T.C
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SOSYAL AĞLARDA ÖRTÜŞEN TOPLULUKLARIN TESPİT EDİLMESİ

YÜKSEK LİSANS TEZİ
BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

Esra KARADENİZ

ARALIK - 2016

Tezin Başlığı : SOSYAL AĞLARDA ÖRTÜŞEN TOPLULUKLARIN TESPİT EDİLMESİ

Tezi Hazırlayan : Esra KARADENİZ

Sınav Tarihi : 16.12.2016

Yukarıda adı geçen tez jürimizce değerlendirilerek Bilgisayar Bilimleri Ana Bilim Dalında Yüksek Lisans Tezi olarak kabul edilmiştir.

Sınav Jüri Üyeleri

Prof. Dr. Ali KARCI (Jüri Başkanı, Tez Danışmanı)
İnönü Üniversitesi

Doç.Dr. M.Fatih TALU (Üye)
İnönü Üniversitesi

Doç.Dr. Galip AYDIN (Üye)
Fırat Üniversitesi

Prof.Dr. Halil İbrahim ADIGÜZEL
Enstitü Müdürü

ONUR SÖZÜ

Yüksek Lisans Tezi olarak sunduğum “Sosyal Ağlarda Örtüşen Topluluk Tespiti” başlıklı bu çalışmanın bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığını ve yararlandığım bütün kaynakların, hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuğunu belirtir, bunu onurumla doğrularım.

Esra KARADENİZ

ÖZET

Yüksek Lisans Tezi

SOSYAL AĞLARDA ÖRTÜŞEN TOPLULUKLARIN TESPİT EDİLMESİ

Esra KARADENİZ

İnönü Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

52 + viii sayfa

2016

Danışman: Prof. Dr. Ali KARCI

Sosyal medya ve ağ yapılarının artan önemi bu konudaki çalışmaları da arttırmıştır. Sosyal ağlar toplulukların bir araya gelmesiyle oluşan yapılardır. Sosyal ağların en genel özelliği, topluluk yapılarıdır. Gerçek ağ yapılarında bir elemanın birden fazla topluluğa dâhil olma olasılığı vardır ve bu duruma örtüşme (overlapping) denir.

Bu çalışmada örtüşen topluluk keşfi problemine iki çözüm önerilmiştir. İlk yönteme göre sosyal ağ bir graf olarak modellenmiştir ve bu graftaki her bir tam bağlı alt graf topluluk olarak kabul edilmiştir. Elde edilen sosyal ağın bitişiklik matrisine Bron-Kerbosch algoritması uygulanmış ve yönsüz graftaki tüm maksimal-klikler bulunmuştur. Ardından bu maksimal-klikler revize edilmiş ve önerilen yöntem eşliğinde kesişen toplulukların keşfi sağlanmıştır. Diğer bir yöntemde ise sosyal ağ yine bir graf olarak modellenmiştir. Grafın Laplace matrisi hesaplanmış ve graf özdeğer ve özvektörlerine göre iki gruba ayrılmıştır. Daha sonra minimum kesen ayrıt işlemleri uygulanarak iki grupta da olma ihtimali olan elemanlar tespit edilmiştir.

ANAHTAR KELİMELER: Sosyal ağlar, örtüşen topluluk tespiti, ortak birey tespiti

ABSTRACT

Master Thesis
OVERLAPPING COMMUNITY DETECTION IN SOCIAL NETWORKS
Esra KARADENİZ
İnönü University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering
52 + viii pages
2016
Supervisor: Prof. Dr. Ali KARCI

The growing importance of social media and networking has increased the efforts in this area. Social networks are structures formed by the communities which are came together. The main common feature of all kind of social networks is community structures. In real network structures, an element is likely to be included in multiple groups and this situation is called as overlapping.

In this paper, we have two methods for solving the problem of identifying overlapping groups. According to the first method, social network was modeled as a graph and each fully connected subgraphs in this graph has been accepted as a community. Bron-Kerbosch algorithm has been applied to the adjacency matrix of social network modelled as graph and all maximal cliques in undirected graphes has been found. Then, with the suggested method, these maximal cliques was revised so that overlapping communities could be found. In another method, the social network is modeled as a graph again. The Laplacian matrix of graph is calculated and divided into two groups according to its eigenvalues and eigenvectors. Then, the possibility of elements being in two groups is identified by applying the minimum cutting edges process.

KEYWORDS: Social networks, overlapping community detection, identify common individual

TEŞEKKÜR

Bu tez çalışmasının her aşamasında bilimsel tecrübeleri, farklı bakış açıları ve katkılarıyla beni aydınlatan, öneri ve desteğini esirgemeden beni her konuda yönlendiren, sürekli motivasyonumu sağlayan ve bana değerli zamanını ayıran danışman hocam Sayın Prof. Dr. Ali KARCI'ya;

Manevi destekleriyle bana güç veren değerli hocalarım Kenan İnce, Ahmet Karadoğan'a, kıymetli ağabeyim Murat Ergül'e, değerli dostum Melike Merve Temel'e;

Başta bu tezi yazmamı benden çok dert edinen sevgili annem ve dualarıyla sürekli beni destekleyen yeğenim Elif Ece olmak üzere tüm hayatım boyunca olduğu gibi bu çalışmalarım süresince de benden her türlü desteklerini esirgemeyen değerli aileme

sonsuz teşekkürü bir borç bilirim.

İÇİNDEKİLER

ŞEKİLLER DİZİNİ

SİMGELER VE KISALTMALAR

A= $[a_{ij}]$	: Bitişiklik matrisi
Adj (v)	: v düğümüne bitişik düğümler kümesi
G=(V,E)	: V düğüm ve E ayrıtlardan oluşan graf
D=$[d_{ij}]$	: Derece matrisi
d(v_i)	: v_i düğümünün derecesi
E	: Ayrıtlar kümesi
ϕ	: grafın izoperimetrik sayısı
L=$[l_{ij}]$	: Bir grafın Laplace matrisi
λ	: Bir matrisin öz değeri
V	: Düğümler kümesi
$\vec{x}$	: vektör
VLSI	: very large scale integration (çok geniş ölçekli tümleşim)
ZKK	: Zachary'nin Karate Kulübü
RSB	: Recursive Spectral Bisection

1. GİRİŞ

“Ağ” kavramı Bilişim Terimleri Sözlüğünde; “Birçok nokta ile bunlar arasındaki bağlantılarla gösterilebilen bir dizgeye ilişkin yapı” olarak tanımlanmaktadır [1]. Ağ yapılarının analiz edilmesi ile düğümler olarak adlandırılan ağda bulunan öğeler arasındaki doğrudan göz önünde olmayan ancak anlamlı bilgiler ihtiva eden ilişkiler görünür hale getirilebilir. Örneğin, protein etkileşim ağı içindeki hastalık ile ilişkili proteinlerin keşfi [2], sosyal ağ kullanarak reklamlarda potansiyel müşteri hedefleme [3], ve kimyasal molekül yapılarında benzer ağ örüntüleri araştırma gibi konular ağ tabanlı analitik metodolojileri kullanarak gerçekleştirebilir.

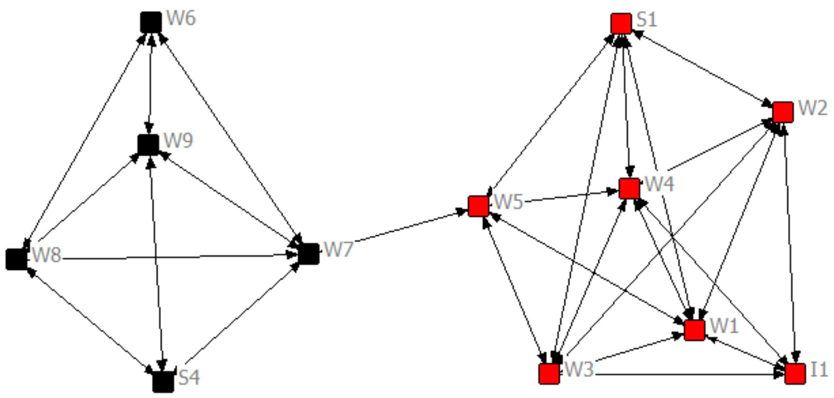
Günlük hayatta çevremizde çeşitli ağ yapıları mevcuttur. Bu ağ yapılarına maddelerin atomları arası bağlardan oluşan yapılar, güneş sistemi, biyolojik ağlar (hücre, doku, dolaşım sistemi, sindirim sistemi vb.), canlılar ve çevre arasındaki ekolojik ağlar (besin zincirleri, bitki örtüsü vb.), kültürel ağlar (edebiyat, inançlar vb.), sosyal ağlar (ilişkiler, ırklar, toplumlar vb.), ulaşım ağları (Karayolları, demiryolları ve vb.) sayısal ağlar (internet, telefon, uydu sistemleri vb.) örnek verilebilir [4].

Düğüm etiketli ağ (node-annotated network) düğüm listesi, düğümler arasında ilişkiyi temsil eden bağlantılar ve düğümlerle ilişkili özelliklerden oluşur. Örneğin, Facebook ağı düşünülürse: Facebook, bir düğüm etiketli sosyal ağdır. Her düğüm kullanıcıyı, kullanıcılar arasındaki arkadaşlık aktivitesi bağlantıyı ve sevilen ya da favori edilen parçalar düğüm özelliğini oluşturmaktadır. Facebook içeriklerinin paylaşımları kullanıcıların kendisi tarafından gönüllü olarak gerçekleştirildiğinden dolayı, kullanıcıların hepsi bütün bilgilerini açıklamak istemez. Benzer şekilde başka bir örnek olarak Twitter ağı düşünülürse; Twitter da düğüm etiketli sosyal ağdır. Her bir düğüm kullanıcıyı, kullanıcılar arasındaki takip etme ilişkisi bağlantıyı ve kullanıcının demografik bilgileri veya takip ettiği önemli kişiler düğümlerle ilişkili özellikleri oluşturmaktadır [5].

Sosyal ağ birey ya da elemanların temsilini sağlayan düğümler ve bu düğümlerin arasındaki ilişkilerin temsilini sağlayan bağıntılardan oluşan yapı şeklinde tanımlanabilir. İnternet üzerindeki web siteleri ve diğer uygulamalar, araştırmacılara çok miktarda analiz edilecek veri ve yeni araştırma alanları sunmaktadır. İnsanlar ve nesneler arasındaki ilişkileri gösteren günlük hayatımızdaki verilerin çoğu da sosyal ağlar olarak modellenebilmektedirler. Günümüzde artan internet kullanımının etkisiyle sosyal ağ olarak adlandırılan platformların popülaritesinde akıllara durgunluk verecek bir artış söz konusudur. Öyle ki sosyal ağlar bazı isyanların, iç savaşların, devrimlerin vb. olayların artık ilk adımlarının atıldığı mecra olarak görülmektedir. Sosyal ağ kavramının artan önemi sosyal ağ yapılarının incelenip analiz edilmesi, bağlantı tahmini, duygu analizi, ağlardaki topluluklar ve toplulukların keşfedilmesi gibi konularda araştırmaları da artırmıştır. Sosyal ağ analizi, ağ yapılarının içerisinde yer alan düğümler arasındaki ilişkilerin farklı bilimsel yöntemler vasıtasıyla ayrıntılı olarak irdelenmesiyle elde edilen verilerden anlamlı ve yorumlanabilir sonuçlar üretilmesi işine denir.

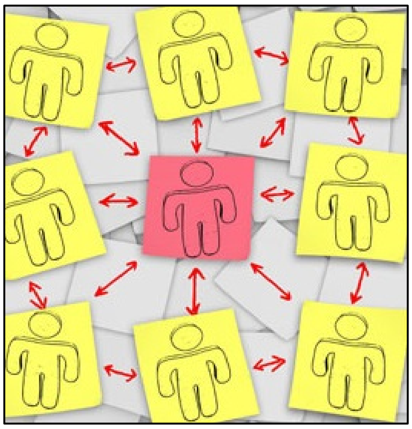
Sosyal ağ analizinin günümüzde kullanım alanı oldukça geniştir. Kullanım alanlarına birey ve sosyal grup yapılarının ve davranışlarının analizi (gruplama, bileşenlerine ayırma, bağıntı tahmini), çevrimiçi reklamcılık ve e-ticaret (müşteri profillerinin belirlenmesi, şahsa özel reklamcılık ve şahsa özel teklif sunma), ulaşım, tesisat, kanalizasyon altyapı vs. yapıların analizi ve büyük veri kümelerinin analizi (sosyal medya analizi, akademik yayın analizi, atıf analizi, genetik incelemeler) örnek olarak gösterilebilir [18].

Sosyal ağların en genel özelliği, düğüm gruplarının kendi içerisindeki bağlantıların ağın geri kalanına göre yoğun ilişkiler içinde olduğu topluluk yapılarıdır. Ağ yapıları içerisindeki topluluklar bize bireylerin ortak benzerlikleri hakkında somut veriler sunmaktadır. Şekil 1.1'de iki topluluktan oluşan bir ağ örneği şematik olarak gösterilmiştir.



Şekil 1.1. İki topluluktan oluşan bir ağ yapısı

Gerçek ağ yapılarında karşılaşılan en büyük sorun düğümlerin birden fazla topluluğa ait olabilme ihtimali olarak adlandırılan “örtüşme (overlapping)” durumudur. Sosyal ağların karakteristik özelliklerinden biri olarak, toplulukların örtüştüğü görülmektedir (Şekil 1.2). Örtüşme (Overlapping) faktörünü hesaba katarak yapılan işlemlerin karmaşıklığı çok fazla olduğundan klasik algoritmalar genellikle her bir düğümü bir gruba dahil edecek şekilde çalışırlar ama bu durumda bazı bilgiler göz ardı edilir [6]. Bu çalışmada amaç, bu genelde göz ardı edilen örtüşen toplulukların minimum maliyet maksimum güvenilirlikle tespit etmektir.



Şekil 1.2. Örtüşme durumundaki bir bireyin gösterimi

Bu tez çalışmasında; ağ yapısı, sosyal ağ analizi, sosyal ağların gösteriminde kullanılan graf teorisi, sosyal ağlarda topluluk bulma yöntemleri ve kesişen toplulukların nasıl tespit edilebileceği konuları incelenmiştir. Bu kavramlarla ilgili litaratür taramalarına yer verilmiştir.

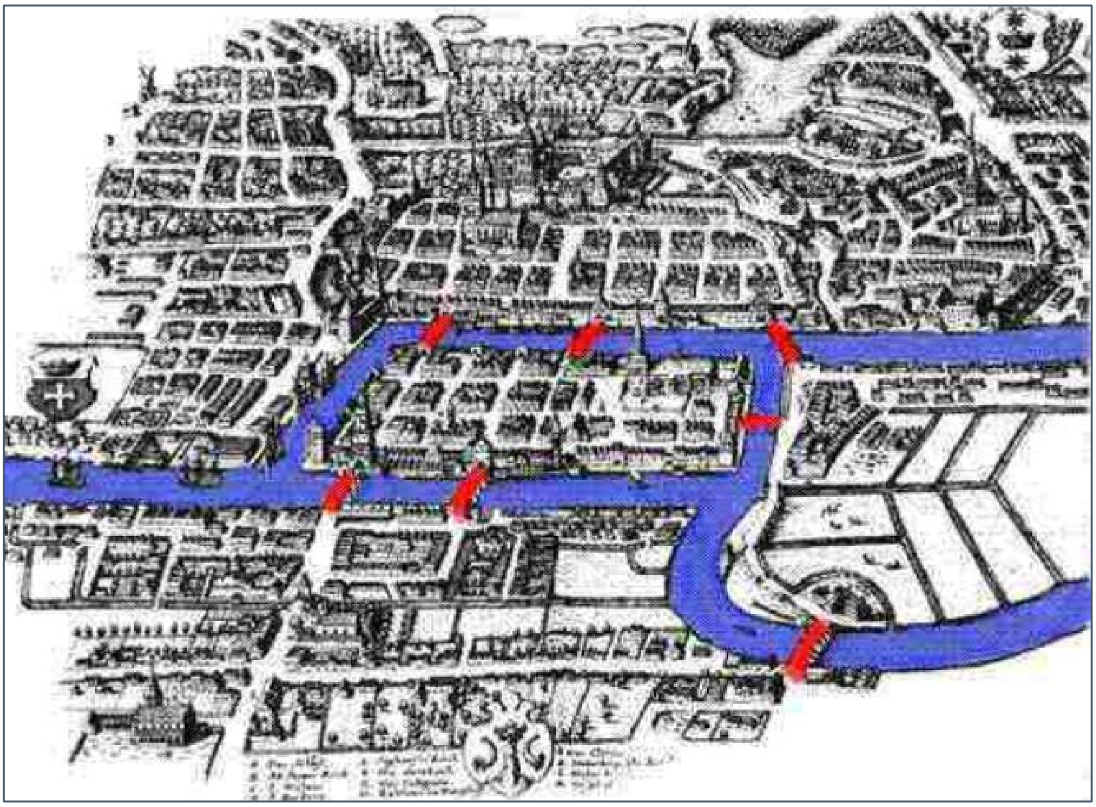
Tezde uygulanan ilk yönteme graftaki tüm tam bağlı alt graflar topluluktur fikri ile yola çıkarak topluluk ve kesişen topluluk tespiti yapılmıştır. Bu yöntem için sosyal ağ bir graf olarak modellenmiştir ve bu graftaki her bir tam bağlı alt graf topluluk olarak kabul edilmiştir. Elde edilen sosyal ağın bitişiklik matrisine Bron-Kerbosch algoritması uygulanmış ve yönsüz graftaki tüm maksimal-klik'ler bulunmuştur. Daha sonra bulunan bu maksimal-klik'leri içeren matrisde sütunları toplayıp iki düğüm içeren bağıntıların olduğu sütunlar sıfır yapılmıştır. Böylece minimum üç düğümü olan maksimal-klik’ler kalmıştır. Bundan sonra elde edilen matrisin satır toplam değerleri elde edilmiştir. Bu değerlerden 1’den büyük olan satırlardaki düğümler örtüşen (en az iki alt-ağ tarafından içerilen) düğümlerdir tespiti yapılmıştır. Böylece tam bağlı alt graflarda kesişen topluluk tespiti yapılmıştır.

Tezde üzerinde çalışılan ikinci yöntem de örtüşen topluluk keşfi problemini çözebilmek için sosyal ağ bir graf olarak modellenmiştir. Grafın Laplace matrisi hesaplanmış, bu graf özdeğer ve özvektörlerine göre iki gruba ayrılmıştır. Daha sonra minimum kesen ayrıt işlemleri yapılarak (Kernighan-Lin algoritması) iki grupta da olma ihtimali olan örtüşen elemanlar tespit edilmiştir.

Tezin ikinci bölümünde graf teorisi ve graf türlerine değinilmiştir. Üçüncü bölümde ağ kavramı ve ağ modelleri ayrıntılandırılmıştır. Dördüncü bölümde sosyal ağlar, sosyal ağların graf olarak modellenmesi, sosyal ağ analizi konuları işlenmiştir. Beşinci bölümde sosyal ağlarda topluluk tespiti konusu incelenmiş ve ilgili konu ile ilgili litaratür taraması sonuçları anlatılmıştır. Altıncı bölümde yukarda da bahsettiğim kesişen toplulukların tespiti işleminde kullandığım iki yöntem ayrı ayrı ayrıntılandırılmış, önerilen yöntemler yapay ve gerçek dünya verileri üzerinde koşulmuş ve sonuçları sunulmuştur.

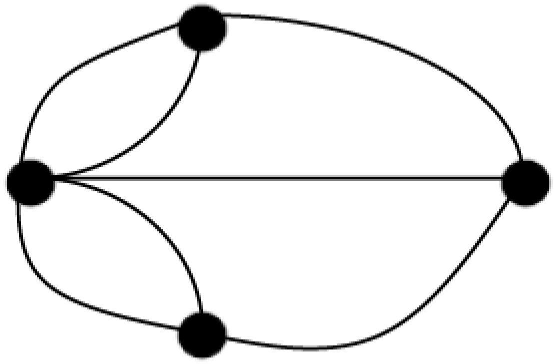
2. GRAF (ÇİZGE) TEORİSİ

Graf (çizge) teorisi 1736 yılında İsviçreli matematikçi Leonhard Euler tarafından Königsberg kasabasının "Yedi Köprü" problemi göz önüne alınarak kaleme alınmış makalesinin yayımlanması ile ortaya çıkmıştır.



Şekil 2.1. Königsberg kasabası ve Königsberg'deki yedi köprü görseli [7]

Königsberg Kasabası'ndaki Pregel nehri, Kneigh isimli adacığın etrafından akarak Şekil 2.1'de görüldüğü gibi iki kola ayrılmaktaydı. Şekildeki yedi köprü, Prusya (Almanya) şehrinde bulunan dört anakarayı birbirine bağlıyor. Hikayeye göre kasaba halkı eğlence amaçlı şehrin farklı noktalarından hareket ederek yedi köprünün her birini yalnızca bir kez geçip yine başladıkları konuma varmayı hedefliyorlarmış ama içlerinden hiçbiri bu amacı gerçekleştirememiş. Şehrin ortak merakı haline gelen bu yedi köprü problemi dönemin ünlü matematikçilerinden Leonhard Euler'in de ilgisini çekmiştir. Euler, problemin çözümlenmesi noktasında, problem üstünde rahatça analiz ve yorumlama yapılabilinecek şekilde temsil edecek grafı çizmiştir (Şekil 2.2) [8,9].



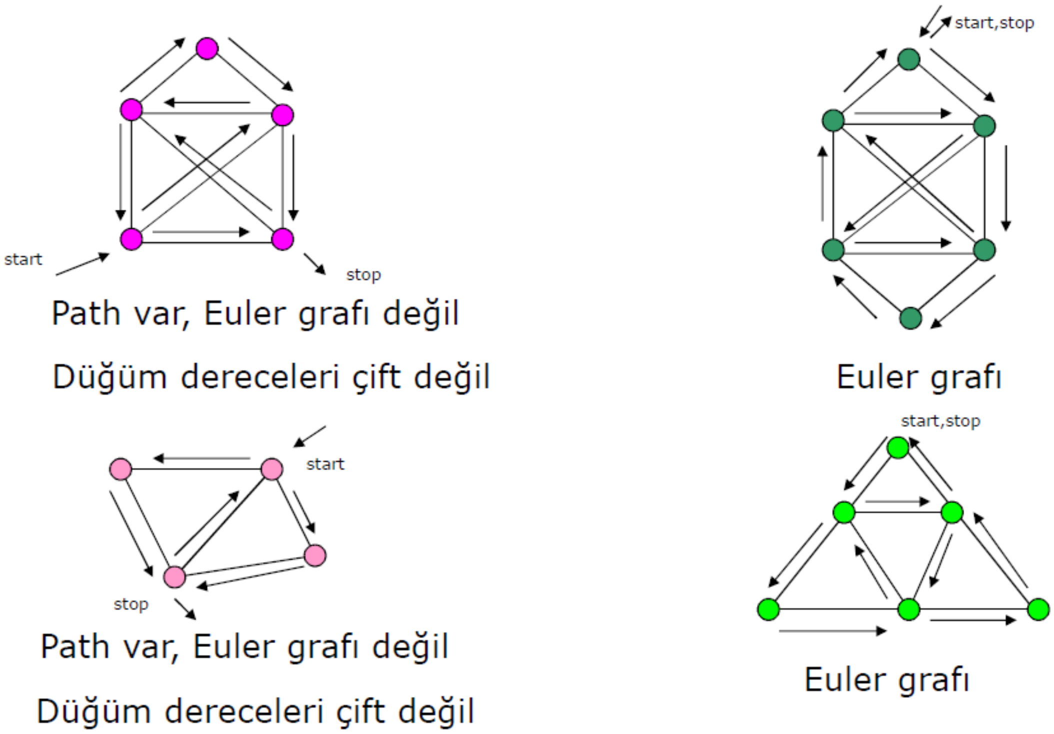
Şekil 2.2 Königsberg'in grafla gösterimi

Euler, kara parçalarının her birini bir noktayla, köprüleri de kenar denilen çizgilerle temsil etmiştir. Problem, graf teorisi ile "Herhangi bir noktadan harekete başlayıp, bütün kenarlardan bir ve yalnız bir defa geçerek, bütün noktaları ziyaret ettikten sonra başlangıç noktasına varabilir miyiz?" şeklinde tanımlanmıştır.

Euler, çalışmalarının sonucunda bunun mümkün olabilmesi için tüm noktaların çift dereceli olması gerektiğini ispatlamıştır. Yukarıdaki grafta da tüm noktaların dereceleri tek olduğundan Königsberg probleminin çözümünün mümkün olmadığı görülür [9].

Burada Euler, noktaların tümünün derecelerinin çift olması gerektiğini şu şekilde elde etmiştir: Bu tür bir probleme bir noktadan başlanıldığında ve herhangi bir noktaya varıldığında bu noktaya bir tane giren bir tane de çıkan kenar olmalıdır. Bu da noktanın derecesi çift olduğu manasına gelir. Bu graftaki bütün noktalar için geçerli bir durum olmalıdır, ama biri gezintiye başladığımız diğeri de gezintiyi sonlandırdığımız nokta ise bu iki noktanın derecesi tek olabilir. Ancak Könisberg köprüleri probleminde, başlanılan noktaya geri dönülmesi amaçlandığından bu iki noktanın da derecesi çift olmalıdır. Buradan hareketle bu problemin çözülebilmesi ve grafiğinin çizilebilmesi için gerek ve yeter koşul tüm noktaların derecelerinin çift olmasıdır.

Bir G grafına Euler grafıdır yorumu yapabilmek için Euler cycle'ına sahip olması gerekir. Bir Euler grafında bütün düğümlerin derecesi çifttir. Konigsberg köprü probleminde düğüm dereceleri çift değildir ve bu graf bir Euler grafı olmadığından bu problemin çözümü yoktur. Euler formülü; tek parça ve düzlemsel bir grafın bölge sayısı b, hat sayısı k, düğüm sayısı n ise $b - k + n = 2$ eşitliğini verir.



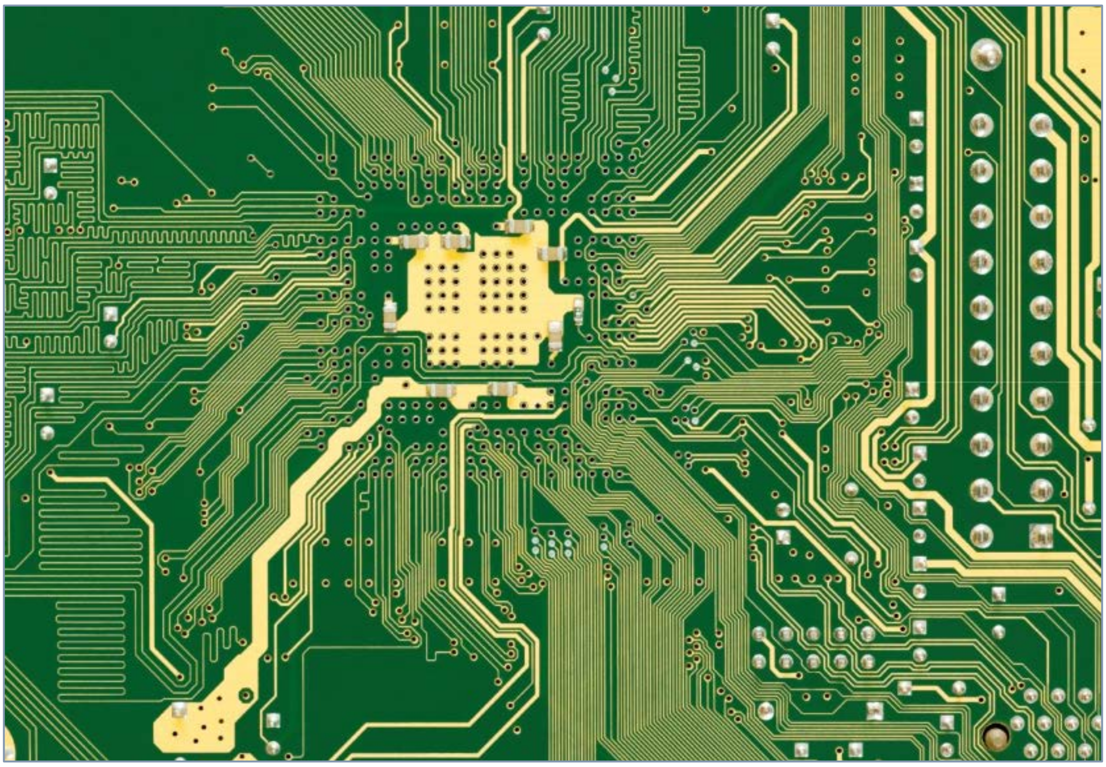
Şekil 2.3 Örnek Euler ve Euler olmayan graflar [10]

Konigsberg bridge problemi, graf teorisinin başlangıç problemi olarak bilinir [9]. Grafların geliştirilme basamakları Şekil 2.4'de verilmiştir.

Yıllar	**Bilim Adamları**	**Graf teorisine katkıları**
1736	Leonard Euler	" Könisberg'in yedi köprüsü" makalesini yayımlayarak graf teorisinin başlangıcını oluşturduğu ve topoloji ile ilişkisini göstermiştir.
1822	J.J. Silvester	İlk kez graf sözcüğünü kullanmıştır.
1845	Gustav Kirchhoff	Elektrik devrelerinde akım ve gerilimleri hesaplamaya yardımcı olan ünlü Kirchhoff devre kuramını yayımlamıştır.
1852	Francis Guthrie	Yanıtlaması zor olan 4 renk problemini ortaya atmıştır.
1927 1930	Pontryagin K.Kuratovski	Sonralar VLSI teknolojisinde önemli kullanım alanı bulan düzlemsel grafların özelliklerini bulunmuştur.
1936	D.König	Graf teorisi ile ilgili ilk kitap yayınlanmıştır.
1976	Kenneth Appel, Wolfgang Haken	1200 saati aşan bilgisayarlı hesaplamalar neticesinde haritanın renklendirmesi için 4 rengin gereklilik ve yeterliliği sağladığı ispatlanmıştır.

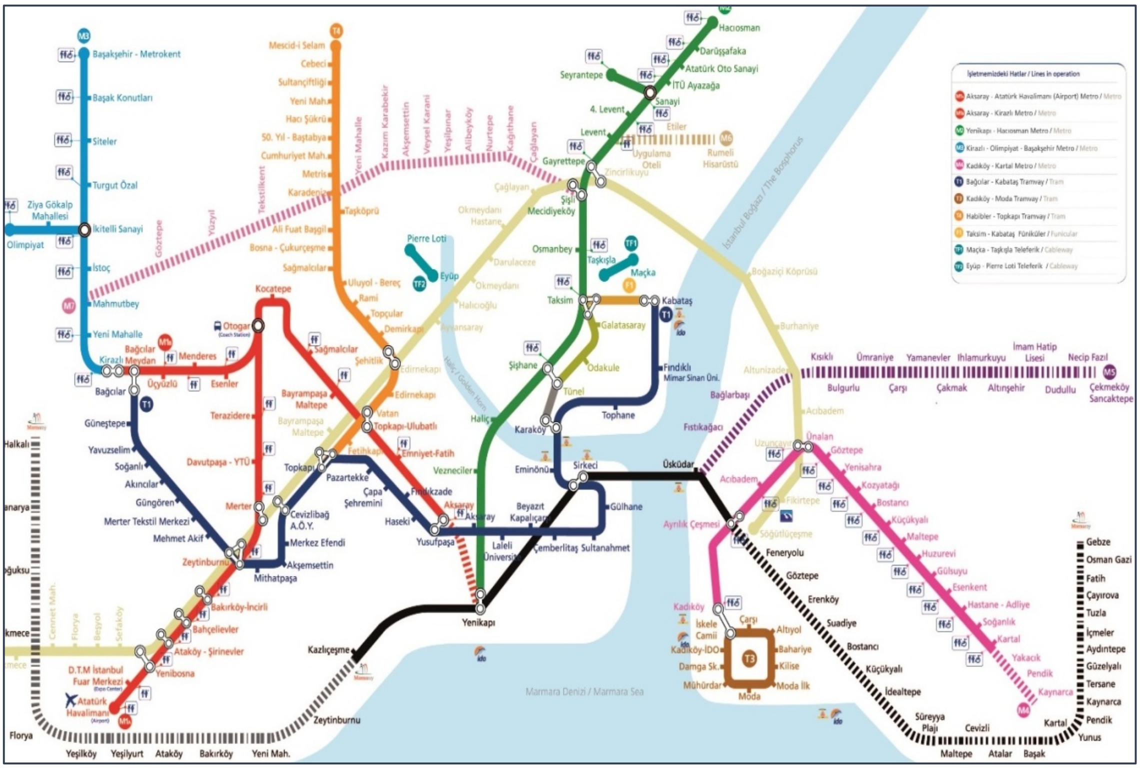
Şekil 2.4. Grafların gelişim adımları

Fizik, kimya, biyoloji, istatistik, ekonomi, işletme, bilgisayar bilimleri, mühendislik gibi bilim alanlarında var olan problemlerin çözümünde, problem unsurları ve aralarındaki ilişki çoğu kez çizgelerle ifade edilir. Böylece problemin çözümüne graf teorisiyle rahatça ulaşılması amaçlanır. Başlıca grafların uygulama sahaları; elektronik devreler (baskı devre kartları (PCB), entegre devreler) (Şekli 2.5), ulaşım ağları (otoyol ağı, havayolu ağı) (Şekil 2.6, Şekil 2.7), bilgisayar ağları (lokal alan ağları, internet) (Şekil 2.9), veritabanları (entity relationship diyagram) (Şekil 2.8), kimya&biyoloji(Şekil 2.10) , soy ağaçları, kan dolaşımı vb. dir.

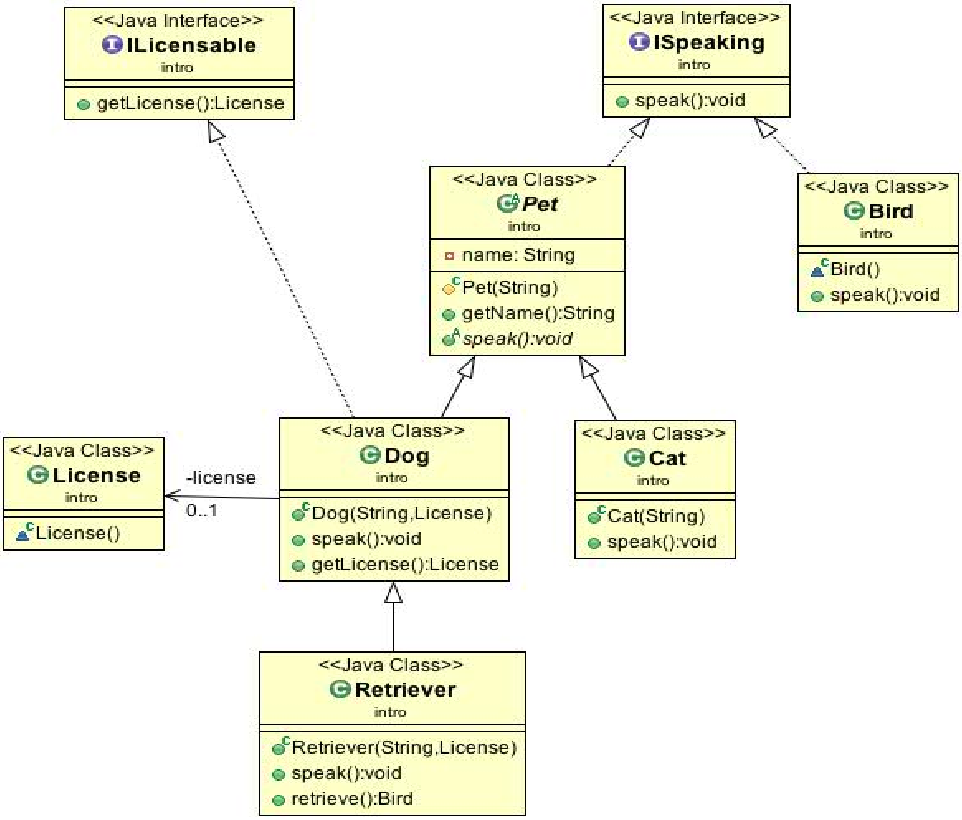


Şekil 2.5. Elektrik devreleri

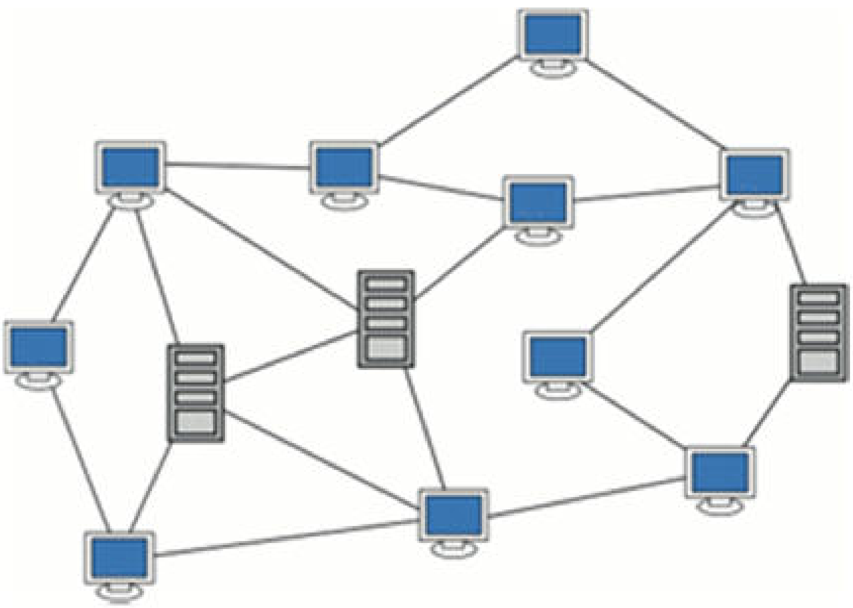
Şekil 2.6. Thy İstanbul-Orlando arası uçuş rotasının graf ile gösterimi



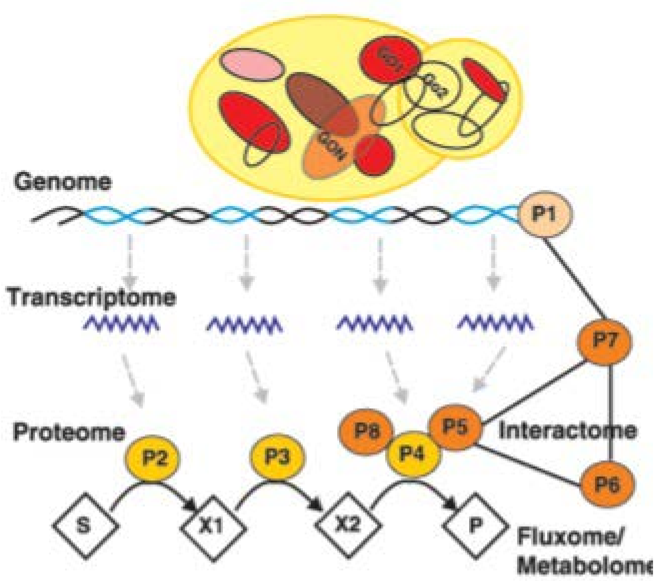
Şekil 2.7. İstanbul metro ağı



Şekil 2.8. UML diyagramları graf veri yapısında tasarlanır.



Şekil 2.9. Bilgisayar ağları



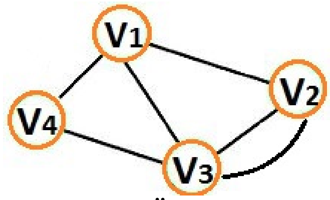
Şekil 2.10. Moleküler Biyolojide ve kimyasal reaksiyonların gösteriminde graf teorisinden faydalanılır.

Sosyal yapının kendisi bir çizge, içerdiği varlıklar bu çizgedeki düğümler ve varlıklar arası ilişkiler de çizgedeki ayrıtlar olarak ele alınmaktadır. Çizgeler ve bu çizgelerden elde edilen matrisler üzerinde gerçekleştirilen çeşitli matematiksel hesaplamalar ile düğümler arasındaki kümelenme (clustering), merkezilik, benzerlik, uzaklık-yakınlık vb. ilişkiler somut olarak hesaplanmakta ve bu değerlere bağlı olarak ağdaki ilişkiler yorumlanmaktadır [4].

Graf noktalar kümesi olan düğümlerden ve bu noktaları ya da sadece noktanın kendisini birleştiren ayrıtlardan oluşur. Örneğin bir yol haritasında şehirleri düğüm (vertice) ve bunlar arası yollar ayrıt (edge) olarak ifade edilebilir.

Bir grafı tanımlamak için düğümlerin ve ayrıtların kümesi tanımlanır. Ardından hangi ayrıtların hangi düğümlerle bağlandığı belirtilir. Bir G grafı V=$\{v_1$, $v_2, \dots, v_n\}$ ile gösterilen düğümlerden (vertices) ve E=$\{e_1$, e_2, ... , $e_m\}$ (E ⊆ VxV) ile gösterilen ayrıtlardan (edges) oluşur ve G = (V,E) şeklinde gösterilir. Her ayrıt iki düğümü birleştirir. Her bir e ayrıtı için $\{v_1 , v_2\}$ kümesi tanımlanır (e ayrıtı v_1 ve v_2 düğümlerini bağlar). $\{v_1, v_2\}$ kümesi δ(e) ile gösterilir ve düğümler kümesinin bir alt kümesidir [11].

Grafları kenarları yönlendirilmiş ise yönlü graf, aksi durumda ise yönsüz graftır. Örnek olarak Facebook sosyal ağı üzerinde arkadaş olan kişilerin oluşturduğu yapı verilebilir. Birisini arkadaş olarak eklediğinizde karşıdaki kişi de isteği kabul ederse arkadaş olunur ve bu ilişki karşılıklı yani yönsüz (e_{ij} ile e_{ji} aynı ayrıt) bir graf ile temsil edilir. Twitter ya da instagram gibi sosyal ağlarda ise takip ettiğiniz kişinin sizi takip etmesi zorunlu değildir ve ilişkiler yönlü graflarla temsil edilir.



Şekil 2.11. Örnek G grafı

Şekil 2.11'de yönsüz örnek bir graf verilmiştir. Bu graf için küme gösterimi aşağıdaki şekilde yapılır (G graf, v düğümler, e ayrıtlar).

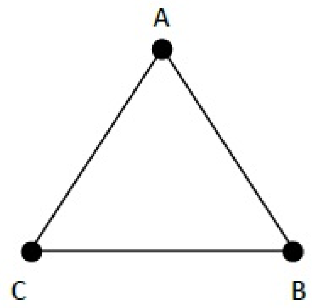
V = { v_1, v_2, v_3, v_4 }
E = {(v_1, v_2), (v_1, v_4), (v_1, v_3), (v_3, v_2), (v_3, v_2), (v_3, v_4)}
G = (V,E)

Bir grafta bir veya daha fazla düğümden ve kenardan geçen rotaya f (path) denir. Graf teorisinde bir düğümden başlanıp herhangi bir güzergah izlenerek aynı düğüme gelinmesi durumuna döngü (cycle) denir. Örneğin yukarıdaki Şekil 2.10’daki v_1 düğümünden başlanarak yine bu düğümde biten { v_1, v_2, v_3 } yoluna döngü denir. Bir grafta döngü olup olmadığı (cycle detection) problemi, literatürde üzerinde araştırma yapılan konulardandır. Yönsüz bir grafta bulunan bir döngünün bulunma karmaşıklığı O(n) olacaktır [12].

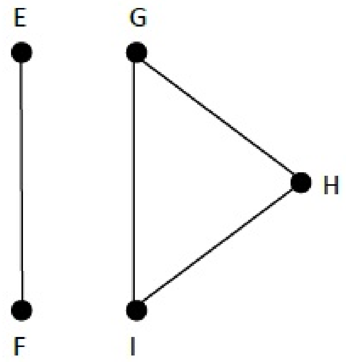
Bir graftaki bütün düğümler diğer tüm düğümler ile ilişkili ise bu grafa tam bağlı graf denir. Grafta tam bağlı graf bulunup bulunmadığı ya da grafın bir alt grafa bölünüp bölünemeyeceği de tıpkı döngü bulundurma problemi gibi üzerinde çalışılan konulardandır. Literatürde tam bağlı alt graflara özel olarak klik (clique) ismi verilir. Klik ifadesi aynı zamanda sosyoloji gibi beşeri çalışma alanlarında bir toplulukta çok yakın arkadaş gruplarını tanımlamada kullanılır (bir okul veya bir iş yerindeki yakın arkadaş grubu). Bu tanım grafın her düğümünün diğer tüm düğümler ile doğrudan doğruya bir bağlantısının olduğu alt grafı ifade etmektedir. Örneğin Kosaraju Algoritması, Tarjan Algoritması veya Dijkstra’nın tam bağlı graf algoritması gibi farklı algoritmalar, tam bağlı alt grafları bulmayı ve bir graf tam bağlı mı değil mi bunu bulmayı amaçlar [13,14].

Graf teorisinde kullanılan diğer terimlerden bazıları da şunlardır:

İki grafın birleşimi olarak yazılamayan bir grafa bağlı graf (Şekil 2.12), aksi halde parçalı graf denir (Şekil 2.13) [15].

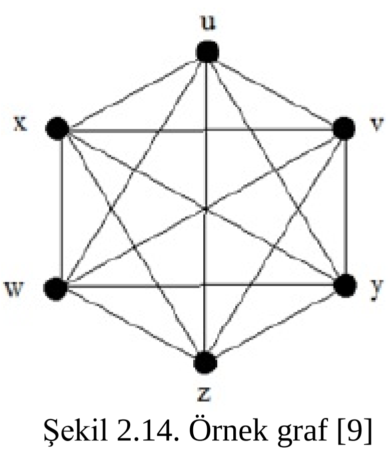


Şekil 2.12. Bağlantılı graf

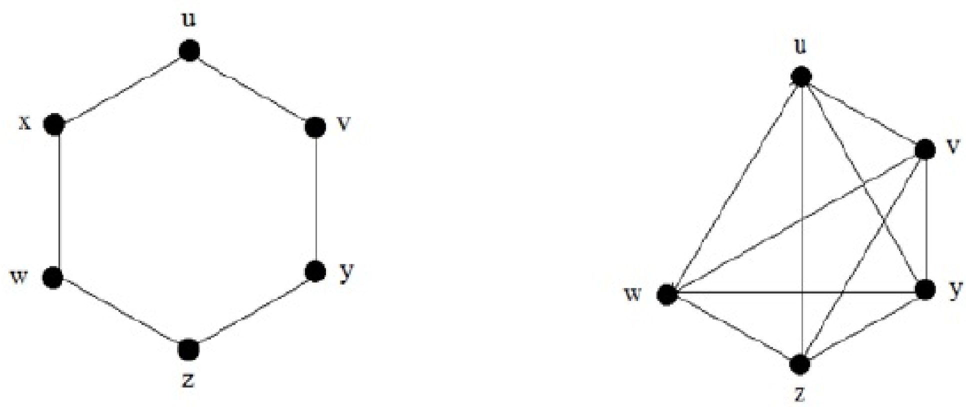


Şekil 2.13. Bağlantısız graf

G, V(G) düğümler kümesi ve E(G) ayrıtlar kümesine sahip bir graf olsun. G'nin bir alt grafı düğümleri V(G)'ye ayrıtları E(G)'ye ait başka bir graftır. Şekil 2.14'de örnek bir graf ve Şekil 2.14'de bu grafının bazı alt grafları gösterilmiştir [9]. Her graf kendisinin alt grafıdır. G grafının herhangi bir noktası tek başına G'nin alt grafıdır. G'deki tek bir ayrıt kendi başlangıç ve bitiş noktaları ile birlikte G'nin alt grafıdır.



Şekil 2.14. Örnek graf [9]



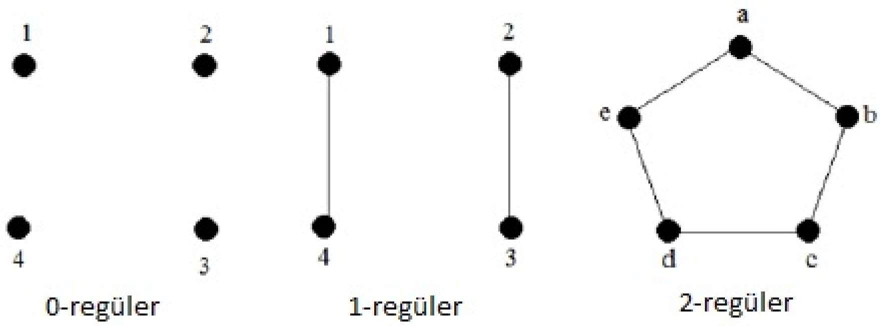
Şekil 2.15. Alt graf örnekleri [9]

Bir grafın düğümler kümesinin eleman sayısı o grafın mertebesini verir. G=(V,E) için eğer |V|=n ise G grafının mertebesi n'dir. v gibi bir düğüme giren ve çıkan ayrıt sayısı o düğümün derecesini verir ve d(v) ile gösterilir.

G=(V,E) , bir yönlü graf olsun. Eğer V = ∅ veya V = ∅ ve E = ∅ ise bu tip graflara boş graf denir ve $\forall v_j \in$ V için d(v_i) = 0 olur. Sadece tek bir düğümden, tek bir döngüden oluşan veya paralel ayrıt içermeyen graflara basit graf (yalın çizge) denir. Sadece çevreleri (döngüleri) içeren graflara sözde graf denir. Bu tez çalışmasında basit graflar esas alınmıştır.

Herhangi bir grafta tüm düğüm dereceleri eşit ise böylesi graflara düzenli graf denir. Eğer bir graf paralel ayrıtlar bulunduruyorsa bu tip graflara çoğul graf (multi-çizge) denir. Her türlü ayrıtı bulunduruyor ise bu graflara karmaşık graftır denir. Tam olan bir grafta her düğüm diğer bütün düğümler ile komşudur ve (2.1) bağıntısı v_j düğümünün komşu düğümler kümesini vermektedir.

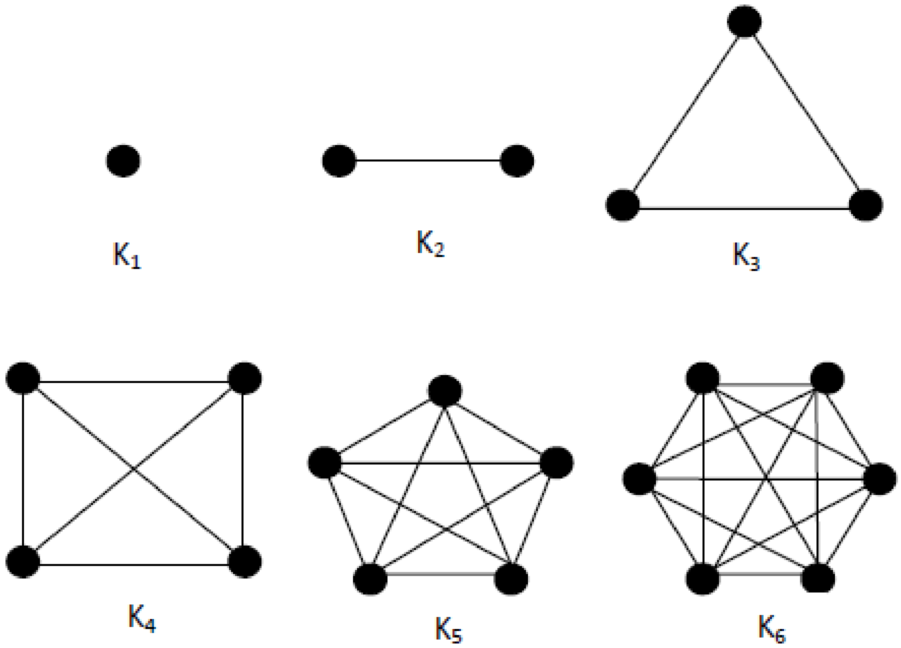
$$\forall v_j \in \mathrm{V},\ \mathrm{Adj}\ (v_j) = \{v_1, v_2, \ldots, v_i, v_{i+1}, \ldots, v_{|v|}\} \tag{2.1}$$



Şekil 2.16. Regüler graflar

2.1. Tam (complete) Graflar

Graftaki her bir düğümün diğer bütün düğümlerle arasında bir ayrıt varsa, yani ihtimal dahilinde ki bütün ayrıtlara sahipse, bu tür graflara tam bağlı (completed) graf denir. n düğüm sayısı olmak üzere K_n ile gösterilir (Şekil 2.17).



Şekil 2.17. Tam Bağlı Graflar

K_4'ün biri hariç tüm K_n'lerin bir düzgün n-genin içine tüm köşegenleri çizilerek elde edildiği görülür. Ayrıca K_n, n-1'inci mertebeden regülerdir. Bu tür bir grafta tüm düğümlerin dereceleri birbirine eşit ve derecesi toplam düğüm sayısının bir eksiği kadardır. Dolayısıyla $\frac{n(n-1)}{2}$ kenarı vardır. Bu çalışmada tam bağlı alt graflar (klikler) kullanılacak graf türlerindendir.

2.2. Grafların Bilgisayar Ortamında Temsili

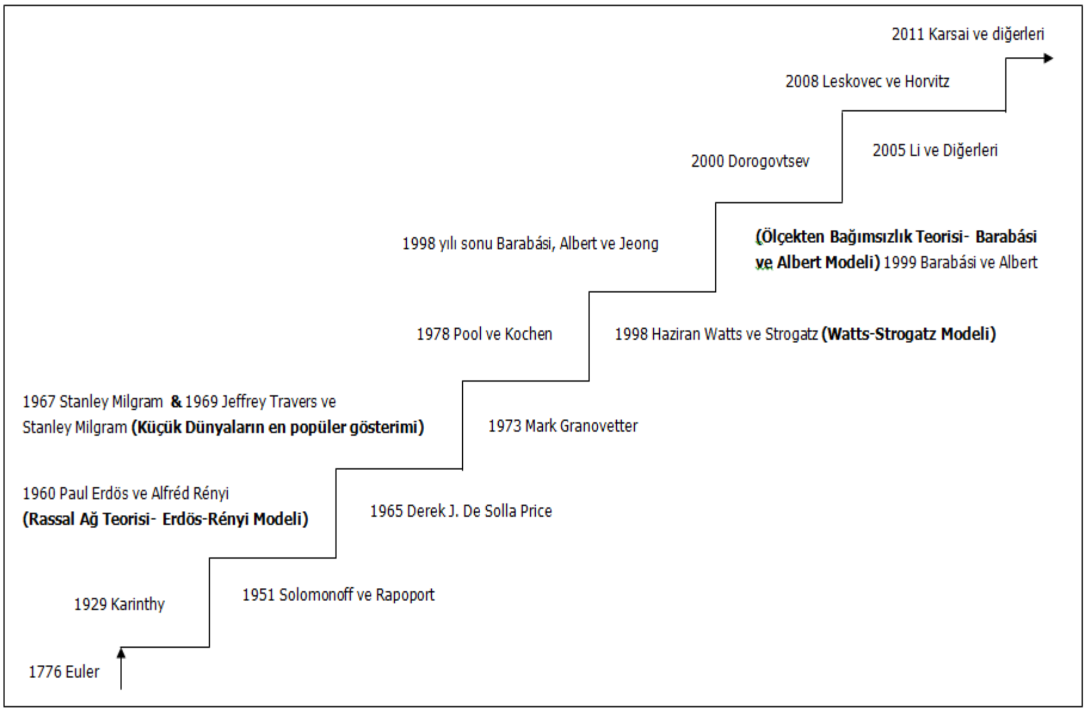
Grafların bilgisayar ortamında ifade edilmesinde kullanılan başlıca sosyal ağ uygulama araçları UCI-NET, Pajek, NetMiner, NetDraw, ORA, SocNet-V, Stat-Net, InFlow ve Keyhubs'tır. Bu yazılımlardan UCI-NET, Pajek, NetMiner sıklıkla kullanılmaktadır.

3. AĞ

“Ağ nedir?” diye sorulunca çoğu insanın aklına örümcek ağı, voleybolda sahaya gerilen file, sinirlerin oluşturduğu sinir ağı, bir şehrin ulaşım veya kanalizasyon ağı, futbol sahasında kaleye gerilen ağ, bilgisayarları birbirine bağlayan internet ağı, yapay sinir ağları gibi yapılar akla gelir. Ağ bilimi ağı, bir karmaşık sistemin, sistemi oluşturan parçaların ve bu parçalar arasındaki etkileşimlerin görsel gösterimi olarak tanımlar.

Günlük hayatta biyolojik ağlar, fiziksel ağlar ve sosyal ağlarla etrafımız çevrilidir. Biyolojik ağlar, ekolojik ağlar, soylar arasında hiyerarşinin tutulduğu gen ağları, gen haritaları ve benzeri ağlardır. Fiziksel ağlar ise, elektrik dağıtım ağları, su dağıtım ağları, kanalizasyon ağları, telefon, internet gibi telekomünikasyon ağları, karayolu, havayolu, demiryolu ve deniz yolu gibi ulaşım ağları ve benzeri ağlardır. Sosyal ağlar, insanların etkileşiminden doğan ağlardır. Aile bireyleri arası ilişki, bir iş yerindeki çalışanların birbirleriyle etkileşiminden oluşan ilişkiler, arkadaşlık bağları gibi ağlar birer sosyal ağdır. Biyolojik ağlar, fiziksel ağlar ve sosyal ağların analizi ile kıymetli veriler elde edilir ve bunlar çeşitli problemlerin çözümünde kullanılır.

Ağ bilimine ilişkin literatürün 1776 yılında Euler ile başladığı ve günümüze kadar gelişerek devam ettiği kabul edilmektedir. Şekil 3.1' de ağ biliminin tarihindeki önemli gelişmelerin kronolojik sıralaması yer almaktadır.



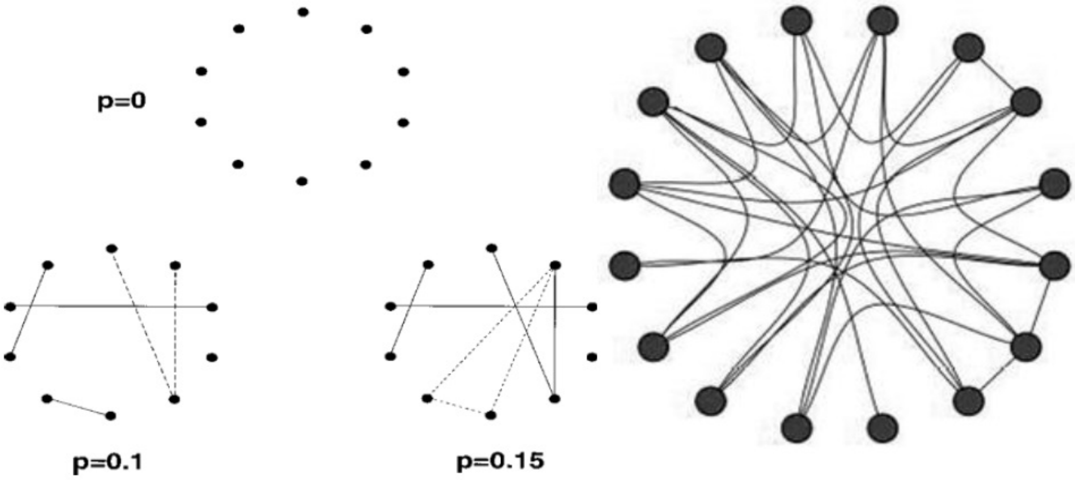
Şekil 3.1 Ağ biliminin tarihindeki önemli gelişmeler [21]

Ağ bilimi; disiplinler arası çalışma, deneysel ve veri güdümlü yapı, matematiksel yapı ve hesaplanabilir yapı özelliklerine sahiptir. Ağların graf teorisi (graph theory), bağlantı (edge), düğüm (nodes), komşuluk matrisi (adjancency matrice), en kısa yol (the shortest path), yarıçap, yoğunluk (density), derece ve derece dağılımı, merkezicilik (centrality), kümelenme katsayısı (clustering coefficient) vb. yapısal özellikleri vardır.

Ağ bilimciler aslında üzerinde çalıştıkları alan ağların oluşum safhaları veya nasıl meydana geldikleri, ağların görünüşlerini ve yapılarını düzenleyen yasaların neler olduklarıdır. Ağ biliminde, ağların işleyiş kurallarını, kanunlarını belirlemek ve bunların ne olduklarını açıklamak amacı ile bilim adamları çeşitli ağ modelleri ileri sürmüştür. Bunlar; Rassal ağ ((ER modeli) Erdös-Renyi modeli), küçük dünya ağı ((WS modeli) Watts-Strogatz modeli) ve ölçekten bağımsız ağ modelleri (scale free networks)(Barabasi-Albert modeli (BA modeli)) dir [21].

3.1. Endös-Renyi Rassal Ağ Modeli

Bu ağ modelinde düğümlerin graftaki diğer düğümlerle bağ kurması eşit olasılıklıdır. Özellikle çok sayıda düğüm ve bağlantı içeren ağların analizinde kullanılan en popüler yaklaşımlardan biri Erdos ve Renyi tarafından geliştirilmiş olan düzensiz grafiktir.



Şekil 3.2. Rassal Ağ'ın oluşum süreci & Rassal Ağ'ın görünümü [22]

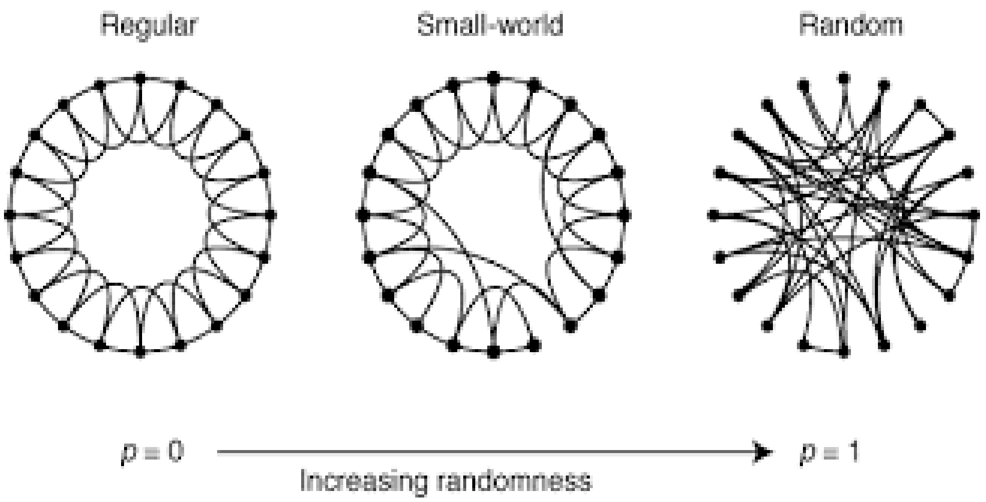
Bir grup rassal noktadan düzensiz bir grafik oluşturma işleminde herhangi iki nokta arası bağlantı olma olasılığı tanımlanır. Tüm noktalar arasında bu olasılığı gerçekleyecek bağlantılar kurulur. Oluşturulan bu grafiğin nokta sayısı çok fazla olsa dahi küçük dünya etkisi vardır. Fakat küçük dünya etkisi hesaplanması gereken tek özellik değildir. Karmaşık sistemlerde görülen yoğun komşuluk ilişkilerini ve aşırı bağlantı içeren noktaları düzensiz grafiklerde göremeyiz. Düzensiz grafiklerin zayıf komşuluk ilişkileri vardır ve bir Poisson eğrisini takip eden bağlantı sayısı dağılımına sahiptir.

3.2. Küçük Dünya Ağları

Dünya üzerinde rastgele seçilen bir kişi ile yine rastgele seçilen diğer bir kişinin asasında herhangi bir sosyal ilişki olma olasılığı ve bu yolun kısalığı küçük dünya ağını tanımlar.

Dünyada rassal olarak seçilen herhangi iki insanın birbirine az sayıda adım ile bağlandığını gösteren küçük dünya etkisi en ünlüsü John Milgram tarafından yapılan deneyler sonucu keşfedilmiştir. Böyle deneylerde bulunan en iyi adım sayısı altıdır.

Omaha ve Wichita’da yaşayan rastgele seçilmiş kişilerden birer mektup yazmalarını isteyen John Milgram, bu mektupları Boston'da yaşayan hiç tanımadıkları birine göndermeleri istenmiştir. Bu kişilere hedefteki kişinin adres bilgisini direk olarak kullanmaları yasaklandı ve bunlardan hedefe ulaştırması kendilerinden daha muhtemel olduğunu düşündükleri arkadaşlarına göndermeleri istendi. Bu deneyde mektuplar 5 veya 6 kişinin elinden geçtikten sonra hedefteki kişiye ulaştığı gözlemlenmiştir. Bu sonuç oldukça ilginçtir, özellikle Amerika gibi büyük bir ülkede bu kadar sınırlı adımda hedefe ulaşılması oldukça şaşırtıcı bir sonuçtur. Sosyal ilişkileri tanımlayan küçük dünya özelliği analizi yapılması gereken ilginç veriler sunduğu ortaya çıkmıştır [23].

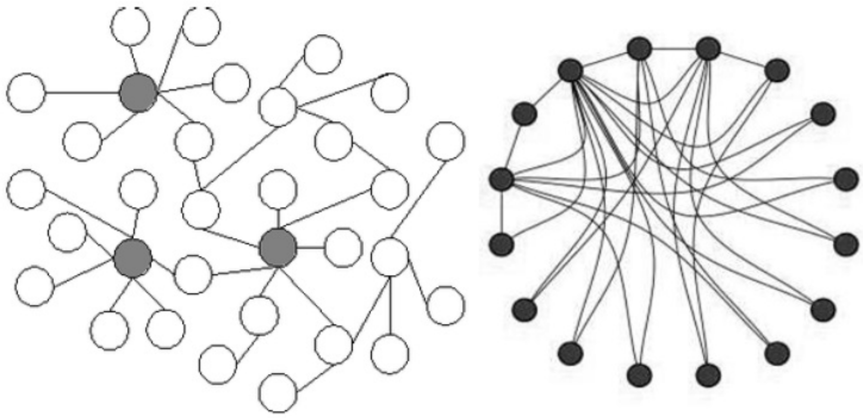


Şekil 3.3. Küçük dünya ağının oluşum süreci [24]

Bu model Duncan J. Watts ve Steven Strogatz tarafından geliştirilmiştir. Düzenli bir grafik ile başlayıp rastgele seçilen noktaları farklı noktalarla birleştirerek grafiği yeniden inşaa eder (Şekil 3.3). Bu yöntem gerçek hayatta karşılaşılan ve Endös Renyi yönteminde de gözlemlenen küçük dünya etkisini korur. Bu karmaşık ağ teorisi açısından oldukça önemli bir gelişme olarak değerlendirilse de beklenmedik bir şekilde bağlı noktalar oluşturamamaktadır.

3.3. Ölçekten Bağımsız Ağlar

Poisson dağılımına sahip rassal ağlar veya küçük dünya ağında düğümlerin çoğunluğu aynı sayıda bağa sahiptir. Barabasi'nin tanımladığı bu ağlarda ters orantı gözlemlenir ve çok sayıda düğümün az sayıda bağlantı ile eşleştiği, az sayıda düğümün ise çok sayıda bağlantı ile eşlendiği gözlemlenmiştir (Şekil 3.4). Bu model kuvvet yasası dağılım eğrisine sahiptir ve Poisson dağılım eğrisi gibi bir tepe noktası yoktur.



Şekil 3.4. Ölçekten bağımsız ağın oluşum süreci ve görünümü [22]

4. SOSYAL AĞLAR

Son yıllarda sanal ortamların kullanımının artmasıyla beraber insanların günlük alışkanlıklarında birçok değişiklik meydana gelmiş ve gelmeye devam etmektedir. Internet üzerinden alışveriş, internet bankacılığı, çevrimiçi haberleşme gibi birçok alanda eski alışkanların yerine internet üzerinden kullanım tercih edilmeye başlanmıştır. Gerçek dünyadan sanal dünyaya geçişin veya sanal dünyadaki ilginin artmasının temel nedeni internet olanaklarının iyileşerek (özellikle web 2.0 teknolojisinin gelişimi) günün herhangi bir anında kolay, hızlı, uygun bir şekilde kullanımının sağlanması ve bağlantılı olarak sosyal ağ uygulamaları ile sosyal medya sitelerinin hızlı çıkışıdır.

Sosyal medya siteleri kendi içerisinde farklı kullanımlara olanak sağlayarak farklı kitledeki kullanıcıların da bu siteleri kullanımını sağlamıştır. Örneğin, kullanıcılar Facebook sitesini daha çoğunlukla ilgi alanları, fotoğraflarını veya videolarını paylaşmak için kullanırken, Linkedin sitesini iş arama, meslektaşlarını ile iletişim kurma amaçları ile kullanmaktadırlar. Benzer şekilde Twitter ise, genelde fikirlerin ve yorumların paylaşıldığı bir ortam olarak kendini göstermektedir [5].

Sosyal medya denilen mecra son yıllarda bilhassa reklam sektörünün yapıtaşı olmuştur. Standart pazarlama ve reklam sektöründe verilmek istenen mesaj üzerinde düşünülür. Bir iş çıkması adına dizayn, baskı, tasdik, dağıtım, montaj, ulaşım ve lojistik gibi eşittir maliyet denilebilecek bir dolu süreç varken sosyal medyada hedef kitleye ulaşmak için iyi bir fikir olması yeterli ve gerek koşuldur. Ulaşmak istenen kitleye zaman geçirmeden ulaşılır ve bir sonuca varılır. Hedef kitlenin tepkisi anında ölçülebilinir, bu doğrultuda kampanya veya verilmek istenen mesaj anında güncellenebilir. Maliyetler geleneksel pazarlama ile karşılaştırılamayacak kadar azdır.

"Sosyal Ağ" terimi ilk kez 1954 yılında J. A. Barnes kullanılmıştır ve sosyal ağı düğümler olarak adlandırılan öğelerden veya bireylerden oluşan ve bu düğümler aralarındaki etkileşimlere birbirine bağlayan bir yapı olarak tanımlanmıştır. David Liben-Nowell'e göre ise sosyal ağ, sosyal bağlamda kişilerin veya diğer çoklukların aralarındaki etkileşimi, yardımlaşmayı, etkileri gösteren bir yapıdır [25]. Teknik olarak günümüzde sosyal ağ kavramı için akla gelen ilk tanım, bireylerin sayısal platformlar üzerinden duygu, düşünce, resim, müzik, video vb. içerik paylaşımında bulunduğu, yakınları, arkadaşları ve diğer dış dünya ile iletişim kurdukları sanal ortamlar olan sosyal medyadır.

Yapılan bu sanal etkinlikler ve düğümler sosyal aktörler olarak adlandırılır, bu düğümleri birbirine bağlayan ilişkiler, çeşitli anlamlarda karşımıza çıkmaktadır. Bunlar, tanışıklık, akrabalık, aynı siyasi, ideolojik ve dini görüşler, yakınlık, benzerlik, ticari etkileşim, fiziki bağlılık, iletişim, yönlendirme, fan kulüpler vb. olabilir. Özellikle internet teknolojisi, mesafe ve zaman kısıdını ortadan kaldırarak sosyal ağların oluşması için mükemmel bir altyapı teşkil etmektedir. Son yıllarda, özellikle genç kuşağın, neredeyse günlük yaşamlarının tamamını sosyal ağlar üzerinde yaşar ve paylaşır hale geldikleri gözlemlenmektedir. En yaygın sayısal sosyal ağlara örnek olarak Facebook, Twitter, Instagram, Periscope, Swarm, Whatsapp Grupları örnek gösterilebilir [4].

Sosyal ağlar bilimsel olarak birbirinden farklı ifade ve modellenme yöntemleri ile temsil edilebilinir. Literatürde en çok kabul görmüş ve uygulamada en çok kullanılan gösterim şekli graf teoremi (graph theory) dir. Graf teorisi grafları kullanılarak ağdaki bireylerin veya varlıkların birer düğüm (node) ve ilişkilerin birer kenar (edge) olarak temsil edildiği modelleme türüdür.

Sosyal medyanın popülaritesindeki artış ile sosyal ağların analiz edilmesi sosyal ağlardan somut veriler elde edilmesi işlemleri de son zamanlarda araştırmacıların ilgisini cezp etmektedir. Sosyal ağ analizini hakkında Barry Wellman " Sosyal ağ analizi sadece bir metod değildir. Toplumsal eylemleri açıklamak için en önemli paradigmadır. Çünkü şu anda hayatımızın her alanında ağlaşmış bir bireysellik var ve bunu ancak sosyal ağ analizi ile inceleyebiliriz." şeklinde tanımlanmıştır [26].

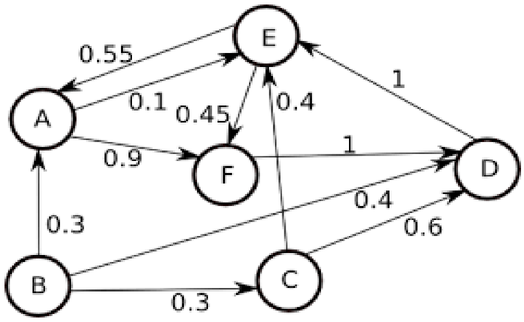
Sosyal ağ analizi, bir sosyal ağdaki düğümler kümesi arasındaki her türlü ilişkilerin analizidir. Sosyal ağları incelemek için birçok farklı teknik, yaklaşım ve perspektif kullanılmaktadır. Bu tekniklerden bazıları ağdaki zayıf veya kuvvetli bağları bulmaya odaklanırken, bazıları topluluk bulmayı amaçlar. Sosyal ağ analizinde ilgilenilen konular bağlantı tahmini (link prediction), eğilim ve fikir tahmini, duygu analizi (sentiment analysis), ağlardaki topluluklar ve toplulukların keşfedilmesi, örtüşen toplulukların keşfi vb. konulardır. Sosyal ağlar, düğümler arasındaki ilişkilere göre ağırlıklandırılmış veya ağırlıklandırılmamış, yönlendirilmiş veya yönlendirilmemiş olarak tanımlanabilir.

4.1. Ağırlıksız Graf

Bütün kenarlarının değeri aynı olan graflardır. Ayrıt ağırlıkları bir olduğu kabul edilir.

4.2. Ağırlıklı (weighted) Graf

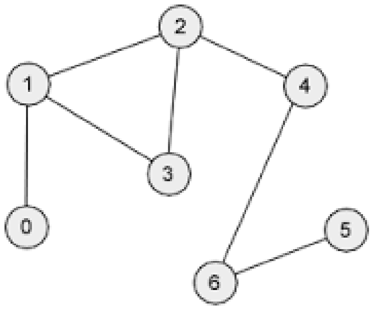
Graf yapısında kenarlar değer alabilir ve bu değerler grafın yapısına katılır. Şekil 4.1'de örnek grafta olduğu gibi kenarların değerleri eşit değilse ve her bir kenar farklı değer alabiliyorsa böylesi graflara maliyetli veya ağırlıklı graf (weighted graph) denir. Şehirlerarası mesafenin kenarlara değer olarak verildiği yol haritalarını gösteren graflar, ağırlıklı graflara örnek olarak gösterilebilinir. Benzer şekilde iş akış şemalarında, her işin bitirilme sürecini gösteren graflar da yine ağırlıklı graflara başka bir örnektir. Graftaki bütün kenarlara ait ağırlığın toplamı o grafın toplam maliyetini verir.



Şekil 4.1 Ağırlıklı graf

4.3. Yönsüz (undirected) Graf

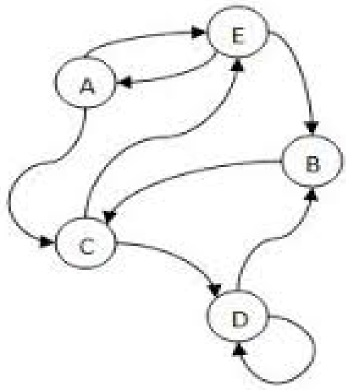
Yönlendirilmemiş ayrıtlardan oluşan graftır (Şekil 4.2). Sınırsız düğüm çiftleriyle temsil edilir. (v_1, v_2) ile (v_2, v_1) aynı şeyi ifade ederler. Uçuş rotası örnek olarak gösterilebilir.



Şekil 4.2 Yönsüz graf

4.4. Yönlü (directed) Graf

G, yönlü bir graf (directed) veya digraph her bir kenarı sıralı bir düğüm çifti ile ilişkilendirilmiştir ve her kenarı yönlüdür (Şekil 4.3). Yönlü kenar sıralı düğüm çiftleriyle ifade edilir. (v_1, v_2) ile (v_2, v_1) aynı değildir. İlk düğüm orijin, ikinci düğüm ise hedeftir. İki nokta arasındaki uçuş örnek olarak verilebilir. Yine yol ağını temsil eden bir grafta trafiğin tek ya da çift yönlü oluşu yönlü graflar için bir örnektir. Yönlü bir çizgenin derecesi giriş ve çıkış dereceleri toplamına eşittir. Bir köşeye yönlenmiş bağların sayısına giriş derecesi, bir köşeden yönlenmiş bağların sayısına da çıkış derecesi denir.



Şekil 4.3 Yönlü graf

5. SOSYAL AĞLARDA TOPLULUK TESPİTİ

Sosyal ağlardaki bireyler düğümler olarak nitelendirilirse, bu düğümlerin aralarındaki bağlantının yoğunlaştığı grupları topluluk (community) olarak tanımlanabilir. Diğer bir deyişle kendi içinde çok fakat dışarısıyla az sayıda bağlantı varsa bu kümeler topluluk adını alırlar. Topluluk (community) ve iletişim (communication) sözcükleri aynı köke sahiptir. Bir iletişim ağı kurulan her yerde bir toplulukta kurulur. Sosyal ağdaki topluluklardan bireylerin ortak yönleri, çalışma konuları, ortak ilgi, bilgi, beceri ve benzerlikleri hakkında veriler elde edilebilinir.

Gruplar ile ilgili yapılan çalışmalardan; düğümlerin grup içerisindeki yapılarına bakıp sınıflandırılması ve grup yapılarının ortaya çıkarılması en önemlileridir. Ağ yapılarında ki topluluklar göz önüne alındığında bir birey birden fazla topluluğa üye olabilir. Böyle yapılarda toplulukların belirlenmesi problemi, sosyal ağ analiz çalışmalarında incelenen ve çözümlenmesi amaçlanan problemlerdendir. Bu olaya "örtüşme (overlapping)" denir ve çoğu topluluk bulma algoritmaları tarafından hesaplama zorluğu gerekçesiyle ihmal edilir. Bu durumda önemli bilgiler ihtiva eden veriler göz ardı edilmiş olunur ki buda istenilmeyen bir durumdur.

Grup tespitinde en çok kullanılan tanımlamada grubun içindeki ayrıt sayısının dışarıya olan bağlantı sayısından fazla olması gerektiği varsayımı kullanılır. Bu tanımdan yola çıkılarak oluşturulan "cut-size" parametresi söz konusu topluluğu çizgenin geri kalanına bağlayan ayrıt sayısı olarak adlandırılmaktadır. İyi bir topluluğun cut-size değerinin düşük olması beklenir.

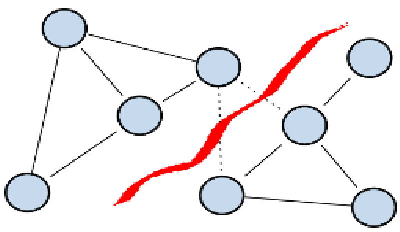
Diğer bir tanımlama olan "Düğüm benzerliği (vertex similarity)" ise düğümlerin bir uzay düzleme yerleştirildiklerinde aralarında kalan mesafenin bir benzerlik ölçütü olarak ele alınmasıdır. Klasik gruplama yöntemleri sıklıkla bu yaklaşımdan faydalanmaktadırlar. Düğümler bir uzay düzleme yerleştirilemiyor ise bu durumda komşuluk matrisi (adjencency matrix) kullanılabilir. Komşuları aynı ise kendileri komşu olmasalar bile benzerdirler denilebilir. Atıf analizi yaklaşımlarında da bu mantıktan faydalanılmaktadır. Bunun dışında iki düğüm arasındaki bağımsız yolların sayısının ölçülmesi, en kısa yolun mesafesi veya rastgele yürüyüş gibi yöntemlerle de düğümler arası benzerlikler saptanabilir [4,6].

Bir diğer yaygın yaklaşım ise çizge üzerinde yapılan işlemler (bölünme, birleştirme, ayrıt silme, ekleme vb.) sonucunda bir kalite fonksiyonunun iyileştirilmesine dayanmaktadır. En yaygın kullanılan kalite fonksiyonu modülaritedir. Toplulukların tespitinde halen belirsizlikler olsa da etkili ve verimli topluluk keşif yöntemleri de geliştirilmiştir. Bu yöntemler içerisinde en yaygın ve etkili olarak kullanılanlar aşağıda açıklanmıştır [4,6].

5.1. Geleneksel Yöntemler

5.1.1. Graf bölütleme (graph partitioning)

Graftaki düğümlerin, kaç gruba ayrılacağını ifade eden k önceden bellidir. Düğümlerin arası bağıntı minimum olacak şekilde bu k adet gruptan hangisinde olacağı belirlenir (Şekil 5.1). Başlıca algoritmaları Iterative Bisectioning [27], KernighanLin keşfi [28], Max-Flow Min-Cut Theoremi [29]'dir.



Şekil 5.1 Graf Bölütleme

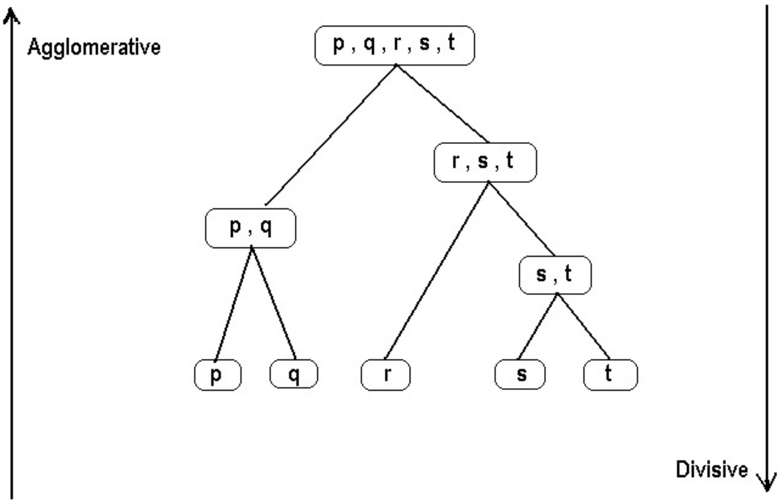
5.1.2. Hiyerarşik gruplama

Sosyal ağların geneli iç içe hiyerarşik yapılı gruplardan oluşur. Benzer olan düğümlerin birleştirilmesi ve grup yapılması ve düşük benzerlikli düğümlerin silinerek grupların bölünmesi fikrinin uygulandığı yöntemdir. Sonuçlar belirlenecek olan benzerlik ölçütüne bağlı olarak değişkenlik gösterecektir [6].

Ele alınan veri setinde kaç grup bulunduğu ilk başta bilinmediği durumlar için oldukça uygun bir yöntemdir. Yöntem veri setinde ilk bakışta keşfedilemeyen verileri irdeleme ve tespit etme imkanı verir. Genelde düğüm sayısı ikiyüz elli den az olan küçük grup yapılarının analizinde işlevseldir.

Veri matrisindeki değişkenlerin başlangıçta oluşturduğu grup sayısına ve grup elemanlarını belirlemede başlangıçta seçilen kriterlere göre iki ana grupta incelenir. Bunlar; Agglomerative (Bottom-up), Divisive (Top-down) dır (Şekil 5.2).

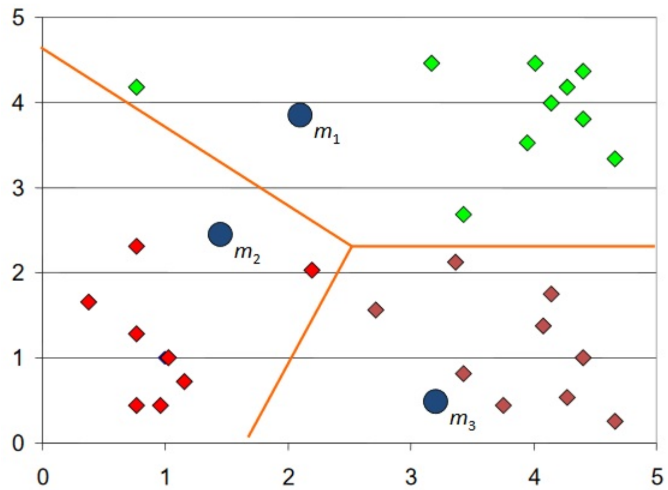
Grupların grafiksel olarak gösterimi işleminde hiyerarşik gruplamada ağaç diyagramlarından (dendrogram) yararlanılır. Grup sayısı bölütleme işlemi öncesi biliniyorsa veya araştırmacı bu graf şu kadar gruba ayrılacak şeklinde bir ön bilgi sunuyorsa böylesi bir durumda hiyerarşik gruplama teknikleri yerine hiyerarşik olmayan yöntemler (k-means vb.) kullanılır. Böylece işlem süresi kısaltılmış olunur.



Şekil 5.2 Hiyerarşik Bölütleme

5.1.3. Bölütlemeli kümeleme (partitional clustering)

Grup sayısı olarak nitelendirilen k önceden belirlenir ve her düğüm uzayda bir noktadır denir. Bir fonksiyona göre düğümlerin merkeze olan uzaklıkları hesaplanır ve bu uzaklıklara göre düğümler k gruba ayrılır. Bölütlemeli kümelemenin hedefi, kümeleme işlemi sonrasında bulunan grupların, grubun kendi içindeki benzerliklerinin maksimum ve gruplar arası benzerliklerinin minimum olmasını sağlamaktır. Sıklıkla kullanılan fonksiyonlar Minimum k-clustering, k-center, k-median, ve k-means'tir. Dezavantajı grup sayısı olan k nın önceden bilinemiyor oluşudur [6]. NP-hard bir problemdir ancak iteratif (tekrarlayıcı) yaklaşım sağlayan k-means algoritması sayesinde genelde iyi bir çözüme ulaşılır (Şekil 5.3).



Şekil 5.3. Bölütlemeli kümeleme (k-means kümeleme. Turuncu çizgi üç gruba ayırmaktadır. m1, m2, m3 k-means ile bulunmuş üç ayrı merkezi göstermektedir)

5.1.4. Spektral kümeleme (spectral clustering)

Bu yöntemde önce benzerlik matrisinin özvektörleri alınır ve daha sonra k-means gibi bir fonksiyon ile gruplara ayrılır. En sık kullanılan matris Laplace matrisidir. Bu yaklaşım sayesinde özvektörlerin bileşenlerinden grafda kaç adet grubun bulunduğu öngörülebilinir. Kullanılan Laplace matrisinin normalize edilip edilmeyişine göre iki farklı versiyonu mevcuttur [6].

5.2. Bölütlemeli Algoritmalar

Graftaki toplulukları, grupları bağlayan ayrıtları silerek ayrıştırma fikrine dayanan yöntemdir. Bu yöntemde asıl mesele grupları bağlayan ayrıtların nasıl tespit edileceğidir. En çok kullanılan algoritma Girvan-Newman algoritmasıdır. Bu algoritmada ayrıt merkeziyeti (edge centrality) ölçüt olarak baz alınır ve ayrıtlar seçilir. Tüm ayrıtlar için merkezilik değeri hesaplanır. En yüksek merkezilik değerine sahip ayrıt silinir. Tekrar birinci adım gerçekleştirilir ve en yüksek değere sahip ayrıt silinerek bu işlem devam eder. Ayrıt merkeziyeti ölçütü dışında ayrıt bitişikliği (edge betweenness), rastgele yürüyüş ayrıt bitişikliği (random walk edge betweenness) ve akım akışı bitişikliği (current flow betweenness) gibi ölçütler de kullanılmaktadır [6].

Bölütlemeli algoritmalar hiyerarşik bölütlemenin aksi usulde çalışır. Bu algoritmalarda merkezden uzak ve düşük ağırlıklı bağlantıları olan düğümlerden başlanır, merkeze doğru gidilir ve gidilirken her adımda kenardaki düğümler silinir. Belirttiğimiz gibi hiyerarşik bölütlemenin tam tersi mantığıyla çalışır. Hiyerarşik bölütlemede boş bir ağdan başlanıp ve her adımda ağa yeni bir birey eklenirken, bölütlemeli algoritmanın her adımında ağdaki bir birey çıkartılır.

5.3. Modülarite Easlı Yöntemler

Modülarite esaslı yöntemler graf analizinde en çok bilinen ve kullanılan kalite fonksiyonudur. Tam olarak ispatlanmamış olsa da yüksek modülarite değerinin iyi grupları işaret ettiği kabul edilmektedir [6]. Eğer bir graf aynı boyuttaki ve derecedeki bir rastgele grafa göre daha yüksek modülarite değerine sahipse o grafın grup yapısına sahip olduğu kabul edilir. Ancak bazı rassal graflarda grup yapısı olmamasına karşın yüksek modülarite değerleri ile karşılaşılabilmektedir. Bundan dolayı da yüksek modülarite grupları işaret eder tabiri her zaman söylenebilecek bir ifade değildir. Modülarite fonksiyonunun iyileştirilmesi NP-Complete bir problemdir ve doğrusal bir zamanda çözümü yoktur. Ancak çeşitli yakınsamalar ile başarılı sonuçlar elde eden algoritmalar geliştirilmiştir. Graf da bir birleştirme, ayrılma veya ayrıt silme gibi değişikliklerden kalite fonksiyonunda en iyi iyileştirmeyi yapan değişiklikler yapılır [4].

5.4. Dinamik Algoritmalar

Dinamik olarak gerçeklemeler yapan en yaygın algoritmalardan biri rastgele yürüyücü algoritmasıdır [30]. Bir gruptaki düğümlerin ve düğümler arası ilişkilerin daha yoğun olacağı fikriyle bir rastgele yürüyücünün belli bir grupta kat edeceği yol daha fazladır ve grup içinde daha fazla zaman geçirilir mantığıyla çalışır. Bu fikirden hareketle iki düğüm arasındaki mesafe hesaplanmaktadır [6].

5.5. Diğer Yöntemler

Şimdiye kadar bahsedilen ve sıklıkla kullanılan yöntemlerin dışında istatistiksel çıkarımı temel alan (Bayes vb.) yöntemler, düğümleri etiketleyen ve her bir iterasyonda komşuları tarafından en çok paylaşılan etiketi alan ve bu şekilde grupları ayıran yöntemler, klik filtreleme yöntemleri, örtüşmeyi önlemeyi amaçlayan yöntemler ve çok çözünürlüklü yöntemler şeklinde yöntemler de mevcuttur [4].

6. MATERYAL VE YÖNTEM

Graflar ile ifade edilebilen bilimsel hesaplamalarda ortaya çıkan kombinatorik yapıya sahip problemler vardır. Bu problemlerin çözümünün daha az zaman alması için problem alt kümelere parçalanmalıdır. Bunun için problemi temsil eden graf alt graflara bölünmelidir. Graf bölmeleme, VLSI devre tasarımında, yük dengelemede, seyrek matrislerin düzenlenerek sistemlerin direk olarak çözümlenmesinde kullanılır. Örneğin bazı bilimsel işlemlerde verilerin işlemcilere dengeli olarak dağıtılması işlemi bir graf bölütleme problemidir. Graftaki düğümler yapılacak işi ve düğümler arası ayrıtlar da bu işlerin atandığı işlemciler arası ilişkiyi temsil eder. En iyi bölütleme işlemci başına düşen işin mümkün olduğunca dengeli ve işlemciler arası iletişimin de minumum olduğu durumdur.

Problemi oluşturan veriseti işlem sırasında değişmiyorsa, bölütleme işlemine başlamadan evvel verinin graf ile temsili yapılır. Ve bu grafın bölütleme işlemi çeşitli algoritmalara göre bir kez yapılır. Eğer veriseti işlem sırasında farklılaşıyor ise bölmeleme işlemi ilk verilere göre yapılır. Ardından yeni veriler eklendikçe bölümlenen graflarda düzenlemeler yapılır. Bölütleme işlemi verilerin durumuna göre süreklilik arz eden bir oluştur [17].

Graf bölütleme işlemi NP-complete bir problemdir. Bundan ötürü bölütleme işini gerçekleyen algoritmanın matematiksel karmaşıklığı (complexity) bir polinom değil, eksponansiyel bir bağlantıdır. Graf bölütleme algoritmalarının hiçbiri verilen grafın optimum bölütleme işlemini yaptığını garanti edemiyor.

Verilen grafda düğümlerin kordinatları varsa, bölütleme işlemi kordinatları olmayan grafa göre biraz daha kolaydır [31]. Böyle graflarda düğümlerin eksen kordinatlarına göre öncelikle bir sıralaması yapılır ve daha sonra bölütleme işlemi yapılır (böl ve yönet).

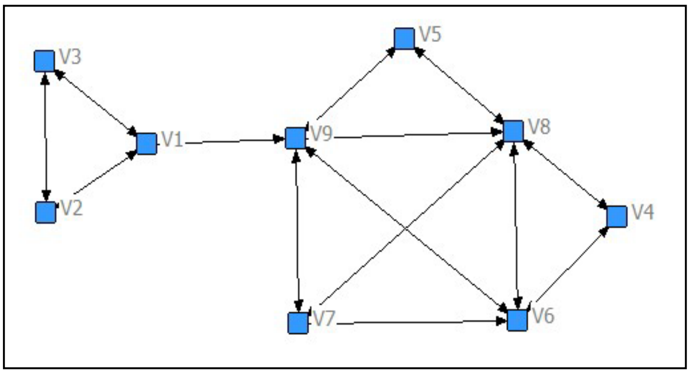
Eğer düğümlerinin kordinatları yoksa, graf bölütleme işlemi için genel olarak üç farklı yöntem kullanılır. Bunların ilki rassal olarak başlangıç bölütleme işleminin yapılması, ikincisi grafın öz değerleri ve bu öz değerlere denk düşen öz vektörleri bularak bölütlemenin yapılması, üçüncü olarak da eğer grafda ki düğüm sayısı fazla ise ilk olarak grafda rassal eşleştirmeler yapılır ve eşleşen düğümler bir düğüme indirgenir (coarsening). Böylece istenilen sayıda düğüm elde edilmiş olunur ve yukarda belirtilen yöntemlerden biri kullanılarak bölütleme işlemi gerçekleştirilir. Daha sonra graf geri açılarak (uncoarsening) orijinal grafın bölütlenmiş hali elde edilir. Başlangıç parçalar elde edildikten sonra bu parçalara iyileştirme metotları kullanılarak hata oranı en aza indirgenmeye çalışılır (Kernighan-Lin) [17].

Graf homojen bir yapıya sahipse (tüm düğümlerin ağırlıkları eşit ve bütün tüm ayrıt ağırlıkları eşit), bu durumda çoğu kez bölütleme işleminden sonra iyileştirme yapmaya gerek kalmaz, eğer graf homojen değilse, bu durumda da çoğu zaman iyileştirme işlemlerine ihtiyaç duyulur [17]. İyileştirme algoritmaları; iki optimal iyileştirme algoritması, Kernighan-Lin iyileştirme algoritması, ısı düşüşünün taklidi (simulated anneling - SA), enerji durum taklidi ve uyumlu ısı düşüşü taklidi (simulated tempering and adaptive simulated tempering - AST)'dir.

6.1.Tam Bağlı Alt Graflardan Hareketle Kesişen Topluluk Tespiti

Gerçek ağ yapılarında karşılaşılan en büyük sorun düğümlerin birden fazla topluluğa ait olabilme ihtimali olan "örtüşme (overlapping)" durumudur. Örtüşen topluluk keşfi için sosyal ağ bir graf olarak modellenmiştir ve bu graf üzerinde ağ graflarındaki toplulukların yapıtaşı olarak kabul edilen tam bağlı alt graflar bulunmuştur. Tam bağlı alt graf, bütün düğümleri arasında en az bir ayrıt olan düğüm kümelerine denir. Daha sonra grafın bitişiklik matrisi (adjancency matrix) bulunur. Bitişiklik matrisi graftaki (ağırlıksız) n tane düğüm için oluşturulan $n \times n$ boyutundaki matrisde birbiri ile komşu olan düğümlerin ağırlıklarının 1 olarak kabul edilip yazılması ve eğer komşu değilse 0 değerinin atanması sonucu oluşturulan matrisdir. Şekil 6.1'de graf olarak temsili yapılan basit bir sosyal ağ gösterilmiştir. Şekil 6.2'de bu basit grafın düğüm ve ayrıtlarından oluşan bitişiklik matrisi gösterilmiştir. Bundan sonraki adımda Bron-Kerbosh algoritması çalıştırılır

.



Şekil 6.1. Örnek bir graf

	V1	V2	V3	V4	V5	V6	V7	V8	V9
V1	0	1	1	0	0	0	0	0	1
V2	1	0	1	0	0	0	0	0	0
V3	1	1	0	0	0	0	0	0	0
V4	0	0	0	0	0	1	0	1	0
V5	0	0	0	0	0	0	0	1	1
V6	0	0	0	1	0	0	1	1	1
V7	0	0	0	0	0	1	0	1	1
V8	0	0	0	1	1	1	1	0	1
V9	1	0	0	0	1	1	1	1	0

Şekil 6.2. Örnek grafın bitişiklik matrisi

A) Bron-Kerbosh Algoritması

Bron-Kerbosch algoritması yönsüz bir graftaki tüm maksimal klikleri bulmak için kullanılır.

Algoritma ilk çağrıldığında R ve X boş küme olarak ayarlanır ve P graftaki tüm düğümleri içeir. R geçici sonuçtur, P olası adayların veri setidir ve X dışlanan veri setini temsil eder. N(v) düğümlerin komşularını gösterir (Şekil 6.3).

```
BronKerbosh (R,P,X)
Eğer P ve X boşsa
  R' yi max clique olarak kaydet
Her bir düğüm için
  BronKerbosch(R∪v, P∩N(v), X∩N(v))
  P ← P\v
  X ← X∪v
```

Şekil 6.3. Bron-Kerbosch Algoritması

P yi genişletmek için bir tepe düğümü (v) seçilir. v R'ye eklenir ve komşusu olmayan düğümler P den ve X den silinir. Ardından yeni P kümesinden başka bir tepe seçilir ve süreç P boşalıncaya kadar tekrar edilir. P boşaldığı zaman eğer X de boşsa R nin içeriği yeni maksimal-klik'dir denir. Eğer X boş değilse R zaten bulduğumuz bir klikin alt kümesidir. Daha sonra son alınan düğümden başlanarak aynı yoldan geri dönülür ve P,R ve X yeniden inşa edilir. Örnek bir graf üstünde uygulanmasını gösterilebilir [32, 33].

G[1] = { v_1, v_2, v_3 }

G[2] = { v_1, v_9 }

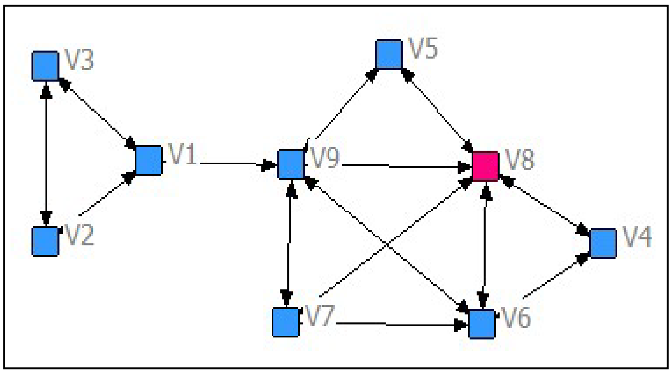
G[3] = { v_4, v_6, v_6 }

G[4] = { v_5, v_8, v_9 }

G[5] = { v_6, v_7, v_8, v_9 }

Şekil 6.4. Örnek grafın (Şekil 6.2'deki) bron kerbosh algoritmasından dönen değerleri

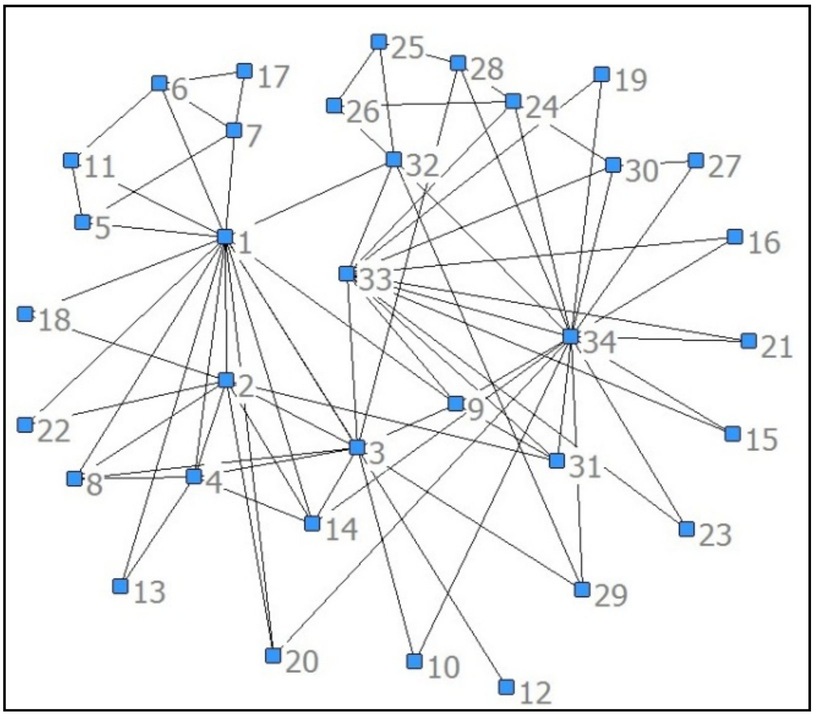
Buraya kadar sosyal ağ graf olarak modellendi, modelin bitişiklik matrisi çıkartıldı ve bu matrisde Bron-Kerbosh algoritması çalıştırılıp maksimal-klik'li bağıntılar tespit edildi. Daha sonra bulunan bu maksimal-klik'leri içeren matris de sütunları toplayıp iki düğüm içeren bağıntıların olduğu sütunlar sıfır yapıldı. Böylece minimum üç düğümü olan maksimal-klik'ler kaldı. Bundan sonra elde edilen matrisin satır toplam değerleri elde edilir. Bu değerlerden 1'den büyük olan satırlardaki düğümler örtüşen (en az iki alt-ağ tarafından içerilen) düğümlerdir. Şekil 6.2'deki örnek grafa yöntem uygulandığında sonuç v8 olarak tespit edilmiştir (Şekil 6.5).



Şekil 6.5. Yöntem uygulandıktan sonraki graf

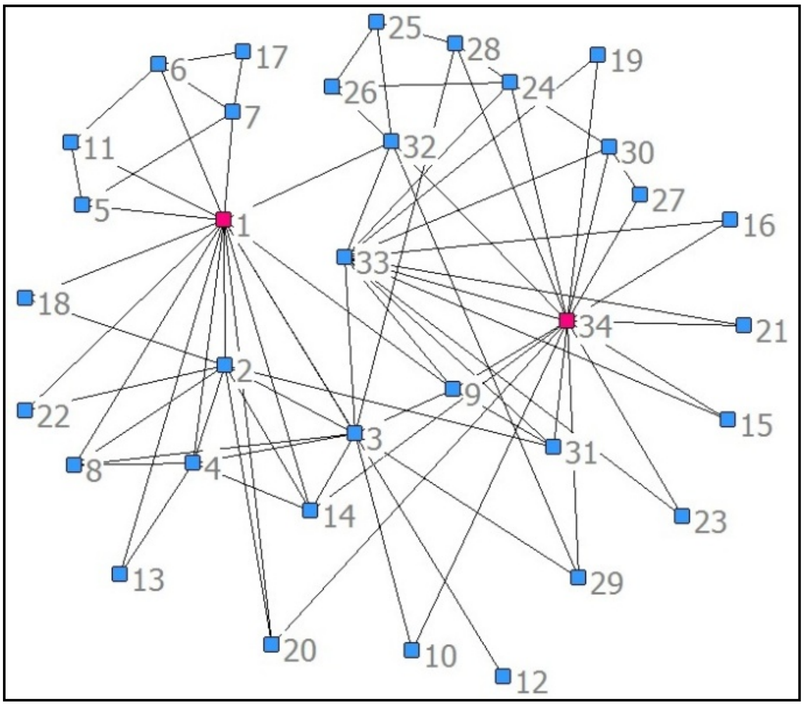
B) Çalışmanın Zachary'nin Karate Kulübü Sosyal Ağ Örneğine Uygulanması

1977 yılında W. W. Zachary tarafından yapılan bir çalışmada sunulan Zachary'nin Karate Kulübü sosyal ağ örneği sosyal ağ analizi konusunda en çok kullanılan örneklerden biridir. Karate kulübünde 34 adet öğrenci bulunmaktadır ve birbiri ile arkadaş olan öğrenciler için graf üzerindeki düğümler arasında bir ayrıt mevcuttur (Şekil 6.6).



Şekil 6.6 Zachary’nin Karate Kulübü sosyal ağ örneği

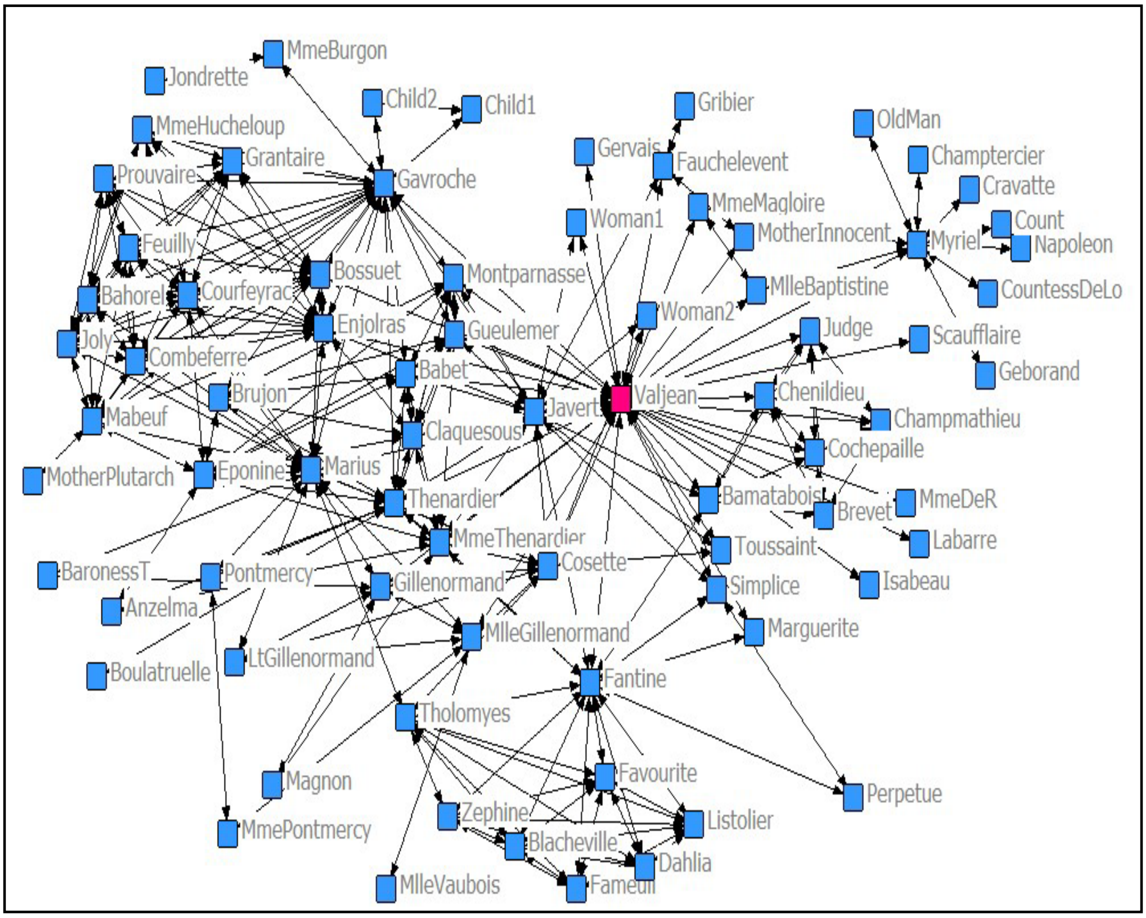
Bu sosyal ağa önerilen yöntem uygulanmış ve örtüşen elemanların {1, 34} olduğu görülmüştür (Şekil 6.7).



Şekil 6.7 ZKK'ne sosyal ağına yöntem uygulandığında elde edilen sonuç

C) Çalışmanın Lesmis Sosyal Ağ Örneğine Uygulanması

Lesmis, Victor Hugo'nun "Les Miserable" romanının karakterlerinin birbiriyle ilişkisini gösteren yönsüz bir sosyal ağdır. Her bir düğüm romandaki bir karakteri, iki düğüm arasındaki bağlantıda iki karakterin kitapta aynı bölümde görünüp görünmediğini ifade eder. Her bağlantının ağırlığı, böyle bir ortak görünüm oluşumunun sıklığını verir. Bu veri seti 77 karakter ve 254 bağıntıdan oluşmaktadır. Öngörülen yöntem bu veri setine uygulandığında "Valjean" isimli karakter (id:11) kesişen eleman olarak dönmüştür (Şekil 6.8).



Şekil 6.8. Lesmis sosyal ağına yöntem uygulandığında elde edilen sonuç

6.2. Yaklaşım Yöntemi ile Kesişen Topluluk Tespiti

Graf bölütleme problemi için günümüze kadar etkili bir algoritma geliştirilememiştir. Genellikle yaklaşımla çözüm yapılır ve bulunan çözüm bazen en iyi çözüm olabileceği gibi bazen de en iyi çözüme yaklaşık bir çözüm olabilir.

Optimum çözümü bulan yöntemler de vardır fakat bu yöntemlerle graf bölütleme işlemleri yaptığımızda zaman kaybı çok fazla olmaktadır. Bu yöntemlerle çözüm yapılması durumunda, belki de verilen problemin ardışıl çözümünde harcanacak zamandan çok daha fazla zaman harcanacaktır. Graf bölütleme işleminde amaçlanan problemi çözüme kavuşturmaktan ziyade problemin çözümü sırasında harcanacak zamanı minimize etmektir.

Bu çalışmada sosyal ağ bir graf olarak modellenmiş ve bu grafın Laplace matrisi bulunmuştur. Daha sonra Laplas matrisinin özdeğer ve özvektörleri bulunup en küçük ikinci özdeğere karşılık gelen özvektör (Fiedler Vectors) tespit edilmiştir. Bu vektör değerleri sınıflandırılmış ve spektral kümeleme tekniğiyle grafın iki gruba ayrıştırılması sağlanmıştır. Ardından elde edilen bu iki topluluk yapısındaki örtüşen bireyler Kernighan–Lin algoritması kullanılarak tespit edilmiştir.

A) Laplace Matrisi (Laplacian Matrix)

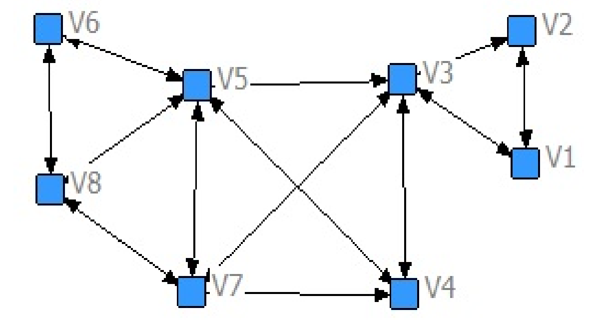
Laplace matrisi graf teorisinde sıkça kullanılan bir matrisdir. Laplace matrisine giriş matrisi (admittance matrix) veya Kirchhoff matrisi (Kirchhoff matrix) isimleri de verilmektedir.

Bir grafın bitişiklik matrisi (A=adjancency matrix) graftaki n tane düğümden $n \times n$ boyutunda ve komşu olan düğümlerin girdilerinin 1 olduğu komşu olmayan düğümlerin girdilerinin 0 olduğu bir matrisdir. Yani bitişiklik matrisinde her hangi bir satır ve sütünun kesiştiği yerde 1 varsa o satır ve sütunun temsil ettiği düğümler arası bir bağıntı vardır. Eğer graf ağırlıklı bir graf ise bitişiklik matrisinde, ağırlıklar 1 değerinin yerine yazılır. Simetrik bir matris olan bitişiklik matrisinde bir düğümün derecesi, ilgili satır veya sütundaki 1'ler toplanarak bulunabilir.

Bir grafın derece matrisi (D=degree matrix) diagonal bir matrisdir ve her bir diyagonal girdi yani dii, i. düğümün derecesini verir. Laplace matrisi;

$$L = D - A \tag{6.1}$$

olarak ifade edilir. Yukarıdaki kuraldan anlaşılacağı gibi matrisin köşegeninde (diagonal) düğüm dereceleri ve matrisin geri kalanında ise -1 ve 0’dan oluşan komşuluk matrisi bulunacaktır. Şekil 6.9'da örnek bir grafın laplas matrisi gösterilmiştir.

Graph (Graf)	Adjacency Matrix (Bitişiklik Matrisi)
	$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$
Degree Matrix (Derece Matrisi)	Laplacian Matrix (Laplas Matrisi)
$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 5 & -1 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & -1 & 5 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & 0 & -1 \\ 0 & 0 & -1 & -1 & -1 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & 3 \end{pmatrix}$

Şekil 6.9. Örnek bir graf ve bu grafın derece, bitişiklik ve laplas matrisi

Laplace matris öz değerlerin biri 0 ve bu öz değere karşılık gelen öz değer vektörü ise tüm elemanları 1 olan bir vektördür. Çünkü Laplace matrisi pozitif yarı tanımlı bir matrisdir ve bundan dolayı diğer bütün öz değerleri negatif olmayan değerlerdir. Bundan sonraki işlemlerde Laplace matrisinin ikinci en küçük öz değere denk düşen öz vektörü bulunur. Daha önce belirttiğim gibi bu öz değere Fiedler değeri ve bu öz değere karşılık bulunan öz değer vektörüne de Fiedler vektörü denir. Her hangi bir $\vec{x} \in R^n$ vektörü için;

$$\vec{x}^T L(G)\vec{x} = \sum_{(i,j)\in E}(x_i - x_j)^2 \tag{6.2}$$

ve bununla birlikte, G çizgesinin Fiedler değeri de;

$$\lambda_2 = \min_{x \perp (1,1,\dots,1)} \frac{(\vec{x}^T L(G)\vec{x})}{\vec{x}^T\vec{x}} \tag{6.3}$$

olur.

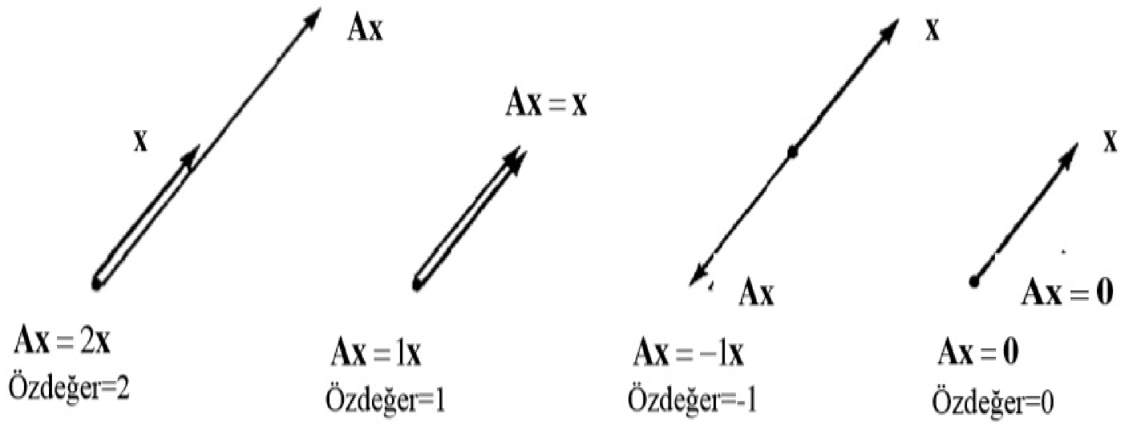
G grafının Laplace matrisinin Fiedler vektörü $\vec{u}(u_1, \dots, u_n)$ olsun. Spektral graf bölütlemenin ana fikri bir tane bölütleme değeri s bulunup, Fiedler vektörü iki parçaya bölünür. Bir parça için $u_i > s$ ve diğer parça için $s \leq u_i$ olur. Bu tip bir graf kesmesine Fiedler kesmesi denir. Genel olarak Fiedler kesmesini bulmak için, $(u_1, \dots, u_n)$'nin medyanı s olarak alınır ve bu şekilde seçilen s kesme ayıracına eş iki parça kesmesi denir. s'nin en iyi kesme oranını verdiği durumda ise, buna kesme oranı, s'nin değerinin 0 olması durumunda s'ye işaret kesmesi, s sıralanmış Fiedler vektörü elemanları içinde en uzun aralık içinde bir değere sahipse buna da aralık kesmesi denir.

B) Özdeğer ve Özvektörler

Bir matrisin ana yapısını görmek için özdeğerler kullanılabilir. Özdeğer kavramından önce özvektör kavramına değinelim. Bir A matrisi ile bir vektör çarpıldığında ilgili vektörün yönü bazı durumlarda değiştiği gibi bazı durumlarda da değişmeyecektir. İşte bu yön değiştirmeyen ve Ax vektörü ile aynı yönde kalan özel x vektörlerine özvektör denir. Ax vektörü bir özvektörün A matrisi ile çarpımı olup aslında x vektörünün λ katıdır.

Sonuçta ana denklem; Ax = λx şeklindedir. A matrisinin bir özdeğeri λ sabitidir. Bu skaler değer, özvektörün A matrisi ile çarpılınca oluşan yeni vektörün uzunluğundaki değişimi vermektedir. Yeni vektör orijinal vektöre göre büyüdü mü, küçüldü mü sabit mi kaldı vs. bilgiler bu skaler ile yorumlanabilir. Denklem Ax = 0x şeklinde olması halinde özdeğer 0 değerini almıştır ve özvektör x, sıfır uzayında tanımlıdır.

A'nın birim matris olması durumunda, $Ix = x$ olur. Bu durumda $n \times 1$ boyutlu tüm vektörler özvektördür ve A matrisinin tüm özdeğerleri $\lambda = 1$'dir. Bir doğrusal dönüşümün tanım matrisinin A matrisi olduğu varsayıldığında $Ax = \lambda x$ eşitliği sağlanıyorsa $T(x) = \lambda x$ olur. Bu denklemin anlamı, şayet x vektörü, A matrisinin özvektörü ise T dönüşümünün sonucunda x vektörünün görüntüsü bir skalerle çarpımı olan λx vektörüdür.



$Ax = \lambda x$ eşitliğini sağlayan λ, A matrisinin öz değeridir ve x'te λ'ya karşılık gelen öz vektördür. Bu eşitliği kullanılarak;

$$(A - \lambda I)X = 0 \tag{6.4}$$

eşitliği elde edilir. A matrisinin karakteristik polinomu (6.4)'deki eşitliğin determinantı alınınca elde edilir. Bu öz değerler kullanılarak öz değer vektörleri elde edilebilir. Bilgisayar ortamında bunun çözümü biraz daha farklı gerçekleşir. Bu işlem için geliştirilen yöntemler vardır. Bunlar; Kuvvet yöntemi, Jacobi yöntemi, Householder yöntemi gibi yöntemlerdir. Kuvvet yöntemi en küçük öz değere karşılık gelen öz değer vektörünü bulur. Jacobi yöntemi bütün öz değerleri ve bunlara karşılık gelen öz vektörleri bulur. Householder yöntemi ise reel simetrik olan bir matrisi üç-diyagonal matrisine dönüştürmektedir.

Jacobi yönteminde bütün öz değerlere ve öz vektörlere erişilebilindiğinden, bu çalışmada Jacobi yöntemi kullanılmıştır. Jacobi yöntemi simetrik matrislerin öz değerlerini ve öz vektörlerini bulmada kullanılır. Bir grafın Laplace matrisi de simetrik bir matris olduğundan bu yöntem kullanılabilir. Bu yöntem bütün simetrik reel matrislerin öz değer ve öz vektörler probleminin çözümünü garanti etmektedir [17].

B.1. Düzlem döndürme

Esas algoritmaya geçilmeden evvel geometrik bilgilerin verilmesi gerekir. X n-boyutlu bir vektör olsun ve $Y = RX$ lineer dönüşümü yapıldığını varsayalım.

$$R = \begin{array}{cc}
\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos\phi & \cdots & \sin\phi & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & -\sin\phi & \cdots & \cos\phi & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \leftarrow p \\ \\ \leftarrow q \\ \\ \\ \end{array} \\
\begin{array}{ccccccc} & & \uparrow & & \uparrow & & \\ & & p & & q & & \end{array} &
\end{array}$$

(6.5)

R, $n \times n$ boyutunda bir matris olup diyagonal elemanları bir, diğer elemanları sıfırdır. p ve q sütun ve satırlarına tekabül eden diyagonal elemanları $\cos\phi$, $r_{pq} = -sin\phi$ ve $r_{qp} = sin\phi$ bağıntısındaki lineer dönüşümü (6.6) bağıntı gibi olur.

$$y_j = x_j \ eğer \ j \neq p$$
$$y_p = x_p \cos\phi + x_q \sin\phi \qquad (6.6)$$
$$y_q = -x_p \sin\phi + x_q \cos\phi$$

Bu dönüşüm n boyutlu bir uzayın $x_p x_q$ düzleminde ϕ açısıyla döndürülmesidir. Aynı uzayın aynı düzlemde -ϕ açısı ile döndürülmesi ise bunun ters dönüşümüdür ve ters dönüşüm $X = R^{-1}Y$ ile ifade edilir. R matrisi ortogonal bir matris olduğundan

$$R^{-1} = R^T \text{ veya } R^T R = I \qquad (6.7)$$

olur.

B.2. Benzerlik ve ortogonal dönüşümler

$AX = \lambda X$ bu öz değer probleminin ele alalım. K matrisi, $B = K^{-1}AK$ bağıntısı ile tanımlanmış olsun. Bağıntının her iki tarafını $K^{-1}X$ ile çarparsak;

$$BK^{-1}X = K^{-1}AKK^{-1}X = K^{-1}AX = K^{-1}\lambda X = \lambda K^{-1}X \qquad (6.8)$$

elde edilir. Değişken değişimi $Y = K^{-1}X$ veya $X = KY$ şeklinde yapıldığında ve (6,8)'de yerine yazalım;

$$BY = \lambda Y \qquad (6.9)$$

şeklinde yeni bir öz değer problemi elde edilmiş olunur. $AX = \lambda X$ ile $BY = \lambda Y$ bağıntılarını karşılaştırırsak, öz değerleri birbirinin aynı ancak öz vektörleri aynı olmak zorunda olmayan A ve B şeklinde iki matris elde edildiği görülür. Böylece benzerlik dönüşümlerinin öz değere etki etmediği gösterilmiştir.

B.3. Dönüşümlerin jacobi serileri

Dönüşümlere, reel ve simetrik olan A matrisi ile başlanır ve ortogonal olan $R_1,\ R_2, \dots, R_n$ matrislerinin elde edilişi;

$$D_0 = A$$
$$D_j = R_j^T D_{j-1} R_j \quad j = 1, 2, \dots \qquad (6.10)$$

Uygulamada D matrisinin diyagonal üzerinde olmayan elemanları sıfıra yaklaşınca işlem durdurulur ve

$$\lim_{j\to\infty} D_j = D = diag\,(\lambda_1, \lambda_2, \dots, \lambda_n) \qquad (6.11)$$
$$D_n = D$$

olur. Buradan;

$$D_n = R_n^T R_{n-1}^T \dots R_1^T A R_1 R_2 \dots R_{n-1} R_n \tag{6.12}$$

olur. Eğer R matrisi için $R = R_1, R_2, \dots, R_{n-1} R_n$ tanımlaması yapılırsa, bu tanımlamadan yola çıkarak, $R^{-1}AR = D$ bağıntısı (6.13)'deki sonucu verir.

$$AR = RD = R\, diag(\lambda_1, \lambda_2, \dots, \lambda_n) \tag{6.13}$$

R matrisi sütunları $X_1, X_2, \dots, X_n$ vektörleri ile gösterilir ise,

$$R = [X_1, X_2, \dots, X_n]$$

olur.

Jacobi yöntemi her işlem adımında A matrisinin a_{pq} ve a_{qp} elemanları sıfır yapılır. R_1 ilk ortogonal matris ve $D_1 = R_1^T A R_1$ olsun. R_1 matrisi (6.8)'de görülen matrisin aynıdır. Yalnızca bir fark vardır. $c = cos\phi$ ve $s = sin\phi$ olarak alınacaktır. Her adımda değişiklik p ve q satırları ile p ve q sütunlarında meydana gelir. Bunun ispatını yapalım. $B = AR_1$ olsun.

$$B = \begin{bmatrix} a_{11} & \cdots & a_{1p} & \cdots & a_{1q} & \cdots & a_{1n} \\ a_{p1} & \cdots & a_{pp} & \cdots & a_{pq} & \cdots & a_{pn} \\ a_{q1} & \cdots & a_{qp} & \cdots & a_{qq} & \cdots & a_{qn} \\ a_{n1} & \cdots & a_{np} & \cdots & a_{nq} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & c & \dots & s & \dots & 0 \\ 0 & \dots & -s & \dots & c & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \tag{6.14}$$

(6.14)'deki çarpım işlemi gerçekleştirilirse;

$$\begin{aligned} b_{jk} &= a_{jk} && \text{eğer } k \neq p \text{ ve } k \neq q \\ b_{jp} &= ca_{jp} - sa_{jp} && j = 1, 2, \dots, n \\ b_{jp} &= sa_{jp} + ca_{jp} && j = 1, 2, \dots, n \end{aligned} \tag{6.15}$$

olur. $D_1 = R_1^T A R_1$ dönüşümü A matrisinin sadece p ve q sütunları ile p ve q satırlarının değişimine neden olacaktır.

$$b_{jp} = ca_{jp} - sa_{jp} \quad \text{eğer } j \neq p \text{ ve } j \neq p$$
$$b_{jp} = sa_{jp} + ca_{jp} \quad \text{eğer } j \neq p \text{ ve } j \neq p$$
$$d_{pp} = c^2 a_{pp} + s^2 a_{pp} - 2csa_{pq} \tag{6.16}$$
$$d_{qq} = s^2 a_{pp} + c^2 a_{pp} + 2csa_{pq}$$
$$d_{pq} = (c^2 - s^2)\, a_{pp} + cs(a_{pp} - a_{qq})$$

D_1 matrisinin simetri ile diğer elemanları bulunur.

Fiedler, özvektör ile bir grafın nasıl bölünebileceğini ilk kez gösteren kişidir. Bu işlemi graflar arası ayrıtların ağırlıkları toplamının minimum olması gerekliliği fikrine dayanarak gerçeklemiştir. Grafın iki gruba bölündüğü varsayımıyla yola çıkan Fiedler, bir gruptaki düğümlere +1 değeri atamış diğer gruptaki düğümlere ise -1 değerini atamıştır.

C) Spektral bölütleme

Spektral graf bölütleme yönteminde, nümerik metotlar kullanılmaktadır. Bunlardan biri de reel-simetrik matrislerin öz değerleri ve öz vektörlerinin bulunmasıdır. Bir garafın Laplace matrisi reel ve simetrik olan bir matristir. Bu matrisin öz değer ve öz vektörleri bulunur. Laplace matrisinin ikinci en küçük öz değere tekabül eden öz değer vektörü grafın ikiye ayrılmasında kullanılır. Bu yöntemi ilk kez Fiedler kullanmıştır [35]. Bu ikinci en küçük değerin öz değer vektörü grafın cebirsel bağlılığı olarak nitelendirilir. Fiedlerin bu çalışmasından ötürü, bu öz değere matrisin Fiedler değeri ve öz değer vektörüne de Fiedler vektörü denilmiştir. Bu vektör grafın en küçük ayıracını verir [17].

Bu yöntem nümerik algoritmaların çoğunun algoritma yapısında bulunur. Spektral bölütleme yöntemi, graf ve matris bölmelemede sıkça kullanılmanın yanı sıra devre dizaynı ve simülayonlarında da kullanılmaktadır.

G=(V,E) bağlı ve yönsüz bir graf olsun. Grafın bölütlenmesi işlemi grafın düğümlerinin ayrık birden fazla alt kümeye bölünmesidir. Bu alt kümeler A ve $\overline{\mathrm{A}}$

olsun $(A \cap \bar{A} = \emptyset$ ve $A, \bar{A}\ V)$. Başlangıçta $|A| \leq |\bar{A}|$ kabul edilmesiyöntemin genelliğini bozmaz. $E(A, \bar{A})$ grafın ayıracı olsun. $(A, \bar{A})$ bölütlemenin kesme boyutu $|E(A, \bar{A})|$ olur ve buna grafın ayıracının boyutu denir ve genelde bölütleme işlemlerinde minimum olması hedeflenir. Kesme oranı $\phi(A, \bar{A})$ şu formül ile hesaplanır;

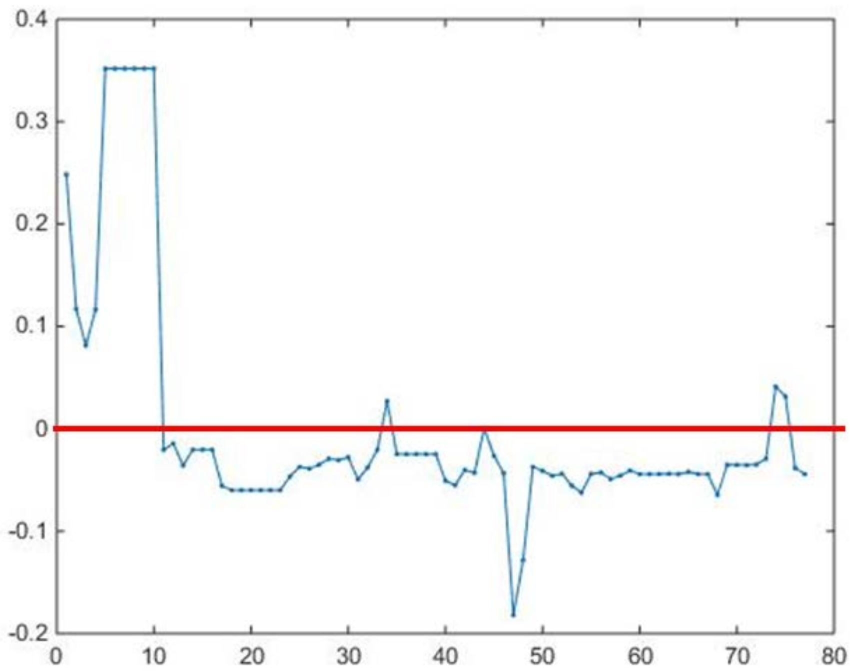
$$\phi(A, \bar{A}) = \frac{|E(A,\bar{A})|}{\min(|A|,|\bar{A}|)} \tag{6.17}$$

Bir grafın beklenen en iyi kesme oranına o grafın izoperimetrik sayısı denir ve şu eşitlik ile hesaplanır;

$$\phi(G) = \frac{\min |E(A,\bar{A})|}{|A| \leq \frac{n}{2} |A|} \tag{6.18}$$

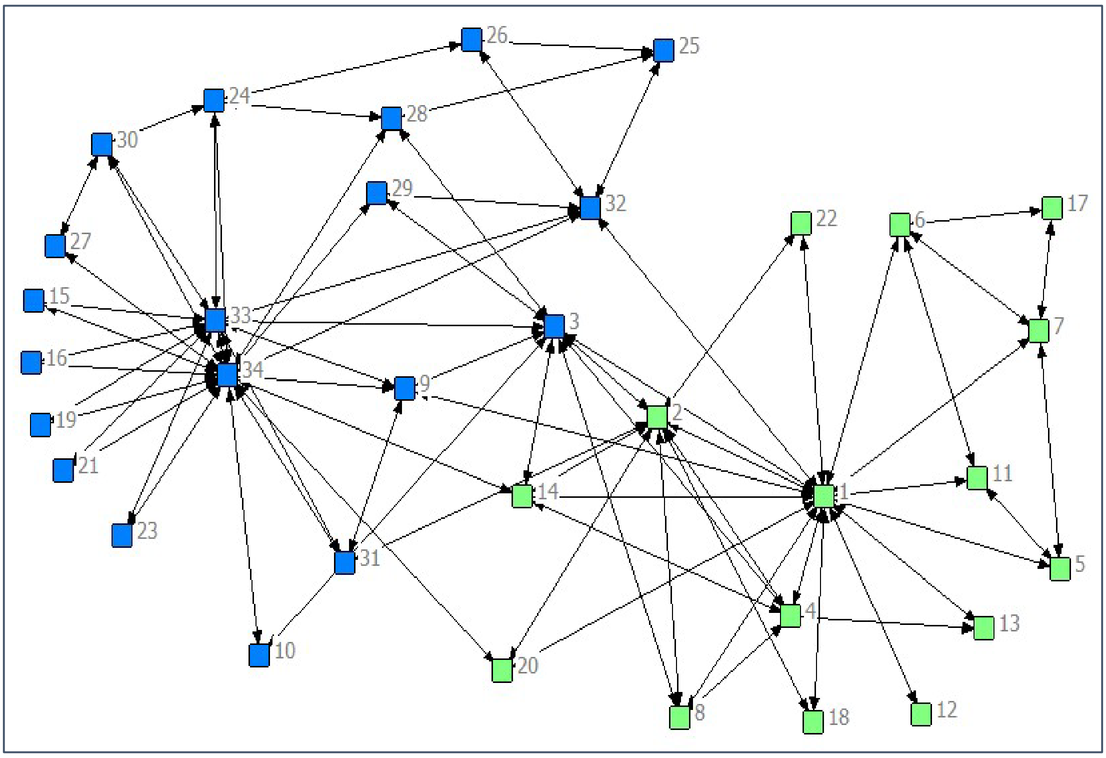
Bir grafın alt grafları arasında en fazla bir düğüm fark varsa bu alt graflara, grafın eş iki parçaları denir. Bir grafın alt grafları ve izoperimetrik değeri bulunduktan sonra o grafın eş iki parça boyutunun $O(\phi n)$ olduğu görülür.

Simon, Pothen ve Liu'nun [20], geliştirdikleri metot olan "Recursive Spectral Bisection" (RSB) da grafın Laplace matrisi ve bu Laplace matrisinin özdeğer ve özvektörleri bulunur. Laplace matrisinin özvektörlerinden en küçük ikinci özdeğere sahip özvektör alınır. "Fiedler vektörü" olarak adlandırılan bu özvektör graf hakkında yararlı veriler içerir. Düğümler arası mesafe hakkında bilgi, Fiedler vektörünün koordinatları arasındaki fark ile bulunur. Recursive Spectral Bisection metodu düğümlerin Fiedler vektörü koordinatlarının sınırları açısından grafı iki topluluğa ayırır. RSB'yi kullanarak grafı istenilen sayıda topluluğa ayırmak mümkündür [19]. Şekil 6.10'da Lesmis verisetinin Fiedler vektörü gösterilmiştir. 77 düğümden oluşan veriseti sıfırdan küçük değerler bir grup ve sıfırdan büyük değerler bir gruptur mantığıyla iki gruba ayrılmıştır.



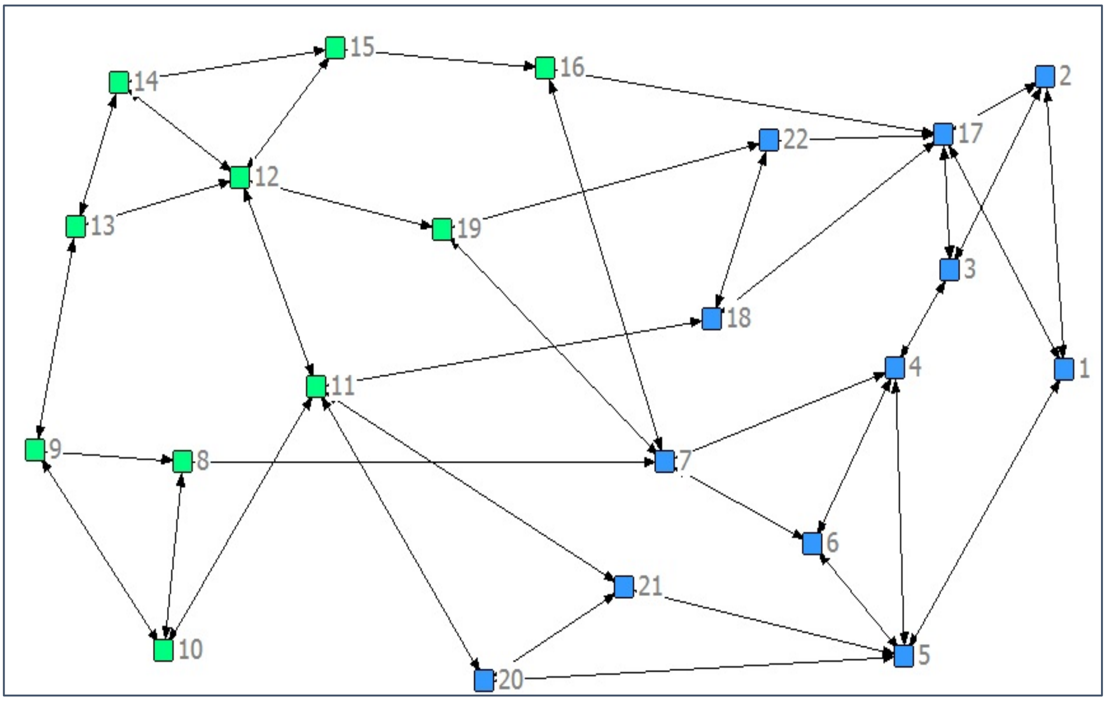
Şekil 6.10. Lesmis veriseti fiedler vektörünün gösterimi ve sosyal ağın iki gruba ayrılması

Zachary’nin karate kulübü sosyal ağ örneğine spektral bölütleme işlemleri uygulanmış Şekil 6.11'deki gibi sosyal ağ iki gruba ayrılmıştır.



Şekil 6.11 Spektral bölütleme uygulanmış Zachary Karate Klübü veriseti

Ucinet programının sunduğu datasetlerden biri olan "Taro" dataseti üstünde bahsedilen spektral bölme işlemi uygulanmıştır. Graf iki topluluğa ayrılmıştır (Şekil 6.12).



Şekil 6.12. Spektral bölütleme uygulanmış Taro veriseti

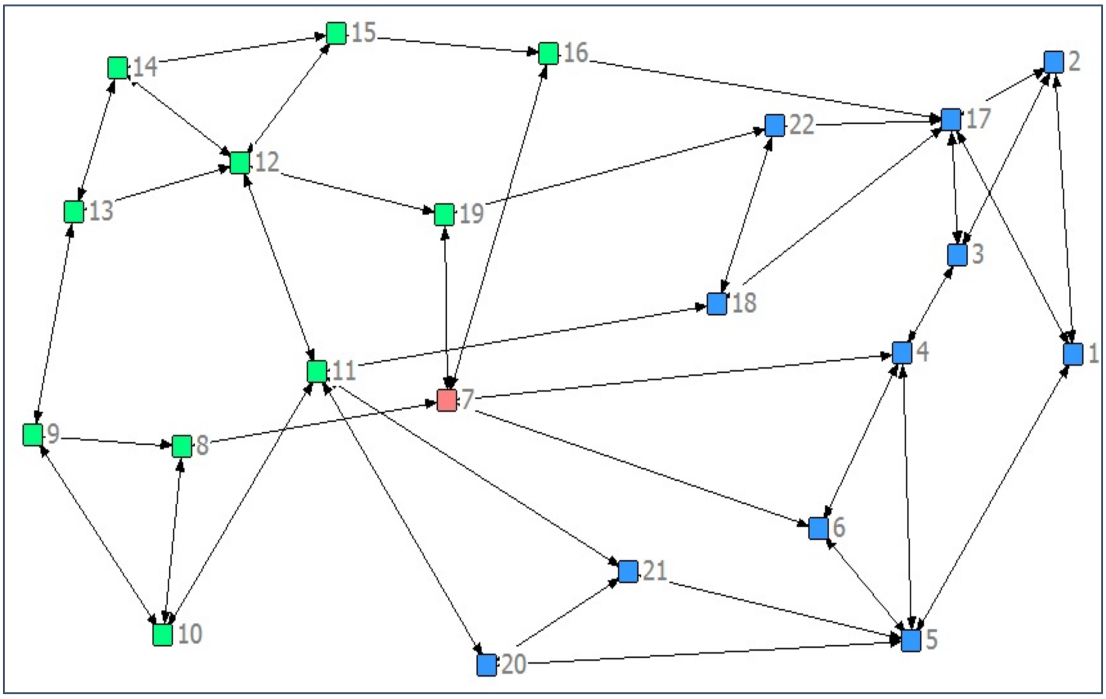
D) Kernighan–Lin

Brian Wilson Kernighan ve Shen Lin ilk graf bölütleme metotlarından biri olan Kernighan-Lin algoritmasını tasarlamışlardır. Geliştirilen graf bölütleme metotlarından birçoğu yerel bölütleme metodu Kerninghan-Lin metodundan esinlenmiştir ve bu metodun farklı versiyonları şeklindedir. Bu algoritmada graf rastgele iki gruba bölünür. Bu iki grup algoritmamızın girdilerini oluşturur. Kernighan-Lin grupların ayrıştırılması esnasında kesilen ayrıtları hesaplar ve kesilen ayrıt sayısını minimize etmeyi amaçlar [19]. Farklı gruplardaki düğümler ardışıl bir şekilde bir diğer gruba alınır ve kesilen ayrıt sayısı yeniden hesaplanarak en iyi çözüme yani minimum kesen ayrıt sayısına ulaşmak hedeflenir.

Bir graf rassal bölütleme algoritmaları ile ikiye bölünür ve bu parçalara iyileştirme algoritmaları uygulanır. İyileştirme adımı oldukça yavaştır. Çünkü graf gelişi güzel iki parçaya bölünmektedir ve bu parçalar birbiriyle tamamen alakasız parçalar da olabilir.

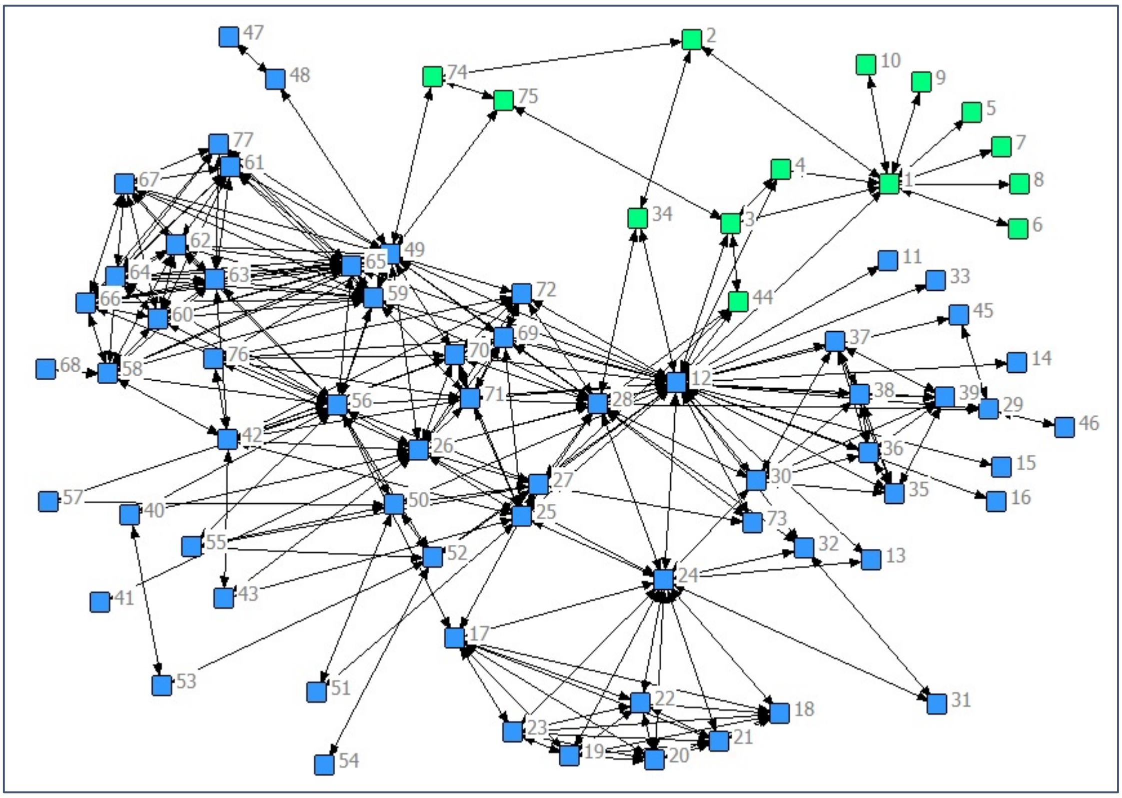
Bu çalışmada uygulanan metotta Kernighan-Lin algoritmasında rassal olarak belirlenen iki grup değilde spektral bölütlemeyle ayrıştırılan gruplar koşulmuştur. Bu sayede graftaki tüm ayrıtlar incelenmemiş ve işlem maliyeti azaltılmıştır. Grafta bir elemanın grubu değiştirildiğinde grupları ayırırken kesilen ayrıt sayısında değişme olmaması veya daha az ayrıtın kesilmesi durumlarında bu eleman için kesişen elemandır yorumu yapılmıştır.

Spektral bölütlemeyle iki gruba ayrılan Taro datasetinde Kerninghan-Lin algoritmasını çalıştırdı. 7 nolu düğüm diğer gruba alındığında kesen ayrıt sayısında azalma görülmüştür. Böylece bu eleman kesişen elemandır yorumu yapılmıştır (şekil 6.13).

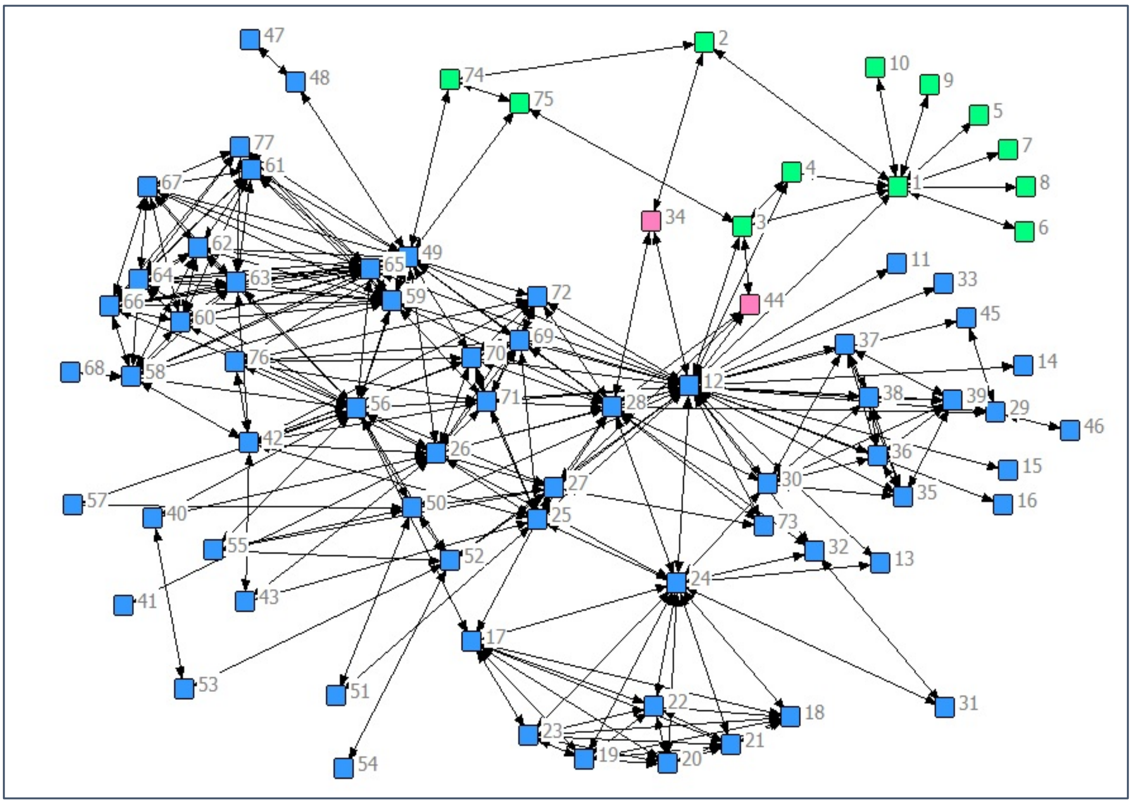


Şekil 6.13. Spektral bölütleme uygulanmış Taro verisetine Kernighan-Lin algoritması uygulandıktan sonraki durumu

Fransız yazar Victor Hugo'nun Les Miserables adlı roman karakterlerinin düğümlerini oluşturduğu ve bu karaterlerden aynı chapter da yer alma ilişkilerinin ayrıtlarla ifade edildiği veri setinin Lesmis veri seti olduğundan daha önce bahsetmiştik. Yapılan çalışma bu veri setine uygulandığında " Woman1" adlı karakter (id:34) ve "Woman2" adlı karakterler kesişen eleman olarak tespit edilmiştir. Şekil 6.14'de spektral bölütlemeyle iki gruba ayrılan lesmis datasetini Şekil 6.15'de de özelleştirdiğimiz Kernighan-Lin algoritmasının ilgili veri setine uygulandıktan sonraki son şekli gözlemliyoruz.



Şekil 6.14. Spektral bölütleme uygulanmış Lesmis veriseti



Şekil 6.14. Kesişen topluluk tespiti yapılmış lesmis veriseti (Pembe renkli elemanlar kesişenleri temsil eder.)

7. SONUÇ

Sosyal ağlarda topluluk keşfi ve örtüşen topluluk tespiti konusunda literatür taraması yapılmış ve ilgili problem için iki farklı çözüm üzerinde çalışılmıştır. Graf bölütleme işleminde problemin çözümünden ziyade problemin çözümü için harcanacak zamanın mümkün olduğunca minimum olması üzerine odaklanılmıştır.

İlk yöntemde tam bağlı graflarda kesişen düğümlerin tespiti etkili bir şekilde gerçekleşmiştir. İkinci yöntemde ise yaklaşım ile kesişen toplulukların tespiti yapılmıştır.

Genelde Kernighan-Lin algoritmasında rassal olarak belirlenen iki grup üzerinde kesen minimum ayrıt prensibiyle iyileştirme yapılarak grupların ayrıştırılması hedeflenir. Bu çalışmada ise grafın Laplace matrisi bulundu. Bu matrisin özdeğer ve özvektörlerine göre gruplar ayrıştırılıp ayrıştırılan bu gruplar Kerrnighan-Lin algoritmasında çalıştırılmıştır. Böylece graftaki tüm ayrıtlar birer birer incelenmemiş yalnızca grubupların ayrıştığı noktada olan düğümler işleme katılmıştır. Buda işlem maliyetini azaltılmıştır. Bir elemanın grubunu değiştiğimizde grupları ayırırken kesilen ayrıt sayısında değişme olmaması veya daha az ayrıtın kesilmesi durumlarında bu eleman için overlapping durumundadır yani kesişen elemandır tespiti yapılmıştır.

Uygulanan metot Matlab ortamında kodlanmış ve örnek veri setleri üzerinde (Zachary's Karate Club, American Collage Footbool, Lesmis etc.) çalıştırılmıştır. Uygulama sonunda graflar da kesişen düğümlerin tespiti etkili bir şekilde gerçekleşmiştir.

8. KAYNAKLAR

[1] A. Köksal, " *Bilişim Terimleri Sözlüğü*", Türk Dil Kurumu Yayınları, 1981, p. 126.

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

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

[4] Onur Boy. *"Sosyal Ağlarda Topluluk Yapılarının Analizi"*, İnönü Üniversitesi Bilgisayar Mühendisliği Yüksek Lisans Tezi, İnönü Üniversitesi, Malatya, 2012

[5] Burak Işıklı. *Sosyal Ağlarda Etkileşim Örüntüleri Kullanılarak Bağlantı Tahmini*, Gebze Yüksek Teknoloji Enstitüsü Yüksek Lisans Tezi, Kocaeli, 2014.

[6] S. Fortunato, Community detection in graphs. arXiv:0906.0612, 2009.

[7] L Euler, "Solutio problematis ad geometriam situs pertinentis." Comment. Acad. Sci. U. Petrop. 8, 128-140, 1736. Reprinted in Opera Omnia Series Prima, Vol. 7. pp. 1-10, 1766.

[8] N.L.G Biggs, E.K. Lloyd, R.J.Wilson, Graph Theory, Oxford, 1986

[9] Fikriye Ersoy. *"Grafların Komşuluk matrisleri",* Uludağ Üniversitesi Matematik A.B.D. Yüksek Lisans Tezi, Bursa, Eylül 2013

[10] B. Diri, YTÜ Yazılım Anabilim Dalı Web Sayfası, https://www.ce.yildiz.edu.tr/personal/banud/file/1320/Graflar.pdf (Erişim tarihi 10 Ekim, 2016).

[11] K.Thulasiraman, M.N.S. Swamy, "Graphs: Theory and Algorithms", Concordia University, Montreal, Canada, 1992.

[12] R. Sedgewick, Graph algorithms, Algorithms. AddisonWesley, 1983.

[13] T.H. Cormen, C.E. Leiserson, R.L. Rivest & C. Stei, Introduction to Algorithms (Vol. Second Edition). MIT Press and McGraw-Hill, 2001.

[14] S.E. SEKER, *Çizge Teorisi (Graph Theory)*, YBS Ansiklopedi, v.2, is.2, pp. 17-29, 2015

[15] J. Gross, J. Yellen, Graph Theory and Its Applications, CRS Press, 1999

[16] B.W. Douglas, Introduction to Graph Theory, second ed., Prentice Hall, 2001

[17] Ali Karcı. *"Çizge Algoritmaları ve Çizge Bölmeleme",* Fırat Üniversitesi Bilgisayar Mühendisliği Yüksek Lisans Tezi, Elazığ, 1998.

[18] A. Karcı, O. Boy, "Sosyal Ağların Web Madenciliği Teknikleri ile Analizi ve Ortak Atıf Analizi ile Benzerlik Tahmini", Elektrik-Elektronik ve Bilgisayar Sempozyumu, pp:154-161, 2011, Elazığ, Türkiye.

[19] M. K. Beşer, " En Kısa Yol Probleminde Çizge Parçalama Yöntemi Kullanılarak Yeni Bir Yaklaşım ", Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi9(1), 06,2008

[20] A. Pothen, H.D. Simon, K.P.Liu, K.P. (1990), Partitioning sparse matrices with eigenvectors of graphs, SIAM J. on Matrix Analysis and Applications 11(3), 430-452.

[21] S. Tüzüntürk, *Ağ Bilimi,* Dora Yayıncılık, Bursa, 2012

[22] N. Gürsakal, S. Tüzüntürk, F. Sert, "Sosyal Ağ Verilerinin Kuvvet Yasası Olasılık Dağılımına Uygunluk Analizi: Twitter Örneği", 15. Uluslararası Ekonometri, Yöneylem Araştırması ve İstatistik Sempozyumu, p. 501-523, 2014

[23] Wikipedia: Karmaşık ağ. FL: Wikipedia Foundation, Inc. Retrieved October, 2016, from https://tr.wikipedia.org/wiki/Karmaşık_ağ

[24] J.D. Watts, S.H. Strogatz, "Collective Dynamics of Small-World Networks", Nature, Vol.393, pp. 440-442, 1998.

[25] D. Liben-Nowell. An Algorithmic Approach to Social Networks, Massachusetts Institute of Technology, 2005.

[26] Barry Wellman, "Structural Analysis: From Method and Metaphor to Theory and Substance." Pp. 19-61 in Social Structures: A Network Approach, edited by Barry Wellman and S.D. Berkowitz. Cambridge: Cambridge University Press, 1988

[27] A. Pothen, Graph partitioning algorithms with applications to scientific computing, Tech. rep., Norfolk, VA, USA, 1997.

[28] B.W. Kernighan, S. Lin, An efficient heuristic procedure for partitioning graphs, Bell Syst. Tech. J. 49 (1970) 291307.

[29] L.R. Ford, D.R. Fulkerson, Maximal flow through a network, Canad. J. Math. 8 (1956) 399404.

[30] B.D. Hughes, Random Walks and Random Environments: Random Walks, vol. 1, Clarendon Press, Oxford, UK, 1995.

[31] G.V. LAszewski, A Collection of Graph Partitioning Algorithms, Tecnical Report, Northeast Parallel Architectures CEnter at Syracuse University

[32] C.Bron, J.Kerbosch. Algorithm457: finding all cliques of an undirected graph. Comm.ACM,16(9):575–577, 1973

[33] A. Conte, Review of the Bron-Kerbosch algorithm and variations, School of Computing Science Sir Alwyn Williams Building University of Glasgow,05.2013

[34] D.A. Spielman, S.H. Teng, Spectral Partitioning Work: Planar graphs and finite element meshes, Technical Report, University of California, Berkeley, 1996

[35] B. Hendrickson, R. Leland, A Multilevel Algorithm for Partitioning Graph, Proc. Supercomputing-95, ACM, 1995

ÖZGEÇMİŞ

Ad Soyad : Esra Karadeniz
Doğum Yeri ve Tarihi : Malatya / 23.05.1990
E-Posta : esrakaradeniz@yandex.com
Lisans : İnönü Üniversitesi Bilgisayar Mühendisliği Bölümü (2013)

TEZDEN TÜRETİLEN YAYINLAR/SUNUMLAR

- Esra Karadeniz, Ali Karcı, “Overlapping Community Detection In Social Networks “ ICNASE-2016: International Conference On Natural Science And Engineering, pp:3137-3144, 2016, Kilis, Türkiye.

- Esra Karadeniz, Ali Karcı, “Definition of Subcommunity In Social Networks and Identify Common Individual“ IDAP-2016: International Artificial Intelligence and Data Processing Symposium, pp:556-560, 2016, Malatya, Türkiye.