

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**ÇİZGE VERİ TABANI İLE TÜRK MUTFAĞI LEZZET AĞI'NIN
İNCELENMESİ**

Beyza ÖZHAN

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**ANKARA
2022**

Her hakkı saklıdır

ÖZET

Yüksek Lisans Tezi

ÇİZGE VERİ TABANI İLE TÜRK MUTFAĞI LEZZET AĞI'NIN İNCELENMESİ

Beyza ÖZHAN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Bülent TUĞRUL

Teknolojinin ilerlemesi, veri miktarı ve çeşitliğinin artmasıyla hayatımıza giren NoSQL kavramı yapısal olmayan verilerin de depolama ve analiz edilebilmesine olanak tanımıştır. NoSQL veri tabanı türlerinden olan çizge veri tabanları veriler ile birlikte aralarındaki bağlantıları da depolayıp analiz edebilmekte ve bu yönüyle öne çıkan teknolojiler arasında yer almaktadır. Ayrıca çevik gelişime uygun bir yapı sunabilmesi, derinlik arttıkça performans sorunu yaşamaması gibi özellikleri de diğer güçlü yönlerine örnek olarak verilebilmektedir. Bu avantajlara sahip olmasıyla eğitim, sağlık, finans, gıda gibi farklı alanlarda yapılan çalışmalarda tercih edilmiştir. Yapılan tez çalışmasında zengin bir kültüre sahip olan Türk Mutfağı üzerinde çizge veri tabanı teknolojisi kullanılarak analiz çalışması gerçekleştirilmiştir. Bu çalışma için Türk Mutfağına ait tarifler, malzemeler ve malzeme bileşenleri (protein, demir ve kalsiyum gibi çeşitli besin değerleri, vitamin, mineraller) arasındaki ilişkiler incelenerek bir lezzet ağı modellenmiştir. Analiz aşaması öncesinde Karadeniz, Akdeniz, İç Anadolu, Güneydoğu Anadolu ve Doğu Anadolu Mutfaklarına ait tarifleri ve bu tariflerde kullanılan malzemelere ilişkin bileşenleri içeren bir veri seti hazırlanmıştır. Hazırlanan veri seti üzerinde hem bölgesel bazda hem de genel Türk Mutfağı üzerinde çeşitli analiz çalışmaları gerçekleştirilmiştir. Veri setinde bileşen bilgisinde bulunmasıyla tariflerde alınan bileşen türleri ve miktarları, kalori bilgileri gibi çeşitli incelemelerin yapılmasına olanak sağlanmıştır. Önemli malzemelerin tespiti, malzemeler arasında topluluk yapısının incelenmesi gibi çalışmalarda merkezilik ve topluluk tespiti algoritmalarından yararlanılmıştır. Yapılan analizler sonucu elde edilen bulgulara göre Türk Mutfağında et ürünlerinden büyükbaş hayvan eti, meyvelerden limon, süt ürünlerinden yoğurt, yağlardan tereyağı ve tahıl ürünlerinden buğday ununun yüksek oranda tercih edildiği gözlemlenmiştir.

Eylül 2022, 79 sayfa

Anahtar Kelimeler: NoSQL, Çizge Veri Tabanları, Türk Mutfağı

ABSTRACT

Master Thesis

ANALYSIS OF TURKISH CUISINE FLAVOR NETWORK WITH GRAPH DATABASE

Beyza ÖZHAN

Ankara University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

Supervisor: Dr. Öğr. Üyesi Bülent TUĞRUL

The NoSQL concept, which has entered our lives with the advancement of technology and the increase in the amount and variety of data, allows unstructured data to be stored and analyzed. Graph databases, one of the NoSQL database types, can store and analyze the connections between them and the data. In addition, its features, such as providing a structure suitable for agile development and not experiencing performance problems as the depth increases, can be given as examples of its other strengths. With these advantages, it has been preferred in studies conducted in different fields such as education, health, finance and food. In the thesis, an analysis study was carried out on Turkish Cuisine, which has a rich culture, using graph database technology. For this study, a flavor network was modelled by examining the relationships between Turkish Cuisine recipes, ingredients and components (various nutritional values, vitamins, minerals such as proteins, iron and calcium). Before the analysis phase, a data set containing the recipes of the Black Sea, Mediterranean, Central Anatolia, Southeastern and Eastern Anatolian Cuisines and the ingredients used in these recipes were prepared. Various analysis studies were carried out on the prepared data set, both on a regional basis and the general Turkish Cuisine. With the components information in the data set, it has been possible to make various analyzes such as components types and amounts taken in recipes and calorie information. Centrality and community detection algorithms were used in studies such as detecting essential ingredients and examining community structure among ingredients. According to the findings obtained as a result of the analysis, it has been observed that in Turkish Cuisine, cattle meat from meat products, lemon from fruits, yoghurt from dairy products, butter from oils and wheat flour from cereal products are highly preferred.

September 2022, 79 pages

Key Words: *NoSQL, Graph Databases, Turkish Cuisine*

TEŞEKÜR

Yüksek lisans tez çalışmamı hazırlamamda bana her zaman destek olup, çalışma sürecimde beni yönlendirip, gelişimime katkı sağlayan danışman hocam sayın Dr. Öğr. Üyesi Bülent Tuğrul'a ve maddi manevi anlamda her zaman desteklerini hissettiren aileme en içten duygularıyla teşekkür ederim.

Beyza ÖZHAN
Ankara, Eylül 2022

İÇİNDEKİLER

TEZ ONAY SAYFASI

ETİK.....	i
ÖZET.....	ii
ABSTRACT	iii
TEŞEKÜR.....	iv
SİMGELER DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
ÇİZELGELER DİZİNİ	x
1. GİRİŞ	1
2. KAYNAK ÖZETLERİ	3
3. MATERYAL ve YÖNTEM.....	6
3.1 Materyal	6
3.2 Yöntem	7
3.2.1 NoSQL veri tabanı teknolojisi nedir?	7
3.2.2 NoSQL veri tabanı çeşitleri	9
3.2.3 Büyük şirketler tarafından desteklenen NoSQL veri tabanları	10
3.2.4 Çizge veri tabanı nedir?	11
3.2.5 Çizge veri tabanlarının ana özellikleri ve avantajları	12
3.2.6 Çizge veri tabanları kullanım alanları	13
3.2.7 İlişkisel veri tabanları ile çizge veri tabanlarının performans karşılaştırması.....	14
3.2.8 Neo4j çizge veri tabanı.....	15
3.2.9 Cypher sorgu dili.....	15
3.3 Merkezilik Algoritmaları.....	16
3.3.1 Derece merkeziliği algoritması.....	16
3.3.2 Arasındalık merkeziliği algoritması	19
3.3.3 Yakınlık merkeziliği algoritması	21
3.3.4 PageRank merkezilik algoritması.....	22
3.4 Topluluk Tespiti Algoritmaları.....	24
3.4.1 Üçgen sayma algoritması ve yerel kümelenme katsayısı	24
3.4.2 Louvain algoritması	27
3.4.3 Etiket yayılımı algoritması	29

3.4.4 Zayıf bağlı bileşenler algoritması	30
3.5 İki Parçalı Çizgelerde Merkezilik ve Topluluk Tespiti Algoritmalarının Kullanılması.....	32
4. DENEYSEL ÇALIŞMA VE ARAŞTIRMA BULGULARI.....	37
4.1 Veri Tabanı Oluşturulması	37
4.2 Türk Mutfağı Üzerinde Yapılan Genel Analiz Çalışmaları.....	38
4.2.1 Malzemelere göre tarif önerisi	38
4.2.2 Bileşenlere göre tarif önerisi	38
4.3 Tariflerde Kullanılan Malzemelerin Merkezilik Algoritmalarına Göre Değerlendirilmesi	39
4.3.1 Derece merkeziliği algoritması sonuçlarının değerlendirilmesi	41
4.3.2 Arasındalık merkeziliği algoritması sonuçlarının değerlendirilmesi	43
4.3.3 Yakınlık merkeziliği algoritması sonuçlarının değerlendirilmesi	44
4.3.4 PageRank merkezilik algoritması sonuçlarının değerlendirilmesi	46
4.4 Tariflerde Kullanılan Malzemeler için Topluluk Tespiti Algoritmalarının Çalıştırılması	46
4.4.1 Üçgen sayma algoritması ve yerel kümelenme katsayısına göre malzemelerin değerlendirilmesi	47
4.4.2 Louvian algoritmasına göre malzemelerin gruplandırılması	48
4.4.3 Etiket yayılımı algoritmasına göre malzemelerin gruplandırılması	49
4.4.4 Zayıf bağlı bileşenler algoritmasına göre malzemelerin gruplandırılması.....	50
4.5 Bölgesel Bazda Yapılan Analiz Çalışmaları	51
4.5.1 Bölgesel bazda tariflerde kullanılan ortalama malzeme sayısı	51
4.5.2 Bölgesel bazda bileşen değerlerine göre yapılan analiz çalışması	53
4.5.3 Bölgesel bazda en çok tercih edilen malzemeler.....	54
5. TARTIŞMA VE SONUÇ	61
KAYNAKLAR	63
EKLER.....	67
EK 1 Bileşen İsimleri ve Birimleri.....	67
EK 2 Malzemelere Ait Merkezilik Algoritması Sonuçları	70
EK 3 Malzeme Çiftlerinin Bölgelere Göre Göreceli Yaygınlık Değerleri (Şekil 4.14)	76
ÖZGEÇMİŞ.....	79

SİMGELER DİZİNİ

g	Gram
mg	Miligram

Kısaltmalar

ACID	Atomicity, Consistency, Isolation, Durability
NOSQL	Not Only SQL
AGPLv3	Affero Genel Kamu Lisansı
PR	PageRank
IMDB	Internet Movie Database

ŞEKİLLER DİZİNİ

Şekil 3.1 ACID prensipleri.....	8
Şekil 3.2 NoSQL veri tabanı çeşitleri	9
Şekil 3.3 Veri boyutu ile veri karmaşıklığı yönleri arasında NoSQL veri modelleri (Neubauer 2010)	10
Şekil 3.4 İlişkisel veri tabanlarının çizge veri tabanlarına dönüştürülmesi	11
Şekil 3.5 Çizge veri tabanı veri tipleri	12
Şekil 3.6 Yönsüz çizge örneği.....	17
Şekil 3.7 Yönlü çizge örneği.....	18
Şekil 3.8 Yönsüz çizge örneği – 2.....	20
Şekil 3.9 Yönsüz çizge örneği – 3.....	21
Şekil 3.10 Üçgen sayma algortiması örneği.....	26
Şekil 3.11 Louvain topluluk tespiti algoritması (Blondel vd. 2008).....	28
Şekil 3.12 Bağlı yönsüz çizge örneği	31
Şekil 3.13 Bağlı olmayan yönsüz çizge örneği	31
Şekil 3.14 Yönlü çizgeyi yönsüz çizgeye çevirme işlemi.....	32
Şekil 3.15 İki parçalı çizge örneği.....	32
Şekil 3.16 İki parçalı çizgenin tek parçalı ağırlıksız çizgeye dönüştürülmesi	34
Şekil 3.17 İki parçalı çizgenin tek parçalı ağırlıklı çizgeye dönüştürülmesi	35
Şekil 4.1 Veri tabanı modeli.....	37
Şekil 4.2 Tarif ve Malzeme arasındaki iki parçalı çizgenin malzemeler arasında kurulan tek parçalı, ağırlıklı ve yönsüz çizgeye dönüştürülmesi	41
Şekil 4.3 Türk Mutfağında su ve baharatlar dışında en çok kullanılan 20 malzeme	42
Şekil 4.4 Üçgen sayma algoritması ile tespit edilmiş üçgen ilişki örnekleri	47
Şekil 4.5 Louvain algoritması sonucu.....	48
Şekil 4.6 Etiket yayılımı algoritması sonucu	49

Şekil 4.7 Zayıf bağlı bileşenler algoritması sonucu	50
Şekil 4.8 İlişki ağırlığı 1 olan ilişkilerin önemsenmediği durumda zayıf bağlı bileşen algoritmasından elde edilen sonuç	51
Şekil 4.9 Karadeniz bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları	55
Şekil 4.10 Akdeniz bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları	56
Şekil 4.11 İç Anadolu bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları	56
Şekil 4.12 Doğu Anadolu bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları.....	57
Şekil 4.13 Güneydoğu Anadolu bölgesinde en çok kullanılan 5 malzemeye ait yaygınlık değerleri, bu malzemelerin kullanım durumları ve sahip oldukları ortak bileşen sayıları	57
Şekil 4.14 Bölgesel bazda en çok kullanılan malzeme çiftlerinin bölgelere göre göreceli yaygınlığı.....	58

ÇİZELGELER DİZİNİ

Çizelge 3.1	Veri setinde bulunan tariflerin bölgelere göre dağılım pasta grafiği	6
Çizelge 3.2	MySQL ve Neo4j veri tabanı sistemleri performans karşılaştırması (Partner vd. 2014)	14
Çizelge 3.3	Şekil 3.6’da verilen çizgedeki düğümlere ait derece merkeziliği değerleri	17
Çizelge 3.4	Şekil 3.7’de verilen yönlü çizgede düğümlere ait derece merkeziliği değerleri.	18
Çizelge 3.5	Şekil 3.8’de verilen yönsüz çizgedeki düğümler için arasındalık merkeziliği değerleri	20
Çizelge 3.6	Şekil 3.9’da verilen yönsüz çizgedeki düğümler için yakınlık merkeziliği değerleri	22
Çizelge 3.7	V1 grubundaki düğümlerin V2 grubunda ilişkili olduğu düğümler	34
Çizelge 4.1	Düğüm türleri ve bu türlere ait özellikler	37
Çizelge 4.2	İlişki türleri ve bu türlere ait özellikler	38
Çizelge 4.3	Çeşitli bileşenler yönüyle zengin tarifler ve bileşenlerin tariflerde bulunma miktarları.....	39
Çizelge 4.4	Türk Mutfağında su ve baharatlar dışında en yüksek dereceye sahip ilk 20 malzeme için merkezilik algoritması sonuçları	40
Çizelge 4.5	Şekil 4.3’e göre kategori bazında en çok tercih edilen malzemeler	43
Çizelge 4.6	Çizelge 4.4’te verilen malzemelere ait yakınlık merkeziliği algortması sonucu ve kaç farklı malzeme ile birlikte kullanıldığı bilgisi	45
Çizelge 4.7	Şekil 4.5’te verilen Louvain algoritması sonucuna göre bulunan topluluklar ve içerdikleri malzeme sayıları	49
Çizelge 4.8	Bölgelere göre tariflerde kullanılan malzeme sayılarının dağılımı	52
Çizelge 4.9	Bileşenlerin bölgedeki tariflerde ortalama bulunma miktarı	53

1. GİRİŞ

Günümüzde teknolojinin hızlı bir şekilde ilerlemesi ve kullanımının artmasıyla beraber dijital bir veri yığını oluşmaya başladı. Özellikle web sayfaları, bloglar, sosyal medya uygulamaları sürekli olarak veri toplar hale geldi. Toplanan bu verilerin alışılan veri yapısının dışında, farklı çeşit ve boyutlarda olması mevcut ilişkisel veri tabanı sistemlerinde depolanmasını olanaksız hale getirdi. Bu nedenle ortaya veri dünyasına farklı bir boyut kazandıracak “NoSQL” (Not Only SQL) kavramı atıldı.

Son yıllarda önem kazanan NoSQL kavramı ile birlikte pek çok yeni veri tabanı teknolojisini de ortaya konmuştur. Özellikle sosyal bağlantıları temel alan Facebook, Instagram, Twitter, LinkedIn gibi uygulamaların popülerliği sadece verileri değil veriler arasındaki ilişkilerinde analiz edilmesi gerekliliğini ortaya koymuş ve bu bağlamda NoSQL veri tabanı çeşidi olan “çizge veri tabanları” teknolojileri geliştirilmiştir.

Çizge veri tabanları verilerin yanı sıra veriler arasındaki ilişkileride depolayıp, kullanıcıya gerçek zamanlı, daha yönetilebilir ve erişilebilir bir sistem sunduğundan Google, Facebook, LinkedIn gibi büyük şirketlerde kendi veri yönetimlerinde çizge veri tabanlarını kullanmış ve bu teknolojinin gelişimine katkı sağlamışlardır (Patil vd. 2014).

Çizge veri tabanlarında veriler düğümler ve düğümler arasında kurulan ilişkiler ile modellenmektedir. Oluşturulan bu model sayesinde düğümler arasında doğrudan veya dolaylı olarak bulunan ilişkiler, geleneksel veri tabanı yöntemlerine göre daha hızlı, verimli şekilde analiz edilebilmektedir (Guia vd. 2017). Ayrıca ilişkiler üzerinde özellik bilgileri bulundurulabilmesi, görsel açıdan kullanıcılar için daha anlaşılır bir yapı ortaya koyması gibi avantajlara sahiptir.

Pek çok gerçek dünya probleminin çözümüne uygun bir yapı sunan çizge veri tabanları farklı alanlarda yapılan çalışmalarda tercih edilmektedir. Bu alanlardan biri de gıda sektörü üzerine yapılan çalışmalardır. Dünya üzerinde yiyecekler ve aralarındaki ilişkiler kişilerin dikkatini çekmiş ve bu konu üzerinde çeşitli çalışmalar

gerçekleştirilmiştir. Bu araştırmalardan Ahn ve arkadaşlarının (2011) Lezzet Ağı (Flavor Network) çalışması ile besinler ve besine ait aroma bileşenleri arasındaki ilişkiler çizge yaklaşımı kullanılarak ortaya konmuştur. Yine bu tür araştırmaları ülke mutfağı bazında özelleştirmek, tarifler, malzemeler ve malzeme bileşenleri arasındaki ilişkileri incelemek, analiz etmekte mümkündür. Bu amaçla yapılan tez çalışmasında Türk Mutfağına özgü tarifler, bu tariflerde kullanılan malzemeler ve malzeme bileşenleri arasındaki ilişki ortaya konarak Türk Mutfağının lezzet ağı oluşturulmuştur.

Tez aşamasında yapılan çalışmalar Giriş, Kaynak Özetleri, Materyal ve Yöntem, Deneysel Çalışma ve Araştırma Bulguları ile Tartışma ve Sonuç kısımları olmak üzere 5 ana başlık altında anlatılmıştır. İlk kısımda tez konusundan bahsedilip NoSQL ve çizge veri tabanları kavramlarına değinilmiştir. İkinci kısımda literatür taraması yapılarak konu ile ilgili yapılan çalışmalar özetlenmiştir. Üçüncü kısımda hazırlanan veri seti tanıtılmış ve analiz çalışmalarında kullanılan algoritmalar anlatılmıştır. Dördüncü kısımda veri seti üzerinde yapılan analiz çalışmalarına ait elde edilen bulgulara yer verilmiştir. Son kısımda ise tez süreci boyunca yapılan çalışmalar özetlenerek genel bir değerlendirme yapılmıştır.

2. KAYNAK ÖZETLERİ

Veriler ile birlikte veriler arasındaki ilişkilerin incelenmesi, bu ilişkilere bağlı çeşitli gözlemlerin yapılabilmesi amacıyla literatürde farklı alanlarda yapılan çalışmalarda çizge veri tabanlarının tercih edildiğini görmek mümkündür. Örneğin Afandi ve Wahyuni (2020) akademi dünyasının parçası olan araştırmacı, yayıncı kuruluş, organizasyonlar (fıkri mülkiyeti onaylayan, hibe desteği alınan gibi çeşitli kuruluşları ifade eder), yapılan çalışmalar, yayınlanan makaleler gibi değerleri ve bu değerler arasındaki bağlantıları içeren bir ağ oluşturmuşlar. Oluşturdukları bu ağ sayesinde daha küresel boyutta iş birliklerini sağlayabilmek, araştırmacıların yer aldığı tüm çalışmalara ulaşabilmek, güçlü bağlantıları olan organizasyonları tespit etmek, araştırmacıların üniversite içerisinde ve dışarısında bireysel ve ortak performanslarını değerlendirmek gibi çeşitli analiz çalışmaları yapmak mümkün hale gelmiştir. Çalışmalarında birincil kaynak olarak Google Scholar, ikincil kaynak olarak internet üzerindeki çeşitli siteler, üniversitelerin veri tabanları gibi kaynaklardan yararlanmışlardır. Topladıkları verileri çizge veri tabanı üzerinde modellemişlerdir. İlişkisel veri tabanları ile karşılaştırma yapabilmek için aynı sorguları ilişkisel veri tabanları üzerinde de çalıştırmışlar ve çizge veri tabanlarının performansının çok daha iyi olduğunu tespit etmişlerdir.

Popüler bir film veri seti olan Internet Movie Database (IMDB) veri seti üzerinde Maxwell ve arkadaşları (2021) kişilerin ve filmlerin etkilerini ölçmek istemişlerdir. Bunun üzerine oyuncu, film, film türü, bölüm gibi veri türlerini içeren bir model oluşturmuşlardır. Bu türler arasındaki ilişkileri yönetebilmek için çizge veri tabanı olan Neo4j'i tercih etmişlerdir.

11.5 milyon dosyayı içeren, büyük bir veri seti olan Panama Kağıtları üzerinde dolandırıcılık tespiti için Srivastava ve Singh (2022) bir çalışma yapmışlar, çizge veri tabanı üzerinde modelledikleri veri seti ile daha hızlı sorgulama yapmak için bir yöntem geliştirmişlerdir. Panama Kağıtları veri setinde memur, kuruluş ve aracı olmak üzere 3 ana tür vardır ve çizge veri tabanı üzerinde modelleme buna göre yapılmıştır. Bu yöntemde PageRank (PR) yaklaşımına dayalı algoritma ile önemsiz düğümleri (finansal

işlemlerde aktif olmayan düğümler) çıkarmışlar sonrasında bu düğümler üzerinde dolandırıcılık tespiti için memurun aracı gibi davrandığı ilişkilere odaklanmışlardır. Geliştirdikleri yöntemde dağıtık mimari kullanımı ile sorguları paralelleştirmişler ve bu sayede dolandırıcılık tespiti algılama süresini %18.2 oranında azaltmışlardır.

Çizge tabanlı yaklaşımların kullanıldığı diğer farklı bir alan olarak gıda sektörü örnek verilebilir. Yemekler insanların günlük hayatının önemli bir parçası olduğundan bu durum literatürde yapılan çalışmalara da etki etmiş, konu ile ilgili çeşitli bilimsel çalışmalar yapılmıştır. Örneğin Ahn ve arkadaşları (2011) dünya üzerinde çok fazla malzeme ve bu malzemelerle yapılabilecek çok çeşitli kombinasyonlar olmasına rağmen insanların tariflerinde sadece belirli malzeme ve malzeme kombinasyonlarını neden kullandıkları üzerine bir araştırma yapmışlardır. Gıda bilimcilerin görüşüne göre kullanılan malzemelerin ortak aroma bileşen sayısı arttıkça tat alma olasılığı da o kadar artmaktadır. Örneğin bazı çağdaş restoranlarda havyar ve beyaz çikolatanın birlikte kullanıldığı, bu durum incelendiğinde bu malzemelerin ortak en az 73 aroma bileşeni içerdiği görülmüştür. Bu hipotez üzerine yaptıkları çalışmada malzeme kombinasyonları seçiminde ortak bileşen sayısının bir etkisi olup olmadığını incelemişlerdir. Çalışmalarında 381 malzeme ve 1021 aroma bileşeni arasında iki parçalı bir çizge tanımlayıp malzeme çiftlerinin ne kadar ortak bileşen paylaştıklarını gösteren bir lezzet ağı ortaya koymuşlardır. Araştırmalarında Doğu ve Batı Mutfaklarına ait tariflere yer vermişlerdir. Kullandıkları veri setinde Batı Mutfağından Kuzey Amerika, Latin Amerika, Güney Avrupa ve Batı Avrupa Mutfakları, Doğu Asya Mutfağından ise Kore Mutfağına ait tarifler bulunmaktadır. Kuzey Amerika ve Doğu Asya Mutfağı özelinde yaptıkları karşılaştırmada; Kuzey Amerika Mutfağında süt, tereyağı, vanilya, yumurta, şeker kamışı pekmezi ve buğday malzeme kombinasyonları ile yapılan çok sayıda tarif olduğu, Doğu Asya Mutfağında ise daha çok soya sosu, yeşil soğan, susam yağı, pirinç, soya fasulyesi ve zencefil malzemelerinden oluşan kombinasyonların tercih edildiğini belirlemişler. Bunun üzerine bu malzeme kombinasyonlarına ilişkin ortak bileşen sayısına baktıklarında, Kuzey Amerika Mutfağında kullanılan malzeme bileşenlerinin paylaştığı ortak bileşen sayısı fazla iken Doğu Asya Mutfağında kullanılan malzemelerin ortak bileşen sayısının az olduğunu gözlemlemişlerdir. Tüm veri seti üzerinde yaptıkları çalışmada vardıkları ortak çıkarım ise Batı Mutfaklarında tercih

edilen malzeme kombinasyonlarının paylaştıkları ortak bileşen sayısının fazla, buna karşılık Doğu Asya Mutfağında tercih edilen malzeme kombinasyonlarına ait ortak bileşen sayısının az olduğu yönündedir.

Bajaj ve arkadaşları (2018) ise çizge veri tabanı teknolojisine dayalı bir tarif öneri sistemi geliştirmişlerdir. Farklı tarifler denemek isteyen kişilere fikir vermek, mevcutta sahip oldukları malzemeler ile ne yapacağı konusunda kararsızlık yaşayan kişilere tarif önerisinde bulunmak için geliştirdikleri bir sistemdir. Oluşturdukları veri seti tarif ve malzemeleri içermekte ve türler arasında yoğun ilişkiler bulundurmaktadır. İlişkisel veri tabanları bu ilişkileri yönetmekte pahalı bir işlem olan birleşim (join) yapısından faydalandığından ve veri boyutu arttıkça sorguları daha da verimsiz çalıştıracığından dolayı ilişkisel veri tabanı yerine NoSQL çizge veri tabanı teknolojisi olan Neo4j'i tercih etmişlerdir. Benzer bir çalışma olarak Huang ve arkadaşları da (2022) bir tarif önerme sistemi oluşturmuşlardır. Ancak Bajaj ve arkadaşlarının (2018) oluşturduğu sistemden farklı olarak bu sistemde kişilerin hastalığı, diyet yapma durumu, sahip olduğu malzemeler gibi farklı parametreleri dikkate alarak tarif önerilerinde bulunmaktadırlar. Sistemi oluştururken bilgi grafiği (knowledge graph) teknolojisinden yararlanılmış, veriler ise Neo4j veri tabanında depolanmıştır.

3. MATERYAL ve YÖNTEM

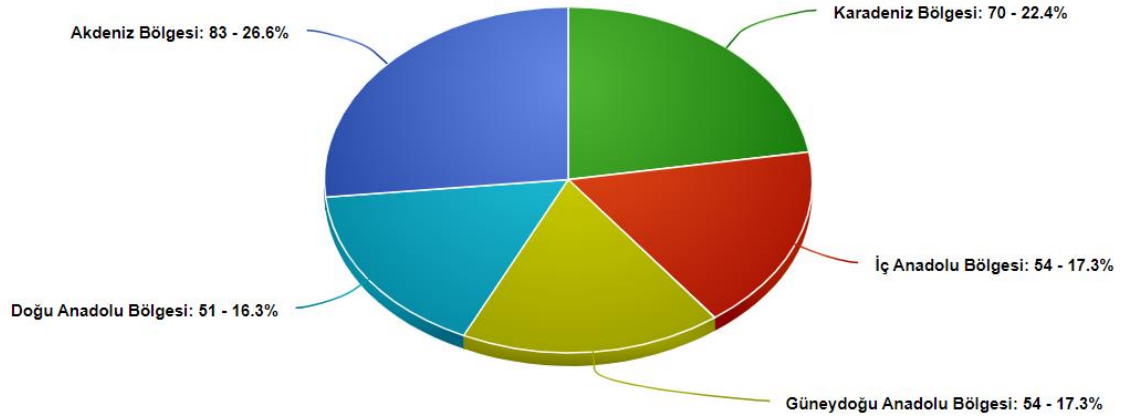
Tez çalışmasında kullanılan Türk Mutfağı veri seti ve analiz aşamasında kullanılan yöntemler bu bölümde anlatılmıştır.

3.1 Materyal

Türk Mutfağındaki tarifleri, malzemeleri, malzeme bileşenlerini, yemeğin ait olduğu bölge bilgisi gibi bilgileri içeren hazır bir veri seti bulunmamaktadır. Bu nedenle analiz çalışması için veri seti oluşturulmuştur.

Hazırlanan veri setinde Akdeniz, Karadeniz, İç Anadolu, Güneydoğu Anadolu ve Doğu Anadolu olmak üzere 5 farklı bölgeye ait toplam 312 tarife yer verilmiştir. Veri setindeki tariflerin bölge bazında dağılımı Çizelge 3.1 de bulunmaktadır.

Çizelge 3.1 Veri setinde bulunan tariflerin bölgelere göre dağılım pasta grafiği



Veri seti hazırlarken genelde Türk halkının günlük yaşamında tükettiği, karmaşık olmayan malzeme içerikleri ile yapılan tarifler olması nedeniyle yemek kitapları yerine kaynak olarak internet sitelerindeki tarifler tercih edilmiştir (Anonim 2002, Anonim 2007).

Tariflerdeki malzemelere ait bileşen bilgileri için Tübitak tarafından oluşturulmuş TürKomp veri seti kullanılmıştır (Anonim 2017). Bu veri setinde Türkiye’de üretilen ve tüketilen, 14 gıda grubuna ait toplam 645 malzeme için besin öğeleri ve enerji değerleri yer almaktadır. Veri setinde bir malzeme için verilen bileşen bilgileri malzemenin yenilebilir 100 g’ı içindir. Tariflerde kullanılan ancak TürKomp veri setinde bulunmayan malzemeler için farklı kaynaklardan yararlanılmıştır (Anonim 2012).

Hazırlanan veri seti içerisinde 231 farklı malzeme bulunmaktadır. Malzemelerin içerdiği besin bileşenleri incelendiğinde ise tüm malzemeler içerisinde 98 farklı bileşen bulunduğu tespit edilmiştir (EK 1).

Toplanan tariflerin analiz edilebilmesi için ön işleme sürecine tabi tutulmuştur. Tariflerde malzemeler için su bardağı, çay bardağı, aldığı kadar un gibi ifadeler ile ölçüler verildiğinden gerçekçi bir analiz yapılması için her bir malzeme miktarı gram cinsine çevrilmiştir. Ayrıca farklı sayıda kişiler için tarifler verildiğinden ortak bir anlayış olması için tüm miktarlar tek kişi için hesaplanmıştır.

3.2 Yöntem

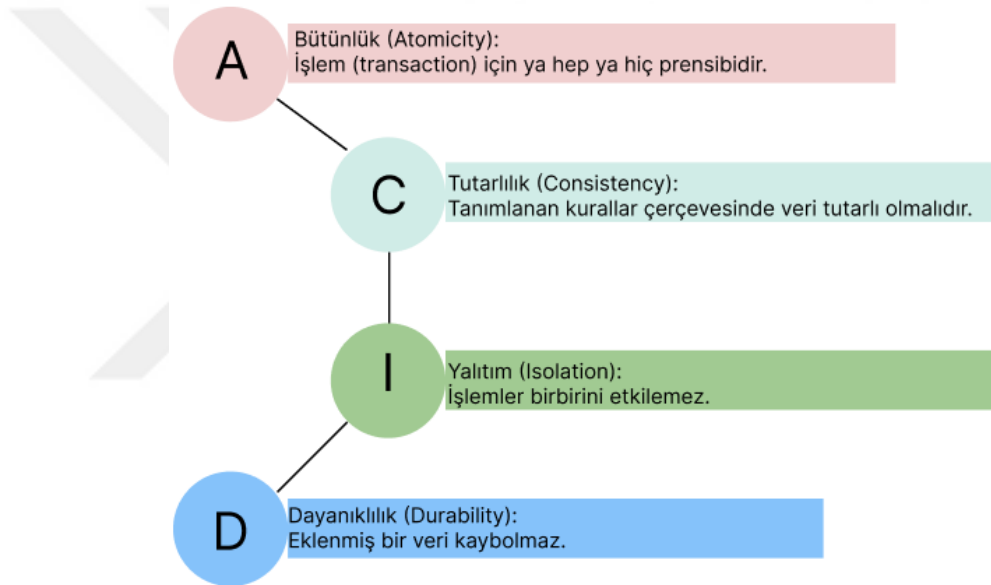
3.2.1 NoSQL veri tabanı teknolojisi nedir?

NOSQL ilişkisel veri tabanı sistemlerine alternatif olarak türetilmiş bir veri tabanı sistemidir. NOSQL ismi bu veri tabanlarında sql kullanılmadığından “Not Only SQL” olarak isimlendirilmiştir.

Karmaşık bağlantılı bilgileri içeren büyük veri problemleri bilimlerde giderek yaygınlaşmaktadır. Karmaşık veriler yapısal olmayan farklı tip (doküman, fotoğraf, konum verisi) ve boyutlardaki yapısal olmayan verilerdir. Geleneksel ilişkisel veri tabanı yaklaşımları bu tür karmaşık verilerin depolanması, alınması ve işlenmesinde yetersiz hale gelmiştir (Miller 2013).

Geleneksel veri tabanlarının, mevcut ilişkisel modeller üzerindeki gereksinimleri karşılamadaki kısıtlamaları, NOSQL veri tabanları adı verilen yeni teknolojilerin geliştirilmesine öncülük etmiştir (Anonymous 2009).

İlişkisel veri tabanının temel prensipleri olan bütünlük (atomicity), tutarlılık (consistency), yalıtım (isolation) ve dayanıklılık (durability) (ACID) ilkeleri veri tabanının güvenilirliğini garanti eder. NoSQL veri tabanları genel olarak ACID ilkelerini reddederken, bu aileden Neo4j çizge veri tabanı bu ilkeleri benimser (Vicknair vd. 2010).



Şekil 3.1 ACID prensipleri

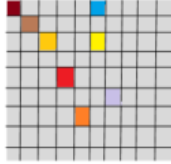
- **Bütünlük:** İşlemin (transaction) herhangi bir anında hata ile karşılaşırsa tüm yapılan işlemler geri alınır.
- **Tutarlılık:** Veri tabanı işlem sonrasında yine geçerli durumda olmalıdır.
- **Yalıtım:** Veri üzerinde bir işlem yapılıyorsa aynı veri üzerinde farklı bir işlem yapılması engellenir.
- **Dayanıklılık:** İşlem başarı ile tamamlanmış ise verilerdeki değişiklikler kalıcı olur.

3.2.2 NoSQL veri tabanı çeşitleri

Sütun tabanlı (column based), döküman tabanlı (document based), anahtar-değer tabanlı (key-value based) ve çizge veri tabanı olmak üzere 4 farklı tip NoSQL veri tabanı çeşidi bulunmaktadır (Anonymous 2021).

- **Sütun tabanlı veri tabanı:** Özel kolonlar tanımlanır ve bu alanların altına her kayıt için farklı bir kolon gelebilir. Örnek olarak Google'ın BigTable modeli, Cassandra verilebilir.
- **Döküman tabanlı veri tabanı:** Döküman tipindeki verileri saklar. Örnek olarak MongoDB verilebilir.
- **Anahtar – değer tabanlı veri tabanı:** Her bir key'e karşılık bir array veya veri yapısı bulunur. Key üzerinden index ile hızlı arama yapılır. Örnek olarak BerkeleyDB verilebilir.
- **Çizge veri tabanları:** Çizge yapısında (ilişkiler ile) verileri saklar. Örnek olarak Neo4j verilebilir.

SÜTUN TABANLI VERİTABANI



ÇİZGE VERİTABANI



DÖKÜMAN TABANLI VERİTABANI

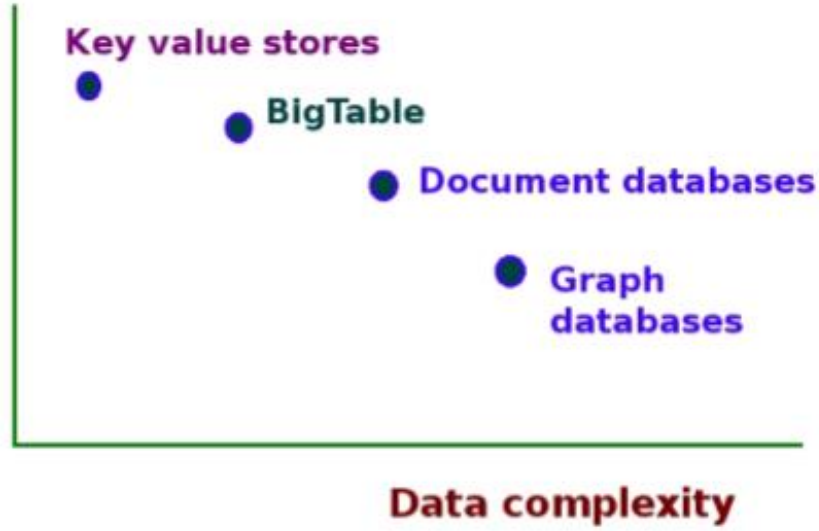


ANAHTAR - DEĞER TABANLI VERİTABANI



Şekil 3.2 NoSQL veri tabanı çeşitleri

Data size



Şekil 3.3 Veri boyutu ile veri karmaşıklığı yönleri arasında NoSQL veri modelleri (Neubauer 2010)

3.2.3 Büyük şirketler tarafından desteklenen NoSQL veri tabanları

Google, Facebook, Amazon, LinkedIn gibi büyük şirketler kendi verilerini depolamak ve analiz edebilmek için NoSQL veri tabanlarını kullanmış ve bu teknolojinin gelişimine katkı sağlamıştır. Bu şirketler tarafından geliştirilen bazı önemli NoSQL veri tabanlarına BigTable, Dynamo, Casandra ve Project Voldemort verilebilir.

BigTable:

- Google tarafından geliştirilmiştir.
- Hızlı, büyük ölçekli ve dağıtık.
- Sütun tabanlı veri tabanıdır (Chang vd. 2008).

Dynamo:

- Amazon tarafından geliştirilmiştir.
- Ölçeklenebilir, güvenilir, dağıtık.

- Anahtar – değer tabanlı veri tabanıdır (DeCandia vd. 2007).

Cassandra:

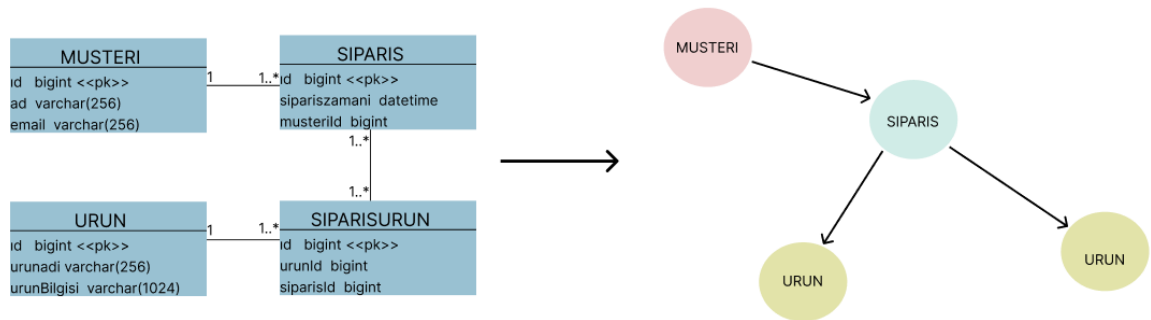
- Facebook tarafından geliştirilmiştir.
- Açık kaynaklıdır.
- Sütun tabanlı veri tabanıdır.

Project Voldemort:

- LinkedIn'in büyük ölçekli kalıcı hash tablosudur.
- Dağıtık ortamda çalışır (Anonymous 2008).

3.2.4 Çizge veri tabanı nedir?

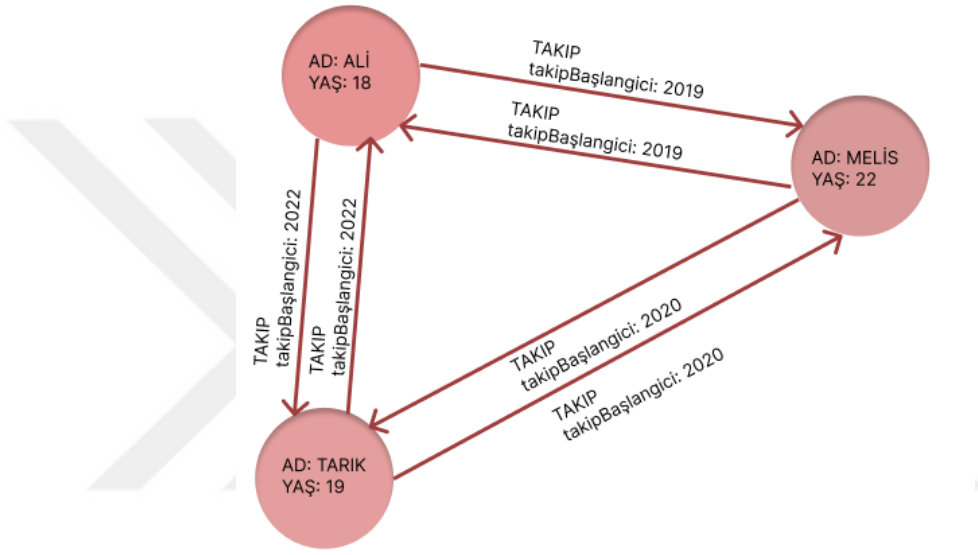
Çizge veri tabanları, verilerin çok güçlü bir şekilde birbirine bağlandığı mevcut bağlamda bilgi depolama için etkili ve verimli bir çözüm sağlayan bir tür ilişkisel olmayan veri tabanıdır. Verileri depolamak için düğümleri, kenarları ve özellikleri olan grafik yapılarını kullanır (Guia vd. 2017).



Şekil 3.4 İlişkisel veri tabanlarının çizge veri tabanlarına dönüştürülmesi

Çizge veri tabanları düğüm, ilişkiler, özellikler ve etiketler olmak üzere 4 tip veri içerir.

- **Düğüm(Node):** Varlıkları, nesneleri temsil eder.
- **İlişkiler (Relationships):** Düğümler arasındaki bağlantıyı tanımlar.
- **Özellikler (Properties):** İlişkilere, düğümlere ait özelliklerdir.
- **Etiketler (Labels):** Düğümleri, ilişkileri kümeler halinde gruplamak için kullanılır.



Şekil 3.5 Çizge veri tabanı veri tipleri

3.2.5 Çizge veri tabanlarının ana özellikleri ve avantajları

Çizge veri tabanlarının özellikleri ve güçlü yanları bu veri tabanlarının kullanılmasındaki başlıca sebeplerdir. Bu özellikler (Guia vd. 2017);

- Çevik yazılım geliştirme yaklaşımına uygundur.
- Birbiri ile bağlantılı veriler için uygundur, genelde gerçek dünya problemleri bu şekildedir.
- Veri madeciliği uygulamaları için optimize edilebilir.
- İlişkisel veri tabanları ile karşılaştırıldığında birleşim (join) sayısı fazla olan sorgularda yüksek performans göstermektedir.

3.2.6 Çizge veri tabanları kullanım alanları

Grafik veri tabanlarının dünyanın önde gelen şirketleri tarafından kullandıkları bazı alanlar şunlardır (Patil vd. 2014):

- **PR:** Google, arama sonuçlarının görüntülenme sırasını belirlerken çizge veri tabanı teknolojisinden faydalanmaktadır. Çizge veri tabanı üzerinde uygulanan merkezilik algoritması türü olan PR algoritmasını geliştirerek önemli sayfaların üst sıralarda görünmesini sağlamıştır.
- **Veri Yönetimi (Data Management):** Cisco gibi önemli bir ağ şirketi, son zamanlarda Neo4j tabanlı bir veri yönetim anlayışı benimsemiştir. Oracle RAC ile karşılaştırıldığında verilere erişim çok daha hızlıdır.
- **Sosyal Ağ:** Facebook, Twitter, LinkedIn, Viadeo, Glassdoor gibi web siteleri bağlantılarını ilişki olarak çizge veri tabanları üzerinde depolamaktadır. Böylece çizgedeki ilişkilere göre çeşitli parametrelere bakarak kişilere uygun arkadaşlık önerisi gibi bazı özellikler kullanıcılara sunulabilmektedir. Ayrıca ilişkiler, ilişkisel veri tabanlarına kıyasla gerçek zamanlı olarak çok daha iyi yönetilebilir ve erişilebilirdir.
- **Ağ Yönetimi (Network Management):** SFR, Telenor, Huawei, Just Dial gibi telekomünikasyon şirketleri, birbirine bağlı planlar, müşteriler ve gruplardan oluşan ağlarını, daha hızlı kimlik ve arıza tespiti gibi durumlar için çizge veri tabanı teknolojisi ile modellemeye geçmişlerdir.
- **Bioinformatik:** Era7, DNA dizilimi, genler, proteinler, enzimler vb. konularda bilgi depolayan bir şirkettir. Verileri ve bu veriler arasındaki karmaşık ilişkileri saklamak için Bio4j isimli Neo4j veri tabanına benzeyen bir çizge veri tabanı teknolojisini kullanmaktadır.

3.2.7 İlişkisel veri tabanları ile çizge veri tabanlarının performans karşılaştırması

Partner ve arkadaşları (2014) çizge veri tabanlarının ilişkisel veri tabanlarından çok daha iyi performans gösterdiğini kanıtlamak için bir çalışma yapmıştır. Bu çalışmalarında 1 milyon kullanıcı bir arkadaşlık ilişkisinin bulunduğu veri seti kullanmışlardır. Çalışmada ilişkisel veri tabanı olarak MySQL, çizge veri tabanı olarak Neo4j karşılaştırılmıştır.

Veri üzerinde çalıştırılan sorgular çizge içerisinde dolaşma (graph traversal) sorgularıdır. Çizelge 3.2’de MySQL ve Neo4j üzerinde, 1000 kullanıcıdan oluşan bir veri ile çalıştırılan derinlik sorgularının çalışma süresi saniye cinsinden verilmiştir. Derinlik sorguları arkadaşının kaçınıcı dereceden arkadaşlarına bakıldığı sorgulardır. Örneğin derinlik değerinin 2 olması arkadaşının arkadaşlarını bulmaya yönelik olduğudur.

Çizelge 3.2 MySQL ve Neo4j veri tabanı sistemleri performans karşılaştırması
(Partner vd. 2014)

Derinlik	Çalışma Süresi (s) – MySQL	Çalışma Süresi (s) – Neo4j
2	0.016	0.010
3	30.267	0.168
4	1543.505	1.359
5	1 saatte bitmedi	2.132

Sonuçlar incelendiğinde MySQL’de derinlik arttıkça sorguların oldukça yavaş çalıştığını, Neo4j’in ise derinlik artmasına rağmen performansında ciddi bir düşüş olmadığı ve MySQL’e göre oldukça iyi bir performans gösterdiği görülmektedir.

3.2.8 Neo4j çizge veri tabanı

Tez çalışmasında çizge veri tabanı teknolojisi olan Neo4j tercih edilmiştir. Neo4j, Neo Technology tarafından geliştirilen lider grafik veri tabanı türüdür. Affero Genel Kamu Lisansı (AGPLv3) ile açık kaynaklıdır. Profesyonel destek seçeneği için ticari olarak lisanslanmıştır (Patil vd. 2014).

Özellikleri (Patil vd. 2018):

- Neo4j'de hem düğümler hem de ilişkiler özellikler içerebilir.
- Bir ağdaki milyarlarca düğüm ve ilişki ölçeklendirebilir.
- Kullanıcı dostu bir arayüze sahiptir.
- ACID ilkelerini benimsediğinden güvenilir bir depolama sistemi sunmaktadır.

3.2.9 Cypher sorgu dili

Neo4j'deki sorgu yöntemleri, Neo Technology tarafından geliştirilen sorgu dili olan Cypher dilinde yazılmıştır (Guia vd. 2017).

- Cypher dili, grafik veri tabanının anlamlı ve verimli bir sorgu çalıştırmasına ve güncellenmesine izin verir.
- Nispeten basit, ama çok güçlü bir dildir.
- İlişkisel bir veri tabanında çalışan çok karmaşık sorgular, Cypher'da kolayca yürütülebilir. Bu, kullanıcıların veri tabanı erişiminde kaybolmak yerine alan bilgisine (domain) odaklanmalarına olanak tanır.
- Bu sorgu dile ile düğüm ve ilişkiler için ekleme (insert) / güncelleme (update) / seçim (select) / silme (delete) işlemleri yapılabilir.

3.3 Merkezilik Algoritmaları

Bir çizgede bazı düğümler komşuları ile olan ilişkisi, bilgiyi yayma hızı, düğümler arasında köprü görevi görmesi gibi pek çok farklı açıdan çizge için önemli bir hale gelebilmektedir. Çizgede bu önemli, merkezi düğümlerin belirlenmesi pek çok açıdan fayda sağlayabilir. Örneğin salgın bir hastalığın yayıldığı bir ağda yayılımı hızlandıracak merkezi kişilerin belirlenmesi hastalığın yayılma hızının azalmasına katkı sağlayacaktır.

Merkezilik algoritmaları bir çizgede önemli olan düğümlerin belirlenmesinde kullanılan algoritmalarlardır. Bu algoritmaların farklı yönlerden düğümün önemi belirleyen çeşitleri bulunmaktadır. Tez çalışmasında bu algortimalardan en çok tercih edilen derece merkeziliği (degree centrality), arasındalık merkeziliği (betweenness centrality), yakınlık merkeziliği (closeness centrality) ve PR algortimalarına yer verilmiştir.

3.3.1 Derece merkeziliği algoritması

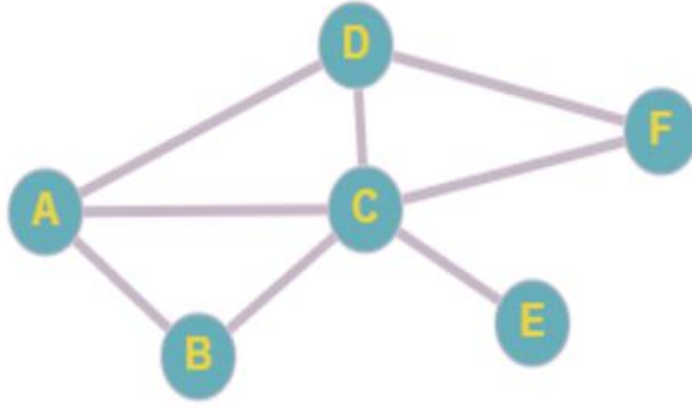
Derece merkeziliği algoritması bir çizgede düğümün sahip olduğu komşuların sayısını ifade etmektedir. Bu kavram Freeman (1979) tarafından yapılan sosyal ağ analizi çalışmasında ortaya konmuş ve bir düğümün çizge içerisindeki önemini ifade eden ölçümlerden biri olarak kabul görmüştür. Algoritmanın çalışma mantığına göre komşu sayısı fazla olan düğüm önemli konumdadır. Derece merkeziliği algoritması eşitlik 3.1 ile hesaplanmaktadır.

$$(D_i) = \sum_{j=1, i \neq j}^n I_{ij} \quad (3.1)$$

$D_i \rightarrow i$ düğümüne ait derece

$n \rightarrow$ Çizgedeki toplam düğüm sayısı

$I \rightarrow$ Komşuluk matrisi



Şekil 3.6 Yönsüz çizge örneği

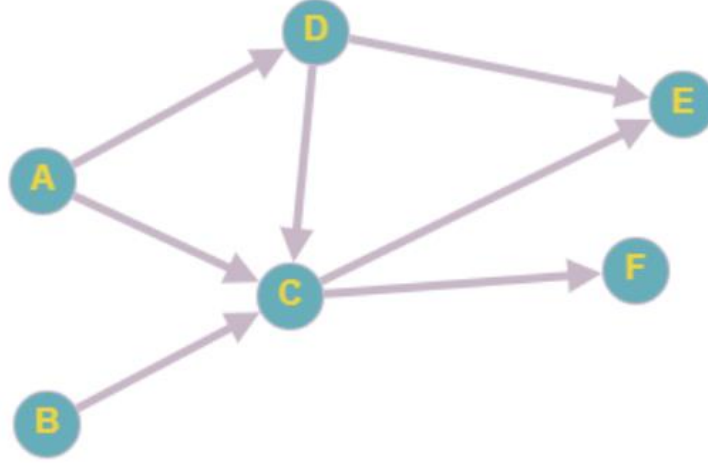
Çizelge 3.3 Şekil 3.6’da verilen çizgedeki düğümlere ait derece merkeziliği değerleri

Düğüm	Derece Merkeziliği
A	3
B	2
C	5
D	3
E	1
F	2

Şekil 3.6’da verilen çizgedeki düğümler için Çizelge 3.3’de derece merkeziliği değerleri verilmiştir. Buna göre çizgede en fazla komşuluk ilişkisi bulunan C düğümünün en yüksek derece merkeziliğine sahip olduğu görülmektedir.

Yönlü ilişkiler içeren bir çizgenin düğümlerine ilişkin derece merkeziliği hesaplanırken düğüme gelen ilişkileri ifade eden iç derece ve düğümden çıkan ilişkileri ifade eden dış derece kavramları dikkate alınmaktadır. Düğümdeki toplam derece ise iç derece ve dış

derecenin toplamına eşittir. Düğüme gelen ilişkiler çizgedeki diğer düğümler tarafından ne kadar desteklendiğini işaret ederken düğümden çıkan ilişkiler düğümün diğer düğümler üzerindeki liderliğini ifade etmektedir (Mrvar 2015).



Şekil 3.7 Yönlü çizge örneği

Çizelge 3.4 Şekil 3.7’de verilen yönlü çizgede düğümlere ait derece merkeziliği değerleri

Düğüm	İç Derece	Dış Derece	Toplam Derece
A	0	2	2
B	0	1	1
C	3	2	5
D	1	2	3
E	2	0	2
F	1	0	1

Şekil 3.7’de verilen çizgedeki düğümlere ait iç derece, dış derece ve toplam derece bilgileri Çizelge 3.4’de yer almaktadır. Çizelge 3.4’deki değerler incelendiğinde en

fazla iç derece merkeziliği değerinin C düğümüne, en fazla dış derece merkeziği değerinin ise A, C ve D düğümlerine ait olduğu görülmektedir. Toplam dereceye bakılacak olursa C düğümünün ağdaki en yüksek derece sahip olduğunu söylemek mümkündür.

3.3.2 Arasındalık merkeziliği algoritması

Arasındalık merkeziliği algoritması düğümün en kısa yol üzerinde bulunma durumuna göre önem kazanması anlayışına dayalıdır. Bir düğümün, birbiri ile bağlantılı olmayan diğer düğümler arasındaki en kısa yol üzerinde bulunma sayısına göre önem kazandığını ifade eder. Arasındalık değeri yüksek düğümler çizgedeki bilgi akışında etkilidirler (Freeman 1979). Bu nedenle çizgeden koparıldıklarında bilgi yayılımı zarar göreceğinden arasındalık merkeziliği yüksek olan düğümler, diğer düğümler için önemli olmaktadır.

Algoritma mantığı eşitlik 3.2’de verildiği gibi çizgedeki bir düğümün düğüm çiftleri arasındaki en kısa yollardan kaçında bulunduğu hesaplamasına dayanmaktadır.

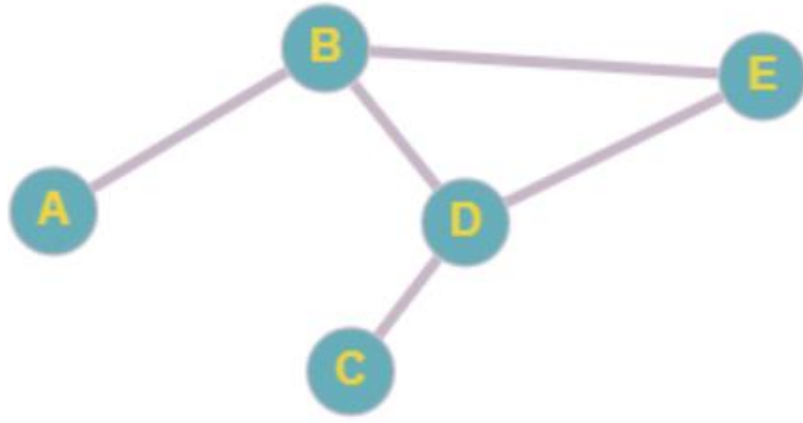
$$A(i) = \sum_{i \neq j \neq k} \frac{y(i)}{Y} \quad (3.2)$$

$A(i) \rightarrow i$ düğümüne ait arasındalık merkeziliği değeri

$y(i) \rightarrow j$ ve k düğümleri arasındaki en kısa yollardan i düğümü üzerinden

geçenlerin sayısı

$Y \rightarrow j$ ve k düğümleri arasındaki en kısa yolların sayısı



Şekil 3.8 Yönsüz çizge örneği – 2

Çizelge 3.5 Şekil 3.8’de verilen yönsüz çizgedeki düğümler için arasındalık merkeziliği değerleri

Düğüm	Arasındalık Değeri
A	0
B	3
C	0
D	3
E	0

A ve C düğümleri herhangi iki düğüm arasında yer almadığından bu düğümlerin arasındalık değeri 0 olur. A-E, A-D, A-C düğümleri arasında tek bir kısa yol vardır ve B üzerinden geçer. Bu durumda B’nin arasındalık değeri 3 olur. D düğümü ise A-C, B-C ve C-E arasındaki en kısa yol üzerinde bulunur. Bu nedenle arasındalık değeri 3’tür. E düğümü ise herhangi bir kısa yol üzerinde bulunmadığından 0 arasındalık değerini alır.

3.3.3 Yakınlık merkeziliği algoritması

Yakınlık merkeziliği algoritması bir düğümün diğer düğümlere olan yakınlığı ile ilgilenmektedir. Bu algoritmaya göre bir düğüm diğer düğümlere ne kadar yakınsa o kadar merkezi ve bilginin yayılma hızında o kadar etkili olur (Freeman 1979). Yani bir düğüme ait yakınlık merkeziliği değeri, çizge içerisinde bilgiyi yayma hızındaki etkisini göstermektedir. Aynı şekilde düğümlerin iletilen bilgiyi edinme hızları da yakınlık merkeziliği değeri ile doğru orantılıdır (Borgatti 2005).

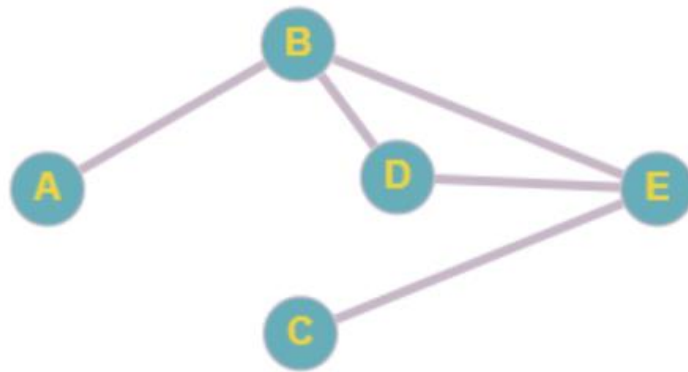
n büyüklüğündeki bir çizgede i düğümüne ait yakınlık merkeziliği değeri eşitlik 3.3 ile hesaplanmaktadır.

$$Y(i) = \frac{n-1}{\sum y(i,j)} \quad (3.3)$$

$Y(i) \rightarrow i$ düğümüne ait yakınlık merkeziliği değeri

$n \rightarrow$ çizgedeki düğüm sayısı

$y(i,j) \rightarrow i$ ve j arasındaki en kısa yol



Şekil 3.9 Yönsüz çizge örneği - 3

Çizelge 3.6 Şekil 3.9’da verilen yönsüz çizgedeki düğümler için yakınlık merkeziliği değerleri

	A	B	C	D	E	Yakınlık Merkeziliği
A	0	1	3	2	2	$4/8 = 0.5$
B	1	0	2	1	1	$4/5 = 0.8$
C	3	2	0	2	1	$4/8 = 0.5$
D	2	1	2	0	1	$4/6 = 0.66$
E	2	1	1	1	0	$4/5 = 0.8$

Düğümlere ait yakınlık merkeziliği değerleri incelendiğinde B ve E düğümlerinin 0.8 değeri ile çizgede en yüksek yakınlık merkeziliği değerine sahip olduğu görülmektedir. Buna göre B ve E düğümlerinin merkeziliği diğer düğümlere göre daha fazladır ve bilgiyi iletme hızları diğer düğümlere kıyasla daha yüksektir.

3.3.4 PageRank merkezilik algoritması

PR algoritmasının çalışma mantığına göre bir düğümün önemini ona bağlanan diğer düğümlerin önemi belirlemektedir. Google çalışanı Page ve Brin (1998) tarafından ilk olarak ortaya konmuştur. Google site aramalarında popüler sayfaları üst sıralara yerleştirmek amacıyla oluşturulmuştur. Arama yapıldığında PR algoritmasına bağlı sitelere puan verilir ve bu puana göre sitelerin yüksekten düşük puana doğru sıralaması yapılır. Google orjinal makalesinde (Page ve Brin 1998) yer alan PR algoritmasına eşitlik 3.4’de yer verilmiştir. Bu algoritma PR algoritmasının ilk ortaya konmuş halidir.

$$PR(A) = (1 - d) + d \left(\frac{PR(T_1)}{C(T_1)} \right) + \dots + d \left(\frac{PR(T_n)}{C(T_n)} \right) \quad (3.4)$$

$PR(A) \rightarrow A$ düğümüne ait PR değeri

$d \rightarrow$ sönümlleme factörü (*damping factor*)

$T \rightarrow A$ düğümüne bağlantı veren düğüm

$C(T) \rightarrow T$ düğümünden çıkan bağlantıların sayısı

Bu formüle göre $T_1 \dots T_n$ düğümleri A düğümüne bağlanan düğümleri temsil etmektedir. Formülde yer alan sönümlleme faktörü 0 ile 1 arasında değerler alabilmekte ve genelde 0.85 olarak kullanılmaktadır (Page ve Brin 1998).

Algoritmada ilk olarak tüm düğümlere bir PR ataması yapılır ve eşitlik 3.4'e göre her bir düğüm için PR değeri hesaplanmaya başlanır. Bu adım sayfaların PR değeri büyük oranda değişmeye kadar devam eder. Buna yakınsama denir.

PR algoritması ağırlıklı ilişkilere sahip ağlar üzerinde de kullanılabilir. Ağırlık değerinin dikkate alınması ile uygulanan formül eşitlik 3.5'de (Jaganathan ve Desikan 2015) yer almaktadır.

$$PR(v) = (1 - d) + d \sum_{v \in B(u)} PR(v) * W_{(v,u)}^{in} * W_{(v,u)}^{out} \quad (3.5)$$

$$W_{(v,u)}^{in} = \frac{I_u}{\sum_{p \in r(v)} I_p}$$

$$W_{(v,u)}^{out} = \frac{O_u}{\sum_{p \in r(v)} O_p}$$

$B(u) \rightarrow u$ düğümüne bağlantı veren düğümler kümesi

$I \rightarrow$ Gelen ilişkinin sayısı

$O \rightarrow$ Giden ilişkinin sayısı

$r(v) \rightarrow v$ düğümüne bağlantı veren düğümler kümesi

PR algoritması arama motorunda arama yapılması sonucu sayfaları puanlandırıp sıralama yapması dışında pek çok farklı alan üzerinde kullanılmıştır. Örneğin Twitter’da önemli hesapların tespiti, kamusal alanda trafik akışı ve insan hareketlerini ön görebilme gibi farklı alanlarda yapılan çalışmalarda PR tercih edilmiştir (Gleich 2014).

3.4 Topluluk Tespiti Algoritmaları

Düğüm grupları bazı özelliklere, davranışlara göre kümelenme eğilimi gösterebilirler. Küme içerisindeki düğümlerin birbiri ile olan ilişkileri küme dışındaki düğümler ile olan ilişkilerine göre daha yoğundur.

Çizgelerde bu düğüm gruplarını tespit edebilmek için topluluk tespiti algoritmalarından yararlanılmaktadır. Literatürde pek çok farklı topluluk tespiti algoritmaları bulunmaktadır. Tez çalışmasında üçgen sayma (triangle count), yerel kümelenme katsayısı (local clustering co-efficient), Louvain, etiket yayılımı (label propagation) ve zayıf bağlı bileşenler (weakly connected components) algoritmaları kullanılmıştır.

3.4.1 Üçgen sayma algoritması ve yerel kümelenme katsayısı

Çizgedeki düğümlerin içinde bulundukları üçgen toplulukları sayar. Bir üçgen topluluğu içerisindeki tüm düğümlerin birbiri ile ilişkisi bulunmaktadır (Needham ve Hodler 2019). Bu anlamda çizgenin yerel yapısının ne kadar bağlı olduğunu ifade eder (Menczer vd. 2020).

Spam filtreleme, sosyal ağ analizi, ilişki tahmini (link prediction) gibi konularda tercih edilmesinden dolayı üçgen sayma algoritması dikkat çekmiş ve önem kazanmıştır. Bu nedenle büyük çizgelerde verimli çalışması için çeşitli algoritmalar önerilmiştir (Hasan ve Dave 2017).

$G = (V, E)$ olan bir çizge tanımı için algoritmanın temel yaklaşımında;

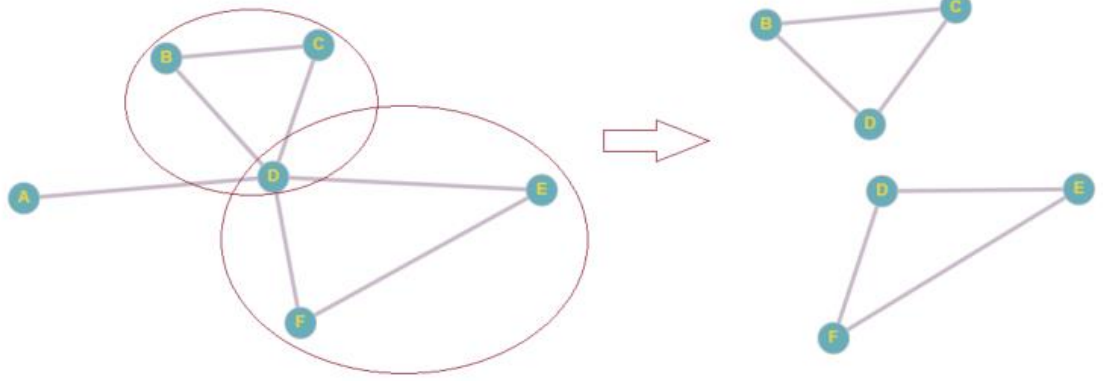
$u, v \in V,$

$uv \in E,$

$S(u) \rightarrow u$ düğümün ilişkili olduğu düğümler kümesi,

$S(v) \rightarrow v$ düğümün ilişkili olduğu düğümler kümesi,

olmak üzere u ve v düğümlerinin bulunduğu üçgen sayısı $|S(u) \cap S(v)|$ şeklinde hesaplanır. Bu durumda u düğümünün bulunduğu toplam üçgen sayısı $\sum_{v \in S(u)} |S(u) \cap S(v)|$ şeklindedir (Becchetti vd. 2008).



Şekil 3.10 Üçgen sayma algoritması örneği

Yerel kümelenme katsayısı bir düğümün komşularının birbiri ile ne kadar ilişkili olduğunu göstermek için kullanılır. Yönsüz bir çizgenin kümelenme katsayısı çizgede bulunan üçgen topluluklarının sayısına karşılık gelmektedir (Gürsakal 2018). Buna göre üçgen sayma ve kümelenme katsayısı birbiri ile ilişkili olan algoritmalarıdır.

Yerel kümelenme katsayısı eşitlik 3.6 ile hesaplanmaktadır. Bu eşitlik 0 - 1 arasında değer alır.

$$C_i = \frac{2L_i}{k_i(k_i-1)} \quad (3.6)$$

$k_i \rightarrow i$ düğümünün derecesi

$L_i \rightarrow i$ düğümünün komşularının birbiri arasındaki ilişki sayısı

C_i değeri 1'e yaklaştıkça i düğümünün komşularının birbiri arasında ilişki bulunma durumu artmaktadır.

3.4.2 Louvain algoritması

Blondel ve arkadaşları (2008) tarafından ortaya konmuş, büyük ölçekli çizgelerde kullanılabilen bir topluluk tespiti algoritmasıdır. Topluluk tespiti algoritmalarından hiyerarşik kümeleme anlayışına dayanmaktadır.

Louvin algoritmasında her düğüm başlangıçta ayrı bir topluluktur. Düğümleri ait oldukları topluluklara yerleştirirken modülerlik kavramını dikkate alır. Bu kavram sayesinde topluluk dışındaki kenarlara göre topluluk içerisindeki kenarların yoğunluğunu ölçülür (Anonymous 2020a). Modülerlik $[-0.5, 1]$ aralığındadır ve -0.5 olması modüler bir topluluk bulunmadığını 1 ise tamamıyla modüler bir topluluk olduğunu ifade etmektedir. Modülerlik formülü eşitlik 3.7 de yer almaktadır.

$$Q = \frac{1}{2m} \sum_{ij} [A_{ij} - \frac{k_i k_j}{2m}] \delta(c_i, c_j) \quad (3.7)$$

$Q \rightarrow$ modülarite

$m \rightarrow$ Çizgedeki tüm ağırlıkların toplamı

(ağırlıksız çizge ise ilişki sayılarının toplamı)

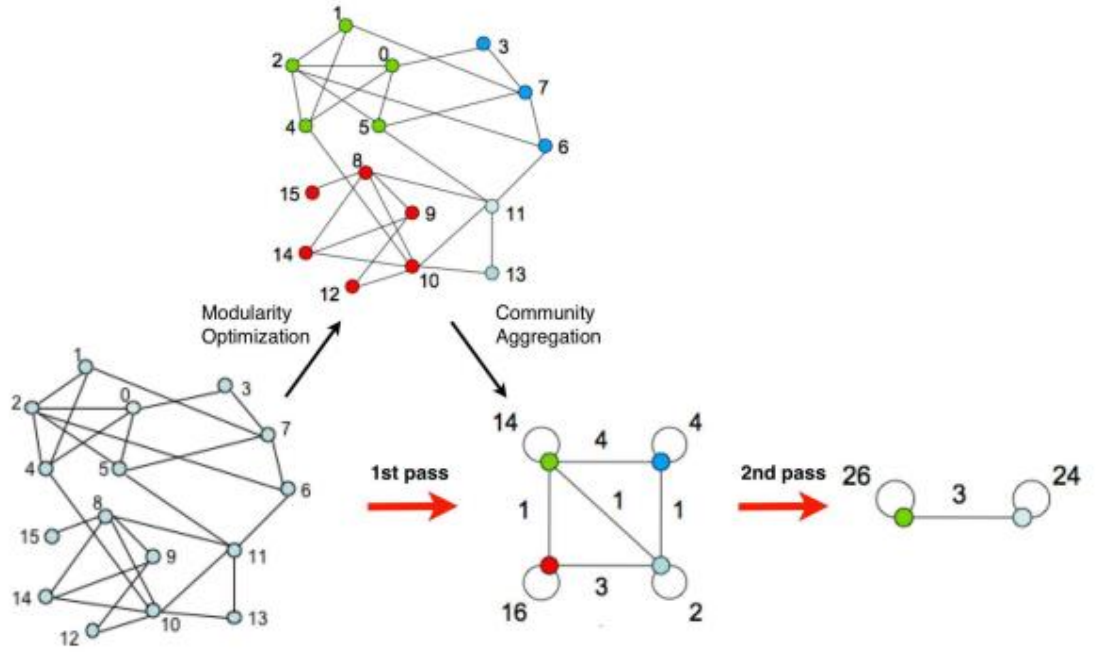
$A_{ij} \rightarrow i$ ve j düğümleri arasındaki ilişkinin ağırlığı,

(ağırlıksız çizgelerde i ve j arasında ilişki varsa 1 yoksa 0 değerini alır)

$k_i \rightarrow i$ düğümüne ait derece

$k_j \rightarrow j$ düğümüne ait derece

$\delta(c_i, c_j) \rightarrow i$ ve j aynı toplulukta ise 1 değilse 0 'dır.



Şekil 3.11 Louvain topluluk tespiti algoritması (Blondel vd. 2008)

Louvain algoritma adımları:

1. Başlangıçta her düğümü farklı bir topluluk olarak atar. Sonrasında her i düğümü komşusu olan j düğümünün bulunduğu topluluğa bağlanır ve modülerite yeniden hesaplanır. Yeni oluşan çizgede modülerlik iyileşmiş ise i düğümü j düğümünün bulunduğu topluluğa dahil olur, iyileşme yoksa bir işlem yapılmaz. Bu işlem iteratif şekilde modülerlikte bir iyileşme olmayana kadar devam eder.
2. 1. adım sonunda belirlenen topluluklara göre yeni bir çizge oluşturulur. Bu çizgede aynı topluluk içerisinde bulunan düğümler tek bir düğüm olarak birleştirilir. Topluluklar arasındaki ilişkiler ise toplanır ve toplulukları temsil eden yeni düğümler arasına eklenir. Bu yeni ilişkinin ağırlığı iki topluluk arasındaki tüm kenar ağırlıklarının toplamı kadardır. Çizgedeki ilişkilerin ağırlıksız olması durumunda, ağırlık değeri topluluklar arasındaki ilişkinin sayısı

kadar olur. Topluluk içerisindeki ilişkilerde aynı şekilde toplanarak öz döngü (self loop) olarak topluluğu temsil eden düğüme eklenir.

3. 2. adım sonunda oluşan yeni çizge için tüm bu adımlar yeniden yapılır. Süreç topluluk yapısında bir değişiklik görülmeğe kadar tekrarlanır.

3.4.3 Etiket yayılımı algoritması

Raghavan ve arkadaşları (2007) tarafından geliştirilen etiket yayılımı algoritması, çizge içerisindeki toplulukları belirlemek için kullanılan hızlı bir algoritmadır. Milyonlarca düğüm ve ilişkisi olan büyük boyutlu çizgeler bu yöntem ile topluluklara bölünebilmektedir (Menczer vd. 2020).

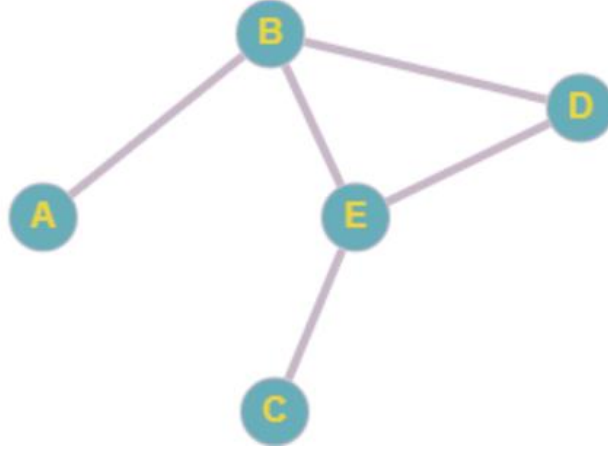
Algoritmada düğümler komşularına bakarak hangi topluluğa dahil olacaklarına karar verirler. Temel anlayış çizgenin yoğun ilişkilerinin bulunduğu kısımda etiketin hızla yayılacağı, daha seyrek ilişkilerin bulunduğu kısımda ise etiketin yayılmasının zorlaşacağı yönündedir. Bu şekilde çizgedeki topluluklar elde edilmektedir. Aynı etiket değerini alan düğümler aynı topluluk içerisinde yer almaktadır (Needham ve Hodler 2019).

Algoritma başlangıcında çizgedeki bazı düğümler tohum (seed) düğüm olarak verilebilir. Bu durumda algoritma yarı denetimli öğrenme (semi supervised) modeli yaklaşımına göre davranır.

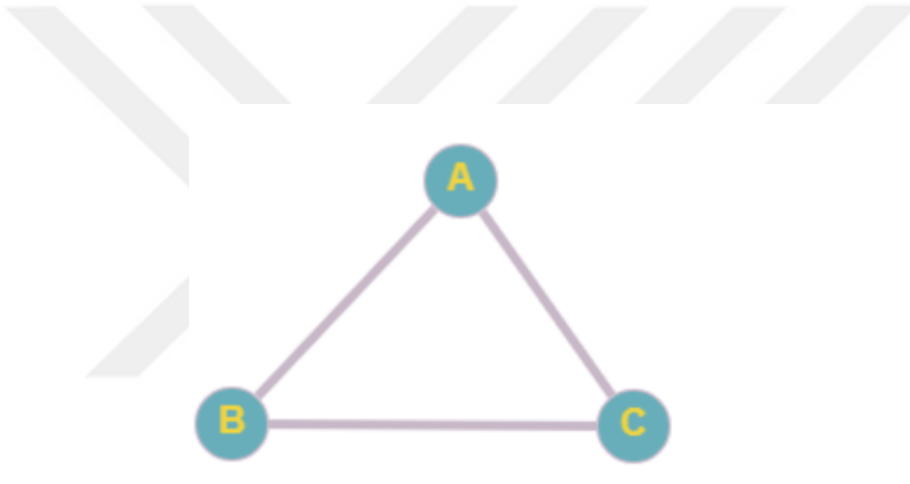
Etiket yayılımı algoritmasında itme (push) ve çekme (pull) olmak üzere iki adet yaklaşım bulunmaktadır (Needham ve Hodler 2019).

İtme yaklaşımına göre algoritma adımları şu şekildedir:

1. Her bir düğüm için ayrı bir etiket atanır veya bu aşamada bazı düğümler tohum olarak etiketlenirken bazıları etiketsiz bırakılabilir.



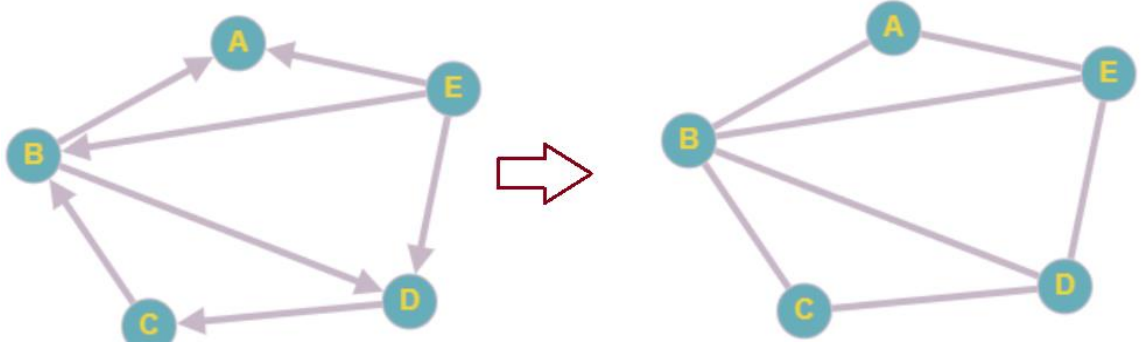
Şekil 3.12 Bağlı yönsüz çizge örneği



Şekil 3.13 Bağlı olmayan yönsüz çizge örneği

Şekil 3.12'deki çizge örneğinde bir düğümden diğer düğümlere ulaşılabilirdi. Ancak Şekil 3.13'te verilen çizge incelendiğinde böyle bir durum söz konusu değildir. Örneğin D düğümünden A düğümüne ulaşmak imkansızdır. Bu tür çizgelere bağlı olmayan (disconnected) çizge denilmektedir.

Yönlü bir çizgenin zayıf bağlı olup olmadığının anlaşılması için yönsüz çizgeye çevrilmiştir. Oluşan yeni çizgede her düğüm çifti arasında bir yol varsa bu çizgeye zayıf bağlı denilmektedir. Bu şekilde yönlü çizgelerin içerisinde bağlı olmayan bir topluluk olup olmadığı kolaylıkla öğrenilebilmektedir.

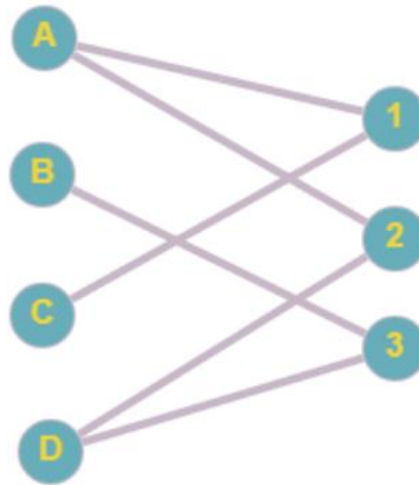


Şekil 3.14 Yönlü çizgeyi yönsüz çizgeye çevirme işlemi

Zayıf bağlı bileşenler algoritmasının düğümler arasındaki ilişki ağırlığına dikkat eden versiyonu da bulunmaktadır. Buna göre belirli bir ağırlık değeri altında olan ilişkilerin dikkate alınmaması sağlanabilmektedir (Anonymous 2020b).

3.5 İki Parçalı Çizgelerde Merkezilik ve Topluluk Tespiti Algoritmalarının Kullanılması

Çizge içerisinde 2 farklı düğüm grubu varsa ve aynı grup içerisinde yer alan düğümler arasında ilişki bulunmuyorsa bu çizgelere iki parçalı çizge (bipartite graph) denmektedir.



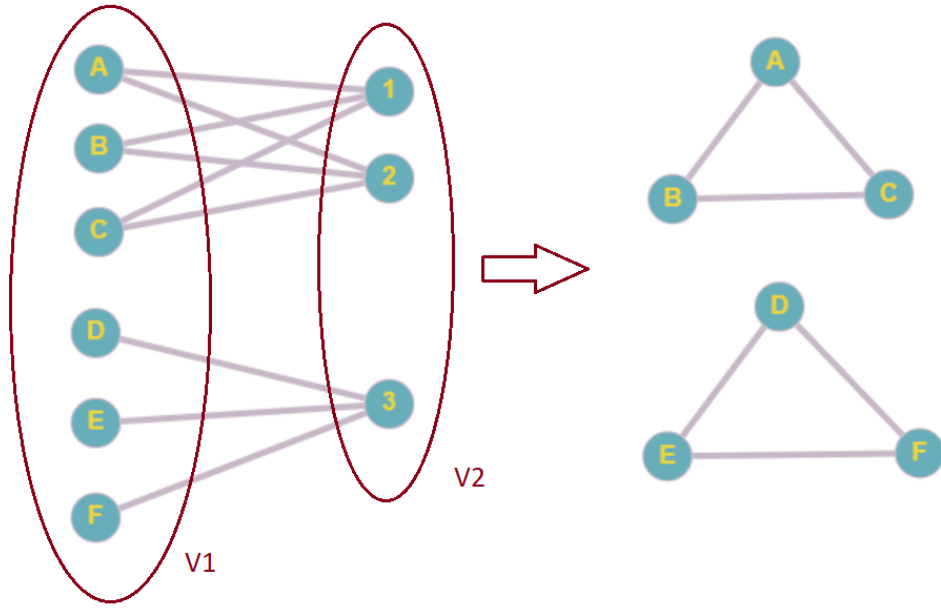
Şekil 3.15 İki parçalı çizge örneği

Analiz işlemlerinde aynı grup içerisindeki düğümlerin birbirleri ile olan ilişkileri gözlemlenmek istediğinde iki parçalı çizgenin tek parçalı çizgeye dönüşmesi yöntemine başvurulabilir.

$G = (V, E)$ tek parçalı çizgesinde, V düğüme E düğümler arasındaki ilişkiye karşılık gelmektedir. İki parçalı çizge tanımı bu durumda $G = (V1, V2, E)$ şeklindedir. İki parçalı çizgenin tek parçalı çizgeye dönüştürülmesi ile analiz işlemindeki ihtiyaca göre $G = (V1, E1)$ veya $G = (V2, E2)$ tek parçalı çizge tanımları elde edilebilir.

Tez çalışmasında tarifler ve malzemeler arasındaki ilişki iki parçalı çizgeyi temsil etmektedir. Yani tarif ve malzeme düğümlerinin kendi arasında bir ilişkisi bulunmamaktadır. Çizgede ilişki, tarif ve malzeme düğümleri arasındadır. Malzemeler üzerinde merkezilik ve topluluk tespiti algoritmalarının uygulanması için bu yöntemle başvurulmuş böylelikle malzemelerden oluşan tek parçalı çizge üzerinde analiz çalışmaları yapılmıştır.

İki parçalı çizgeden tek parçalı çizgeye dönüştürülme işlemi aynı grupta bulunan düğümler arasında ilişki kurulmasıyla yapılır. Analiz işlemi yapılmak istenen düğüm grubu tek parçalı çizgeye belirli kurallar dahilinde çevrilir.

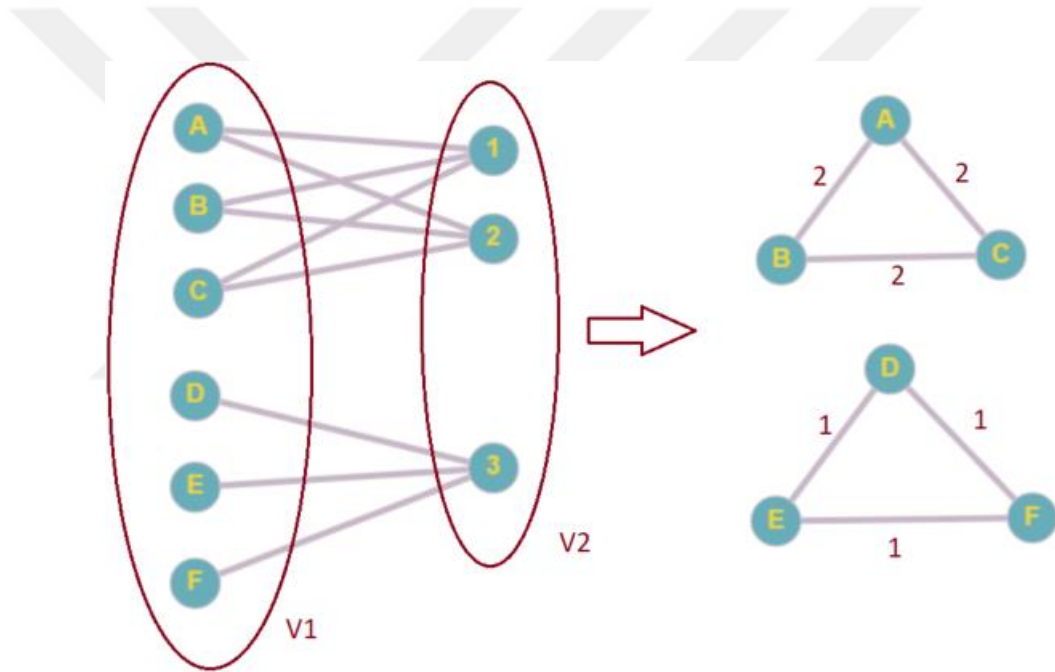


Şekil 3.16 İki parçalı çizgenin tek parçalı ağırlıksız çizgeye dönüştürülmesi

Çizelge 3.7 V1 grubundaki düğümlerin V2 grubunda ilişkili olduğu düğümler

V1 grubu	V2 Grubunda İlişkili Olduğu Düğümler
A	1,2
B	1,2
C	1,2
D	3
E	3
F	3

Şekil 3.16’da verilen çizgede V1 ve V2 olmak üzere iki farklı düğüm grubunun olduğu ve V1 düğüm grubunu içeren tek parçalı ağırlıksız çizgeye dönüştürülme işleminin uygulanıldığı görülmektedir. V1 grubundaki düğümler arasında ilişki kurulurken V2 grubundaki hangi düğümler ile ilişkili olduklarına bakılır. V2 grubundan aynı düğümler ile ilişkisi bulunan V1 grubuna ait düğümler arasında ilişki kurulur (Menczer vd. 2020). Örneğin Çizelge 3.7 incelendiğinde A, B ve C düğümleri V2 düğüm grubundan 1 ve 2 düğümleri ile ilişkili olduğundan oluşacak tek parçalı çizgede A – B, A – C, B – C ilişkileri kurulur. Aynı şekilde D, E ve F düğümleri V2 grubundan 3 ile ilişkilidir ve oluşan tek parçalı çizgede bu düğümler arasında D-E, D-F ve E-F ilişkileri oluşturulur. Böylece tek parçalı çizgeye çevirme işlemi tamamlanır.



Şekil 3.17 İki parçalı çizgenin tek parçalı ağırlıklı çizgeye dönüştürülmesi

Şekil 3.17’de V1 düğüm grubundan oluşan tek parçalı ağırlıklı çizgeye dönüştürme işlemi yapılmıştır. Bu işlem yapılırken ağırlıksız çizgeye dönüştürme işlemi ile aynı süreç takip edilir. Aradaki tek fark oluşan ilişkiler ağırlığa sahiptir. Ağırlık değeri belirlenirken V1 grubundaki düğümlerin V2 grubundaki düğümler ile ne kadar ortak ilişkisinin olduğu durumuna göre belirlenebilir (Kronmueller vd. 2018). Örneğin A, B ve C düğümleri V2 grubundan 1 ve 2 düğümleri ile ilişkisi bulunmaktadır. Yani V2

grubundan 2 ortak ilişkileri vardır. Bu nedenle A, B ve C düğümleri arasındaki oluşan ilişkinin ağırlığı 2 olarak belirlenmiştir. D, E ve F düğümlerinin V2 grubundan yalnızca 3 ile ilişkisi bulunduğundan bu düğümler arasındaki ilişkinin ağırlığı 1 olmaktadır.



4. DENEYSEL ÇALIŞMA VE ARAŞTIRMA BULGULARI

Bu bölümde hazırlanan Türk Mutfağı veri seti üzerinde yapılan depolama ve analiz çalışmalarına yer verilmiştir.

4.1 Veri Tabanı Oluşturulması

Hazırlanan veri seti üzerinde çeşitli analiz çalışmalarının yapılabilmesi için ilk olarak verilerin çizge veri tabanı üzerinde modellenmesi sağlanmıştır. Modelde Tarif, Malzeme ve Bileşen olmak üzere 3 düğüm türü bulunmaktadır. Düğüm türlerinin kendi içlerinde ilişkisi bulunmazken ilişkiler Tarif - Malzeme düğümleri ile Malzeme - Bileşen düğümleri arasında yer almaktadır.

Düğümlere ait özellikler Çizelge 4.1’de, düğümler arasındaki ilişki ve ilişkilere ait özellikler ise Çizelge 4.2’de yer verilmiştir.



Şekil 4.1 Veri tabanı modeli

Çizelge 4.1 Düğüm türleri ve bu türlere ait özellikler

Düğüm	Özellikleri
Tarif	✓ Tarif ismi ✓ Bölge ismi
Malzeme	✓ Malzeme ismi ✓ Ait olduğu kategori
Bileşen	✓ Bileşen ismi

Çizelge 4.2 İlişki türleri ve bu türlere ait özellikler

İlişki	Özellikleri
Tarif -> Malzeme	✓ Miktar ✓ Birim (g)
Malzeme -> Bileşen	✓ Birim (g, mg, vb.) ✓ Minimum miktar ✓ Ortalama miktar ✓ Maksimum miktar

Ek1’de veri tabanındaki malzemelerin içerdiği bileşen türlerinin tamamı bulunmaktadır.

4.2 Türk Mutfağı Üzerinde Yapılan Genel Analiz Çalışmaları

4.2.1 Malzemelere göre tarif önerisi

Tariflerden oluşan veri setinde kişiler sahip oldukları malzemelere göre bir tarif seçebilirler. Örneğin soğan, sarımsak, ceviz, domates salçası, yumurta gibi malzemeler ile yemek yapmak isteyen bir kişiye veri setinde bulunan tariflerden Kallili Köfte önerilmektedir. Bu durumun terside mümkündür örneğin kişi alerjisinin bulunduğu bir malzemenin olmadığı tarifleri listeleterek isteği bir tarifi yapabilmektedir.

4.2.2 Bileşenlere göre tarif önerisi

Kişilerin çeşitli günlük gıda takviyelerine ihtiyacı bulunabilmektedir. Bu ihtiyaçları doğal olarak besinlerden almak isteyen kişiler, hazırlanan veri seti ile ihtiyacı olan bileşene göre tarif belirleyebilmektedir. Çizelge 4.3’te çeşitli bileşenler yönüyle en zengin tariflere yer verilmiştir. Bileşen miktarları bileşenin malzemede ortalama bulunma miktarına göre hesaplanmıştır.

Çizelge 4.3 Çeşitli bileşenler yönüyle zengin tarifler ve bileşenlerin tariflerde bulunma miktarları

Bileşen	Bileşen Yönüyle En Zengin Tarif	Bileşen Miktarı	Bileşen Birimi
B-12 Vitamini	Ciğer Taplemesi	167.73	µg
C Vitamini	Akdeniz Usulü Sebzeli Salata	220.31	mg
Kalsiyum	Akciğerli Yaprak Sarma	2178.40	mg
Demir	Madımak Yemeği	31.33	mg
D vitamini	Deniz Börülceli Çipura Salatası	2592.50	iu

Veri setindeki tarifler üzerinde kalori hesaplaması da yapılabilmektedir. Örneğin veri seti içerisinde kalori hesaplamasında en yüksek kalorili tarifin 2079.17 kcal değeri ile “Katmer / Gaziantep”, en düşük kalorili tarifin ise 27.79 kcal ile Mırza olduğu tespit edilmiştir.

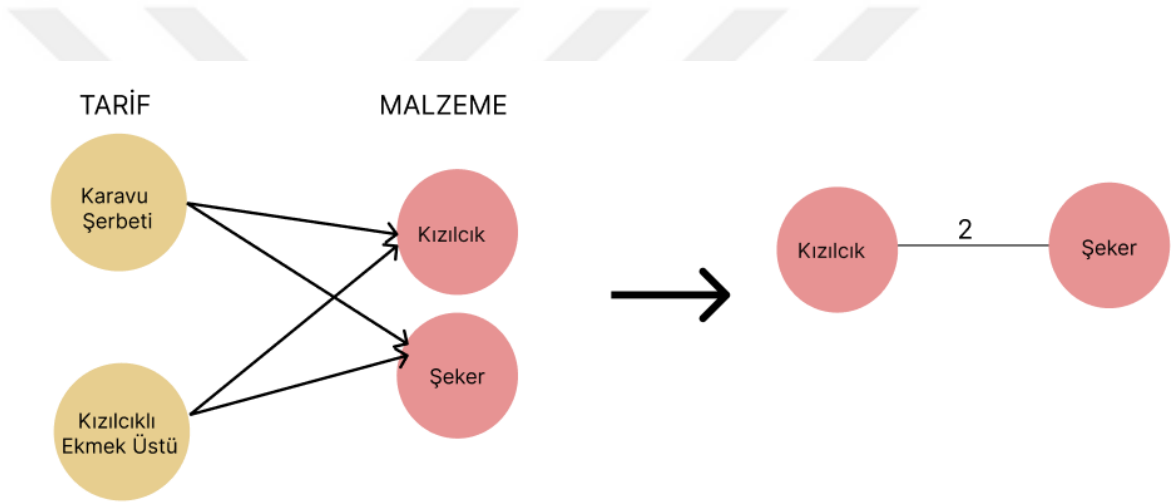
4.3 Tariflerde Kullanılan Malzemelerin Merkezilik Algoritmalarına Göre Değerlendirilmesi

Malzemeler üzerinde merkezilik algoritmaları çalıştırılarak önem derecesi en yüksek olan malzemelerin bulunması amaçlanmıştır. Bu doğrultuda derece, arasındalık, yakınlık ve PR algoritmaları kullanılmıştır. Çizelge 4.4’te su ve baharatlar dışında Türk Mutfağında kullanılan malzemeler arasında derece merkeziliği algoritmasına göre en yüksek değere sahip 20 adet malzemeye yer verilmiştir. Ayrıca çizelgede derece merkeziliği algoritması dışında, arasındalık, yakınlık ve PR algoritmalarından elde edilen sonuçlar yer almaktadır. Türk Mutfağındaki tüm malzemelere ait derece, arasındalık, yakınlık ve PR algoritma sonuçları ise EK 2’de bulunmaktadır.

Çizelge 4.4 Türk Mutfağında su ve baharatlar dışında en yüksek dereceye sahip ilk 20 malzeme için merkezilik algoritması sonuçları

Malzeme İsmi	Derece Merkeziliği Algoritma Sonucu	Arasındalık Merkeziliği Algoritma Sonucu	Yakınlık Merkeziliği Algoritma Sonuuc	PageRank (PR) Merkezilik Algoritma Sonucu
Soğan	142	1244,864	0,744	8,831
Tereyağı	138	1586,528	0,737	7,762
Buğday Unu	117	811,871	0,678	6,711
Yumurta (Tavuk)	93	631,545	0,657	5,459
Ayçiçek Yağı	92	623,619	0,680	5,595
Kırmızı Et (Büyükbaş)	85	333,836	0,641	5,155
Zeytinyağı	83	1586,873	0,764	5,677
Sarımsak	78	799,965	0,708	5,140
Domates	75	532,913	0,671	4,778
Şeker (Beyaz)	71	835,818	0,663	4,159
Yoğurt	66	435,363	0,648	4,017
Domates Salçası	63	419,701	0,655	4,340
Maydanoz	61	521,886	0,669	4,095
Limon	56	1050,009	0,697	3,861
Nane	50	357,210	0,635	3,383
Bulgur	45	121,305	0,594	2,786
Sivri Biber	41	142,113	0,602	2,510
Biber Salçası	40	132,913	0,599	2,707
Ceviz	38	303,778	0,617	2,355
Pirinç	35	183,452	0,605	2,342

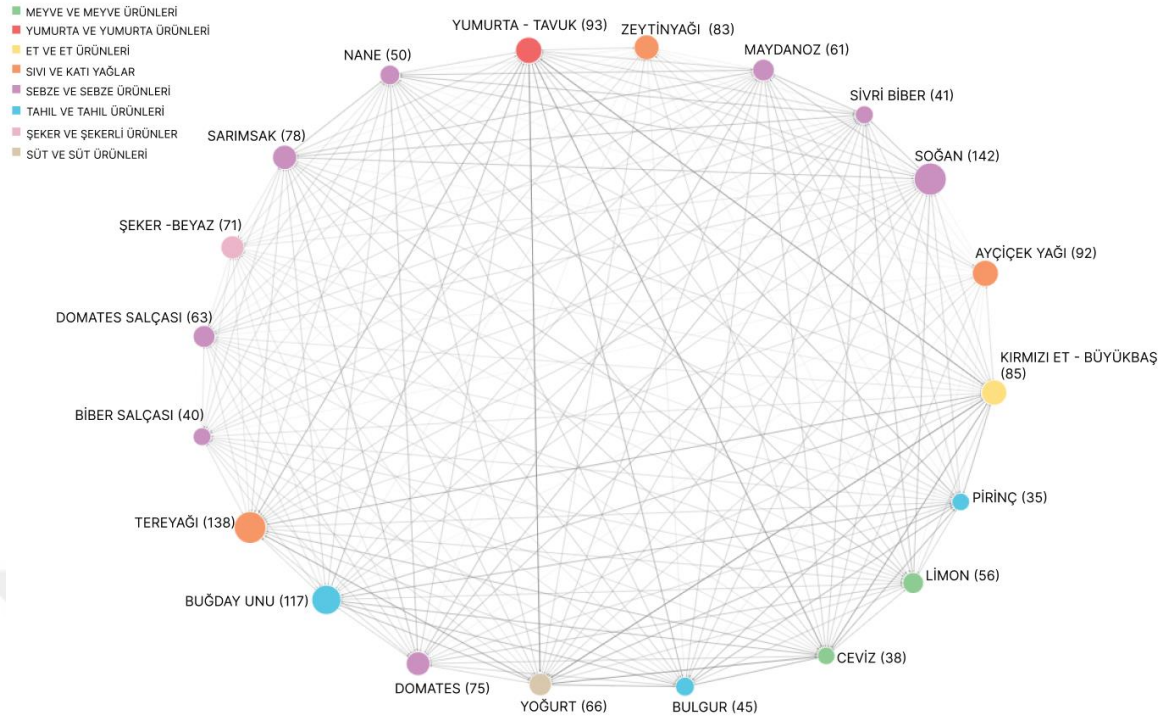
Arasındalık, yakınlık ve PR merkeziliği algoritmalarında malzemelerin merkezilik değerlerini ölçmek için ilk olarak Tarif ve Malzeme düğümleri arasındaki iki parçalı çizge malzemeler arasında kurulan tek parçalı, yönsüz ve ağırlıklı çizgeye dönüştürülmüştür. Malzemeler arasında ilişkinin ağırlığı malzeme çiftinin tariflerde birlikte bulunma sayısına karşılık gelmektedir. Bu durumda malzemenin komşuları tariflerde birlikte kullanıldığı diğer malzemeleri ifade etmektedir. Örneğin kızılıcık ve şeker aynı tarifte birlikte bulunmuş ise kızılıcık ve şeker arasında bir yönsüz ilişki kurulmuş ve birlikte bulundukları tarif sayısı ilişkiye ağırlık olarak eklenmiştir. Ancak bu ağırlık değeri arasındalık ve yakınlık merkeziliği algoritmalarında göz ardı edilmiştir. Şekil 4.2’de yapılan işleme dair örnek bir süreç bulunmaktadır.



Şekil 4.2 Tarif ve Malzeme arasındaki iki parçalı çizgenin malzemeler arasında kurulan tek parçalı, ağırlıklı ve yönsüz çizgeye dönüştürülmesi

4.3.1 Derece merkeziliği algoritması sonuçlarının değerlendirilmesi

Malzemelerin tüm tarifler içerisinde kaç adet tarifte bulunduğu tespit etmek amacıyla kullanılmıştır.



Şekil 4.3 Türk Mutfağında su ve baharatlar dışında en çok kullanılan 20 malzeme

Çizelge 4.4'te Türk Mutfağındaki su ve baharatlar dışında en çok tercih edilen ilk 20 malzemeye göre Şekil 4.3 elde edilmiştir. Şekil 4.3'te düğümlerin büyüklüğü derecesi ile orantılı olarak görselleştirilmiştir. Malzeme isimleri yanında bulunan sayılar malzemenin derecesini ifade etmektedir. Düğümlerin renkleri ise malzemenin kategorisini belirtmektedir. Malzemeler arasındaki ilişkiler sahip oldukları ortak bileşenlere göre kurulmuştur. İlişkilerin kalınlığı ortak bileşen sayısı ile orantılı olarak artmaktadır.

Şekil 4.3'e bakıldığında en fazla tercih edilen malzemenin soğan olduğu gözlemlenmektedir. Besin gruplarına göre en çok tercih edilen malzemelere ise Çizelge 4.5'te yer verilmiştir.

Çizelge 4.5 Şekil 4.3'e göre kategori bazında en çok tercih edilen malzemeler

Kategori	Malzeme
Et ve Et Ürünleri	Kırmızı Et (Büyükbaş)
Meyve ve Meyve Ürünleri	Limon
Sebze ve Sebze Ürünleri	Soğan
Süt ve Süt Ürünleri	Yoğurt
Sıvı ve Katı Yağlar	Tereyağı
Tahıl ve Tahıl Ürünleri	Buğday unu
Yumurta ve Yumurta Ürünleri	Yumurta (Tavuk)
Şeker ve Şeker Ürünleri	Şeker (Beyaz)

Derece merkeziliği algoritması ile Türk Mutfağında en çok tercih edilen malzemeler tespit edilmiştir. Bu durum hangi malzemelerin üretimine önem verilmeli veya bu malzemelerin tedariğinde oluşacak bir sıkıntının etkisinin büyüklüğü ne olabilir gibi çeşitli konularda önem kazanabilir.

4.3.2 Arasındalık merkeziliği algoritması sonuçlarının değerlendirilmesi

Arasındalık algoritması uygulaması öncesinde malzemeler arasında oluşturulan tek parçalı çizgede ağırlık bilgisi ihmal edilmiştir. Bu şekilde malzeme çiftinin birlikte kaç tarifte kullanıldığı önemsenmeyip birbirleri ile olan kullanım uyumu dikkate alınabilmiştir.

Çizelge 4.4 incelendiğinde zeytinyağının arasındalık değerinin diğer malzemelere göre daha fazla olduğu görülmektedir. Bu da çizgede birbiri ile aynı tarifte bulunmayan diğer

malzeme çiftleri arasındaki en kısa yollarda en fazla zeytinyağının bulunduğunu ifade etmektedir.

Malzeme çiftinin en kısa yol üzerinde bulunan malzemeler ile birlikte kullanılması, en kısa yol dışındaki diğer yollarda bulunan malzemeler ile birlikte kullanımından daha uyumlu olarak değerlendirilebilir. Bu durum birbiri ile aynı tarifte olmayan malzemelerden oluşturulan bir tarifte, Çizelge 4.4'te verilen malzemeler arasında en yüksek zeytinyağının bulunma ihtimalinin olduğunu göstermektedir.

4.3.3 Yakınlık merkeziliği algoritması sonuçlarının değerlendirilmesi

Yakınlık merkeziliği algoritması çizgede bir düğümün diğer düğümlere ne kadar yakın olduğunu ifade etmektedir. Malzemeler arasında kurulan tek parçalı çizgede malzemenin komşuları, tarifte birlikte kullandığı malzemeleri içermektedir. Yani en yakınında bulunanlar tarifte birlikte kullanıldığı malzemeler olmaktadır. Bu nedenle uygulanan yakınlık algoritması sonuçları bir malzemenin ne kadar farklı malzeme ile birlikte kullanıldığı veya kullanılabileceği konusunda fikir verebilmektedir.

Malzemeler arasında kurulan tek parçalı çizge üzerinde yakınlık algoritması çalıştırılmıştır. Algoritma uygulanırken malzemeler arasındaki ilişki ağırlığı dikkate alınmamıştır. Çizelge 4.4'te yakınlık merkeziliği algoritmasına göre en yüksek değerin zeytinyağı olduğu görülmektedir. Bu da zeytinyağının çizgede daha fazla malzeme çeşidi ile kullanılma ihtimalinin yüksek olduğunu göstermektedir. Çizelge 4.4'te verilen malzemelerin yakınlık algoritmasına göre sonuçları ve kaç farklı malzeme ile birlikte kullanıldığı bilgisi Çizelge 4.6'da yer almaktadır. Burada farklı malzemeler ile kullanım sonuçlarının yakınlık merkeziliği algoritması ile uyumlu olduğu görülmektedir. Örneğin Çizelge 4.6'da en yüksek yakınlık merkeziliği değerine sahip zeytinyağının yine farklı malzemeler ile kullanım sayısının en fazla olduğu görülmektedir.

Çizelge 4.6 Çizelge 4.4’te verilen malzemelere ait yakınlık merkeziliği algortiması sonucu ve kaç farklı malzeme ile birlikte kullanıldığı bilgisi

Malzeme İsmi	Birlikte Kaç Farklı Malzeme ile Kullanıldığı Bilgisi	Yakınlık Merkeziliği Algoritma Sonuçları
Zeytinyağı	159	0,764
Soğan	151	0,744
Tereyağı	148	0,737
Sarımsak	135	0,708
Limon	130	0,697
Ayçiçek Yağı	122	0,680
Buğday Unu	121	0,678
Domates	117	0,671
Maydanoz	116	0,669
Şeker (Beyaz)	113	0,663
Yumurta (Tavuk)	110	0,657
Domates Salçası	109	0,655
Yoğurt	105	0,648
Kırmızı Et (Büyükbaş)	101	0,641
Nane	98	0,635
Ceviz	87	0,617
Pirinç	80	0,605
Sivri Biber	78	0,602
Biber Salçası	76	0,599
Bulgur	73	0,594

4.3.4 PageRank merkezilik algoritması sonuçlarının değerlendirilmesi

Malzemelerden oluşan tek parçalı yönsüz çizge üzerine ilişki ağırlığı dikkate alınarak PR algoritması çalıştırılmıştır. PR değerine göre bir çizgede düğümün önem derecesi ona bağlanan düğümün önem derecesi ile doğru orantılıdır.

Çizelge 4.4'te bulunan malzemeler arasında en yüksek PR değerinin soğan olduğu görülmektedir. Bu sonuca göre soğan malzemesinin ilişkisinin yine çizgede ilişki sayısı ve ağırlığı fazla olan diğer malzemeler ile yani daha popüler malzemeler ile bulunduğu bir göstergesidir. Yine Çizelge 4.4'te zeytinyağı ve kırmızı ete bakılacak olursa, zeytinyağın derecesinin kırmızı etin derecesinden az olmasına rağmen PR değerinin kırmızı ete göre daha fazla olduğu görülmektedir. Bu durumda zeytinyağının ilişkili olduğu malzemelerin kırmızı etin ilişkili olduğu malzemelerden daha popüler, önemli olduğunu söylemek mümkündür.

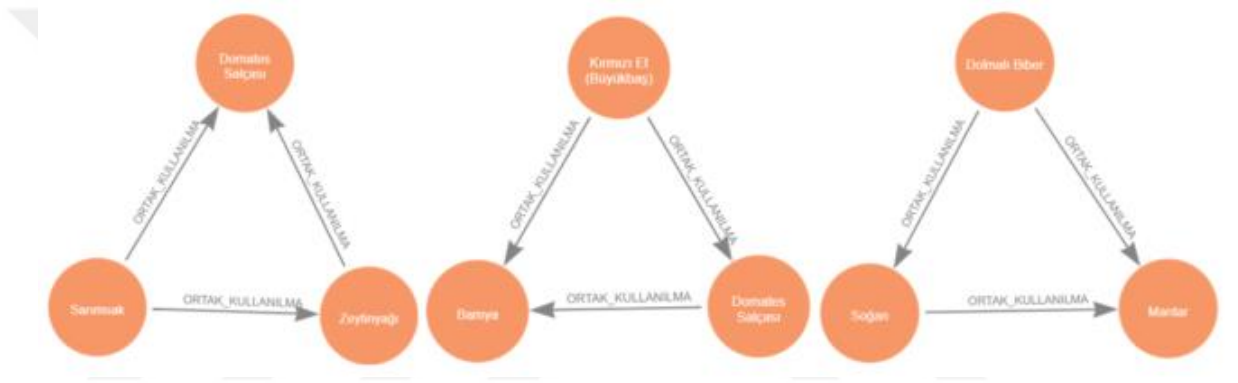
4.4 Tariflerde Kullanılan Malzemeler için Topluluk Tespiti Algoritmalarının Çalıştırılması

Tariflerde kullanılan malzemeler arasında bir topluluk olup olmadığının tespiti için yine malzemeler arasında tek parçalı, yönsüz, ağırlıklı bir çizge oluşturulmuştur. Oluşturulma yöntemi Şekil 4.2'de özetlenen yöntem ile aynıdır.

Topluluk tespiti algoritmalarından üçgen sayma (triangle count), yerel kümelenme katsayısı (local clustering coefficient), Louvain, etiket yayılımı (label propagation) ve zayıf bağlı bileşenler (weakly connected components) algoritmaları kullanılmıştır. Algoritmalarından elde edilen sonuçlar alt başlıklar halinde detaylandırılmıştır.

4.4.1 Üçgen sayma algoritması ve yerel kümelenme katsayısına göre malzemelerin değerlendirilmesi

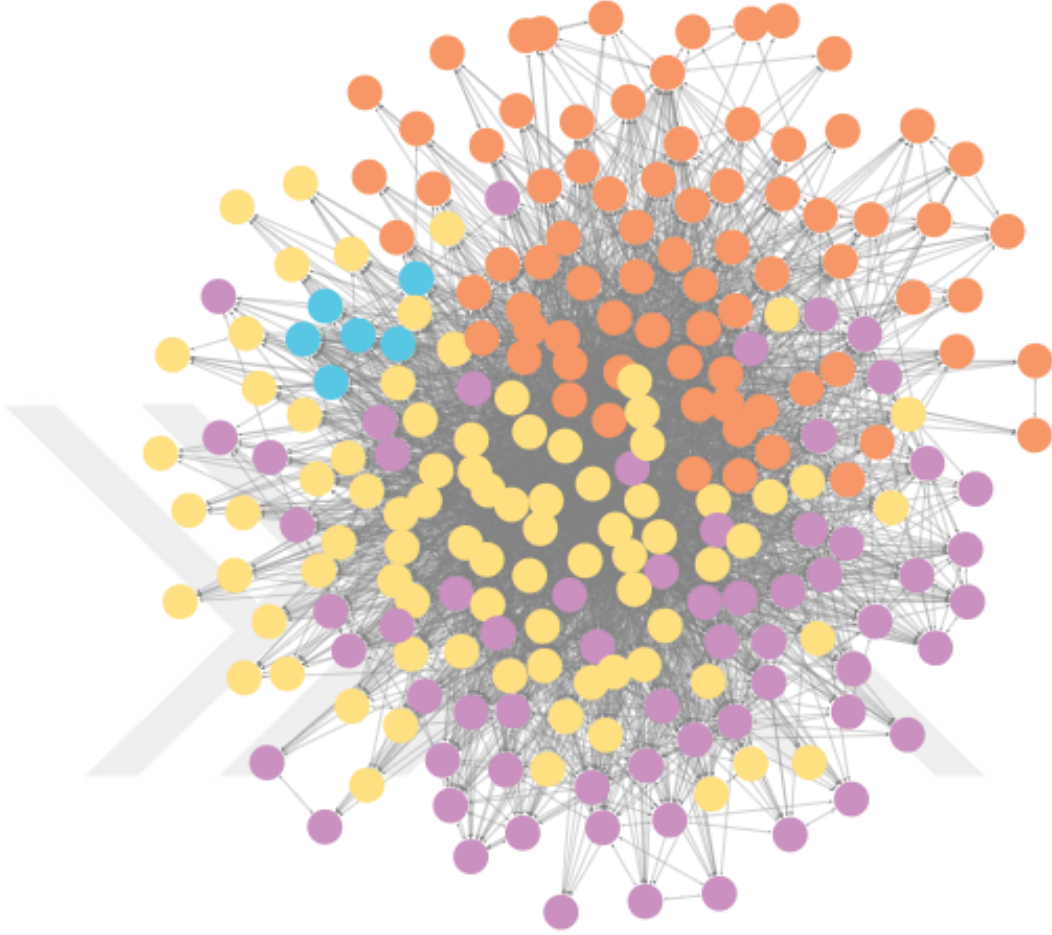
Üçgen sayma algoritması malzemelerden oluşan tek parçalı yönsüz çizge üzerinde çalıştırıldığında 35.757 adet üçgen ilişki olduğu tespit edilmiştir. Üçgen sayma algoritması yerel olarak çizgedeki düğümlerin ne kadar bağlı olduğunu ifade etmektedir. Buna göre malzemelerden oluşan çizgenin birbiri ile ilişkilerinin güçlü olduğunu, malzemelerin birbirleri ile kullanım durumunun yoğun olduğunu göstermektedir. Şekil 4.4'te tespit edilen üçgen ilişkilerden örnekler bulunmaktadır.



Şekil 4.4 Üçgen sayma algoritması ile tespit edilmiş üçgen ilişki örnekleri

Malzemelerin komşularının birbirleri ile olan ilişkilerini gözlemlemek için yerel kümelenme katsayısına da bakılmıştır. Çizge üzerinde yerel kümelenme katsayı algoritması çalıştırıldığında elde edilen ortalama değer 0.79'dur. Bu değer her bir malzemenin komşularının birbirlerine ortalama %79 oranında bağlı olduğunu göstermektedir. Üçgen sayma algoritması ile uyumlu olarak malzemelerin birbirleri ile olan ilişkilerinin kuvvetli olduğu tespit edilmiştir.

4.4.2 Louvian algoritmasına göre malzemelerin gruplandırılması



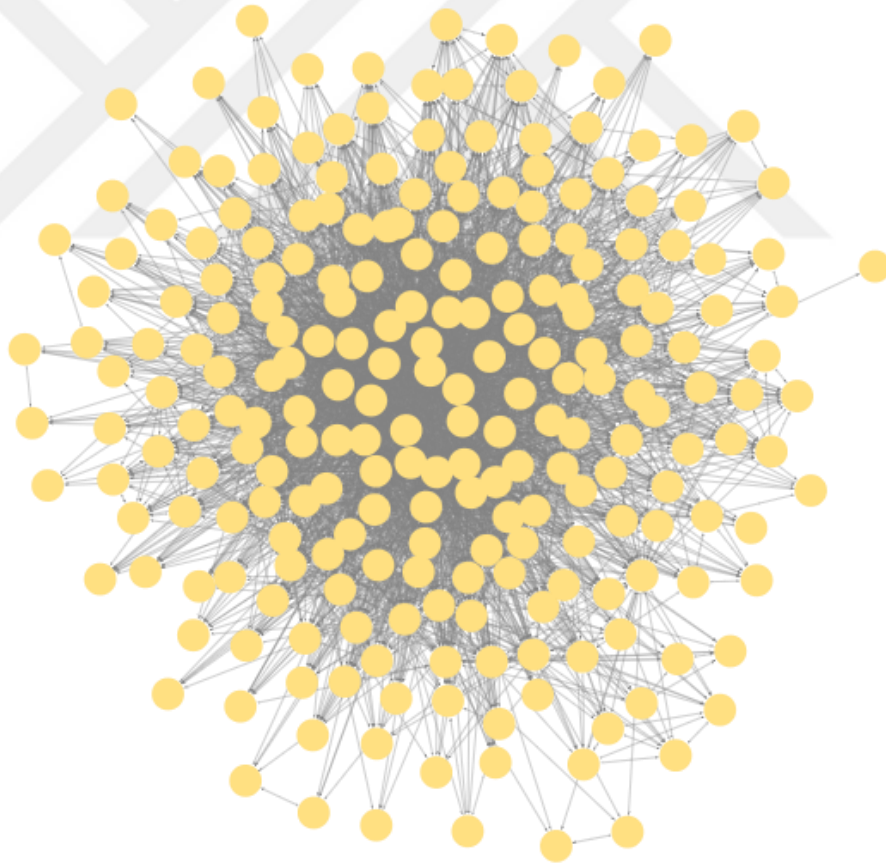
Şekil 4.5 Louvain algoritması sonucu

Şekil 4.5'te malzemelerden oluşan ağ üzerinde çalıştırılan Louvain algoritması sonucu verilmiştir. Buna göre 4 topluluk bulunmuştur. Her bir topluluk farklı renk ile gösterilmiştir. Oluşan topluluklar incelendiğinde 3 büyük topluluk ve 1 küçük topluluk bulunduğu görülmektedir. Toplulukların içerdiği malzeme sayısı Çizelge 4.7'de yer almaktadır.

Çizelge 4.7 Şekil 4.5’te verilen Louvain algoritması sonucuna göre bulunan topluluklar ve içerdikleri malzeme sayıları

Topluluk	Topluluk İçerisinde Bulunan Malzeme Sayısı
Turuncu	78
Sarı	83
Mor	64
Mavi	6

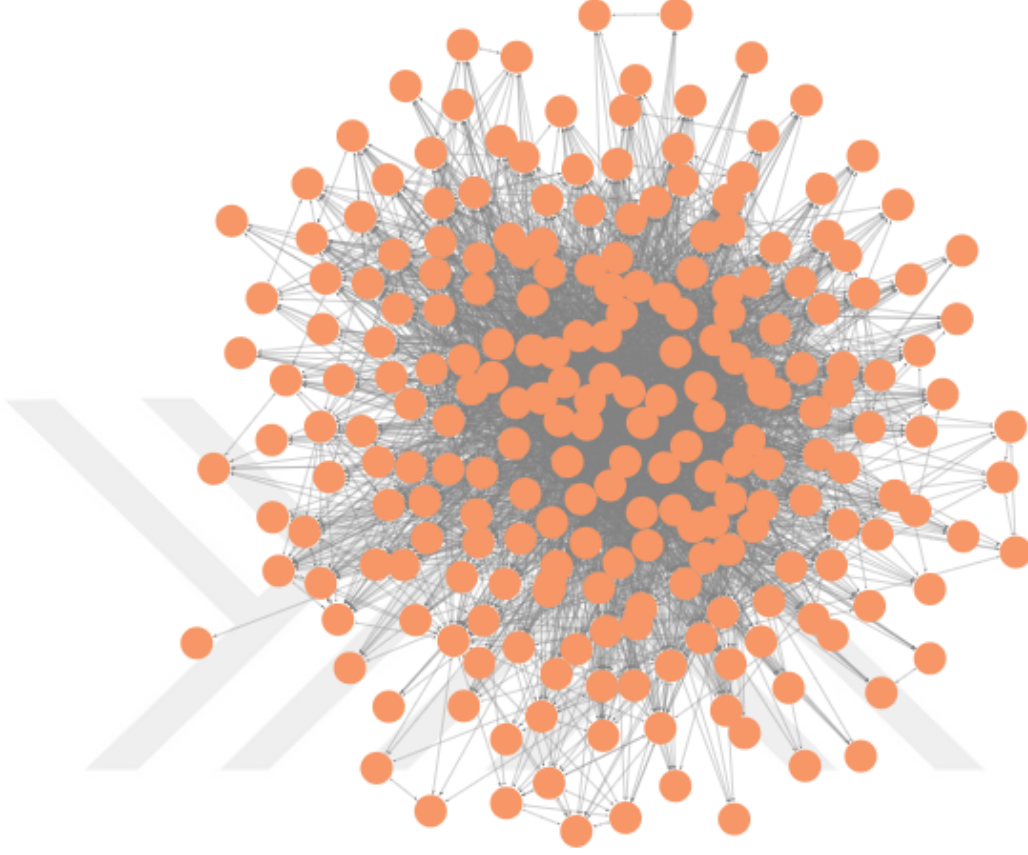
4.4.3 Etiket yayılımı algoritmasına göre malzemelerin gruplandırılması



Şekil 4.6 Etiket yayılımı algoritması sonucu

Şekil 4.6’da malzemelerden oluşan çizge üzerine etiket yayılımı algoritması uygulamasının sonucu verilmiştir. Bu algoritmaya göre tek bir topluluk bulunmuştur.

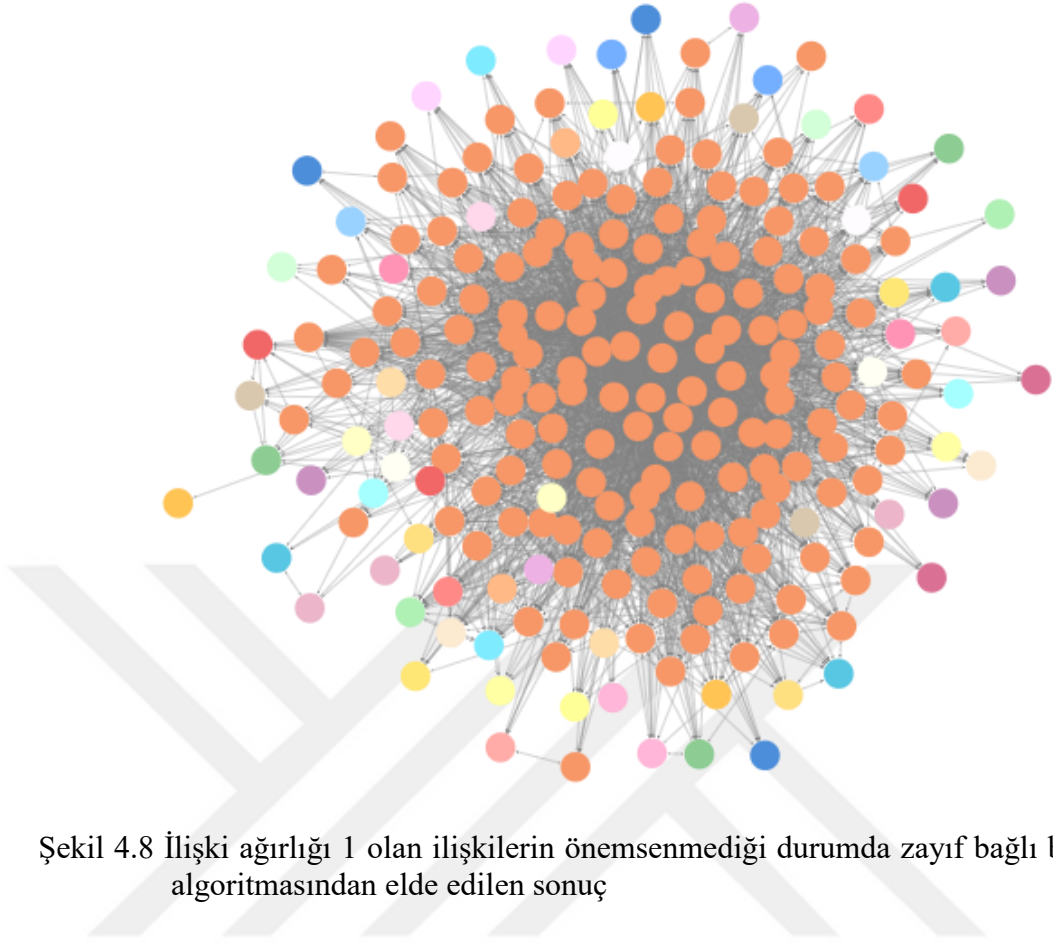
4.4.4 Zayıf bağılı bileşenler algoritmasına göre malzemelerin gruplandırılması



Şekil 4.7 Zayıf bağılı bileşenler algoritması sonucu

Zayıf bağılı bileşenler algoritması çizge içerisinde bağılı olmayan toplulukları tespit etmektedir. Malzemelerden oluşan çizge üzerine uygulandığında tek bir topluluk bulunmuştur. Bu da çizgenin tümüyle birbirine bağılı ve her malzeme çifti arasında bir yol olduğunu göstermektedir.

Algoritma ile belirli bir eşik değerinin altında ağırlığı olan ilişkiler göz ardı edilebilir. Bu şekilde önemsenmeyecek, eşik değerinin altında bulunan ilişkiler dikkate alınmayarak çeşitli çalışmaların ana topluluk üzerinde yapılması tercih edilebilir. Örneğin Şekil 4.8’de ağırlığı 1 olan ilişkiler dikkate alınmamıştır. Algoritma bu durumda 76 farklı topluluk bulmuştur. Her renk bir topluluğu göstermektedir. Ana topluluk dışındaki tüm topluluklar tek bir malzemedan oluşmaktadır.



Şekil 4.8 İlişki ağırlığı 1 olan ilişkilerin önemsenmediği durumda zayıf bağlı bileşenler algoritmasından elde edilen sonuç

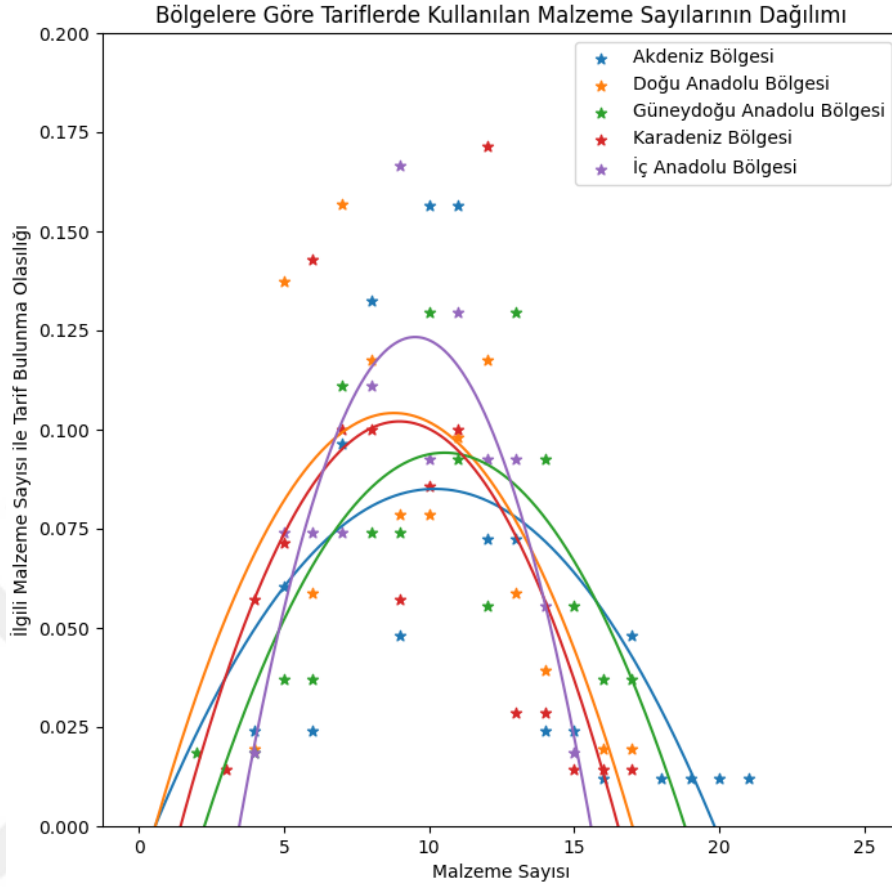
4.5 Bölgesel Bazda Yapılan Analiz Çalışmaları

Türk Mutfağı veri seti Akdeniz, Karadeniz, İç Anadolu, Doğu Anadolu ve Güneydoğu Anadolu olmak üzere 5 farklı bölgeden tarifler içermektedir. Bölgesel bazda çeşitli çıkarımlar yapabilmek için tarifler üzerinde analiz çalışmaları yapılmıştır. Yapılan bu analiz çalışmaları alt başlıklar halinde verilmiştir.

4.5.1 Bölgesel bazda tariflerde kullanılan ortalama malzeme sayısı

Tarif içerisinde kullanılan malzeme sayısı bir tarifin malzeme yönünden zenginliğini göstermektedir. Bölgesel bazda tariflerde kullanılan ortalama malzeme sayısı bilgisi ile ilgili bölgeye ait tariflerdeki lezzet çeşitliliği hakkında fikir sahibi olunabilir.

Çizelge 4.8 Bölgelere göre tariflerde kullanılan malzeme sayılarının dağılımı



Çizelge 4.8’de bölgelere ait tariflerde kullanılan malzemelerin dağılımları verilmiştir. Çizelgedeki y eksenini bölgedeki tarif başına kullanılan malzeme sayısı tekrarının bölgedeki tüm tarif sayısına bölümü ile elde edilmiştir. Örneğin Akdeniz bölgesinde 20 adet malzeme içeren 1 tarif bulunmaktadır ve Akdeniz bölgesindeki toplam tarif sayısı 83’dür. Bu durumda $1/83 = 0.0120$ oranında Akdeniz Bölgesinde 20 malzeme ile tarif bulunabileceği çıkarımı yapılabilir.

Genel olarak tüm bölgeler değerlendirildiğinde ise tariflerin 10 veya 10’a yakın malzemeden oluştuğunu söylemek mümkündür.

4.5.2 Bölgesel bazda bileşen değerlerine göre yapılan analiz çalışması

Bölgelerdeki tariflerde tercih edilen malzemelere göre ortalama olarak hangi bileşenden ne kadar alınabildiği öngörülebilir. Çizelge 4.9’da örnek olarak B12 vitamini, c vitamini, kalsiyum, demir ve d vitaminlerinin bölgedeki tariflerde ortalama bulunma miktarları verilmiştir. Bu değerler bölgeye ait tariflerdeki bileşenlerin toplamının bölgedeki toplam tarif sayısına bölünmesiyle elde edilmiştir. Hesaplamalarda bileşenlerin bir malzemede ortalama bulunma miktarı dikkate alınmıştır.

Çizelge 4.9 Bileşenlerin bölgedeki tariflerde ortalama bulunma miktarı

	Akdeniz	Karadeniz	İç Anadolu	Güneydoğu Anadolu	Doğu Anadolu
B12 Vitamini (μg)	3.30	0.88	1.41	1.17	5.73
C Vitamini (mg)	40.47	24.21	19.78	26.69	14.82
Kalsiyum (mg)	161.16	191.28	161.50	118.20	122.28
Demir (mg)	3.55	3.31	3.64	3.48	3.56
D Vitamini (iu)	49.87	42.48	23.96	23.44	34.81

Çizelge 4.9 incelendiğinde D ve C vitaminlerinin Akdeniz, kalsiyumun Karadeniz, B12 vitamininin Doğu Anadolu bölgelerine ait tariflerde daha fazla bulunduğu görülmektedir. Demir miktarı en fazla İç Anadolu Bölgesinde görünmesine rağmen diğer bölgelere ait sonuçlar ile karşılaştırıldığında miktarların birbirlerine yakın olduğunu söylemek mümkündür.

4.5.3 Bölgesel bazda en çok tercih edilen malzemeler

Her bölge için bölgesel bazda en çok tercih edilen 5 malzeme incelenmiştir. İnceleme yapılırken su ve baharatlar göz ardı edilmiştir. En çok tercih edilen malzemeler bölge içerisindeki yaygınlık değeri bulunarak tespit edilmiştir. Yaygınlık değeri hesaplanırken eşitlik 4.1'den (Ahn vd. 2011) yararlanılmıştır.

$$Y_m^b = \frac{n_m^b}{N_b} \quad (4.1)$$

$Y_m^b \rightarrow m$ malzemesinin b bölgesinde yaygınlık değeri

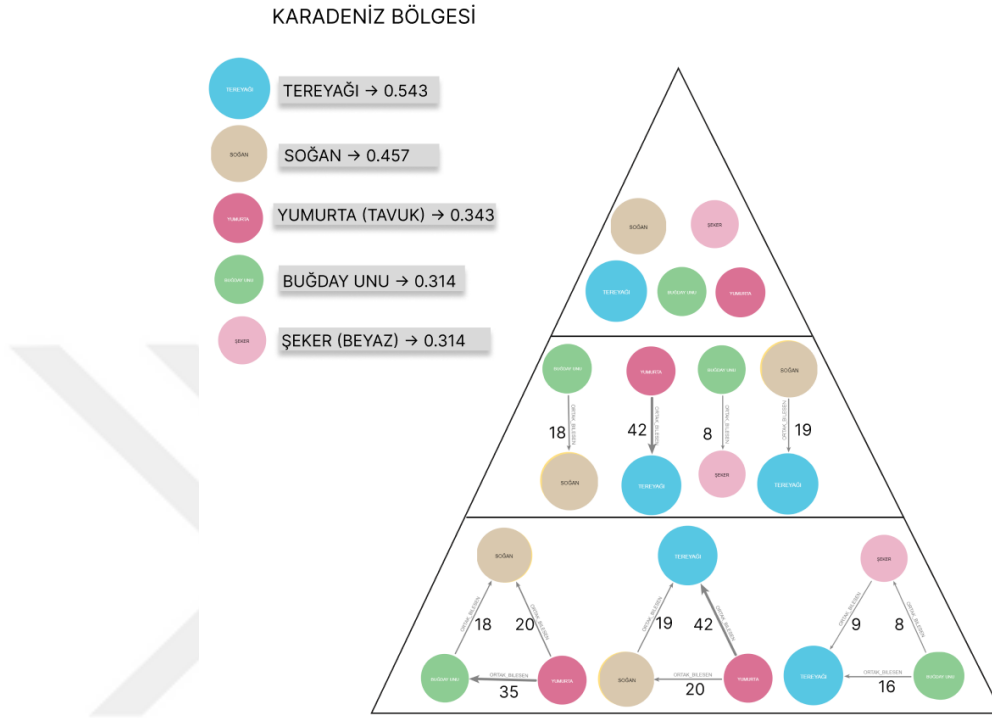
$n_m^b \rightarrow m$ malzemesinin b bölgesindeki tariflerde kullanılma sayısı

$N_b \rightarrow b$ bölgesindeki toplam tarif sayısı

Hesaplama sonuçları bölge bazında Şekil 4.9 – 4.13'te bulunmaktadır. Her bir düğümüne ait yaygınlık değeri düğüm adının yanında yer almaktadır. Düğüm büyüklükleri yaygınlık değeri ile orantılı olarak verilmiştir. Görselleştirmeler piramit şeklinde yapılmıştır. Piramidin en üst kısmında bölge içerisindeki tariflerde en çok tercih edilen 5 malzeme bulunmaktadır. Orta ve taban kısımlarında ise bu malzemeler arasındaki 3'lü ve 2'li gruplara yer verilmiştir. Bu gruplar arasındaki ilişkiler malzeme çiftlerinin sahip olduğu ortak bileşen değerine göre oluşturulmuştur. Sahip oldukları ortak bileşen sayısı ilişkilere ağırlık değeri olarak eklenmiştir. İlişkilerin kalınlığı ise ilişki ağırlığı ile orantılı olarak görselleştirilmiştir.

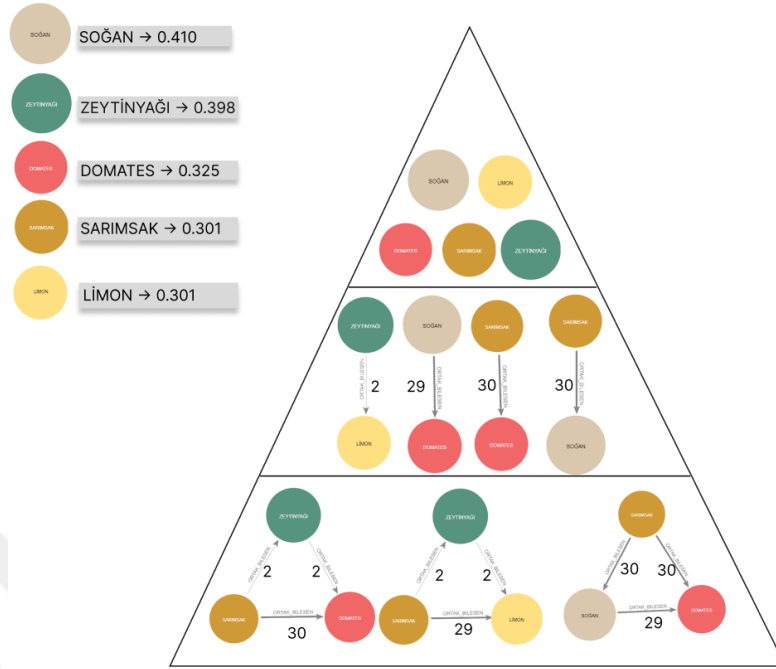
Şekillerde 3'lü ilişkiler bulunurken bölgeye ait tarif ve malzeme arasındaki iki parçalı çizge, malzemelerden oluşan tek parçalı, yönsüz çizgeye dönüştürülmüştür. Yapılan işlem Şekil 4.2'de özetlenen yöntemle aynıdır yalnızca ilişkiler arasında bir ağırlık değeri bulunmamaktadır. Malzemeler arasında tek parçalı çizgeye dönüştürülme işlemi tamamlandıktan sonra üçgen sayma algoritması çalıştırılmıştır. Elde edilen sonuçlar aynı tarifte birlikte kullanılan 3'lü malzeme gruplarını göstermektedir. Şekillerde

bulunan üçgen ilişkilerden örnek olarak 3 tanesine yer verilmiştir. Piramidin orta kısmında yer alan 2'li gruplar ise aynı tarifte kullanılan malzeme çiftlerinden oluşmaktadır. Görselleştirmelerde malzeme çifti örneklerinden 4'üne yer verilmiştir.



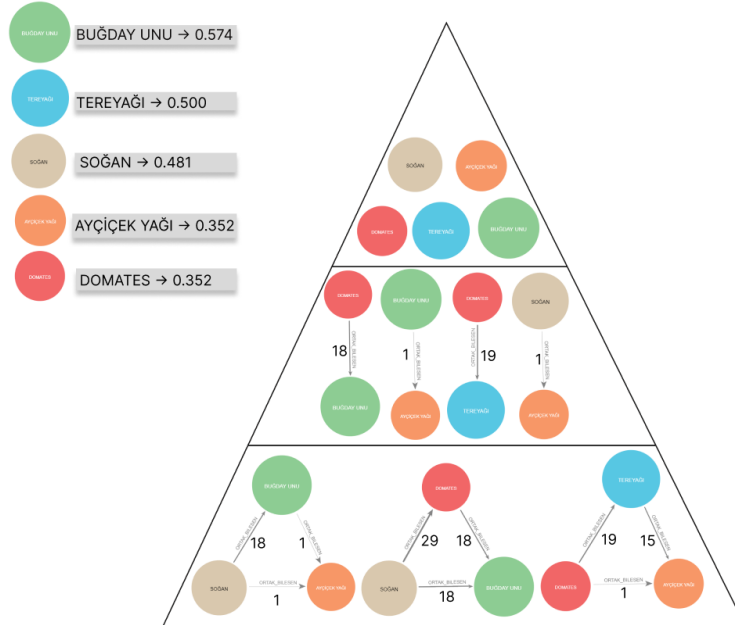
Şekil 4.9 Karadeniz bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları

AKDENİZ BÖLGESİ



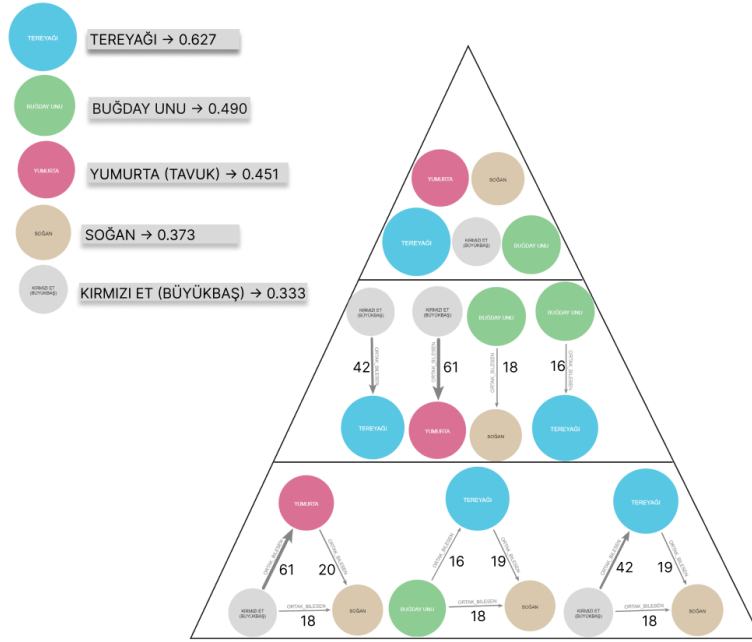
Şekil 4.10 Akdeniz bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları

İÇ ANADOLU BÖLGESİ



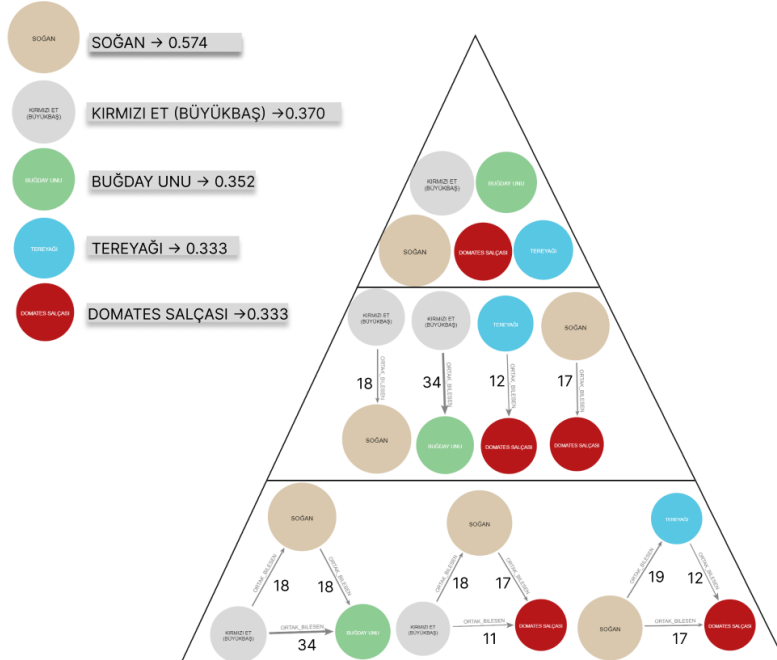
Şekil 4.11 İç Anadolu bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları

DOĞU ANADOLU BÖLGESİ



Şekil 4.12 Doğu Anadolu bölgesinde en çok tercih edilen 5 malzemeye ait yaygınlık değerleri, bu malzemelerin birlikte kullanım durumları ve sahip oldukları ortak bileşen sayıları

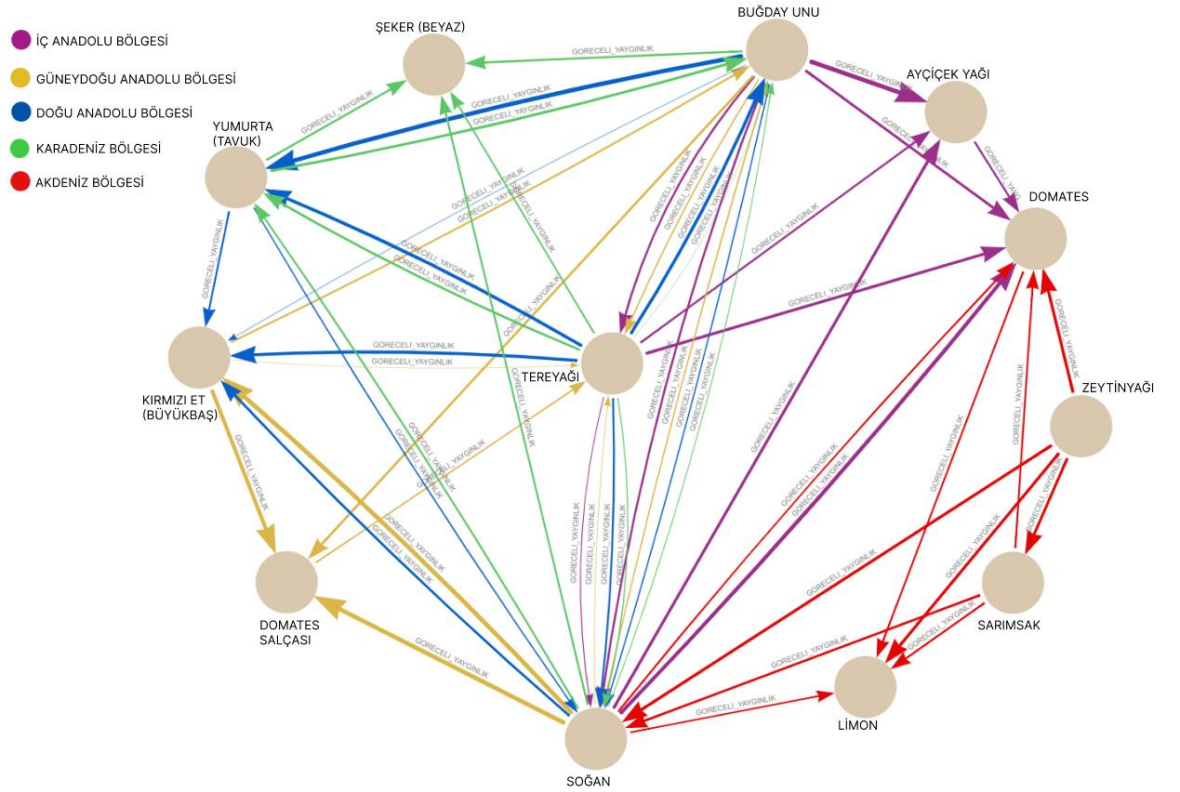
GÜNEYDOĞU ANADOLU BÖLGESİ



Şekil 4.13 Güneydoğu Anadolu bölgesinde en çok kullanılan 5 malzemeye ait yaygınlık değerleri, bu malzemelerin kullanım durumları ve sahip oldukları ortak bileşen sayıları

Değerlendirme:

- 1- Tüm bölgelerde soğan en çok tercih edilen malzemeler arasında yer almaktadır.
- 2- Yağ kategorisinden Akdeniz bölgesinde en çok zeytinyağı, diğer bölgelerde tereyağı kullanımı baskındır. Ayçiçek yağı kullanımında ise İç Anadolu bölgesi öne çıkmaktadır.
- 3- Kırmızı et kullanımı diğer bölgelere oranla Doğu Anadolu ve Güneydoğu Anadolu bölgesinde daha fazladır.
- 4- Buğday unu Akdeniz bölgesi dışında yaygın olarak tercih edilen malzemeler arasında bulunmaktadır. En yaygın kullanımı ise İç Anadolu bölgesinde görülmektedir.
- 5- Sarımsak ve limon kullanımı en çok Akdeniz bölgesinde görülürken, domates salçası Güneydoğu Anadolu’da, şeker kullanımı ise Karadeniz bölgesinde yaygın olarak kullanıldığı tespit edilmiştir.



Şekil 4.14 Bölgesel bazda en çok kullanılan malzeme çiftlerinin bölgelere göre göreceli yaygınlığı

Tüm bölgeler incelendiğinde bölgesel bazda en çok kullanılan 12 farklı malzeme olduğu tespit edilmiştir. Şekil 4.14 bu 12 malzemeye ait malzeme çifti kullanımlarının bölgelere özgü yaygınlık değeri hesaplanarak oluşturulmuştur. Burada amaçlanan tariflerde birlikte kullanılan malzeme çiftinin bölge içerisindeki kullanımının diğer bölgelere kıyasla ne kadar özgün olduğunu belirleyebilmektir.

İlişkiler bölgede yaygın kullanıldığı tespit edilen malzemeler arasında kurulmuştur. Her bir renk malzeme çiftinin birlikte kullanıldığı bölgeyi ifade etmektedir. İlişkilerin kalınlığı ise malzeme çifti kullanımının bölgeye olan özgünlüğü ile doğru orantılıdır. Bölgeye özgü yaygınlık değeri “Göreceli Yaygınlık” eşitliği ile belirlenmektedir. Göreceli yaygınlık formülü eşitlik 4.2’de (Ahn vd. 2011) yer almaktadır. Bu formüle göre bölge içerisinde malzeme çiftinin özgünlük değeri, malzeme çiftinin bölgedeki yaygınlığından diğer bölgelerdeki yaygınlık değeri ortalaması çıkarılarak bulunmaktadır. Şekil 4.14’te malzeme çiftleri arasında hesaplanan bölgelere ait göreceli yaygınlık değerleri EK 3 kısmında yer almaktadır.

$$y_{m_1 m_2}^b = Y_{m_1 m_2}^b - \langle Y_{m_1 m_2}^{b'} \rangle_{b' \neq b}$$

(4.2)

$y_{m_1 m_2}^b$ = m_1 ve m_2 malzemelerinin birlikte kullanımlarının
b bölgesi için göreceli yaygınlık değeri

$Y_{m_1 m_2}^b$ = m_1 ve m_2 malzemelerinin b bölgesindeki tariflerde
birlikte kullanılma yaygınlığı

$$Y_{m_1 m_2}^b = \frac{n_{m_1 m_2}^b}{N_b}$$

$n_{m_1 m_2}^b$ = b bölgesinde m_1 ve m_2 malzemelerinin birlikte kullanıldığı tarif sayısı

N_b = b bölgesindeki toplam tarif sayısı

$\langle Y_{m_1 m_2}^{b'} \rangle_{b' \neq b} = m_1$ ve m_2 malzemelerinin b bölgesi dışındaki diğer bölgelerde birlikte kullanım yaygınlığının ortalaması

Değerlendirme:

- 1- Tereyağı - soğan, tereyağı – buğday unu, soğan – buğday unu malzeme çiftlerinin Akdeniz bölgesi dışındaki tüm bölgelerde yaygın olduğu görülmektedir.
- 2- Zeytinyağı, sarımsak ve limon malzemelerinin birlikte kullanımının Akdeniz bölgesinde sık görüldüğü, kullanım oranın yüksek olduğu gözlemlenmektedir.
- 3- Buğday unu ve yumurta kullanımının Doğu Anadolu ve Karadeniz bölgelerinde baskın olduğu, Doğu Anadolu bölgesindeki yaygınlığının ise Karadeniz bölgesine oranla daha fazla olduğunu görülmektedir.
- 4- Tereyağı ve kırmızı et kullanımının Doğu Anadolu ve Güneydoğu Anadolu bölgelerinde yaygın olarak görüldüğü, Doğu Anadolu bölgesinde kullanımının ise daha yüksek olduğu tespit edilmiştir.
- 5- Ayçiçek yağı kullanımının İç Anadolu bölgesinde yaygın olarak tercih edildiği ve bölge içerisindeki diğer malzemeler ile birlikte kullanım oranın yüksek olduğu gözlemlenmiştir.

5. TARTIŞMA VE SONUÇ

Geleneksel ilişkisel veri tabanı yaklaşımının farklı tür ve boyuttaki verileri saklamada yetersiz kalması NoSQL veri tabanlarının ortaya çıkmasında etkili olmuştur. NoSQL ailesinin en bilinen türlerinden olan çizge veri tabanları verileri ve ilişkileri depolayabilmesi, çevik gelişime uygun olması, veriler arasında bağlantı derinliği fazla olan sorgularda ilişkisel veri tabanları ile kıyaslandığında daha performanslı çalışması gibi avantajlarıyla gerçek dünya problemlerinin çözümünde tercih edilen bir teknoloji haline gelmiştir. Bu yönleriyle tez kapsamında oluşturulan Türk Mutfağı Lezzet Ağı çalışmasında verilerin depolanması ve analiz edilmesi için çizge veri tabanı teknolojisi tercih edilmiştir.

Türk Mutfağı Lezzet Ağı tarifler, malzemeler ve bu malzemelere ilişkin bileşenlerden oluşmaktadır. Bu ağ oluşturulurken çalışmaya ilk olarak veri setinin hazırlanması ile başlanmıştır. Bu kapsamda oluşturulan veri seti Akdeniz, Karadeniz, İç Anadolu, Güneydoğu Anadolu ve Doğu Anadolu olmak üzere 5 bölgeye ait toplam 312 tarif içermektedir. Hazırlanan veri setinde malzemelere ilişkin bileşenlere de yer verilmiş ve bu sayede tariflere ait kalori miktarları, bu tarifler ile alınan ve/veya alınamayan bileşenler gibi bilgiler tespit edilebilmiştir. Ayrıca veri setinde farklı yörelere ait farklı malzemelerden oluşan tarif seçeneklerinin bulunmasıyla kişilere istenilen malzeme ve bileşenler özelinde tarif önerileri sunulabilmektedir.

Çalışmada malzemeler arasında önemli olan, öne çıkanların bulunması, birlikte kullanılma durumlarına göre oluşan toplulukların tespiti gibi analiz çalışmalarına da yer verilmiştir. Bu kapsamda tüm Türk Mutfağı genelinde tariflerde kullanılan malzemeler arasında bir çizge oluşturularak merkezilik ve topluluk tespiti algoritmaları çalıştırılmıştır. Merkezilik algoritmalarından olan derece merkeziliğine göre en fazla kullanılan malzemelerin soğan, tereyağı, buğday unu gibi malzemeler olduğu tespit edilmiş ve en çok kullanılan yani derece merkeziliği değeri en yüksek olan ilk 20 malzeme için arasındalık, yakınlık ve PR algoritma sonuçları değerlendirilmiştir. Bu sayede farklı açılardan önem kazanan malzemelerin tespit edilmesi sağlanmıştır.

Topluluk tespiti algoritmalarından ise Louvain, etiket yayılımı, zayıf bağılı bileşenler, üçgen sayma ve yerel kümelenme katsayısı kullanılmış böylelikle malzemeler arasında oluşan topluluk yapısı incelenmiştir. Buna göre Louvain algoritması ile 4 topluluk, etiket yayılımı algoritması ile tek bir topluluk bulunduğu gözlemlenmiştir. Zayıf bağılı bileşenler algoritması ile malzemeler arasında bağılı bulunmayan bir kümelenmenin olmadığı tespit edilirken üçgen sayma ve yerel kümelenme katsayısı algoritma sonuçlarına göre ise yerel olarak malzemelerin ilişkilerinin genel anlamda yoğun olduğu gözlemlenmiştir. Son olarak bölgesel bazda yapılan çeşitli analiz çalışmalarına yer verilmiştir. Bu sayede veri setini oluşturan bölgeler farklı ve benzer yönleri açısından değerlendirilebilmiştir. Yapılan çalışma kapsamında tüm bölge mutfaklarında tariflerin genel anlamda 10 malzemedan oluştuğu gözlemlenmiştir. Çeşitli bileşenlerin bölge mutfaklarına ait tariflerde ortalama bulunma miktarları incelenmiş ve bileşen bazında karşılaştırmalar yapılmıştır. Bu analizlere ek olarak her bir bölge için tariflerde sıklıkla kullanılan malzemeler belirlenmiş ve en yaygın kullanılan 5 malzeme incelenmiştir. Bu malzemeler özelinde birlikte kullanılan malzeme çiftlerinin, bölgedeki tariflerde kullanımlarının diğer bölgelere kıyasla ne kadar özgün olduğu tespit edilmiştir.

Tez kapsamında yapılan tüm bu analiz çalışmaları ile birlikte Türk Mutfağı hem genel hem bölgesel olarak farklı yönleriyle incelenebilmiştir. İlerleyen aşamalarda çalışmaya katkı olarak veri setine yeni tarifler eklenip daha fazla veri üzerinde farklı merkezilik ve topluluk tespiti algoritmaları uygulanarak çeşitli çıkarımların yapılması sağlanabilir.

KAYNAKLAR

- Afandi, M.I., Wahyuni, E.D. (2020). University Research Graph Database For Efficient Multi-Perspective Data Analysis Using Neo4j, 2020 6th Information Technology International Seminar (ITIS), s.286–290.
- Ahn, Y.Y., Ahnert, S.E., Bagrow, J.P., Barabasi, A.L., (2011), Flavor network and the principles of food pairing, <https://arxiv.org/abs/1111.6074>.
- Anonim. 2002. Web Sitesi: <https://www.lezzet.com.tr/>, Erişim Tarihi: 10.10.2021
- Anonim. 2007. Web Sitesi: <https://www.nefisyemektarifleri.com/>, Erişim Tarihi: 05.03.2022
- Anonim. 2012. Web Sitesi: <https://www.diyetkolik.com/>, Erişim Tarihi: 12.04.2022
- Anonim. 2017. Web Sitesi: <http://www.turkomp.gov.tr/>, Erişim Tarihi: 25.07.2022
- Anonymous. 2008. Web Sitesi: <https://www.project-voldemort.com/voldemort/>, Erişim Tarihi: 19.04.2022.
- Anonymous. 2009. Web Sitesi: <http://nosql-database.org/>, Erişim Tarihi: 10.02.2022.
- Anonymous. 2020a. Web Sitesi: https://en.wikipedia.org/wiki/Louvain_method, Erişim Tarihi: 26.06.2022
- Anonymous. 2020b. Web Sitesi: <https://neo4j.com/docs/graph-data-science/current/algorithms/wcc/>, Erişim Tarihi: 06.07.2022
- Anonymous. 2021. Web Sitesi: <https://www.blazecan.com/blog/dive-deep-types-nosql-databases/>, Erişim Tarihi: 04.04.2022.
- Bajaj, V., Panda, R., B., Dabas, C., Kaur, P. 2018. Graph Database for Recipe Recommendations. 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 1-6, doi: 10.1109/ICRITO.2018.8748827.
- Becchetti, L., Boldi, P., Castillo, C. 2008. Efficient Semi-streaming Algorithms for Local Triangle Counting in Massive Graphs. August, The 14th ACM SIGKDD international conference. DOI:10.1145/1401890.1401898
- Blondel, V.D., Guillaume, J.L., Lambiotte, R. ve Lefebvre, E. 2008. Fast unfolding of communities in large networks, Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008, <https://doi.org/10.1088/1742-5468/2008/10/p10008>
- Borgatti, S. P. 2005. Centrality and network flow. Social Networks. 27(1). 55–71. <https://doi.org/10.1016/j.socnet.2004.11.008>

- Chang, F., Dean, J., Ghemawat, S., Hsieh, W.C., Wallach, D.A., Burrows, M., Chandra, T., Fikes, A., Gruber, R.E. 2008. Bigtable: A distributed storage system for structured data. *ACM Trans. Comput. Syst.*, 26(2):1–26.
- DeCandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., Sivasubramanian S., Voshall P., Vogels, W. 2007. Dynamo: amazon’s highly available key-value store. *SIGOPS Oper. Syst. Rev.*, 41(6):205–220
- Freeman L.C. 1979. Centrality in Social Networks Conceptual Clarification, Elsevier Sequoia S.A., 215-239, Lausanne
- Gleich, D.F. 2014. PageRank beyond the Web, <https://arxiv.org/abs/1407.5107>.
- Guia, J., Gonçalves, V., Bernardino J. 2017. Graph Databases: Neo4j Analysis, *Proceedings of the 19th International Conference on Enterprise Information Systems*, 351-356.
- Gürsakal N. 2018. Sosyal Ağ Analizi, Anadolu Üniversitesi Yayını, 50 – 52.
- Hasan, M., Dave, V. 2017. Triangle Counting in Large Networks: A Review. <https://doi.org/10.1002/widm.1226>
- Huang, B., Shi, X., Wang, R., Wang, C., Han, Y. 2022. A Novel Recipes Recommendation System Based on Knowledge-Graph. *7th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 1408-1412, doi: 10.1109/ICSP54964.2022.9778762.
- Jaganathan, B., Desikan, K. 2015. Weighted page rank algorithm based on in-out weight of webpages. 8. 10.17485/ijst/2015/v8i34/86120.
- Kronmueller, M., Chang, D., Hu, H., A, Desoky, A. 2018. Graph Database of Yelp Dataset Challenge 2018 and Using Cypher for Basic Statistics and Graph Pattern Exploration. *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*.135-140, doi: 10.1109/ISSPIT.2018.8642700.
- Maxwell, G., Kronmueller, M., Chang, D., Desoky, A. 2021. Measuring the “Impact” of People, Films and TV using the IMDb Graph Database. *SoutheastCon*. 01-08, doi: 10.1109/SoutheastCon45413.2021.9401938.
- Menczer, F., Fortunato, S., Davis, C. A. 2020. *A First Course in Network Science*, Cambridge Univeristy Press, 105-108, İngiltere. doi:10.1017/9781108653947
- Miller, J.J. 2013. *Proceedings of the Southern Association for Information Systems Conference on Graph Database Applications and Concepts with Neo4j*, March 23-24, 2013; 141-147, Atlanta, GA, USA.
- Mrvar, A. 2015. Centrality and Prestige. *Network Analysis using Pajek*. <http://mrvar.fdv.uni-lj.si/sola/info4/uvod/part4.pdf>

- Needham, M., Hodler, A., E. 2019. Graph Algorithms Practical Examples in Apache Spark and Neo4j. O'Reilly Media, USA.
- Neubauer, P. 2010. Web Sites: <http://www.infoq.com/articles/graph-nosql-neo4j>, Eriřim Tarihi: 05.04.2022.
- Page, L., Brin, S. 1998. The anatomy of a large scale hyper-textual web search engine. Computer Networks and ISDN Systems. Apr; 30(1-7):107–17.
- Partner, J., Vukonic A., Watt N., Abedrabbo, T., Fox, D. 2014. Neo4j in Action. Manning Publications.
- Patil, N.S., Vaswani, G., Bhatia, A. 2014. Graph Databases- An Overview, International Journal of Computer Science and Information Technologies, 657-660.
- Patil, N. S., Kiran, P., Kavva, N. P., Naresh, K. M. 2018. Survey on Graph Database Management Techniques for Huge Unstructured Data, International Journal of Electrical and Computer Engineering, 1140-1149
- Raghavan, U.N., Albert, R., Kumara, S. 2007. Near linear time algorithm to detect community structures in large-scale networks, Phys. Rev. E, 76, <https://doi.org/10.48550/arXiv.0709.2938>.
- Srivastava, S., Singh, A.K. 2022. Fraud detection in the distributed graph database. Cluster Comput. <https://doi.org/10.1007/s10586-022-03540-3>
- Vicknair, C., Macias, M., Zhao, Z., Nan, X., Chen, Y., Wilkins, D. 2010. Proceedings of the 48th Annual Southeast Regional Conference on A Comparison of a Graph Database and a Relational Database, April 15-17, 2010; Oxford/England.

EKLER

EK 1 Bileşen İsimleri ve Birimleri

Bileşen İsmi	Birim
a vitamini	re
alanine	mg
alfa-tokoferol	mg
arjinin	mg
aspartik asit	mg
azot	g
b-12 vitamini	µg
b-6 vitamini, toplam	mg
beta-karoten	µg
c vitamini	mg
d vitamini, iu	iu
d-3 vitamini (kolekalsiferol)	µg
d-mannitol	g
demir, fe	mg
e vitamini	α-te
e vitamini, iu	iu
enerji	kj
enerji	kcal
fenilalanin	mg
folat, gıda	µg
fosfor, p	mg
fruktoz	g
glisin	mg
glukoz	g
glutamik asit	mg
histidin	mg
iyot, i	µg
izolosin	mg
k-1 vitamini	µg
k-2 vitamini	µg
kalsiyum, ca	mg
karbonhidrat	g
kolesterol	mg
ksilitol	g
kül	g
l-askorbik asit	mg
laktoz	g

Bileşen İsmi	Birim
lif, suda çözünmeyen	g
lif, suda çözünür	g
lif, toplam diyet	g
likopen	µg
lizin	mg
lutein	µg
lösin	mg
magnezyum, mg	mg
maltoz	g
metiyonin	mg
niasin	mg
niasin eşdeğerleri, toplam	ne
nişasta	g
potasyum, k	mg
prolin	mg
protein	g
retinol	µg
riboflavin	mg
sakaroza	g
selenyum, se	µg
serin	mg
sistin	mg
sodyum, na	mg
sorbitol	g
su	g
tiamin	mg
tirozina	mg
treonin	mg
triptofan	mg
tuz	mg
valin	mg
yağ asidi 10:0 (kaprik asit)	g
yağ asidi 12:0 (laurik asit)	g
yağ asidi 14:0 (miristik asit)	g
yağ asidi 14:1 n-5 cis (miristoleik asit)	g
yağ asidi 15:0 (pentadesilik asit)	g
yağ asidi 16:0 (palmitik asit)	g
yağ asidi 16:1 n-7 cis (palmitoleik asit)	g
yağ asidi 17:0 (margarik asit)	g
yağ asidi 18:0 (stearik asit)	g
yağ asidi 18:1 n-9 cis (oleik asit)	g
yağ asidi 18:1 n-9 trans (elaidik asit)	g
yağ asidi 18:2 n-6 cis,cis	g

Bileşen İsmi	Birim
yağ asidi 18:3 n-3 all-cis	g
yağ asidi 18:3 n-6 all-cis	g
yağ asidi 20:0 (araşidik asit)	g
yağ asidi 20:1 n-9 cis	g
yağ asidi 20:4 n-6 all-cis	g
yağ asidi 20:5 n-3 all-cis	g
yağ asidi 22:0 (behenik asit)	g
yağ asidi 22:1 n-9 cis (erüsik asit)	g
yağ asidi 22:6 n-3 all-cis	g
yağ asidi 24:0 (lignoserik asit)	g
yağ asidi 24:1 n-9 cis	g
yağ asidi 4:0 (bütirik asit)	g
yağ asidi 6:0 (kaproik asit)	g
yağ asidi 8:0 (kaprilik asit)	g
yağ asitleri, toplam doymuş	g
yağ asitleri, toplam tekli doymamış	g
yağ asitleri, toplam çoklu doymamış	g
yağ, toplam	g
çinko, zn	mg

EK 2 Malzemelere Ait Merkezilik Algoritması Sonuçları

Malzeme İsmi	Derece Merkeziliği Algoritma Sonucu	Arasındalık Merkeziliği Algoritma Sonucu	Yakınlık Merkeziliği Algoritma Sonuuc	PageRank (PR) Merkezilik Algoritma Sonucu
Tuz	243	3225,370	0,875	14,283
Su	177	2797,625	0,788	10,433
Soğan	142	1244,864	0,744	8,831
Tereyağı	138	1586,528	0,737	7,762
Karabiber	129	1036,209	0,730	8,079
Buğday Unu	117	811,871	0,678	6,711
Yumurta (Tavuk)	93	631,545	0,657	5,459
Ayçiçek Yağı	92	623,619	0,680	5,595
Kırmızı Et (Büyükbaş)	85	333,836	0,641	5,155
Zeytinyağı	83	1586,873	0,764	5,677
Sarımsak	78	799,965	0,708	5,140
Domates	75	532,913	0,671	4,778
Pul Biber	73	290,732	0,635	4,542
Şeker (Beyaz)	71	835,818	0,663	4,159
Yoğurt	66	435,363	0,648	4,017
Domates Salçası	63	419,701	0,655	4,340
Maydanoz	61	521,886	0,669	4,095
Limon	56	1050,009	0,697	3,861
Nane	50	357,210	0,635	3,383
Bulgur	45	121,305	0,594	2,786
Sivri Biber	41	142,113	0,602	2,510
Biber Salçası	40	132,913	0,599	2,707
Ceviz	38	303,778	0,617	2,355
Pirinç	35	183,452	0,605	2,342
Süt	34	140,916	0,594	1,985
Kırmızı Toz Biber	31	156,144	0,605	2,102
Kırmızı Et (Küçükbaş)	30	62,127	0,575	1,722
Kimyon	28	103,283	0,591	2,107
Nohut	23	139,069	0,587	1,733
Patlıcan	22	25,603	0,560	1,293
Kapya Biber	21	90,846	0,587	1,463
Kabartma Tozu	20	132,895	0,567	1,423
Sumak	18	31,227	0,567	1,268
Kekik	17	114,387	0,590	1,378
Margarin	16	70,797	0,567	1,064
Patates	15	61,329	0,569	1,014
İrmik	14	25,381	0,553	0,948

Malzeme İsmi	Derece Merkeziliği Algoritma Sonucu	Arasındalık Merkeziliği Algoritma Sonucu	Yakınlık Merkeziliği Algoritma Sonucu	PageRank (PR) Merkezilik Algoritma Sonucu
Beyaz Peynir	14	81,325	0,572	0,952
Et Suyu	13	29,270	0,557	0,908
Pekmez	13	114,295	0,541	0,821
Elma Sirkesi	12	167,104	0,585	1,125
Antep Fıstığı	12	62,620	0,554	0,930
Yenibahar	12	49,888	0,561	0,941
Buğday Nişastası	12	23,937	0,552	0,860
Yeşil Soğan	11	63,267	0,565	0,896
Yarma	11	38,904	0,549	0,793
Mısır Unu	11	15,034	0,541	0,583
Yaş Maya	10	6,833	0,537	0,726
Yeşil Mercimek	10	11,466	0,545	0,713
Fesleğen	10	113,516	0,578	0,917
Siyah Zeytin	10	80,086	0,567	0,928
Pudra Şekeri	9	10,525	0,534	0,694
Havuç	9	61,406	0,571	0,865
Nar Ekşisi	9	23,537	0,549	0,724
Ekmek	9	182,818	0,569	0,908
Kıvırcık	9	43,792	0,539	0,881
Tavuk Eti	8	12,885	0,550	0,608
Çarliston Biber	8	11,564	0,545	0,634
Vanilya	8	58,792	0,540	0,636
Iceberg	8	29,562	0,534	0,776
Badem	8	87,259	0,558	0,705
Dolmalık Biber	7	12,809	0,545	0,618
Kabak	7	36,251	0,557	0,657
Dereotu	7	49,577	0,557	0,653
Roka	7	34,565	0,541	0,665
Mısır	7	22,396	0,548	0,572
Yumurta Sarısı (Tavuk)	7	14,187	0,546	0,595
Karanfil	6	31,450	0,560	0,645
Kişiş	6	9,667	0,541	0,505
Tavuk Suyu	6	1,966	0,534	0,421
Eritme Peyniri	5	1,863	0,520	0,329
Trabzon Hurması	5	38,293	0,528	0,535
Taze Fasulye	5	12,791	0,536	0,433
Karbonat	6	20,566	0,542	0,563
Bal	6	132,616	0,548	0,686
Susam	6	85,485	0,552	0,626
Asma Yaprığı	6	15,136	0,541	0,531

Malzeme İsmi	Derece Merkeziliği Algoritma Sonucu	Arasındalık Merkeziliği Algoritma Sonucu	Yakınlık Merkeziliği Algoritma Sonucu	PageRank (PR) Merkezilik Algoritma Sonucu
Lavaş	5	2,013	0,529	0,446
Böreklik Yufka	5	5,727	0,528	0,401
Kaşar Peynir	5	72,361	0,557	0,686
Çörek Otu	5	34,130	0,531	0,478
Mantar	5	30,026	0,549	0,578
Erişte	4	1,587	0,530	0,401
Mor Soğan	4	15,777	0,522	0,440
Tahin	4	18,984	0,537	0,446
Yeşil Zeytin	4	23,871	0,526	0,513
Kaymak	4	6,398	0,522	0,377
Defne Yaprağı	4	14,530	0,541	0,492
Fındık	4	20,810	0,524	0,416
Kırmızı Mercimek	4	1,946	0,532	0,417
Salatalık	4	15,794	0,535	0,433
Eski Kaşar	4	6,306	0,530	0,363
Haşhaş	4	21,589	0,539	0,437
Kuyruk Yağı	4	1,504	0,525	0,377
Tırnak Pide	4	0,329	0,523	0,369
Kuş Üzümlü	4	7,531	0,534	0,488
Reyhan	4	8,911	0,531	0,400
Mahlep	3	0,541	0,520	0,355
Kavurma	3	0,330	0,520	0,290
Ispanak	3	0,000	0,523	0,309
Pazı	3	2,373	0,529	0,355
Kara Lahana	3	4,391	0,525	0,348
Çam Fıstığı	3	18,534	0,542	0,463
Galet Unu	3	2,165	0,512	0,344
Arpacık Soğan	3	0,789	0,511	0,283
Tarhana	3	1,429	0,530	0,330
Ciğer (Küçükbaş)	3	1,189	0,510	0,300
Mayonez	3	6,372	0,509	0,396
Sade Yağ	3	3,356	0,511	0,281
Semizotu	3	1,623	0,524	0,338
California Wonder Tipi Sarı Biber	3	8,814	0,518	0,437
Krema	3	6,820	0,524	0,328
Deniz Börülcesi	2	0,000	0,485	0,227
Bamya	2	0,220	0,508	0,233
Safran	2	4,800	0,529	0,333
Çubuk Tarçın	2	2,343	0,526	0,315

Malzeme İsmi	Derece Merkeziliği Algoritma Sonucu	Arasındalık Merkeziliği Algoritma Sonucu	Yakınlık Merkeziliği Algoritma Sonucu	PageRank (PR) Merkezilik Algoritma Sonucu
Portakal	2	10,076	0,526	0,382
Bitter Çikolata	2	0,480	0,500	0,240
Baklavalık Yufka	2	2,896	0,515	0,274
Dil Peyniri	2	5,926	0,507	0,298
Tel Kadayıf (Hammadde, Erzurum Kadayıf Dolması)	2	0,046	0,507	0,232
Taze Sarımsak	2	0,134	0,520	0,264
Keçi Peyniri	2	0,227	0,498	0,274
Kayısı	2	8,141	0,525	0,395
Kereviz	2	3,201	0,515	0,353
California Wonder Tipi Kırmızı Biber	2	1,075	0,499	0,342
Balık Suyu	2	2,285	0,509	0,339
Levrek	2	2,644	0,510	0,275
Gül Yapağı	2	0,030	0,511	0,242
Milföy Hamuru	2	5,273	0,515	0,275
Balkabağı	2	0,206	0,479	0,213
Ekmek Hamuru	2	0,158	0,472	0,213
Bezelye	2	2,771	0,522	0,288
Labne Peyniri	2	4,672	0,503	0,344
Rezene	2	0,485	0,519	0,305
İsot	2	0,565	0,519	0,301
Buz	2	2,063	0,508	0,301
Firik	2	0,158	0,516	0,258
Muskat	2	2,425	0,525	0,300
Üzüm sirkesi	2	0,051	0,515	0,231
Mısır Nişastası	2	0,575	0,523	0,304
Fasulye Turşusu	2	0,409	0,515	0,245
Kızılıcık	2	1,811	0,484	0,302
Van Otlı Peynir	2	0,240	0,512	0,232
Ayran	2	0,855	0,520	0,261
Siyez Bulguru	2	3,192	0,515	0,330
Barbunya	2	3,592	0,523	0,321
Çekirdekli Siyah Üzüm	2	3,449	0,517	0,355
Lahana Turşusu	2	3,293	0,524	0,320
Kuru Fasulye	2	4,210	0,519	0,299
Ayva	2	0,709	0,503	0,232
Erik Ekşisi	2	4,180	0,522	0,308
Limon Tuzu	2	0,692	0,516	0,240

Malzeme İsmi	Derece Merkeziliği Algoritma Sonucu	Arasındalık Merkeziliği Algoritma Sonucu	Yakınlık Merkeziliği Algoritma Sonucu	PageRank (PR) Merkezilik Algoritma Sonucu
Makarna	1	0,000	0,479	0,189
Koyun Paçası	1	0,000	0,509	0,198
Kapari	1	0,000	0,516	0,322
Kavun	1	0,000	0,516	0,322
Şeftali	1	0,000	0,516	0,322
Greyfurt	1	0,000	0,516	0,322
Et Bulyon	1	0,000	0,510	0,198
Yumurta Akı (Tavuk)	1	0,000	0,477	0,188
Karnabahar	1	0,000	0,500	0,187
Erzincan Tulum Peyniri	1	0,000	0,515	0,223
Hindiba	1	0,000	0,491	0,244
Hindi Salam	1	0,000	0,491	0,244
İç Yağ	1	0,000	0,488	0,199
Çökelek	1	0,000	0,509	0,198
Türk Kahvesi	1	0,000	0,441	0,155
Üzüm Pestili	1	0,000	0,498	0,184
Kar	1	0,000	0,448	0,225
Gül Şurubu	1	0,000	0,448	0,225
Hurma	1	0,000	0,475	0,211
Limon Kabuğu	1	0,000	0,475	0,211
Bağırsak (Küçükbaş)	1	0,000	0,515	0,219
Erik	1	0,000	0,482	0,187
Çağla	1	0,000	0,495	0,197
Aspir	1	0,000	0,515	0,236
Kestane	1	0,000	0,515	0,236
Vanilin	1	0,000	0,478	0,194
Baharat Karışımı, Köfte Harcı	1	0,000	0,494	0,209
Mengen Peyniri	1	0,000	0,498	0,206
Ezine Peyniri	1	0,000	0,467	0,280
Kızılıcak Marmelatı	1	0,000	0,467	0,280
Melisa	1	0,000	0,467	0,280
Beyaz Lahana	1	0,000	0,507	0,214
Zerdeçal	1	0,000	0,498	0,264
Kelek Turşusu	1	0,000	0,498	0,264
Yulaf Unu	1	0,000	0,447	0,246
Hindistan Cevizi	1	0,000	0,447	0,246
Şeker (Esmer)	1	0,000	0,447	0,246
Toz Kahve Kreması	1	0,000	0,472	0,221
Fındık Yağı	1	0,000	0,508	0,195

Malzeme İsmi	Derece Merkeziliği Algoritma Sonucu	Arasındalık Merkeziliği Algoritma Sonucu	Yakınlık Merkeziliği Algoritma Sonucu	PageRank (PR) Merkezilik Algoritma Sonucu
Hamsi	1	0,000	0,500	0,217
Çorum Leblebisi	1	0,000	0,454	0,176
Beyaz Dut	1	0,000	0,466	0,172
Keten Tohumu	1	0,000	0,467	0,178
Ciğer (Büyükbaş)	1	0,000	0,505	0,211
Taze Fasulye Yaprığı	1	0,000	0,486	0,214
Yufka	1	0,000	0,473	0,171
Kaz Eti	1	0,000	0,491	0,198
Börülce Fasulyesi	1	0,000	0,502	0,209
Nar	1	0,000	0,505	0,202
Göce	1	0,000	0,513	0,213
İlibada	1	0,000	0,497	0,208
Pastırma Çemeni	1	0,000	0,487	0,204
Akciğer (Büyükbaş)	1	0,000	0,475	0,195
Maraş Dondurması	1	0,000	0,512	0,232
Mineralli Su	1	0,000	0,507	0,220
Arpa Şehriye	1	0,000	0,503	0,197
Madımak	1	0,000	0,483	0,187
Pastırma	1	0,000	0,493	0,197
Sarı Kuyruk (Balık)	1	0,000	0,502	0,188
Lor Peyniri	1	0,000	0,491	0,182
Darı	1	0,000	0,475	0,166
Siyah Fasulye	1	0,000	0,465	0,205
Çöven	1	0,000	0,484	0,222
Zencefil	1	0,000	0,516	0,273
Kakule	1	0,000	0,516	0,273
Çipura	1	0,000	0,485	0,196
Kolot Peyniri	1	0,000	0,496	0,229
Siyah Zeytin Ezmesi	1	0,000	0,465	0,205
Lolorosso	1	0,000	0,489	0,235
Hardal	1	0,000	0,489	0,235
Tel Kadayıf (Hammadde, Antakya Künefesi)	1	0,000	0,474	0,197
Dana Jambon	1	0,000	0,498	0,229
Somon	1	0,000	0,495	0,272
Bakla	1	0,000	0,505	0,181
Yonca Filizi	1	0,000	0,493	0,215
Krem Peynir	1	0,000	0,488	0,203

EK 3 Malzeme Çiftlerinin Bölgelere Göre Göreceli Yaygınlık Değerleri (Şekil 4.14)

Akdeniz’de Yaygın Olarak Kullanılan Malzemelerin Birlikte Kullanımlarına İlişkin Bölge Özelinde Göreceli Yaygınlık Değerleri

Malzeme Çifti	Akdeniz Bölgesi Göreceli Yaygınlık Değeri
Zeytinyağı Sarımsak	0.157
Zeytinyağı - Soğan	0.169
Zeytinyağı - Domates	0.169
Zeytinyağı - Limon	0.169
Sarımsak - Soğan	0.133
Sarımsak - Domates	0.096
Sarımsak - Limon	0.108
Soğan - Domates	0.092
Soğan - Limon	0.084
Domates - Limon	0.096

Doğu Anadolu’da Yaygın Olarak Kullanılan Malzemelerin Birlikte Kullanımlarına İlişkin Bölge Özelinde Göreceli Yaygınlık Değerleri

Malzeme Çifti	Doğu Anadolu Bölgesi Göreceli Yaygınlık Değeri
Tereyağı – Buğday Unu	0.185
Tereyağı – Yumurta (Tavuk)	0.189
Tereyağı - Soğan	0.104
Tereyağı - Kırmızı Et (Büyükbaş)	0.182
Buğday Unu – Yumurta (Tavuk)	0.257
Buğday Unu - Soğan	0.063
Buğday Unu - Kırmızı Et (Büyükbaş)	0.011
Yumurta (Tavuk) - Soğan	0.066
Yumurta (Tavuk) - Kırmızı Et (Büyükbaş)	0.098
Soğan - Kırmızı Et (Büyükbaş)	0.137

Güneydoğu Anadolu’da Yaygın Olarak Kullanılan Malzemelerin Birlikte Kullanımlarına İlişkin Bölge Özelinde Göreceli Yaygınlık Değerleri

Malzeme Çifti	Güneydoğu Anadolu Bölgesi Göreceli Yaygınlık Değeri
Soğan – Kırmızı Et (Büyükbaş)	0.260
Soğan – Buğday Unu	0.070
Soğan - Domates Salçası	0.259
Soğan - Tereyağı	0.017
Kırmızı Et (Büyükbaş) – Buğday Unu	0.101
Kırmızı Et (Büyükbaş) - Domates Salçası	0.222
Kırmızı Et (Büyükbaş) - Tereyağı	0.007
Buğday Unu - Domates Salçası	0.130
Buğday Unu - Tereyağı	0.070
Domates Salçası - Tereyağı	0.074

İç Anadolu’da Yaygın Olarak Kullanılan Malzemelerin Birlikte Kullanımlarına İlişkin Bölge Özelinde Göreceli Yaygınlık Değerleri

Malzeme Çifti	İç Anadolu Bölgesi Göreceli Yaygınlık Değeri
Buğday Unu - Tereyağı	0.116
Buğday Unu - Soğan	0.117
Buğday Unu – Ayçiçek Yağı	0.259
Buğday Unu - Domates	0.148
Tereyağı - Soğan	0.040
Tereyağı – Ayçiçek Yağı	0.111
Tereyağı - Domates	0.166
Soğan – Ayçiçek Yağı	0.166
Soğan - Domates	0.220
Ayçiçek Yağı - Domates	0.111

Karadeniz’de Yaygın Olarak Kullanılan Malzemelerin Birlikte Kullanımlarına İlişkin Bölge Özelinde Göreceli Yaygınlık Değerleri

Malzeme Çifti	Karadeniz Bölgesi Göreceli Yaygınlık Değeri
Tereyağı - Soğan	0.053
Tereyağı – Yumurta (Tavuk)	0.127
Tereyağı – Buğday Unu	0.0005
Tereyağı – Şeker (Beyaz)	0.085
Soğan – Yumurta (Tavuk)	0.104
Soğan – Buğday Unu	0.036
Soğan – Şeker (Beyaz)	0.1
Yumurta (Tavuk) – Buğday Unu	0.150
Yumurta (Tavuk) – Şeker (Beyaz)	0.114
Buğday Unu – Şeker (Beyaz)	0.114