

AKADEMİK YAZARLARIN YAYINLARI ARASINDAKİ İLİŞKİNİN SOSYAL AĞ BENZERLİK YÖNTEMLERİ İLE TESPİT EDİLMESİ

Furkan ÖZTEMİZ *^{ID}
Ali KARCI **^{ID}

Alınma: 28.02.2019; düzeltme: 12.11.2019; kabul: 28.02.2020

Öz: Sosyal ağlar günümüzde oldukça popüler bir konumda bulunmaktadır. İnsanlar tarafından yoğun olarak kullanılan bir platform halini almıştır. Bu durum yüksek miktarda veri üretimine neden olmaktadır. Bu verilerin anlamlı ve faydalı bir forma dönüştürebilmek için birçok yöntem geliştirilmiştir. Bu yöntemler arasında veri madenciliği teknikleri ilk sıralarda yer almaktadır. Bu çalışmada veri madenciliği ve sosyal ağ yöntemleri kullanılarak yazarların yayınlarında belirtmiş oldukları anahtar kelimelere göre ilgili yazarlar arasındaki çalışma alanı benzerlikleri tespit edilmiştir. Veri seti olarak IDAP 2018(International Conference on Artificial Intelligence and Data Processing) sempozyumunun yayın verileri kullanılmıştır. 536 yazar ve 1188 anahtar kelimedenden oluşan veri setine Jaccard, Euclidean, Cosine benzerlik yöntemleri uygulanmıştır ve çalışmalarına göre yazarların yayınları arasındaki benzerlikler analiz edilerek karşılaştırılmıştır. Yazarların sonraki yayınlarında birbirleri ile çalışma yapabilmeleri açısından yönlendirici sonuçlar elde edilmiştir. Verilerin analize uygun forma getirilmesi için SQL Server kullanılırken, analiz ve görsel öğelerin oluşturulması için ise, R dili ve R Studio IDE kullanılmıştır.

Anahtar Kelimeler: Benzerlik Yöntemleri, Jaccard, İbn-i Sina, Kosinüs, Çizge

The Determination Of The Relationship Between The Publications Of The Academic Authors By The Social Network Similarity Methods

Abstract: Social networks are nowadays very popular. It has become an intensively used platform by people. This status causes a high amount of data production. Many methods have been developed to transform these data into a meaningful and useful form. Data mining techniques are at the top of the list among these methods. In this study, data mining and social networking methods were used to determine the similarity of the study area between the related authors according to the key words that the authors mentioned in their publications. The publication data of the IDAP 2018 symposium was used as a data set. Jaccard, Euclidean, Cosine similarity methods were applied the data set consisted of 536 authors and 1188 keywords and similarities among the authors publications were compared through analyzed. In the subsequent publications of the authors, guiding results were obtained for them to work with each other. while SQL Server was used to convert data to the appropriate format for analysis, for the analysis and creation of visual elements, the R language and R Studio IDE were used.

Keywords: Similarity method, Jaccard, Euclidean, Cosine, Graph

* İnönü University, Department of Computer Engineering, Battalgazi, Malatya / Turkey

** İnönü University, Department of Computer Engineering, Battalgazi, Malatya / Turkey

İletişim Yazarı: Furkan ÖZTEMİZ(furkanoztemiz@gmail.com)

1. GİRİŞ

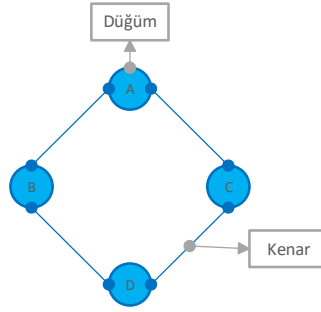
Son dönemlerde sosyal ağlar hayatımızın önemli bir parçası olmuştur. İnternet dünyasında oldukça hızlı yayılan bu akım ile insanlar ekonomi, eğlence, akademik, iş, eğitim vb.. birçok alanda kendilerini sosyal ağların içerisinde bulmaktadır. Oldukça popüler olarak kullanılan bu yapılar beraberinde devasa veriler üreten sistemler haline gelmiştir. Birçok firma bu verileri paraya dönüştürmenin yolları için çalışmıştır. Kullanıcıların ürünler, firmalar, kişiler, kurumlar, olaylar, politikalar vb.. bir çok konudaki eğilimlerini tespit edebilmek için sayısız yöntem ve teknoloji geliştirilmiştir. Verilerin daha sistemli depolanabilmesi ve hızlı sorgulanabilmesi için veritabanı yönetim sistemleri, Analiz işlemleri için Veri madenciliği, yapay zeka ve istatistiksel yöntemler vb.. , Programlama dili olarak R, Python vb.. , Analiz aracı olarak Matlab, Sas, Spss, Weka vb.. programlar geliştirilmiştir. Bu Çalışmayı özgün kılan değerlerden birisi daha önce hiçbir akademik çalışmada kullanılmamış olan IDAP veri setidir. Bu veri seti IEEE Xplore Dijital kütüphanesinde yayımlanmakta olup veri madenciliği ve yapay zeka alanında uluslararası çerçevede yapılan akademik çalışmalardan oluşmaktadır. Bu çalışmada IDAP 'ın 2018 yılına ait akademik verileri kullanılarak yazarlar arasındaki benzerlik bağlantıları sosyal ağ analiz yöntemleri ile tespit edilmiştir. Yaklaşık 550 adet yazarın hazırlamış olduğu yayınlarda bulunan 1188 anahtar kelime kaynak alınarak çizge yapısı oluşturulmuş ve yazarlar arasındaki benzerlik değerleri tespit edilmiştir. Çalışmada Jaccard, Euclid ve Cosine benzerlik değerleri ayrı ayrı tespit edilmiştir. Elde edilen değerler ışığında oransal olarak karşılaştırılmaları yapılmıştır. Çalışmada ayrıntılı olarak bir yazarın kullandığı anahtar kelimeler doğrultusunda diğer yazarlarla olan benzerlikleri ve belirlenen bir anahtar kelimeyi çalışmalarında belirten yazarların birbirleri arasındaki benzerlik değerleri tespit edilmiştir. Ayrıca yazarlar ve anahtar kelimeler arasındaki bağlantılara ait topluluk tespiti işlemleri gerçekleştirilmiştir. Elde edilen sonuçlar neticesinde bu konferansa katılan kişiler arasındaki bağlantının çıkarılması ile ortak çalışma yapabilmeleri için öneri sisteminin temelini oluşturacaktır.

Literatürde sosyal ağ analizi ile ilgili yapılan çalışmalardan bazıları; H.Demir ve F.Taktak hem kamu hem de özel sektörde oldukça yoğun kullanılan konumsal veri alışverişi incelenmiştir. Afyonkarahisar ili Merkez ilçede yapılan sayısal ve görsel analizlerle konumsal veriye ilişkin çalışmalarda en çok hangi kurumun ilgili olduğu, konumsal veri ile ilgili bilgilerin elde ediminde en çok hangi kurumla işbirliği yapıldığı, kurumlardan en çok hangisinin bilindiği, gibi soruların yanıtları bulunmaya çalışılmıştır. Sonuç olarak konumsal veri ile işi olan kurumlar arasındaki ilişkiler sosyal ağ analizi ile değerlendirilerek gözle görülmeyen sosyal ilişkiler ağı ortaya konulmuştur (Demir ve Taktak,2011). M.T.Argan online şikayetler ile ilgili önemli noktalara Sosyal Ağ Analizi bakışını ortaya koymuştur. Bu analiz, hizmetin satın alınması veya kullanımı ile ilgili sosyal ağda yer alan bireyin etkisiyle ilişkilidir. Sosyal ağ analizi, aynı zamanda sosyal ağlardaki tüketiciler arasındaki etkileşim biçimlerini ortaya koymak için dizayn edilir. Bu çalışma, Türkiye'deki mobil telefon markaları bakımından bir şikâyet sitesinin tüketiciler arasındaki şikâyet haritasının geliştirilmesini ortaya koymaktadır. Sosyal ağın geliştirilmesi ve görselleştirilmesinde Ucinet ve Netdraw kullanılmıştır. Araştırma sonuçları ürün değişimi, garanti, işletim sistemi, şarj, hoparlör, sinyal ve ürün iadesinin tüketici şikâyetlerinde önemli bağlantı noktaları olduğunu ortaya koymuştur (Argan,2014). N.Gürsakal, S.Tüzüntürk, F.S.Eteman tarafından NodeXL programı ile Twitter'dan elde edilen sosyal ağ verilerinin kuvvet yasası olasılık dağılımına uygun olup olmadığı test edilmiştir. Sonuçlar, ampirik verilerin kuvvet yasası olasılık dağılımına sahip olduğu gösterilmiştir (Gürsakal ve diğ.,2014). L.Sabah Susceptible-Infectious-Recovered/Removed (SIR) salgın modeli online sosyal ağlar üzerinde bir etkinin yayılımını modellemek için kullanılmıştır. Etki yayılımının dinamiklerini daha iyi kavramak için farklı merkezlik ölçütlerine göre bireyler seçilmiş; bu bireyler başlangıçta aktif hale getirilmiş ve yayılımlar gözlemlenerek detaylı analizler yapılmıştır (Sabah,2018). M. Kaya Keleş ve Ş. Aysel Özel tarafından metinler arasındaki uzaklık metrikleri karşılaştırılmış ve metinler arasındaki benzerlikler tespit edilmiştir. İntihal tespitinin temel aşaması olan bu uygulama da Cosine benzerlik yöntemi ile daha başarılı sonuçlar elde edilmiştir (Keleş ve

Özel,2017). K. İnce ve A.Karcı akademik işbirliğinin tespitine yönelik yaptıkları çalışmada 1980-2015 yılları arasında Web of Science veri tabanında kayıtlı Türkiye’deki kurum ve kuruluşlarda görev yapan yazarın bulunduğu çalışmalar dahil edilmiştir. Çalışmada Bron-Kerbosch (BK) algoritması temel alınarak İstatistiksel Bron-Kerbosch(İBK) algoritması geliştirilmiştir. Her iki algoritma kullanılarak topluluk keşifleri ve karşılaştırılması yapılmıştır. Elde edilen sonuçlar neticesinde Türkiye’deki akademik işbirliklerini etkileyen faktörler hakkında çıkarımlar yapılmıştır. (İnce ve Karcı,2019)

2. ÇİZGE TEORİSİ

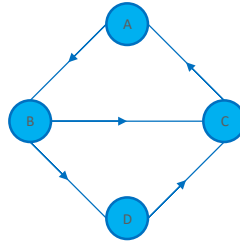
Çizge yapılarının öncelikli kullanım alanları ağ yapılarıdır. Karmaşık veri grupları arasındaki ilişkilerin gösterilmesinde büyük kolaylıklar sağlamaktadır. En kısa yol ve ağ analizi gibi popüler problemlerin çözümü için veri yapılarının önemli parçalarından birisidir. Temel olarak bir çizge Düğüm ve Kenar olmak üzere iki elemandan oluşmaktadır (Demircan,2009).



Şekil 1:
Genel çizge yapısı

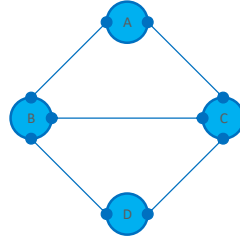
Şekil 1’de genel çizge yapısı verilmiştir. Çizge adını “Graph” kelimesinin baş harfinden almaktadır ve ‘G’ ile gösterilir ve tanıma bağı kalacak biçimde küme olarak $G = (D,K)$ şeklinde tanımlanır. Burada belirtilen D ifadesi düğümlerin kümesini simgelerken K ifadesi ise, düğümler arasındaki bağlantıları sağlayan doğru veya kenar kümesidir (Demircan,2009).

Çizgeler yönlerine göre Yönlü, Yönsüz ve Karışık olmak üzere 3’e ayrılabilir. Yönlü çizge: Şekil 2.’de görüldüğü üzere bütün kenarlar arasında tek yönlü bir bağlantı vardır.



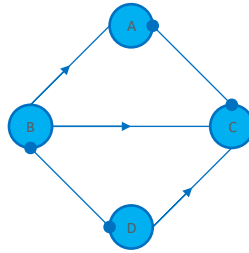
Şekil 2:
Yönlü çizge

Yönsüz çizge Şekil 3’de gösterildiği üzere çizgeyi oluşturan bütün kenarlar yönsüz ise bu çizgeye yönsüz çizge denmektedir.



Şekil 3:
Yönsüz çizge

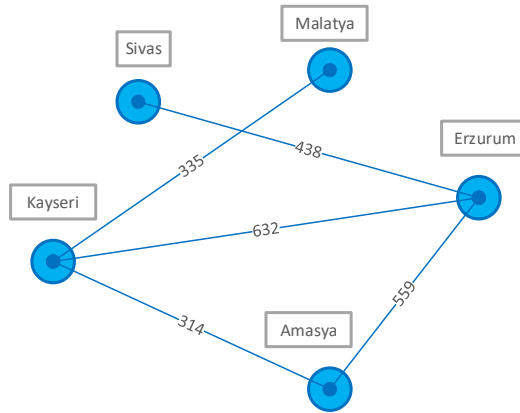
Eğer bir çizgeyi oluşturan kenarlardan bazıları yönlü ve bazıları ise yönsüz ise bu çizgeye karışık çizge denmektedir. Şekil 4.'de karışık çizgeye ait örnek verilmiştir.



Şekil 4:
Karışık çizge

2.1. Ağırlıklı Çizge

Çizge kenarları üzerinde bağlantının yoğunluğu, uzunluğunu, sağlamlığını belirten ifadeler olabilir. Bu tür çizgelere ağırlıklı veya maliyetli çizgeler denmektedir (İnce ve Karcı, 2019). Eğer çizge içerisindeki bütün kenarların maliyeti 1 veya birbirine eşitse bu çizge ağırlıklı çizge olarak adlandırılmaz. Bu çizgelerde yön bilgisi de belirtilmemişse basit çizge olarak isimlendirilir. Çizge üzerinde yapılan modellemeye göre çok çeşitli kullanım alanları mevcuttur. Bunlarda bazıları; Şehirler arasındaki uzaklıkların belirtilmesinde, Yönlendiriciler arası bant genişliklerinin gösteriminde, Nesnelere arasındaki uzaklıkların belirtilmesinde, bağlantılı nesnelere ilişkin kuvvetlerinin belirtilmesinde vb.. birçok uygulama alanı için kullanılabilir. Şekil 5'de iller arasındaki uzaklık değerleri verilen ağırlıklı çizgelere örnek bir yapı verilmiştir.



Şekil 5:
Ağırlıklı çizge

3. MATERYAL VE YÖNTEM

Uygulamadan yapılan bütün sayısal ve grafiksel analiz sonuçları için R programlama dili kullanılmıştır. Yazar benzerlik ve çizge modelleme işlemleri için standart R kütüphanelerine dâhil olmayan 3. Parti R kütüphaneleri projeye dâhil edilmiştir. İlerleyen bölümlerde bu kütüphaneler hakkında bilgi verilecektir.

3.1. R ve Kütüphaneler

R dili ilk 1991 senesinde Aucland Üniversitesi istatistik bölümünde Ross Ihaka ve Robert Gentleman tarafından geliştirilmiş ve 1993 yılında duyurulmuştur. R ismini yazarların isimlerinin baş harfinden almaktadır. R dili bütün işletim sistemlerinde çalışacak şekilde tasarlanmıştır. Açık kaynak bir yazılım olarak yayınlanan R dilinde basitten gelişmişe yönelik birçok analiz çözümü bulunmaktadır (R,2018). R'ın açık kaynak kodlu olması birçok firma (Microsoft, Oracle, SAP, IBM vb..) ve yazılımcı tarafından destek görmesine ve bunun bir sonucu olarak zengin içeriklere sahip 3.parti analiz kütüphanelerinin oluşmasını sağlamıştır (Cran,2018). Çok sayıda paket ve kütüphaneye erişim imkânı olan R kullanıcıları, aynı analiz işlemi için birden fazla farklı yönteme hazır olarak ulaşma imkânına sahiptirler.

Yayın çalışmasında kullanılan 3.parti kütüphanelerden bazıları;

Ggdendro	: Veriyi dendogram ve ağaç diagramlarına dönüştürmek için kullanılmıştır.
Dplyr	: Veri işlemeyi kolaylaştıran bir pakettir.
İgraph	: Graf işlemleri yapabileceğiniz ve görselleştirebileceğiniz bir kütüphane.
Network3d	: Grafın 2 boyutlu olarak modellenmesinde kullanılmıştır.
Treejs	: Grafın 3 boyutlu olarak modellenmesinde kullanılmıştır.

Geliştirme ortamı olarak ise R studio ide kullanılmıştır. Paket ve kütüphanelerin kolay bir şekilde yüklenebilmesi ve ara yüzünün sadeliğinden dolayı tercih edilmiştir.

3.2. Benzerlik Yöntemleri

Bilgisayar üzerinde veriler arasındaki ilişkilerin tespit edilebilmesi için matematiksel formları arasındaki bağlantıların ortaya çıkarılması gerekmektedir. Bunun için birçok benzerlik ölçüsü oluşturan algoritma geliştirilmiştir. Bu çalışmada yazarlar arasındaki benzerliklerin tespit edilmesi için anahtar kelimelere Jaccard, İbn-i Sina (Öklid), Kosinüs (Cosine) benzerlik yöntemleri uygulanmıştır.

3.2.1.Jaccard

Metinler arasındaki ilişkinin ölçülmesi amacıyla geliştirilen başarılı bir benzerlik çıkarma yöntemidir. Kısaca indeks iki metin arasındaki ortak özelliklerin çıkarımlarını yaptıktan sonra bu özelliklerin toplam özellik sayısına bölünmesi ile elde edilir(Jaccard index,2013). Aşağıdaki denklemde Jaccard benzerlik indeksinin formülü denklem 1 de verilmiştir (Dev veri,2014).

$$\text{Benzerlik}(A, B) = |A \cap B| / |A \cup B| \quad (1)$$

Jaccard benzerlik indeksi 0 ile 1 arasında hesaplanan bir değerdir. Bu değer büyüdükçe benzerlik oranının yükseldiğini belirtmektedir. Benzerlik oranı yükseldiğinde benzerlik mesafesi kısılacığından dolayı Jaccard mesafesi, benzerlik indeks değerinin 1 den çıkarılması ile elde edilmektedir. Denklem 2'de Jaccard benzerlik mesafesi formülü verilmiştir (Math overflow,2010).

$$J_{\text{delta}}(A, B) = 1 - J(A, B) \quad (2)$$

Bir örnek ile açıklayacak olursak; Furkan ve Murat iki yazar olsunlar. Bu yazarların çalıştıkları alana ait bilgiler aşağıdaki gibidir.

Furkan : (Büyük veri, Veri madenciliği, Yapay zekâ, Görüntü işleme)

Murat : (Veri madenciliği, Yapay zekâ, Çizge analiz, Bulut bilişim, Görüntü işleme)

İki yazara ait bütün terimlerimize baktığımızda (Büyük veri, Veri madenciliği, Yapay zekâ, Çizge analiz, Bulut bilişim, Görüntü işleme) ifadelerinden oluşur.

Furkan ve Murat isimli yazarların ortak çalıştıkları konu sayısı = (Veri madenciliği, Yapay zekâ, Görüntü işleme) = 3

Yazarların çalıştıkları toplam konu sayısı = (Büyük veri, Veri madenciliği, Yapay zekâ, Çizge analiz, Bulut bilişim, Görüntü işleme) = 6 olarak tespit edilmiştir. Formülü uygularsak $Jaccard(A,B)=|A \cap B|/|A \cup B| = 3/6 = 0,5$ olarak iki yazar arasındaki benzerlik oranı bulunmuştur. Jaccard mesafe ölçüsü ise benzerlikle ters orantılı olduğu için Denklem-2 deki formül ile $1-0,5 = 0,5$ olarak bulunmaktadır.

3.2.2.İbni Sina (Öklid) Uzaklığı

İbn-i Sina mesafe yöntemi çok yaygın olarak kullanılmaktadır. Veri yoğun ve sürekli olduğunda bu en yakın yaklaşımdır. İki nokta arasındaki İbn-i Sina uzaklığı iki noktayı birbirine bağlayan noktadır. Pisagor teoremi bu iki noktayı vermektedir. Denklem 3 de formül verilmiştir (Data aspirant,2015).

$$İbn - i Sina(A, B) = \sqrt{\sum_{i=1}^k (A_i - B_i)^2} \quad (3)$$

3.2.3.Kosinüs (Cosine)

İki metin arasındaki mesafeyi trigonometrideki kosinüs fonksiyonu üzerinden formülize etmeyi amaçlamıştır (Cosine similarty,2012). Metinlerin birer vektör olarak düşünüldüğü bu yaklaşımda iki vektörün birbiri ile ilişkisi bir açı ile ifade edilmektedir. Aynı yönü gösteren iki vektör için kosinüs değeri 1 olacaktır ($\cos(0)=1$). Tamamen birbiri ile ilişkisiz olan vektörler için ise kosinüs değeri 0 olacaktır ($\cos(90) = 0$. Dik vektörler (orthogonal vectors)). Son olarak tamamen birbirini zıddı olan dokümanlar için, ki bu dokümanlar için aralarında ilişki yoktur denemez, kosinüs değeri -1 olacaktır. Bu şekilde iki metindeki bütün değerler +1 ve -1 arasında sayısallaştıracaktır (Cosine similarty,2012). Denklem 4’de kosinüs benzerlik indeksinin formülü verilmiştir (Plerou ve diğ.,2014).

$$Cosine(A, B) = \cos(\theta) = \frac{A * B}{||A|| ||B||} \quad (4)$$

Metinler arasındaki benzerliklerin tespitinde daha anlaşılır olması için kosinüs benzerlik yöntemini bir örnek ile açıklayacak olursak; Furkan ve Murat iki yazar olsunlar. Bu yazarların çalıştıkları alana ait bilgiler aşağıdaki gibidir.

Furkan : (Büyük veri, Veri madenciliği, Yapay zekâ, Görüntü işleme)

Murat : (Veri madenciliği, Yapay zekâ, Çizge analiz, Bulut bilişim, Görüntü işleme)

İki yazara ait bütün terimlerimize baktığımızda (Büyük veri, Veri madenciliği, Yapay zekâ, Çizge analiz, Bulut bilişim, Görüntü işleme) ifadelerinden oluşur. Dokümanlarda bu ifadelerin kaç kere geçtiğini bulacak olursak;

Furkan : [1,1,1,0,0,1]

Murat : [0,1,1,1,1,1]

şeklinde ifade edilebilir. Formül üzerinden bu iki vektör için kosinüs değerini hesaplanırsa.

$A*B = [1,1,1,0,0,1].[0,1,1,1,1,1] = (1)*(0) + (1)*(1) + (1)*(1) + (0)*(1) + (0)*(1) + (1)*(1) = 3$ sonucu elde edilir. Şimdi vektörlerin boyutlarını hesaplanabilir. Bunun için İbn-i Sina uzayından (Euclidean space) faydalanılır ve her boyuttaki değerlerinin karelerini toplayarak toplamın karekökü alınır.

$$\|A\| = \sqrt{(1)^2 + (1)^2 + (1)^2 + (0)^2 + (0)^2 + (1)^2} = 2$$

$$\|B\| = \sqrt{(0)^2 + (1)^2 + (1)^2 + (1)^2 + (1)^2 + (1)^2} = 2.2360679775$$

$$\text{Cosin}(A,B) = (A*B) / (\|A\| \|B\|) = 3 / (2 \times 2.2360679775) = 3 / 4.472135955 = 0.67082039325$$

İki yazarın ilgi alanlarını baz alarak kosinüs benzerliğini bulunmuştur.

4. VERİ SETİ VE VERİ DÖNÜŞÜMÜ

Çalışmada 306 adet bildiriye ait 536 adet yazar ve bu yazarların yayınlarında kullandıkları 1188 adet anahtar kelime kullanılmıştır. Bu verilerden her yazarın yayınlarında kullandıkları anahtar kelimeler filtrelenerek veri tabanına aktarılmış ve benzerlik işlemlerinde kullanılmak üzere matris formuna getirilmiştir. Matris formu satırlar yazarları, sütunlar ise anahtar kelimeleri ifade edecek şekilde oluşturulmuştur.

4.1. IDAP 2018 Veri Seti

IDAP 2018 yılına ait örnek veri setinin bir kısmı Tablo 1’de belirtilmiştir. Tablo 2’de ise matris formuna dönüştürmek için verinin farklı bir formu verilmiştir.

Tablo 1. IDAP 2018 örnek veri seti

Yazar ADI	Anahtar Kelime 1	Anahtar Kelime 2	Anahtar Kelime 3
ALİ KARCI	KIYASLAMA FONKSİYONLARI	OPTİMİZASYON	KSO
ALİ KARCI	SOCİAL NETWORKS	LİNK PREDİCTION	NULL
ALİ KARCI	SU DÖNGÜSÜ ALGORİTMASI	ELEKTROMANYETİK ALAN OPTİMİZASYONU	BÜYÜK PATLAMA-BÜYÜK BÜZÜLME ALGORİTMASI
CELALEDDİN YEROĞLU	KARINCA KOLONİSİ ALGORİTMASI	OPTİMİZASYON ALGORİTMALARI	GEZGİN SATICI PROBLEMİ
CELALEDDİN YEROĞLU	PIDA DENETÇİ	PARÇACIK SÜRÜ OPTİMİZASYONU	BİRİM BASAMAK TEPKİSİ
DAVUT HANBAY	APRİCOT DİSEASE DETECTION	CONVOLUTIONAL NEURAL NETWORKS	ALEXNET MODEL
DAVUT HANBAY	PIDA DENETÇİ	PARÇACIK SÜRÜ OPTİMİZASYONU	BİRİM BASAMAK TEPKİSİ
FURKAN ÖZTEMİZ	KARINCA KOLONİSİ ALGORİTMASI	OPTİMİZASYON ALGORİTMALARI	GEZGİN SATICI PROBLEMİ
FURKAN ÖZTEMİZ	SU DÖNGÜSÜ ALGORİTMASI	ELEKTROMANYETİK ALAN OPTİMİZASYONU	BÜYÜK PATLAMA-BÜYÜK BÜZÜLME ALGORİTMASI
GALİP AYDIN	DEEP LEARNING	TEXT PROCESİNG	TEXT ANALYSİS
GALİP AYDIN	LİCENSE PLATE RECOGNİTION	DEEP LEARNING	TENSORFLOW
GALİP AYDIN	TEXT CLASSİFİCATION	DEEP LEARNING	CONVOLUTIONAL NEURAL NETWORKS
MUHAMMED FATİH TALU	KUMAŞ HATASI TESPİTİ	BANT GEÇİREN FİLTRE	GİRDAP OPTİMİZASYON ALGORİTMASI
MUHAMMED FATİH TALU	SEGMENTATION	THRESHOLD	REGION GROWİNG

Tablo 2. Veri seti farklı form

Yazar	Anahtar Kelimeler
ALİ KARCI	ABC
ALİ KARCI	ARTİFİCİAL İNTELLİGENCE
ALİ KARCI	BÜYÜK VERİ
ALİ KARCI	COMPLEX NETWORKS
ALİ KARCI	DOĞAL DİL İŞLEME
ALİ KARCI	DOKÜMAN BENZERLİĞİ
ALİ KARCI	ENTROPY
ALİ KARCI	GA
ALİ KARCI	KIYASLAMA FONKSİYONLARI
ALİ KARCI	SOCİAL NETWORKS
ALİ KARCI	SU DÖNGÜSÜ ALGORİTMASI
ALİ KARCI	PSO
FURKAN ÖZTEMİZ	KARINCA KOLONİSİ ALGORİTMASI
FURKAN ÖZTEMİZ	SU DÖNGÜSÜ ALGORİTMASI
FURKAN ÖZTEMİZ	OPTİMİZASYON ALGORİTMALARI
FURKAN ÖZTEMİZ	ELEKTROMANYETİK ALAN OPTİMİZASYONU
FURKAN ÖZTEMİZ	GEZGİN SATICI PROBLEMİ
FURKAN ÖZTEMİZ	BÜYÜK PATLAMA-BÜYÜK BÜZÜLME ALGORİTMASI
FURKAN ÖZTEMİZ	YERÇEKİMİ ARAMA ALGORİTMASI

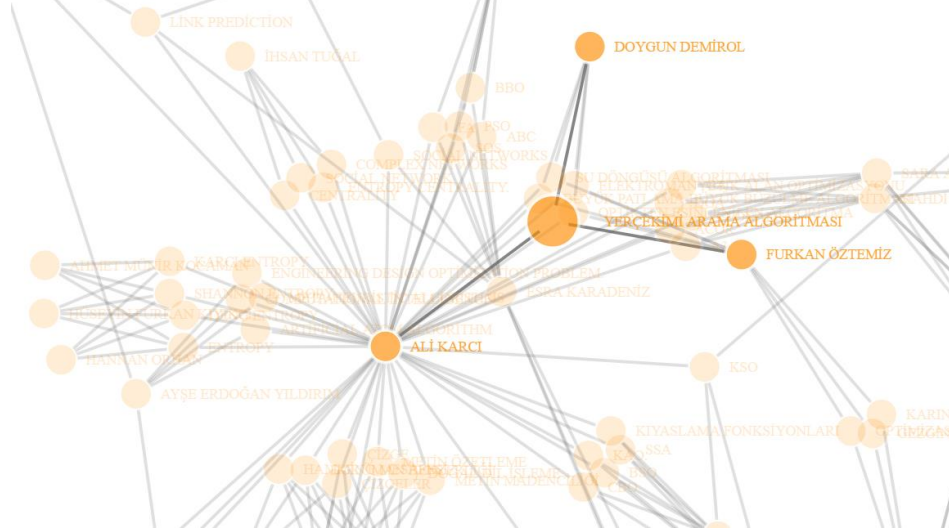
4.2. Matris Formuna Dönüşüm

Yazar ve konular arasındaki bağlantı yoğunluklarının tespit edilmesi için veriler matris formuna dönüştürülmüştür. Yazarlar satırlarda belirtilirken kolon değerlerinde ise anahtar kelimeler verilmiştir. Matris 536 x 1188 boyutunda olup toplam 636.768 adet veri içermektedir. Tablo 3’de matris yapısının bir örneği verilmiştir.

Tablo 3. Yazar-anahtar kelime matris yapısı

Yazar-Anahtar Kelimeler Matrisi	5G	CNN	DEEP LEARNING	DEEP NEURAL NETWORK	DİSTRİBÜTED ALGORİTHMS	GÖRÜNTÜ İŞLEME	MAİZE	RASTGELE SAYILAR	RF
ADNAN FATİH KOCAMAZ	0	0	0	0	0	0	2	0	0
AHMET BEDRİ ÖZER	0	0	0	0	0	0	0	2	0
ASO KHALEEL AMEEN	0	0	1	0	0	0	0	0	0
AYAD ABDUL-MALEK	0	0	1	0	0	0	0	0	0
AYHAN AKBAL	2	0	0	0	0	0	0	0	2
AYŞE ELDEM	0	0	1	2	0	0	0	0	0
BETÜL AY	0	1	2	0	0	0	0	0	0
BİRCAN KAMIŞLIOĞLU	2	0	0	0	0	0	0	0	2
BUKET KAYA	0	0	1	0	0	0	0	0	0
BURHAN ERGEN	0	2	1	0	0	0	0	0	0
ERHAN ERGÜN	0	0	0	0	0	3	0	0	0
GALİP AYDIN	0	1	3	0	0	0	0	0	0
ÖZKAN ARAPOĞLU	0	0	0	0	3	0	0	0	0
PELİN ALTINIŞIK	0	0	0	0	0	3	0	0	0

durum 3 yazarın ayrı ayrı veya ortak olarak yaptıkları en az 1 çalışmada Yerçekimi arama algoritmasını anahtar kelime olarak belirtmiş olduğunu ifade etmektedir.



Şekil 9:

Graf içerisinde yazarlar arasındaki bağlantılar

5. BENZERLİK ANALİZİ

Analiz sürecinde ilk olarak veriler uygun forma getirilmiştir. Yazarların çalıştıkları konuları temsil eden anahtar kelimeler üzerindeki ağırlık değerlerini tespit edebilmek için veriler toplanarak matris formatına getirilmiştir. Bu işlem T-SQL komutları ile yapılmıştır. Kısa bir kod bloğu Tablo 4’de verilmiştir. Bu kod ile Anahtar kelimeler kolon isimleri olarak yazar isimleri ise satır isimleri olarak atanmıştır ve hangi yazar hangi anahtar sözcük le ilişkili olduğu 536*1188 matris olarak hazırlanmıştır.

Tablo 4. Matris forma dönüştürme sql kodu

```
select * from(select [Kisi] as Yazar,[Kisi],[Key] from [NetworkPredict].[dbo].[Sayfa1$])  
as data  
PIVOT ( count([Kisi]) FOR [Key] IN ([ 4. SANAYİ DEVRİMİ], [ DATA MİNNİNG],  
[DRONE SİMULATION.], [ GÜVENİLİRLİK.] ***** bütün anahtar kelimeler  
) ) as PivotTable
```

Elde edilen bu verilere Jaccard, İbn-i Sina ve Kosinüs benzerlik yöntemleri uygulanmıştır. 3 benzerlik yöntemi için de İnönü Üniversitesi Bilgisayar Mühendisliğinde (İBM) görev yapan ve IDAP 2018 konferansına bildiri gönderen akademisyenlere ait benzerlik verileri belirtilmiştir. Bu veriler ışığında sonraki aşamada tek bir akademisyen hedef alınarak Kosinüs ve Jaccard benzerlik değerlerinin karşılaştırılması yapılmıştır. Ayrıca yayınlarında belirli bir anahtar sözcüğü belirten yazarlar arasındaki benzerlik değerleri tespit edilmiştir.

5.1. Jaccard Benzerlik Sonuçları

İBM’de görev yapan akademisyenlerin IDAP konferansına göre birbirlerine olan Jaccard benzerlik değerleri Tablo 5’de verilmiştir. Bu veriler yorumlandığında Adnan Fatih Kocamaz’ın bölümdeki akademisyenlerle bir çalışma yapmadığı görülmektedir. En Yüksek oran ise

Muhammed Fatih Talu –Celaleddin Yeroğlu ve Muhammed Fatih Talu – Sara Altun arasında 0,33 lük bir oranla gerçekleşmiştir. 0,33 değerini %33 benzerlik şeklinde yorumlamak yanlış olmayacaktır. Bu sonuçlar ışığında bu yazarların yaptıkları bildirilerde kullandıkları anahtar kelimelerin %33 aynıdır denilebilir.

Tablo 5. Jaccard benzerlik sonuçları

Jaccard Benzerlik Skoru	Adnan Fatih KOCAMAZ	Ali KARCI	B.Baykant ALAGÖZ	Bilal SENOL	Celaleddin YEROGLU	Davut HANBAY	Muhammed Fatih TALU	Sara ALTUN
Adnan Fatih KOCAMAZ	1	0	0	0	0	0	0	0
Ali KARCI	0	1	0	0	0	0	0	0,10
B.Baykant ALAGÖZ	0	0	1	0,32	0	0	0	0
Bilal SENOL	0	0	0,32	1	0	0	0	0
Celaleddin YEROGLU	0	0	0	0	1	0,15	0,33	0,15
Davut HANBAY	0	0	0	0	0,15	1	0	0
Muhammed FATİH TALU	0	0	0	0	0,33	0	1	0,33
Sara ALTUN	0	0	0	0	0,15	0	0,33	1

5.2. İbn-i Sina Mesafe Benzerlik Sonuçları

Öklid mesafe yöntemi ise Jaccard ve Kosinüs yöntemlerinin aksine benzerlik oranı büyük olan yazarlar arasındaki benzerlik mesafe değeri küçük olmaktadır. Tablo 6’da belirtildiği üzere veriler incelendiğinde, benzerlik değeri yüksek olan yani en kısa mesafe değerleri Muhammed Fatih Talu – Sara Altun arasında 2,8 birim olduğu görülmektedir. Muhammed Fatih Talu-Celaleddin Yeroğlu yazarları arasındaki değer ise 3,8 birim olarak gerçekleşmiştir. Bu tabloya göre birbirine çalışma alanı olarak en uzak olan yazarlar 9,6 birim ile Ali Karcı - Adnan Fatih Kocamaz olarak görülmektedir.

Tablo 6. İbn-i Sina benzerlik sonuçları

İbn-i Sina Benzerlik Skoru	Adnan Fatih KOCAMAZ	Ali KARCI	B.Baykant ALAGÖZ	Bilal SENOL	Celaleddin YEROGLU	Davut HANBAY	Muhammed Fatih TALU	Sara ALTUN
Adnan Fatih KOCAMAZ	0	9,6	7,6	7,2	8	7,2	7,1	7,2
Ali KARCI	9,6	0	8	7,5	8,3	7,6	7,5	6,80
B.Baykant ALAGÖZ	7,7	8	0	3,6	6	5	4,8	5
Bilal SENOL	7,2	7,5	3,6	0	5,5	4,3	4	4,2
Celaleddin YEROGLU	8	8,3	6	5,5	0	4,7	3,70	4,70
Davut HANBAY	7,2	7,5	5	4,3	4,7	0	4	4,2
Muhammed FATİH TALU	7,1	7,4	4,8	4	3,8	4	0	2,80
Sara ALTUN	7,2	6,9	5	4,3	4,7	4,3	2,80	0

5.3. Kosinüs Benzerlik Sonuçları

Tablo 7’de belirtildiği üzere Kosinüs benzerlik değerlerine göre 0,58 birimlik değer ile yaptığı çalışmalarda en çok benzerlik bulunan yazar ikilisi Muhammed Fatih Talu- Celaleddin Yeroğlu

olarak tespit edilmiştir. İkinci sırada ise 0,5 birimlik değeri ile Muhammet Fatih Talu-Sara Altun ve Bilal Şenol – Barış Baykant Alagöz isimli yazarlar bulunmaktadır.

Tablo 7. Kosinüs benzerlik sonuçları

Kosinüs Benzerlik Skoru	Adnan Fatih KOCAMAZ	Ali KARCI	B.Baykant ALAGÖZ	Bilal SENOL	Celaleddin YEROGLU	Davut HANBAY	Muhammed Fatih TALU	Sara ALTUN
Adnan Fatih KOCAMAZ	0	0	0	0	0	0	0	0
Ali KARCI	0	0	0	0	0	0	0	0,24
B. Baykant ALAGÖZ	0	0	0	0,5	0	0	0	0
Bilal SENOL	0	0	0,5	0	0	0	0	0
Celaleddin YEROGLU	0	0	0	0	0	0,29	0,58	0,29
Davut HANBAY	0	0	0	0	0,29	0	0	0
Muhammed Fatih TALU	0	0	0	0	0,58	0	0	0,50
Sara ALTUN	0	0	0	0	0,29	0	0,50	0

Jaccard, İbn-i Sina ve Kosinüs benzerlik yöntemine ait temel işlemleri gerçekleştiren R kodu Tablo 8’de verilmiştir.

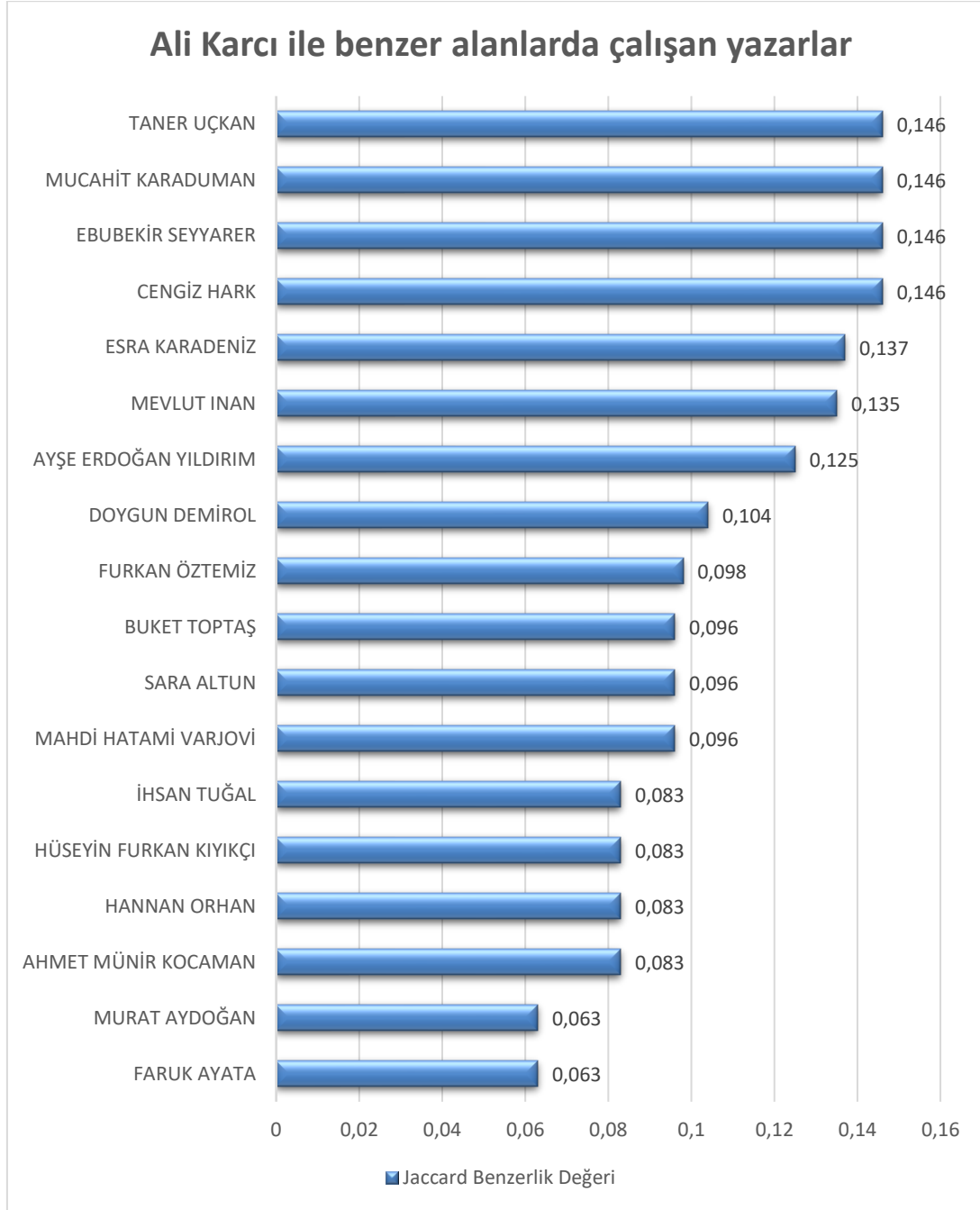
Tablo 8. Benzerlik yöntemine ait kısa kod

```
pairs <- t(combn(1:nrow(A), 2)) # Jaccard benzerlik skoru
for (i in 1:nrow(pairs)){
  num <- sum(sapply(1:ncol(A), function(x)(min(A[pairs[i,1],x],A[pairs[i,2],x]))))
  den <- sum(sapply(1:ncol(A), function(x)(max(A[pairs[i,1],x],A[pairs[i,2],x]))))
  sim.jac[pairs[i,1],pairs[i,2]] <- num/den
  sim.jac[pairs[i,2],pairs[i,1]] <- num/den
}
sim.jac[which(is.na(sim.jac))] <- 0
diag(sim.jac) <- 1
dist.jac <- as.dist(1-sim.jac)
#ağırlıklandırılmış euclidean benzerlik(mesafesi)
dist.eu <- as.matrix(dist(A, method = "euclidean"))
#ağırlıklandırılmış cosine skoru
dist.cos <- cosine(t(A))
diag(dist.cos) <- 0
```

5.4. Jaccard, Kosinüs ve İbn-i Sina Benzerlik Karşılaştırılması

Tablo 9’da İBM’de çalışan akademisyenlerden Ali Karıcı’nın yapmış olduğu yayınlara göre diğer yazarlar arasındaki en yüksek benzerlik değerine sahip 18 yazara ait benzerlik değerleri verilmiştir. Bu veriler incelendiğinde Ali Karıcı ile en çok benzerlik gösteren yazarlar Jaccard için 0,146 ve Kosinüs için 0,382 değerleri ile Cengiz Hark, Ebubekir Seyyarer, Mücahit Karaduman ve Taner Uçkan’dır. Tablo 9 incelendiğinde değerler farklı olmasına rağmen yazar sıralaması Jaccard ve Kosinüs için büyük ölçüde benzerlik göstermektedir.

Tablo 9. Bir yazara ait Jaccard ve Kosinüs benzerlik sonuçları



Tablo 10'da çalışmalarında Data Mining anahtar kelimesini en az bir kere kullanan(9 yazar rastgele seçilmiştir.) yazarların birbirleri arasındaki Jaccard benzerlik değerleri verilmiştir.

Tablo 10. Jaccard Benzerlik Değeri

Jaccard Benzerlik (Data mining)	Hamad	A.Emre KELEŞ	Burhan ERGEN	Erdiñç AVAROĞLU	Eugene C. EZIN	Hasan YETİS	Josky AIZAN	Ümit KILIÇ	Yaşar SAFALI
Hamad	1,00	0,10	0,05	0,06	0,11	0,07	0,11	0,10	0,11
A.Emre KELEŞ	0,10	1,00	0,06	0,07	0,13	0,08	0,13	1,00	0,13
Burhan ERGEN	0,05	0,06	1,00	0,19	0,06	0,05	0,06	0,06	0,29
Erdiñç AVAROĞLU	0,06	0,07	0,19	1,00	0,07	0,05	0,07	0,07	0,36
Eugene C. EZIN	0,11	0,13	0,06	0,07	1,00	0,08	1,00	0,13	0,14
Hasan YETİS	0,07	0,08	0,05	0,05	0,08	1,00	0,08	0,08	0,08
Josky AIZAN	0,11	0,13	0,06	0,07	1,00	0,08	1,00	0,13	0,14
Ümit KILIÇ	0,10	1,00	0,06	0,07	0,13	0,08	0,13	1,00	0,13
Yaşar SAFALI	0,11	0,13	0,29	0,36	0,14	0,08	0,14	0,13	1,00

Tablo 11’de çalışmalarında Data Mining anahtar kelimesini en az bir kere kullanan(9 yazar rastgele seçilmiştir.) yazarların birbirleri arasındaki İbn-i Sina mesafe benzerlik değerleri verilmiştir.

Tablo 11. İbn-i Sina mesafe benzerlik değeri

İbn-i Sina mesafe (Data Mining)	Hamad	A. Emre KELEŞ	Burhan ERGEN	Erdiñç AVAROĞLU	Eugene C. EZIN	Hasan YETİS	Josky AIZAN	Ümit KILIÇ	Yaşar SAFALI
Hamad	0,00	3,00	4,47	3,87	2,82	3,60	2,82	3,00	2,82
A. Emre KELEŞ	3,00	0,00	4,35	3,74	2,64	3,46	2,64	0,00	2,64
Burhan ERGEN	4,47	4,35	0,00	4,35	4,24	4,79	4,24	4,35	3,46
Erdiñç AVAROĞLU	3,87	3,74	4,35	0,00	3,60	4,24	3,60	3,74	2,64
Eugene C. EZIN	2,82	2,64	4,24	3,60	0,00	3,31	0,00	2,64	2,44
Hasan YETİS	3,60	3,46	4,79	4,24	3,31	0,00	3,31	3,46	3,31
Josky AIZAN	2,82	2,64	4,24	3,60	0,00	3,31	0,00	2,64	2,44
Ümit KILIÇ	3,00	0,00	4,35	3,74	2,64	3,46	2,64	0,00	2,64
Yaşar SAFALI	2,82	2,64	3,46	2,64	2,44	3,31	2,44	2,64	0,00

Tablo 12’de çalışmalarında Data Mining anahtar kelimesini en az bir kere kullanan(9 yazar rastgele seçilmiştir.) yazarların birbirleri arasındaki Kosinüs benzerlik değerleri verilmiştir.

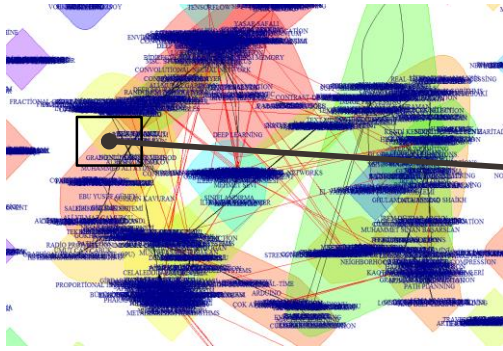
Tablo 12. Kosinüs benzerlik değerleri

Kosinüs Benzerlik (Data Mining)	Hamad	A. Emre KELEŞ	Burhan ERGEN	Erdiñç AVAROĞLU	Eugene C. EZIN	Hasan YETİS	Josky AIZAN	Ümit KILIÇ	Yaşar SAFALI
Hamad	0,00	0,18	0,10	0,12	0,20	0,14	0,20	0,18	0,20
A. Emre KELEŞ	0,18	0,00	0,11	0,14	0,22	0,15	0,22	1,00	0,22
Burhan ERGEN	0,10	0,11	0,00	0,30	0,13	0,08	0,13	0,11	0,50
Erdiñç AVAROĞLU	0,12	0,14	0,30	0,00	0,15	0,10	0,15	0,14	0,60
Eugene C. EZIN	0,20	0,22	0,13	0,15	0,00	0,17	1,00	0,22	0,25
Hasan YETİS	0,14	0,15	0,08	0,10	0,17	0,00	0,17	0,15	0,17
Josky AIZAN	0,20	0,22	0,13	0,15	1,00	0,17	0,00	0,22	0,25
Ümit KILIÇ	0,18	1,00	0,11	0,14	0,22	0,15	0,22	0,00	0,22
Yaşar SAFALI	0,20	0,22	0,50	0,60	0,25	0,17	0,25	0,22	0,00

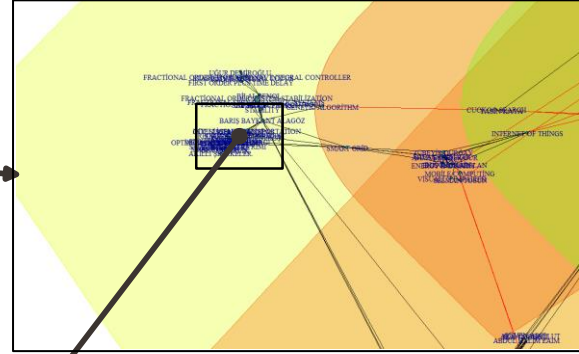
5.5. Topluluk Tespiti (Cominity Detection)

Yazar ve anahtar kelimelerin kullanım yoğunluklarına göre baskın olan gruba dâhil olabilir veya kendi baskın bir özellik gösteriyorsa kendisi bir küme oluşturabilir. Yazar ve anahtar kelimelerin birlikte analiz edildiği çalışmada bağlantı yoğunluklarına göre baskın olan yazar ve anahtar kelimeler kendi etraflarında küme oluşturmaktadırlar. Şekil 10a'da benzer çalışmalarda bulunan yazarların hangi topluluğa ait olduğu belirtilmiştir. Şekil 10a 'da bütün veri setine ait görsel verilmiştir ve tek resimde okunabilirliği zayıf olduğu için yakınlaştırma işlemi uygulanmıştır. Şekil 10b'de ve Şekil 10c'de ise sırası ile 1 ve 2 aşama yakınlaştırma yapılarak verilerin okunabilirliği artırılmıştır. Aynı renk üzerinde bulunan yazarlar incelendiğinde çalışmalarında aynı anahtar kelimeyi kullandıkları için aynı küme içerisinde ifade edildiği görülmektedir.

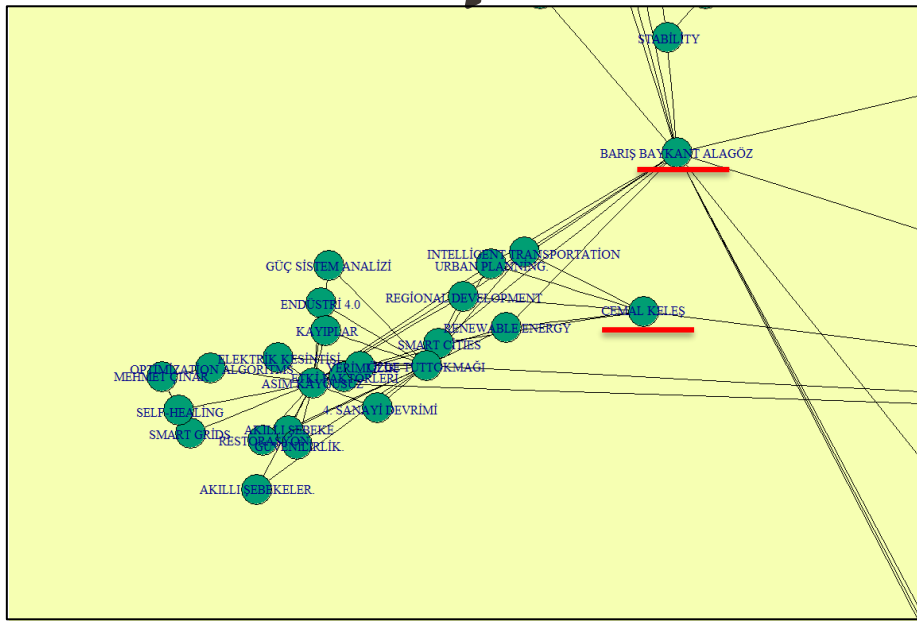
Şekil 10c incelendiğinde Barış Baykant Alagöz ve Cemal Keleş IDAP 2018 sempozyumunda yapmış oldukları yayınlara göre aynı küme içerisinde yer almışlardır. Yazarların aynı küme içerisinde yer almaları çalışma alanları arasındaki benzerliği simgelemektedir. İlgili yazarlar kırmızı çizgi ile altı çizilerek gösterilmiştir. Ayrıca Şekil 10c'de görüldüğü üzere yazarlar çalışmalarında ortak anahtar kelimeler kullanmıştır. Bu anahtar kelimelerden bazıları; Renewable Energy, Intelligent Transportation, Smart Cities oluşturmaktadır. Topluluk tespiti çalışmasında igraph kütüphanesinde tanımlı Cluster Edge Betweenness fonksiyonu kullanılmıştır.



Şekil 10a:
Topluluk tespiti çizge yapısı



Şekil 10b:
1. aşama yakınlaştırma



Şekil 10c:
2. aşama yakınlaştırma

6. TARTIŞMA VE SONUÇ

Yapılan bu yayın çalışmasında daha önce hiçbir akademik çalışmada kullanılmayan IDAP sempozyumuna ait gerçek akademik veriler üzerinden yazarların yapmış oldukları bildirilerdeki anahtar kelimeler kullanılarak yazarlar arasındaki benzerlik bağlantıları tespit edilmiştir. Veriler ön işleme sürecinden geçirilerek matris formuna dönüştürülmüştür. Ağırlıklı çizge yapısı oluşturulmuştur. Benzerlik işlemleri R programlama dilinde Jaccard, Kosinüs ve İbn-i Sina benzerlik yöntemleri kullanılarak gerçekleştirilmiştir. Belirlenen bir yazar ve bir anahtar kelime için ayrı ayrı olarak üç benzerlik yöntemi de uygulanmış ve karşılaştırması yapılmıştır. Değerler incelendiğinde Jaccard yönteminin Kosinüs ve İbn-i Sina yöntemlerine göre daha başarılı sonuçlar verdiği görülmektedir. Yazarlar ve anahtar kelimeler arasındaki bağlantıların tespit edilmesi ile kenar arasındalık (edge betweenness) yöntemi kullanılarak topluluk tespiti işlemi yapılmıştır. Bu şekilde yazarların anahtar kelimeler üzerindeki etki alanları tespit edilmiştir. Sosyal ağ analiz yöntemleri kullanılarak gerçekleştirilen bu yayın çalışmasında ağırlıklı çizge ağları üzerinde benzerlik yöntemleri başarılı bir şekilde gerçekleştirilmiştir ve sonraki analiz süreçleri için temel oluşturacak alt yapıyı sunmaktadır.

KAYNAKLAR

1. Argan, M.T. (2014) E-Şikâyetle İlgili Tanımlayıcı Bir Çalışma: Şikâyet Forumu Olarak Bir Web Sitesinin Sosyal Ağ Analizi, *İnternet Uygulamaları ve Yönetimi Dergisi*, 2014/5 DOI: 10.5505/iuyd.2014.69775
2. Cosine Similarity(2012),
Erişim Adresi: <http://bilgisayarkavramlari.sadievrenseker.com/2012/11/08/kosinus-benzerligi-cosine-similarity-2/> (Erişim Tarihi: 27.02.2019)
3. Cran (2018), Erişim Adresi: <https://cran.r-project.org/> (Erişim Tarihi: 27.02.2019)
4. Data aspirant(2015), Erişim Adresi: <http://dataaspirant.com/2015/04/11/five-most-popular-similarity-measures-implementation-in-python/> (Erişim Tarihi: 27.02.2019)
5. Demir,H ve Taktak, F. (2011) Konumsal Veri Üzerine Sosyal Ağ Analizi (SAA): Afyonkarahisar Örneği, *Harita Teknolojileri Elektronik Dergisi*, Cilt: 3 No: 1 (7-16)
6. Demircan, S.(2009) Çoklu Etmenler Kullanılarak Enerji Nakil Hattı Güzergah Optimizasyonu, *Yüksek Lisans Tezi*, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
7. Dev veri (2014), Erişim Adresi: <http://devveri.com/veri-madenciligi/buyuk-veri-icerisinde-benzer-ogeleri-bulmak> (Erişim Tarihi: 27.02.2019)
8. Gürsakal N.,Tüzüntürk S., Eteman F.S. (2014) Sosyal Ağ Verilerinin Kuvvet Yasası Olasılık Dağılımına Uygunluk Analizi: Twitter Örneği, *Ekonometri, Yöneylem Araştırması ve İstatistik Sempozyumu*, Cilt: 15 ,Mayıs , Isparta.
9. İnce, K. ve Karcı, A.(2019) Akademik işbirliklerinin yeni bir çizge olarak modellenmesi ve istatistiki analizi, *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 2019 Cilt 34 Sayı 1,Ankara Turkey DOI: 10.17341/gazimmfd.416506
10. Jaccard index(2013),
Erişim Adresi: <http://bilgisayarkavramlari.sadievrenseker.com/2013/07/01/jaccard-indeksi-jaccard-index> (Erişim Tarihi: 27.02.2019)
11. Keleş, M.K. ve Özel, S.A. (2017) Mesafeli Metrikler ile Türkçe Metinler Arasında Benzerlik Tespiti, *Uluslararası Bilgisayar Bilimleri ve Mühendisliği Konferansı*, 5-8 Ekim, Antalya.

12. Math overflow(2010), Erişim Adresi: <https://mathoverflow.net/questions/18084/is-the-jaccard-distance-a-distance> (Erişim Tarihi: 27.02.2019)
13. Plerou A. P. , Sioutas S., Salaidis V.G.(2014) CSMR: A Scalable Algorithm for Text Clustering with Cosine Similarity and MapReduce, *IFIP International Conference on Artificial Intelligence Applications and Innovations*, September DOI: 10.1007/978-3-662-44722-2_23
14. R (2018) , Erişim Adresi: <https://www.r-project.org/about.html> (Erişim Tarihi: 27.02.2019)
15. Sabah, L. (2018) Sosyal Ağ Analizi ve Salgın Modelleme, *Yüksek Lisans Tezi*, Düzce Üniversitesi Fen Bilimleri Enstitüsü, Düzce
16. Yıldırım, B.(2012) Unitary Cayley Grafların Baskınlık Sayısının İncelenmesi, *Yüksek Lisans Tezi*, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.

