
3	114669	903	962	25	2014-09-17 14:46:43.403	1	1
4	114875	903	962	28	2014-09-17 14:48:29.063	1	1
5	114891	903	962	10	2014-09-17 14:48:37.323	1	1
6	115048	905	962	6	2014-09-17 14:49:59.517	1	1
7	115072	905	962	28	2014-09-17 14:50:11.947	1	1
8	115096	905	962	40	2014-09-17 14:50:24.343	1	1
9	115164	905	962	20	2014-09-17 14:50:59.570	1	1
10	115180	905	962	42	2014-09-17 14:51:07.883	1	1
11	115358	907	962	7	2014-09-17 14:52:41.677	1	1
12	115371	907	962	27	2014-09-17 14:52:48.430	1	1
13	115488	909	962	36	2014-09-17 14:53:51.477	1	1
14	115510	909	962	34	2014-09-17 14:54:03.067	1	1

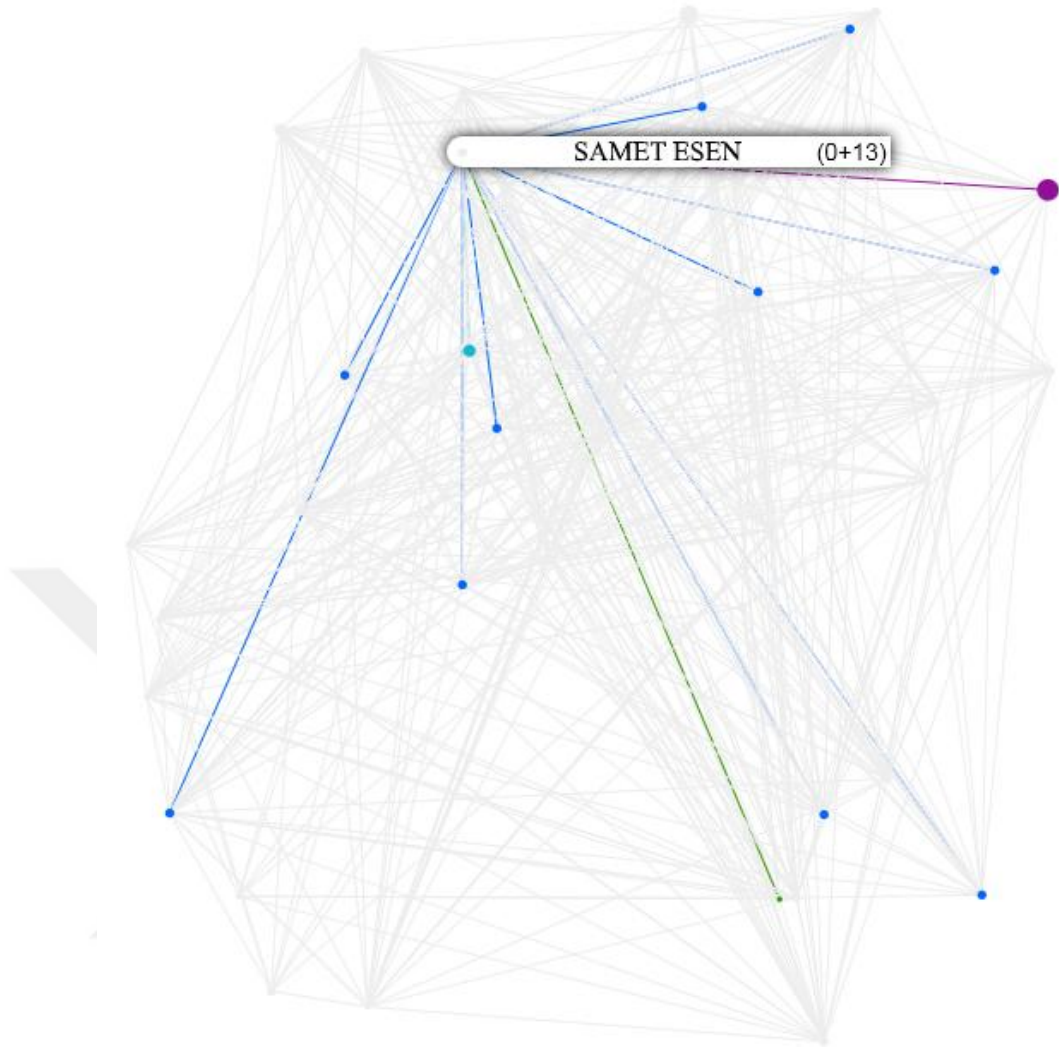
Tablo 1. Oyların veri tabanı tablosundaki kayıt satırları

Sayısal veriler şeklinde tutulan bu değerler ancak görsel bir şekilde sembolize edildiğinde anlam kazanırlar. Oluşturulan ağ kullanıcılar arasında oy verme/alma işlemi sonucu meydana geldiği için en iyi görselleştirme yöntemi bu bağlantıları ve kişileri gösteren dinamik bir ağ grafiği olacaktır. Şekil 4 de MVC yetkinliği için kişiler ve aralarındaki bağlantıyı gösteren grafik verilmiştir.



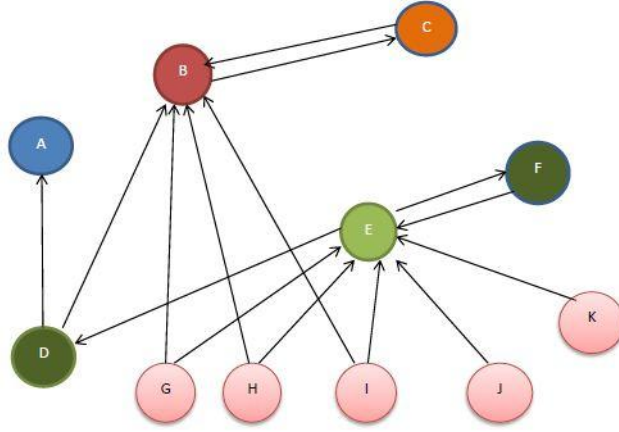
Şekil 4: MVC yetkinliği için temas noktaları (vertex) ve bağlantılar

Bu sistem üzerinde temas noktaları kişileri göstermektedir. Tarayıcı üzerinde çalışan bu uygulama sayesinde belirli bir kişi seçilerek sadece bu kişinin bağlantılarından oluşan grafik görüntülenebilir. Şekil 8 de “Samet ESEN” kişisi için bağlantılar ve bu kişinin baz puanı ile aldığı oyların sayısı gösterilmiştir.



Şekil 5. Grafikte bir kişinin seçilmesi ile elde edilen bağlantılar ve kişinin temel oy bilgileri

Yukarıda tanımlanan bileşenleri içeren ve hedeflenen yapının temsili Şekil 9'daki biçimde verilmektedir. Her düğüm kullanıcıyı, kullanıcılar arasındaki arkadaşlık/yetkinlik aktivitesi bağlantıyı oluşturmaktadır.



Şekil 6: Örnek Veri Kümesi Parçası ¹

4.1.2 Karmaşık Ağdaki Referans Değerlerinin Hesaplanması

Tasarlanan hesaplama işleminde, kullanıcının bir yetkinlik için başka bir kullanıcıya referans verdiği durum göz önünde bulundurularak hesaplama işlemi yapılmıştır.

Kişinin sahip olduğu puanlar: öz değer puanı ve oylardan gelen puan

Eğer kişinin öz değeri sıfırdan farklı ise, yüksek puanlı öz değerlerini minimize etmek amacıyla 0.01 ile çarpılır.

Toplam etki hesaplanırken pagerank algoritmasından [13] esinlenerek önce kişinin puanı 0.18 ile çarpılır ve öz değeri ile toplanır. Kişinin diğer kişilere verdiği oy etkisi ise, ilk adımda 1, daha sonraki her adımda 0.82 ile çarpılır ve etkisinin azaltılması sağlanır.

Bu etki hesaplanırken oy verilen her kişi ve oy veren kişi public bir listede tutulur. Her yeni etki hesaplaması esnasında etkinin oy veren kişiyi etkileyip etkilemediği yukarıda belirttiğimiz liste sayesinde kontrol edilir. Eğer etki oy veren kişiye ulaşır ise puanlama işlemi sonlandırılır.

¹ Şekil 6'da örnek olarak A, B, C, D, E, F, G, H, I, J, K kişileri oluşturulmuş ve bir yetkinlikte ilgili oy vermeleri sağlanmıştır.

```
double baseyuzde = 0.01;  
double pointyuzde = 0.01;  
double pointkatsayi = 0.18;  
double sonraketkikatsayi = 0.82;
```

```
if (item.BaseValue != 0)  
{  
    baseetki = (double)item.BaseValue * baseyuzde;  
}  
double toplametki = baseetki + item.Point * pointkatsayi;
```

Şekil 9’da gösterilen oy verme durumunun puan hesaplaması sonucu Tablo 2’de belirtilmiştir.

	id	label	AldigiOySayisi	point	baseValue	totalPoint
▶	1001	A	1	17,14376799999...	0	17,14376799999...
	1002	B	7	31,64807999999...	0	31,64807999999...
	1003	C	1	19,226568	0	19,226568
	1004	D	1	17,3304	0	17,3304
	1005	E	6	28,23487919999...	0	28,23487919999...
	1006	F	1	16,5104	0	16,5104
	1007	G	0	9,1	0	9,1
	1008	H	0	9,1	0	9,1
	1009	I	0	9,1	0	9,1
	1010	J	0	9,1	0	9,1
	1011	K	1	15,72243999999...	0	15,72243999999...

Tablo 2. Data Set Puan Hesaplama Sonucu-1

Yukarıdaki algoritmayla hesaplandığında özdeğeri yüksek olan kullanıcıların vermiş olduğu referans ve puanların daha değerli olacağı görülmektedir. Bu sayede biriktirilen uzmanlık puanları ve referanların kalitesinin oldukça iyi bir seviyede tutulması sağlanmıştır.

4.2 Power Iteration Yöntemi

Matematikte, **power iteration** yöntemi bir eigenvalue problemidir. Verilen bir A matrisine algoritma λ değerini (eigenvalue) ve sıfır olmayan v vektörünü (eigenvector) üretir. Burada geleneksel eigenvalue probleminde $Av = \lambda v$ eşitliği geçerlidir. Power iteration algoritması basit bir algoritmadır. Matris çözümlenmesi yapılmaz dolayısı ile spars matrislerde kolaylıkla kullanılabilir. Sosyal ağ matrislerine dolayısı ile uygundur. Öte yandan tek bir değer bulması ve yavaş yakınsaması sorunları arasında gösterilmektedir.

Power Iteration metodunda ise, bir kişinin aldığı oy değeri, oy veren kişinin puanın verdiği oy sayısı kadar bölünerek hesaplanır. Her bir kişi(X) için aşağıdaki formül uygulanarak puan hesaplama işlemi gerçekleştirilir.

İlk aşamada oylama başlamadan önce kişilere ortak puan ataması yapılır. (100/11=9.1).

$$PR(x) = \frac{1-\tau}{y} + \tau \sum_{y \rightarrow x} \frac{PR(y)}{out(y)}$$

Puan hesaplama işlemi yapılırken yine pagerank algoritmasından yararlanılarak belirlenen katsayılarla puanlar çarpılır.

`double kendiKatsayisi = 0.18;`

`double digerlerininKatsayisi = 0.82;`

Örneğin, B kişinin puanı:

$$PB(B) = 0.18*9.1 + 0.82* [PR(C) + \frac{1}{3}R(E) + \frac{1}{2}PR(D) + \frac{1}{2}PR(F) + \frac{1}{2}PR(G) + \frac{1}{2}PR(H) + \frac{1}{2}PR(I)] \approx 29$$

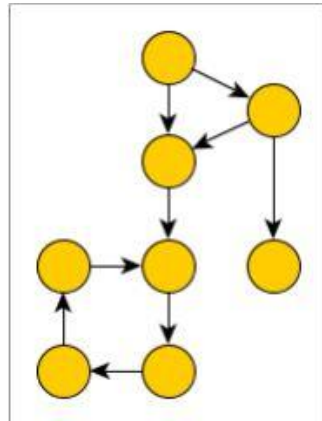
id	label	AldigiOySayisi	point	baseValue	totalPoint
1001	A	1	5,369	0	5,369
1002	B	7	28,998666666666...	0	28,998666666666...
1003	C	1	9,0999999999999...	0	9,0999999999999...
1004	D	1	4,1253333333333...	0	4,1253333333333...
1005	E	6	30,242333333333...	0	30,242333333333...
1006	F	1	4,1253333333333...	0	4,1253333333333...
1007	G	0	1,638	0	1,638
1008	H	0	1,638	0	1,638
1009	I	0	1,638	0	1,638
1010	J	0	1,638	0	1,638
1011	K	1	4,1253333333333...	0	4,1253333333333...

Tablo 3. Data Set Puan Hesaplama Sonucu-2

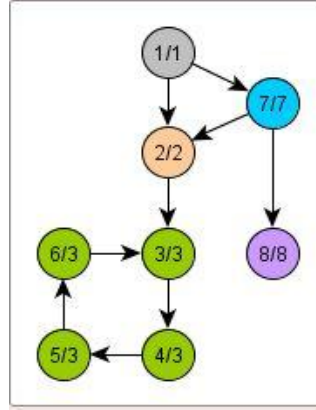
4.3 Puanlama Sistemindeki Anormallikler

Tasarladığımız sistem içerisinde kullanıcılar, istedikleri kullanıcıya, istedikleri yetkinlikte referans olabilirler. Sistem içerisinde ayrıca bir kullanıcı için referans olma yada referans isteme sınırlaması bulunmamaktadır. Bu yüzden sistemin kötü kullanımını engellemek için, birbirine sürekli oy veren grupların tespit edilebilmesi amacıyla Tarjan algoritması kullanılmıştır.

Aşağıda, Şekil 7’de kişileri ve verilen oylar gösterilmektedir. Şekil 8’de tarjan algoritması uygulanarak bulunan kuvvetli bağlı elemanlar aynı renkte belirtilmektedir.



Şekil 7. Kişiler ve oylar



Şekil 8. Kuvvetli bağlı elemanlar

Şekil 7’de ki örnek data setimize Tarjan algoritmasını uygulayarak iki adet kuvvetli bağlı kişiler grubu bulunmuştur.

Bu gruplar; 6/3-3/3 ve 5/3-4/3 kişilerinden oluşan gruplardır.

Tasarladığımız sistemde önce anormalliklerin bulunabilmesi için Tarjan algoritması çalıştırılıp anormallikler en başta bertaraf ediliyor. Daha sonra ise hesaplama işlemi uygulanıyor.

Tarjan algoritması sistem tasarımı sırasında aşağıdaki gibi kullanılmıştır. Rekürsif olarak çalışan algoritma sonrasında birbiri ile kuvvetli bağlı (strongly connected) olan kişi grupları arasındaki bağ koparılıp hesaplama işlemi için uygun hale getirilmiştir.

```

//Rekürsif Cycle detection fonksiyonu
private void ControlConnectedComponents(List<Vertex<int>> graph)
{
    //Gelen graph DetectCycle methoduna gönderilir ve cycle oluşturan listeler bulunur
    var detector = new StronglyConnectedComponentFinder<int>();
    var components = detector.DetectCycle(graph);

    //Eğer cycle var ise
    if (components.Cycles().Count() > 0)
    {
        //bulunan cycle adedi sayılır
        totalCyclesCount += components.Cycles().Count();

        //Cycle oluşturan listelerle işlem yapıldıktan sonra independent node'lar yeni graphımıza
        eklenecektir. Bu nedenle temp bir liste kopyalıyoruz.
        Stack<Vertex<int>> independentTempList = new Stack<Vertex<int>>();
        foreach (StronglyConnectedComponent<int> independent_v in
        components.IndependentComponents())
        {
  
```

```

foreach (Vertex<int> v in independent_v)
{
    independentTempList.Push(v);
}

//Bulunan componentler arasındaki her Cycle incelenir
foreach (StronglyConnectedComponent<int> item in components.Cycles())
{
    //Tüm node'lar bir stack üzerine alınır, bu stack sayesinde cycle oluşturan tüm elemanlar en alt
    kırılımdan en son kırılıma doğru listelenir.
    /*
    A - B - C şeklinde bir cycle {A,B,C} şeklinde stack'e alınacaktır.Bu'da ilk kırılımlın listenin ilk
    elemanı, son kırılımda son elemanı olduğunu gösterir.
    |_____|

    */
    Stack<Vertex<int>> tempList = new Stack<Vertex<int>>();
    foreach (Vertex<int> v in item)
    {
        tempList.Push(v);
    }

    //relatedVertex : üzerinde çalışılan vertex'i gösterir
    Vertex<int> relatedVertex;

    //cycleGraphFirstVertex : cycle'in ilk elemanını tutar
    Vertex<int> cycleGraphFirstVertex = tempList.FirstOrDefault();

    //Eğer oluşan cycle son elemanın ilk elemana bağlantısı yüzünden oluşmuş ise iki eleman
    arasındaki bağlantı kesilir
    //Aksi halde listedeki tüm elemanlara ve o elemanların Dependencies'lerine bakılır.
    if (!tempList.LastOrDefault().Dependencies.Contains(cycleGraphFirstVertex))
    {
        //Dependent node bulunduğu for loop break edilir.
        bool isBreak = false;

        for (int j = 1; j <= tempList.Count() - 1; j++)
        {
            //Tüm listeyi dönüp, listenin ilgili elemanını relatedVertex değişkenine atarız.
            //Örneğin : {A,B,C} listesi için C elemanı alınır.
            relatedVertex = tempList.ElementAt(tempList.Count() - j);

            //RelatedVertex değişkenindeki değeri tüm diğer elemanlar ile kıyaslamak için listeyi
            relatedVertex elemanı DAHİL OLMAYACAK şekilde döneriz.
            //Örneğin : {A,B,C} listesi için B elemanından itibaren diğer tüm elemanlar dönülür.
            for (int i = j + 1; i <= tempList.Count(); i++)
            {
                //C dışındaki tüm elemanları dönüp, önce C'nin dependencies'lerine ardından diğer
                elemanların dependicilerinde C'nin olup olmadığına bakılır.
                cycleGraphFirstVertex = tempList.ElementAt(tempList.Count() - i);
                if (relatedVertex.Dependencies.Contains(cycleGraphFirstVertex) &&
                    cycleGraphFirstVertex.Dependencies.Contains(relatedVertex))
                {
                    //Her iki dependencies bulunduğu iki node arasındaki bağlantı koparılır ve loop'tan
                    çıkarılır.
                    tempList.LastOrDefault().Dependencies.Remove(cycleGraphFirstVertex);
                    isBreak = true;
                    break;
                }
            }
        }
    }
}

```

```

        if (isBreak)
        {
            break;
        }
    }
}
else
{
    tempList.LastOrDefault().Dependencies.Remove(cycleGraphFirstVertex);
}

//Yukarıda oluşturulan independent list, templist'imize eklenir ve yeni graph oluşturulur.
foreach (Vertex<int> independent_v in independentTempList)
{
    tempList.Push(independent_v);
}

//Graph yeniden oluşturulurken Index ve LowLink değerlerinin default haline getirilmesine
dikkat edilmelidir.
graph = new List<Vertex<int>>();
foreach (Vertex<int> new_v in tempList)
{
    new_v.Index = -1;
    new_v.LowLink = 0;
    graph.Add(new_v);
}
//Rekürsif fonksiyon yeniden çağırılır ta ki Cycle bulunamayana dek!!!
ControlConnectedComponents(graph);
}
}
}

```

4.4 Puanlama Sisteminin Matematiksel Modellemesi

Puanlama sistemi, diğer kullanıcıların vermiş oldukları yetkinlik puanları ve kendi yapmış oldukları aktivitelerin ve sistem yöneticisi tarafından belirlenen, meslek grubu ile ilişkiden kazanılan puanların birleşiminden oluşmaktadır. Daha etkin bir sıralama ve ilişki analizi yapılabilmesi amacıyla tasarlanan sistemde eldeki veriler tek bir nihai yetkinliğe çevrilmeden önce vektör olarak işlemlere sokulması, bu sayede farklı yaklaşımlarla parametrelerde oynamak sureti ile yetkinliğin doğru hesaplanması ve sıralama yapılabilmesi hedeflenmiştir. Toplam n adet kullanıcının olduğu bir sistemde i indeksi ile belirtilen i numaraları kullanıcının yetkinlik öz değeri birden çok bileşen içeren bir vektör olarak \mathbf{W}_i ile gösterildiği takdirde aşağıdaki şekilde genel ifade yazılabilir.

$$W_i(R_i, L_i, P_i)$$

Burada vektör bileşenleri ise;

R_i : i numaralı kullanıcının kendine gelen referanslardan elde etmiş olduğu özdeğer olarak tanımlanır.

L_i : i numaralı kullanıcının sistem içine eklemiş olduğu anlamlı linklerin sayısı olarak tanımlanır.

P_i : Kullanıcının sistemdeki genel aktivitesi dolayısıyla almış olduğu puandır.

Sıralama amacı ile bu yetkinlik vektörünün skalar bir değere dönüştürülmesi gerekmektedir.. Bu amaçla r , l ve p katsayıları vektörel bileşenleri tek bir skalar değere dönüştürmede kullanılmaktadır. Dolayısı ile nihai skalar indeks değeri aşağıdaki gibi hesaplanır:

$$W_i = x_i * (rR_i + lL_i + pP_i)$$

Yukarıdaki formülde r, l, p katsayılar tüm sistem için sabit olarak belirlenmiş ve parametre türlerine sistem yöneticisi tarafından atanan ve göreceli önemi belirten katsayılardır. x_i katsayısı ise (0.0-1.0 aralığında) sistemde suistimal ve ağırlık tespiti algoritmalarının vermiş olduğu, kişinin öz değer ağırlık validasyon katsayısıdır. Problemlili olduğu sezgisel, kural tabanlı ya da algoritma ile tespit edilen kullanıcıların skalar yetkinlik değerinin kalibrasyonu ya da gerekirse etki dışına alınması amacıyla kullanılmaktadır. Default değeri 1.0 olarak belirlenir. Pagerank algoritması yürütüldüğü sırada her iterasyonda kolon değerinin normalize edilmesi amacıyla her iterasyonda W_i değerleri $\mathbf{W}(W_1, W_2, W_3, \dots, W_n)$ toplam vektörünün uzunluğu $|\mathbf{W}|$ değerine bölünerek kolon toplamının 1 olması sağlanır. Dolayısı ile burada önemli olan bir olgu, kişiye gelen referansların hesaplanmasında diğer kişilerin nihai skalar değerlerinin kullanılması gerekliliğidir. Bir başka deyişle:

$$R_i = \sum_{j=1}^{j=n} (W_j * k_{ji}) \quad k_j \begin{cases} 1, \text{ eğer referans mevcut ise} \\ 0, \text{ eğer referans mevcut değil ise} \end{cases}$$

olarak hesaplanır. Kullanıcılar, bu değere göre sıralanırlar ve her bir yetkinlik için en yüksek R değerine sahip olan kullanıcı en yetkin olarak belirlenmiş olur.



SONUÇ

Sosyal ağlar insan hayatının vazgeçilmez bir parçası haline gelmiştir ve paylaşılan kişisel bilgilerin, kişinin sahip olduğu yetkinliklerin düzeyini belirleyebilecek bir literatürde sistem mevcut değildir. Mevcut sistemlerde kişiye oy veren diğer kullanıcıların ilgili yetkinlik ile ilgili bilgi düzeyi hesaba katılmaksızın puanlama yapılmaktadır. Kişilerin uzmanlık seviyelerini gösterebilecekleri birer baz değerleri olması ve sistemin bu değerleri baz alarak verilen oyları adil dağıtabilmesi önemli bir durumdur. Bu çalışma ile öncelikle günümüz sosyal ağları incelenmiş, gerekli literatür araştırmalarıyla sosyal ağ kavramının geçmişten günümüze detaylı analizi yapılmıştır. Çalışma kapsamında önerilen yöntem ile, referans verme/olma temeline dayanan sosyal bir ağda dinamik puanlama yapılmasını sağlayacak bir kompleks ağ algoritması sunulmuştur. Çalışmanın temelini oluşturan karmaşık ağ algoritmaları detaylı bir şekilde anlatılmıştır. Google pagerank, centrality ve sıkı bağlı çizge bileşenlerinin (Strongly Connected Components) bulunmasını sağlayan Tarjan algoritması incelenmiştir. Çalışmanın son bölümünde, yöntem basit bir veri seti üzerinde gösterilerek çalışma prensipleri örneklendirilmiştir.

Farklı uygulamalarla tez çalışmasının sonuçlarının gelecekte daha yararlı bir hale gelebileceğine inanılmaktadır.

KAYNAKÇA

- [1] Tarjan R.E. (1972), "**Depth-first search and linear graph algorithms**", SIAM Journal on Computing, s: 146–160
- [2] Brands, R. A. (2013). Cognitive social structures in social network research: A review. *Journal of Organizational Behavior*, 34(S1).
- [3] Knuth, "**The Stanford GraphBase**", s: 512–519.
- [4] Harrison, Knuth. "**Twenty Questions for Donald Knuth**"
- [5] Fienberg S.E., Meyer M. M., Wasserman S., (1985), "**Statistical Analysis of Multiple Socio-metric Relations**", Journal of the American Statistical Association 80, s:51-67
- [6] Katz N, Lazer D., Arrow H., ve Contractor N., (2004), "**Network Theory and Small Groups**", Small Group Research, vol: 35, no:3, s: 307-332.
- [7] Köksal, A., "**Bilişim Terimleri Sözlüğü**", Türk Dil Kurumu Yayınları, 1981, s:126.
- [8] Boy O., "**Sosyal ağlarda topluluk yapılarının analizi**", Yüksek Lisans Tezi, İnönü Üniversitesi, 2012
- [9] Leskovec J., Rajaraman A., Ullman J.D., "**Mining of Massive Datasets**" (2. baskı), Cambridge University Press, Stanford University California, 2014, s:343-345
- [10] Elektronik Sosyal Bilimler Dergisi www.e-sosder.com ISSN:1304-0278 Bahar-2007 C.6 S.20 (229-245)
- [11] Wikipedia, "Sociogram", <https://en.wikipedia.org/wiki/Sociogram>, 20.03.2017

[12] Skeels M.M., (2009), **“When Social Networks Cross Boundaries: A Case Study of Workplace Use of Facebook and LinkedIn”**

[13] Wikipedia, “PageRank”, <https://en.wikipedia.org/wiki/PageRank>, 04.04.2017

[14] A Global Study, **“İş Pazarı Bünyesinde Arz ve Talebin Eşleşmesinde Sosyal Medya Kullanımı İş Arayanların Türkiye’deki Algısı”**, Adecco Group, Türkiye, 2014

[15] Seker S.E., **“Çizge Teorisi (Graph Theory)”**, İstanbul Medeniyet Üniversitesi, 2015

[16] Hsiang S., **“Strongly Connected Components (Tarjan’s Algorithm)”**, 2015, s:1-5

[17] Kirac M., Özsoyoglu G., Yang J., (2006), **“Annotating proteins by mining protein interaction networks”**, Bioinformatics, 22 (14), 260-270

[18] Gupta P., Goel A., Lin J., Sharma A., Wang D., Zadeh R., (2013), **“WTF: The Who to Follow Service at Twitter”**, Proceedings of the 22nd international conference on World Wide Web, 505-514, Rio de Janeiro, Brazil, 13-17 May

ÖZGEÇMİŞ

13 Temmuz 1986 tarihi, İzmir İli Konak ilçesi doğumluyum. İlk ve Ortaokul eğitimimi Konak, Lise eğitimimi ise Güzelyalı ilçesinde tamamladıktan sonra Ege Üniversitesi Fen Fakültesi, Matematik Bölümüne kaydoldum. Bu bölümden 2010 yılında mezun olduktan sonra askerlik hizmetini tamamlayıp çalışmaya başladım. 2014 yılında Beykent Üniversitesi, Bilgisayar Mühendisliği Anabilim Dalında yüksek lisans eğitimime başladım. 2012 yılından bu yana çeşitli özel şirketlerde çalışma hayatıma devam etmekteyim.

Samet ESEN