ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

ON BALANCING SOCIAL NETWORKS

Ph.D. THESIS

Araniyos Terefe WELDEGEBRIEL

Department of Mathematical Engineering

Mathematical Engineering Programme

JANUARY 2019

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

ON BALANCING SOCIAL NETWORKS

Ph.D. THESIS

Araniyos Terefe WELDEGEBRIEL
(509122053)

Department of Mathematical Engineering

Mathematical Engineering Programme

Thesis Advisor: Asst. Prof. Dr. Burak Yıldıran STODOLSKY

JANUARY 2019

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

SOSYAL AĞLARIN DENGELENMESİ

DOKTORA TEZİ

Araniyos Terefe WELDEGEBRIEL
(509122053)

İMatematik Mühendisliği Anabilim Dalı

Matematik Mühendisliği Programı

Tez Danışmanı: Asst. Prof. Dr. Burak Yıldıran STODOLSKY

OCAK 2019

Araniyos Terefe WELDEGEBRIEL, a Ph.D. student of ITU Graduate School of Science Engineering and Technology 509122053 successfully defended the thesis entitled "ON BALANCING SOCIAL NETWORKS", which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Asst. Prof. Dr. Burak Yıldıran STODOLSKY** Istanbul Technical University	....
Jury Members :	**Assoc. Prof. Dr. Ali ERCENGİZ** Istanbul Technical University	....
	Assoc. Prof. Dr. Ergün YARANERİ Istanbul Technical University	....
	Assoc. Prof. Dr. Ayhan GÜNAYDIN Boğaziçi University	....
	Asst. Prof. Dr. Ahmet BAKKALOĞLU Mimar Sinan Güzel Sanatlar Üniversitesi	....

Date of Submission : **05 December 2018**
Date of Defense : **03 January 2019**

To my dearest family

FOREWORD

I would first of all thank The Almighty God for His infinite blessings. Without His guidance and Will nothing would be possible.

This dissertation is dedicated with greatest respects and gratitudes to my father Terefe Weldegebriel, to my mother Senait Yibzu, to my wife Menbere Woldetsadik, to my sweety daughter Afomiya Araniyos, to my sweety son Aron Araniyos and all my sisters and brothers. Many special thanks go out to them for their understanding, support, warmth and unconditional belief in me throughout of my life.

I would like to thank to my esteemed advisor Asst. Prof. Dr. Burak Yıldıran Stodolsky for his assistance, encouragement, patience and guidance during my PhD studies. I would like to present my gratitudes to him for his confidence in me.

I would also like to thank to the Doctoral committee members Assoc. Prof. Dr. Ali Ercengiz from Istanbul Technical University and Assoc. Prof. Dr. Ayhan Günaydın from Boğaziçi University for their supportive and valuable comments and suggestions.

I would also like express my sincere gratitude to Mr. Mustafa Soysal, who programmed our algorithm using Java. His programs helped simplify our calculations immensely.

Finally, I would also like to thank the professors of the Department of Mathematical Engineering at Istanbul Technical University who taught me the courses required for the study, for their help and sincerity.

January 2019 Araniyos Terefe WELDEGEBRIEL

TABLE OF CONTENTS

Page

ABBREVIATIONS

NP	: Non - deterministic Polynomial time
Max-Cut	: Maximum Cut
$V(G)$	: The set of vertices in graph G.
$E(G)$	: The set of edges in graph G.
$\|E(G)\|$	: The number of edges in the edge set $E(G)$.
$H = (V, E, \sigma)$	: Signed graph H with vertex set V, edge set E and a sign function $\sigma : E \to \{\pm 1\}$.
$\lceil x \rceil$	: The least integer that is greater than or equal to any given number x.
$\lfloor x \rfloor$	: The greatest integer that is less than or equal to any given number x.
$l(G)$	: The minimum number of edges of G whose signs have to be negated to balance a signed graph G.
$\partial_S(v)$	: The stability degree of vertex v in the partition $(S, \overline{S})$.
$\partial(v)$	: The stability degree of vertex v when $S = V(G)$.
$\mathscr{D}(S)$	: The stability degree of a partition $(S, \overline{S})$ of signed graph G.
$\mathscr{D}(G)$	: The stability degree of the signed graph G when $S = V(G)$.
$d(v)$	: The degree of a vertex v.
$l(S, \overline{S})$	: The number of bad edges of the bipartition $(S, \overline{S})$
G_{S_k}	: A balanced signed graph with partition $(S_k, \overline{S_k})$ of $V(G)$.
K_n^-	: A complete signed graph with n vertices whose edges all have negative signs.
K_n	: Complete graph on n vertices.
$G \sim H$	: G and H are switching equivalent.
$\partial(S)$	: The stability degree of S for a subset S of $V(G)$ in a signed graph G.
$G[S]$	: The induced graph on S for a subset S of $V(G)$ in a signed graph G.

SYMBOLS

σ	**:** The sigma function.
$\lceil . \rceil$	**:** The ceiling function.
$\lfloor . \rfloor$	**:** The floor function.
$\sim$	**:** The equivalence relation.

LIST OF FIGURES

Page

ON BALANCING SOCIAL NETWORKS

SUMMARY

In this thesis, we study the optimization of complete signed graphs. We develop a graph theoretical algorithm that balances any signed graph with specific attention paid to complete signed graphs. In 1946, Heider observed that the relationships between members of some social networks do not change easily, while relationships in other networks do. He observed that among all networks consisting of three people, the relationships are stable when all members are friendly towards each other or if two friends are mutually hostile towards a third member. He characterized these networks as balanced triads and the remaining networks consisting of three people as unbalanced triads. He postulated that larger networks will be balanced when every triad of such a network is balanced. Later research in the field of structural balance has been concentrated on three main issues: Determining whether or not a given social network is balanced, developing reasonable metrics that quantify how balanced a given social network is, and developing algorithms that balance any social network and examining their correspondence to the transformations of real world social networks.

Cartwright and Harary showed that Heider's observations can be summarized using signed graphs. They defined a *balanced signed graph* as a graph where the product of edge signs of every cycle of G is positive. However, this condition is computationally difficult to check in a general signed graph. To overcome this difficulty Cartwright and Harary proved the Structure Theorem which states that, a signed graph G is balanced if its vertex set can be partitioned into two disjoint sets (one of which could possibly be empty) such that edges lying inside each of the sets are positive edges, while edges with endpoints in different sets are negative. This condition is much easier to verify than the cycle condition since by the Structure Theorem the set of negative edges has to form a bipartite graph, which is a computationally simple condition to check. If the graph of negative edges is bipartite, we can immediately identify the two parts during this process. What remains is answering whether there are any positive edges between the two parts, if not the network is balanced.

To measure the amount of imbalance in a signed graph several metrics have been developed. Harary defined the *line index,* $l(G)$ of a signed graph G as the minimum number of edges whose signs need to be negated so that the resultant signed graph is balanced. The line index $l(G)$ seems like a reasonable metric to consider, however, Barahona showed that the problem of determining the line index of a signed graph is an NP-Complete problem, meaning that most probably it is computationally infeasible even for relatively small networks. In fact even for networks in which all edges are negative, computing $l(G)$ is still of NP-Complete complexity. This case is not difficult, since by the Structure Theorem, to preserve the largest number of the original negative edges, one needs to identify the largest bipartite subgraph of the original network. This is the well known NP- Complete problem of finding a MAX-CUT of a given graph.

Another metric introduced by Harary to measure the amount of imbalance of a signed graph is the ratio of the number of positive cycles to the total number of cycles in that signed graph. Working with this metric, Antal et al. developed two greedy algorithms for balancing complete signed graphs focusing on triangles. Their algorithm negates the sign of an edge as long as the negation reduces the total number of negative triangles. Their algorithms cannot balance all signed graphs. Antal et al. and Strogatz et al. have shown that both of these algorithms may not balance some networks because the algorithms may get stuck in "jammed states", which are local but not global minima of the energy functions used to evaluate the amount of imbalance in a network. It may be said that even for the class of signed graphs, where these algorithms achieve balance, due to the randomness involved in negating edges, the predictive power of these algorithms about real world networks could be questionable.

As a consequence of the perceived difficulties of the approaches mentioned above, new algorithms have been developed which only try to balance signed graphs without much consideration of their predictive power. Marvel et al. developed a dynamic algorithm using a continuous model that balances almost all graphs. They demonstrated that if there are more positive edges than negative edges in the signed graph their algorithm produces a complete signed network by negating the sign of every negative edge. On the other hand, if the majority of the edges have negative sign, their algorithm outputs a balanced signed graph by partitioning the vertex set in to two parts of equal size. We show this is an unreasonable prediction about the evolution of social networks. They also do not quantify the number of edges negated by their algorithm in terms of $l(G)$. The approach in Marvel et al. has been the inspiration of many new research articles that try to balance social networks using continuous models. Although from the perspective of being able to balance networks the work of Marvel et al. is a major achievement, we will show that such an algorithm may fail to correspond with the transformations of actual social networks.

In this thesis we first define *good edges* and *bad edges* for a given bipartition of the vertex set of a signed graph. The set of good edges is composed of positive edges that lie inside each part and negative edges that have each endpoint in a different part. The remaining edges are bad edges. We then develop a new metric, *the stability degree of a vertex*, which is the difference between the number of good edges and bad edges incident to that vertex for a given partition of the vertex set. We define *the stability degree of a partition* $(S,\overline{S})$ as the sum of the stability degree of each vertex of the signed graph for the partition $(S,\overline{S})$.

We proceed to show that the stability degree of a partition and the line index of a partition are equivalent as parameters, with stability degree having the advantage that it is a local parameter to which local maximization algorithms can be applied. We consider one such algorithm based on the greedy Max- Cut algorithm. Our algorithm initializes with the partition $(V(G),\emptyset)$ and moves a vertex with the least stability degree as long as a vertex with negative stability degree exists and terminates after minor modifications when such a vertex does not exist. We then switch the signs of every bad edge of the final partition to yield a balanced signed graph.

We prove that the algorithm is a polynomial time algorithm and that, the number of edges of a complete signed graph whose signs are flipped, is bounded by the line index of the complete signed graph of all negative edges. We also calculate how the stability degree of each vertex and the stability degree of each new partition changes in terms of the stability degree of the vertex which was moved. We show as expected that the algorithm does not always output a partition which achieves $l(G)$. This conclusion holds even when the set of signed graphs is restricted to complete signed graphs. We provide several examples of complete signed graphs, including an infinite family, where our algorithm fails to output a partition which achieves $l(G)$. It is still an open question whether computing the line indices of complete signed graphs is of NP- Complete difficulty.

Using the algorithm we prove that computing the line indices of graphs, where each vertex is the endpoint of at least as many good edges as bad edges, is still NP-Complete. We in fact reduce the problem of computing the line index of any signed graph to computing the line index of a signed graph of the type mentioned above. This point is crucial when considering the algorithm of Marvel et. al. Because their algorithm always yields a balanced signed graph with all positive edges when the input signed graph has a higher density of positive edges, and computing the line index of such signed graphs is NP-Complete as we have proved, their algorithm clearly does not optimize the number of edges whose signs have to be flipped. Indeed we provide examples of signed graphs where the density of positive edges is greater where our algorithm forces much fewer changes on the edge signs of the underlying signed graph.

Since the original goal of structural balance theory was to predict the evolution of actual social networks, the success of any algorithm developed to balance signed graphs will depend on whether or not such an algorithm predicts the evolution of real social networks. As such we have applied our algorithm to two simple networks. We considered the network formed by the five major state participants in the current Syrian War and a social network of members of a 34 person Karate Club. Data on the structure of these two networks was recorded in two academic papers. Our algorithm predicted the final state of these networks with great accuracy.

Lacking data on larger social networks, we have not been able to apply our algorithm to larger networks. There have been a limited number of research articles considering larger networks which make the conclusion that Heider's balance theory may be too simplistic when considering large social networks. Two criticisms made of Heider's research have been that the assumptions that relationships are symmetric and that they lack magnitude are not realistic. The second criticism does not hold for our algorithm since it can be modified slightly to work on weighted graphs. To answer the first point we hope to develop our algorithm in future research to work on directed weighted graphs and apply any such algorithm to real social networks if data is available.

SOSYAL AĞLARIN DENGELENMESİ

ÖZET

Bu tezde, işaretli (+/-) tam çizgelerin optimizasyonunu incelenmektedir. Herhenagi bir işaretli çizgeyi, özellikte tam işaretli çizgeleri, dengleyen bir algoritma geliştirilmiştir. 1946'da, Heider bazı sosyal ağlarda bireyler arasındaki ilişkilerin değişmesinin zor olmasına rağmen diğer ağlarda bireyler arasındaki ilişkilerin kolaylıkla değiştiğini gözlemledi. Üç kişiden oluşan tüm ağları incelediğinde, bireylerinin hepsinin arasında arkadaşça ilişkiler olduğu zaman veya iki arkadaş ortak bir üçüncü kişiye düşman oldukları zaman bu ilişkilerin kalıcı olduğunu fark etti. Bu tür üç kişilik sosyal ağları dengeli üçlüler olarak tanımlarken, geri kalan üç kişilik ağları dengesiz üçlüler olarak tanımladı. Daha büyük sosyal ağlar içinse, eğer ağın içindeki bütün üçlüler dengeli ise ağın dengeli olacağını belirtti. Bu traihten sonra yapısal denge konusu üzerinde yapılan çalışmalarda üç önemli nokta üzerinde durulmuştur: İncelenen herhangi bir ağın dengeli olup olmadığının belirlenmesi, dengesiz ağlardaki dengeszilik miktarını ölçen ölçütlerin geliştirilmesi, ve tüm ağları dengeleyen algoritmaların geliştirilmesi ve bu algoritmaların gerçek sosyal ağlardaki dönüşümleri ne doğrulukta tahmin ettiklerinin irdelenmesi.

Cartwright ve Harary, Heider'in gözlemlerinin işaretli çizgeler kullanılarak özetlenebileceğini gösterdi. Onlar *dengeli işaretli bir çizgeyi* her bir döngüsünün kenar işaretlerinin çarpımının pozitiv olduğu bir çizge olarak tanımladılar. Bununla birlikte, bu koşulun genel işaretli bir çizgede kontrol edilmesi oldukça zordur. Bu zorluğun üstesinden gelmek için Cartwright ve Harary, işaretli bir çizgenin dengeli olabilmesi için çizgenin düğüm kümesinin, her pozitiv kenar parçaların içinde ve her negatif kenar iki parça arasında kalacak şekilde, ikiye bölünmesi gerektiğini gösteren Yapı Teoremini kanıtladılar. Tabiiki bu kümelerden birisinin boş olma ihtimalini de belirttiler. Bu koşulu doğrulamak, döngü koşulundan daha kolaydır, çünkü Yapı Teoremi negatif kenarlar kümesinin iki parçalı (bipartite) bir çizge oluşturduğunu ima etmektedir ki bunu kontrol etmek kolaydır. Geriye kalan işlem positif kenarların parçalar içinde kalıp kalmadığını değerlendirmektir ve eğer böyle ise çizge dengelidir.

İşaretli bir çizgedeki dengesizlik miktarını ölçmek için ise birkaç farklı ölçüt geliştirilmiştir. Harary, $l(G)$ *satır endeksini*, işaretli bir çizge G'nin dengeli bir çizgeye dönüşmesi için işaretlerinin değişmesi gereken en küçük kenar sayısı olarak tanımladı. $l(G)$ satır endeksi, dikkate alınması gereken makul bir ölçüt gibi gözükse de, Barahona, işaretli bir çizgenin satır endeksini belirleme sorununun NP-Tam sınıfına dahil problem olduğunu gösterdi. Bu ise nispeten küçük ağlar için bile $l(G)$ değerinin hesaplanabilmesinin nerdeyse imkansız olduğunu göstermektedir. Aslında, tüm kenarların negatif olduğu ağlar için bile, $l(G)$'yi hesaplamak hala NP-Tam zorluktadır. Bu durum zor değildir, çünkü Yapı Teoremine göre en fazla negatif kenarın işaretinin korunması için çizge içindeki en büyük kesit tespit edilmelidir ve bu da çok iyi bilinen NP-Tam sınıfında bir sorudur.

Harary tarafından işaretli bir çizgenin dengesizlik miktarını ölçmek için geliştirilen diğer bir ölçüt ise, pozitif döngü sayısının o işaretli çizgedeki toplam döngü sayısına oranıdır. Bu metrikle çalışan Antal ve diğerleri, çizgelerin içlerindeki üçgenlere odaklanan ve işaretli çizgeleri dengelemeye çalışan iki açgözlü (greedy) algoritma geliştirdi. Algoritmaları, toplam negatif üçgenlerin sayısını azalttığı sürece bir kenarın işaretini değiştirmektedir. Antal ve diğerleri ve Strogatz ve diğerleri bu algoritmaların her ikisinin de bazı ağları dengeleyemeyebileceğini göstermiştir, çünkü algoritmalar ağdaki dengesizlik miktarını değerlendirmek için kullanılan enerji fonksiyonlarının yerel ancak küresel olmayan minimumlarında sıkışıp kalabilir. Bu algoritmaların dengeyi sağladığı işaretli çizgeler sınıfı için bile, kenar işaretlerinin değişimindeki rastgelelik yüzünden, algoritmaların gerçek dünya ağları hakkındaki tahmin gücü sorgulanabilir.

Yukarıda belirtilen yaklaşımlarda karşılaşılan zorlukların bir sonucu olarak, gerçek sosyal ağlar üzerindeki tahmin gücü göz önüne alınmadan işaretli çizgeleri dengelemeye çalışan yeni algoritmalar geliştirilmiştir. Marvel ve diğerleri hemen hemen tüm işaretli çizgeleri dengeleyen sürekli bir model üzerine inşa edilmiş dinamik bir algoritma geliştirdi. İşaretli bir çizgedeki positif kenar sayısı negatif kenar sayısından fazla ise, algoritmaları her negatif kenarın işaretini positif yaparak dengeli bir çizge oluşturmaktadır. Öte yandan, kenarların çoğunluğu negatif işarete sahipse, algoritmaları iki parçasının büyüklüğü eşit olan dengelenmiş işaretli bir çizge oluşturmaktadır. Bu tezde bu sonuçların sosyal ağların evrimi ile ilgili makul olmayan bir tahmin olduğunu gösteriyoruz. Ayrıca algoritmaları tarafından işareti değiştirilen kenar sayısını $l(G)$ cinsinden nicelememektedirler. Marvel ve diğerlerinin bu algoritmaları, sosyal ağları sürekli modeller kullanarak dengelemeye çalışan birçok yeni araştırma makalesinin ilham kaynağı olmuştur. Her ne kadar ağları dengeleyebilme açısından Marvel ve diğerlerinin araştırmaları büyük bir başarıysa da, böyle bir algoritmanın gerçek sosyal ağların dönüşümüyle uyuşmadığını göstereceğiz.

Bu tezde, ilk önce düğüm kümesinin herhangi bir ikiye parçalanması $(S, \overline{S})$ için *iyi kenar* ve *kötü kenar* kavramları tanımlanmaktadır. İyi kenarlar kümesi, her bir parçanın içinde bulunan pozitif kenarlardan ve uçları farklı parçalarda bulunan negatif kenarlardan oluşur. Kalan kenarlar ise kötü kenarlardır. *Bir düğümün kararlılık derecesi* bu düğümün uç noktası olduğu iyi ve kötü kenarların sayısı arasındaki farktır. *Bir ikiye parçalanmanın kararlılık derecesi* ise, bu parçalanma için bütün düğümlerin kararlılık derecelerinin toplamıdır.

Daha sonra geliştirilen bu yeni ölçütlerin çizgi endeksi ölçütü ile eşdeğer olduğu gösterilmiştir. Karalılık derecesi çizgi endeksine eşdeğer olmasına rağmen, kararlılık derecesi yerel bir ölçüt olması ve bu ölçütü optimize eden algoritmaların geliştirilmesine olanak sağlaması sebebiyle çizge endeksinden üstündür. Tezimizde En Büyük Kesit Algoritması genelleştirilerek oluşturulan böyle bir yerel algoritmayı incelemekteyiz. Algoritmamız başlangıçta $(V(G), \emptyset)$ ikiye parçalanmasından başlıyarak en negatif kararlılık derecesine sahip olan bir düğümü, negatif kararlık derecesine sahip bir düğüm bulunduğu sürece, haraket ettirerek bulunduğu parçayı değiştirir ve basit bazı değişikliklerden sonra sonuç olarak yine farklı bir ikiye parçalanma verir. Bu sonuç parçalanmasındaki bütün kötü kenarların işaretleri değiştirilerek dengeli bir ağ ortaya konur.

Algoritmanın polinom zamanda çalışan bir algoritma olduğunu ve tam çizgelerde algoritmanın işaretini değiştirdiği kenar sayısının, tüm kenarları negatif olan tam çizgenin çizgi endeksiyle sınırlı olduğunu ispatlıyoruz. Ayrıca yeni oluşan her parçalanmanın kararlılık derecesinin ve bu parçalanma için her düğümün kararlılık derecesinin, taşınan düğümün kararlılık derecesi cinsinden nasıl değiştiğini hesaplıyoruz. Fakat beklendiği üzere, algoritmanın her zaman $l(G)$ değerine sahip bir parçalanmayı bulamadığını da gösteriyoruz. Bu sonuç, değerlendirilen işaretli çizgeler kümesi tam işaretli çizgeler kümesi ile sınırlandığında bile geçerlidir. Algoritmamızın $l(G)$ değerine sahip bir parçalanmayı her zaman veremediğini, sonsuz bir aile de dahil olmak üzere, bazı tam işaretli çizgeleri örnek vererek gösteriyoruz. Tam işaretli çizgelerin çizge endeksinin hesaplanması probleminin NP-Tam zorlukta olup olmadığı hala açık bir sorudur.

Algoritmayı kullanarak, her bir düğümünün en az kötü kenarlar kadar iyi kenarların uç noktası olduğu çizgelerin, çizgi endekslerini hesaplamanın hala NP-Tam olduğunu kanıtlıyoruz. Buna ek olarak, genel bir işaretli çizgenin çizgi endeksini hesaplama sorusunun, yukarıda belirttiğimiz cinsten bir işaretli çizgenin çizgi endeksini hasaplama sorusuna indirgenebileceğini de gösteriyoruz. Marvel ve diğerlerinin algoritması dikkate alındığında bu nokta çok önemlidir. Algoritmaları, verilen işaretli çizgede pozitif kenar yoğunluğu yarıdan fazla olduğunda her zaman bütün kenarları pozitif olan dengeli bir işaretli çizge oluşturduğundan ve bu cins işaretli çizgelerin çizgi endeksinin hesaplanması kanıtladığımız gibi NP-Tam zorlukta olduğu için, algoritmaları açıkça işareti değiştirilmesi gereken kenar sayısını optimize etmemektedir. Buna ek olarak, algoritmamızın onların algoritmasından daha az kenarın işaretini değiştirdiği ve positif kenarların yoğunluğunun yarıdan fazla olduğu işaretli çizge örnekleri sunuyoruz.

Yapısal denge teorisinin asıl amacı, gerçek sosyal ağların gelişimini öngörmek olduğundan, işaretli çizgeleri dengelemek için geliştirilen herhangi bir algoritmanın başarısı, böyle bir algoritmanın gerçek sosyal ağların gelişimini öngörme becerisine bağlı olacaktır. Bu yüzden algoritmamızı iki basit ağa uyguluyoruz. Mevcut Suriye Savaşı'nın parçası olan beş büyük devletten oluşan sosyal ağ ile 34 kişilik bir Karate Klübünün üyelerinden olan sosyal ağı inceliyoruz. Bu iki ağın yapısına ilişkin veriler iki akademik makaleden elde edilmiştir. Algoritmamız bu ağların son halini büyük bir isabet oranıyla tahmin etmiştir.

Daha büyük sosyal ağlar hakkında veri eksikliği nedeniyle algoritmamızı daha büyük ağlara uygulayamadık. Daha geniş ağları inceleyen sınırlı sayıda akademik makale mevcuttur ve bu makalelerde büyük sosyal ağlar için Heider'in denge teorisinin varsayımlarının fazlaca basit olabileceği sonucuna varılmıştır. Heider'in araştırmalarına sıklıkla yapılan iki eleştiri, ilişkilerin simetrik olduğu varsayımının ve ilişkilerin herhangi bir ağırlığının olmadığı varsayımının gerçekçi olmadığıdır. Algoritmamız kolaylıkla ağırlıklı çizgeler üzerinde çalışacak şekilde değiştirilebileceğinden yukarıdaki ikinci eleştiri yaptığımız çalışma için geçerli değildir. İlk noktaya cevap vermek için, algoritmamızı yönlü ve ağırlıklı çizgeler üzerinde çalışacak şekilde geliştirmeyi ve eğer veriler bulunabilirse bu tür bir algoritmayı daha büyük sosyal ağlara uygulamayı umuyoruz.

1. INTRODUCTION

In this thesis we develop a graph theoretical algorithm which balances any given signed graph.

This concept of balancing has many practical applications due to the interest in predicting the evolution of social networks. Social networks like face-book, twitter, whatsapp and other social platforms have created increasingly interconnected networks. These networks and others bring together a wide range of people with different backgrounds and agendas, which leads to tension between some of the members of such networks. This tension is the force which continuously transforms the nature of the relationships between members of social networks. Since members of a social network have favorable or unfavorable relationships with other members, it is only natural to consider the role of friends and enemies.

Balance theory, pioneered by Heider [2], is concerned with how an individual's attitudes or opinions coincide with those of others in any social network. If two friends have the same attitude towards a third, then there is a balance; and in the contrary if two friends have different attitudes toward a third, then there is a dissonance. Heider observed that relationships in a network change because members seek to consistently categorize others as friends or enemies. Each member implicitly desires friends of his/her friends to be his/her friends, enemies of his/her friends to be his/her enemies, enemies of his/her enemies to be his/her friends, and friends of his/her enemies to be his/her enemies. Heider concluded that for small networks at least, this desire transformed some relationships until consistency could be achieved and called such networks balanced networks.

Harary and Cartwright [3], described balance theory using the notion of signed graphs and determined conditions which describe all balanced social networks. Harary and Cartwright proved that all balanced network are made up of at most two cliques where members in a given clique are always friends and members of different cliques are always enemies [3]. They developed several parameters describing the degree of balance in signed graphs. Of interest is $l(G)$, which is the size of the smallest set of edges whose signs need to be flipped so that the result is a balanced signed graph.

Unfortunately, Barahona [4] has shown that the problem of computing $l(G)$ is NP-Complete. Consequently, there has been interest in limiting the set of signed graphs to complete signed graphs. It is still not known whether computing $l(G)$ for complete signed graphs is feasible. However, a number of recent studies like Antal et al. [5] and Strogatz et al. [6], have developed polynomial time algorithms that balance some complete signed graphs without achieving $l(G)$. The algorithm in [5] is limited because it cannot balance all complete signed graphs. On the other hand, while the algorithm of [6] balances almost all complete signed graphs, it frequently does so by flipping the signs of considerably more edges than $l(G)$.

In this thesis we develop an algorithm based on the greedy max- cut algorithm, which balances all signed graphs and imposes significantly less change on the signs of the edges than the algorithms described above.

Let $G = (V, E, \omega)$ be a weighted graph where $\omega : E \to \mathbb{R}$. The weight of an edge represents the intensity of the relationship between its two endpoints. If two people have no opinion of each other, we say the corresponding edge has weight 0. Now for the below definitions let us particularly consider the signed graph $G = (V, E, \sigma)$ where $\sigma : E \to \{\pm 1\}$, however, the definitions could be applied for any weighted graph.

Let $(S, \overline{S})$ be a bipartition of $V(G)$ and let $P(v)$ be the set containing v. In other words, $P(v) = S$ if $v \in S$, and $P(v) = \overline{S}$ if $v \in \overline{S}$. We define the *stability degree of v for a partition* $(S, \overline{S})$ as

$$\partial_S(v) = \sum_{x \in P(v)} \sigma(vx) - \sum_{y \notin P(v)} \sigma(vy),$$

.

where $\sigma(vx)$ refers to the weight of edge vx in the weighted graph G. Note that $\partial_S(v)$ is defined even if $v \in \overline{S}$. The *stability degree of a partition* $(S, \overline{S})$ is defined as

$$\mathscr{D}(S) = \sum_{v \in V} \partial_S(v).$$

If $S = V(G)$, instead of $\partial_S(v)$ and $\mathscr{D}(S)$ we will simply write $\partial(v)$ and $\mathscr{D}(G)$ respectively.

For the partition $(S,\overline{S})$, if the edge xv contributes positively/negatively to $\partial_S(v)$, x is called a *good/bad neighbor* of v and the edge xv is called a *good/bad edge*. Also, given a partition $(S,\overline{S})$ of $V(G)$, the set of good edges is composed of the positive edges that lie inside each part and the negative edges that lie between S and $\overline{S}$. Likewise, the set of bad edges is composed of the negative edges that lie inside each part and the positive edges that lie between S and $\overline{S}$. In [7], the number of bad edges of $(S,\overline{S})$ is defined as $l(S,\overline{S})$.

We then prove that a signed graph H is a balanced if and only if $V(H)$ can be partitioned into sets S and $\overline{S}$ such that $\partial_S(v) = d(v)$ for all $v \in V(H)$ and $\mathscr{D}(S) = 2|E(H)|$. Next, we prove that finding the number of bad edges $l(S,\overline{S})$ and the stability degree $\mathscr{D}(S)$ of a partition $(S,\overline{S})$ of $V(G)$ are equivalent, because

$$l(S,\overline{S}) = \frac{2|E(G)| - \mathscr{D}(S)}{4}.$$

We next prove two results for an the upper bound for the line index of signed graphs. They are: $l(G) < \frac{|E|}{2}$, for any signed graph G without isolated vertices and $l(S,\overline{S}) \leq l(K_n^-) = \lceil \frac{n^2-2n}{4} \rceil$, for a complete signed graph G on n vertices and partition $(S,\overline{S})$ identified by our algorithm, where K_n^- is a complete graph of n vertices with all negative edges.

Let G be a signed graph and $S \subseteq V(G)$. If G' is the signed graph obtained from G by negating the signs of the edges between S and $\overline{S}$ and keeping the signs of the remaining edges unchanged, we say G' is obtained from G by *switching* S. If a signed graph H can be obtained from a signed graph G by switching some $S \subseteq V(G)$, G and H are said to be *switching equivalent*, written $G \sim H$, and H is called the *unified representation* of $G(S,\overline{S})$. Note that if $G \sim H$, then $l(G) = l(H)$, which was proved in [7]. Also, since G and H have the same vertex set V, for any vertex v, the stability degree of v in H is exactly the same with the stability degree of v in G. Moreover, each good edge of $(S,\overline{S})$ has positive sign in H, while each bad edge of $(S,\overline{S})$ has negative sign in H.

We also prove that determining the line indices of signed graphs, in which each vertex is incident to at least as many positive edges as negative edges, is still NP-Complete.

For a subset S of $V(G)$ of a signed graph G, the *stability degree of* S is defined as the sum of the weights of the edges which only have one endpoint in S, denoted by $\partial(S)$.

In this thesis, we finally prove the following result containing four equivalent statements for many important reasons.

1. $(S,\overline{S})$ is an optimal partition of G if and only if H cannot be split, where H is the unified representation of the partition $(S,\overline{S})$;

2. H cannot be split if and only if for any $A \subseteq V(H)$, $\partial(A) \geq 0$;

3. For any $A \subseteq V(H)$, $\partial(A) \geq 0$ if and only if for any $A \subseteq V(H)$, $\mathscr{D}(H[A]) \leq \sum_{v \in A} \partial(v)$;

4. For any $A \subseteq V(H)$, and $\mathscr{D}(H[A]) \leq \sum_{v \in A} \partial(v)$ if and only if $(S,\overline{S})$ is an optimal partition of G.

Next, we present some basic definitions in graph theory that will be used throughout the thesis.

1.1 Basic Definitions in Graph Theory

In this section we review the basic definitions that are used in this thesis. We will closely follow the introductory graph theory text of Douglas B. West [8].

A *graph* G is a triple $(V(G), E(G), f)$ where $V(G)$ is the set of vertices of G, $E(G)$ is the set of edges of G, and f is a relation that assigns for every edge two vertices. Note that in $E(G)$, every edge is determined by a pair of vertices which are not necessarily distinct and these vetices are called the endpoints of that edge. Pictorially, we represent a given graph by drawing a point for each vertex and a curve joining the endpoints to represent each edge. The *degree of a vertex* v in a graph G, denoted $d(v)$, is the number of edges in $E(G)$ which have v as an one endpoint. If the vertex set $V(G)$ and edge set $E(G)$ of graph G are finite, then we call G a *finite graph*.

A *simple graph* is a graph where its edges are uniquely determined by distinct endpoints. That is, if e represents an edge in $E(G)$, then e is specified by its unique endpoints, say u and v, written as $e = uv = vu$. Vertices that are endpoints of an edge are adjacent vertices and are also called neighbors. An edge is *incident* to its endpoints. Unless specifically stated otherwise, we consider simple graphs whose vertex and edge sets are finite throughout this thesis.

A set of vertices S in a graph G is said to be an *independent set* if the vertices in S are pairwise non-adjacent.

A *path* P is a simple graph with a finite alternating sequence of vertices and edges formed by consecutive vertices such that every edge and vertex in the sequence is distinct. A path which begins at vertex u and ends at vertex v is called a $u,v-path$. In a $u,v-path$ we start walking from vertex u along an edge adjacent to it to another vertex and continuing until we get to the last vertex v. The length of a path P with edge set $E(P)$ is the size of $E(P)$.

A *cycle* C is a simple graph with a finite alternating sequence of vertices and edges formed by consecutive vertices such that every edge is distinct and the degree of every vertex is two. The length of the cycle C with edge set $E(C)$ is the size of $E(C)$. If the size of $E(C)$ is odd/even, then we call that C is an odd/even cycle.

A graph G is said to be *connected* if for every $u,v \in V(G)$ there exists a $u,v-path$ in G. Otherwise G is called *disconnected.*

A graph H is called a *subgraph* of a given graph G, if $V(H) \subset V(G)$ and $E(H) \subset E(G)$.

Let $S \subset V(G)$, then a subgraph M of G is called an *induced subgraph* if it is obtained from G by deleting the complement of the subset S of $V(G)$, denoted by $G[S]$. Therefore, $M = G[S]$ is made up of $|S|$ number of vertices and all the edges have endpoints only in S.

A *complete graph* is a simple graph in which every pair of vertices are adjacent. The class of unlabeled complete graphs with n vertices is denoted by K_n. A graph G with vertex set $V(G)$ is *bipartite* if $V(G)$ can be partitioned into two disjoint independent sets such that every edge (if there exists) of G connects a vertex in one set to a vertex in the opposite. These independent sets are called *partite sets* of G. A *complete bipartite graph* is a simple and bipartite graph with the property that two vertices u and v in it are neighbors if and only if they are in opposite partite sets. The class of unlabeled complete bipartite graph with partite sets of size m and n is denoted by $K_{m,n}$.

A *signed graph* H, written $H = (V,E,\sigma)$ consists of an underlying graph $G = (V,E)$ together with a sign function σ which assigns to each edge the values $+1$ or -1. More generally, a *weighted graph* is a graph where to each edge a real number is assigned. A given a social network can be modeled using a signed graph, where we represent each person using a different vertex. If the relationship between two people is positive, the edge between the two vertices representing them has weight +1, and if the relationship is negative the edge has weight -1. Furthermore, if two people have no opinion about each other, then there is no edge between the vertices representing them. Similarly, if we would also like to consider the intensity of a relationship between two people, we can instead use a weighted graph as a model.

1.2 Balance in Signed Graphs

In the field of psychology there has been considerable interest in understanding how and why relationships between members of a social network change. In order to simplify this problem some reasonable assumptions can be made:

1. The relationship between two people can be quantified as positive or negative or neither, for example, like or dislike or indifference.

2. The relationship between two people is symmetric, for example, if person **A** likes person **B** with some intensity, then **B** likes **A** with the same intensity.

Heider [2] noticed that in some social networks the relationships between its members do not change easily. He observed that a central characteristic of such networks is that for each member of the network, a friend of a friend is a friend, an enemy of a friend is an enemy, a friend of an enemy is an enemy, and finally an enemy of an enemy is a friend. Such networks are called balanced networks, while networks which are not balanced are called unbalanced. He also pointed out that an unbalanced social network has a tendency to transform to a balanced network because the relationships between some of its members change. In other words, if members of a social network cannot consistently decide who their friends and enemies are, there is some stress on the members of the network, which forces at least some of its members to reevaluate some of their relationships. It is worth noting that most large social networks will initially be unbalanced.

It is only natural to model the question using signed graphs. Cartwright and Harary [3] summarized Heider's observations as follows: A signed graph is *balanced* if any cycle has an even number of edges with weight -1. An equivalent condition is that the product of the edge weights of any cycle is $+1$. Such cycles are called *positive cycles*, while cycles which are not positive are called *negative cycles*. We say a social network is balanced if the corresponding signed graph is balanced. In a signed graph G, $c^+(G)$, $c^-(G)$ and $c(G)$ are the number of positive, negative and all cycles in G, respectively. See Fig. 1.1 for examples of possible social networks consisting of three members. In the remainder of this thesis the term social network refers to both the actual social network and the signed graph representing the social network. And also unless it is specified, solid edges in a given signed graph represent positive relations and broken edges represent negative relations between vertices.

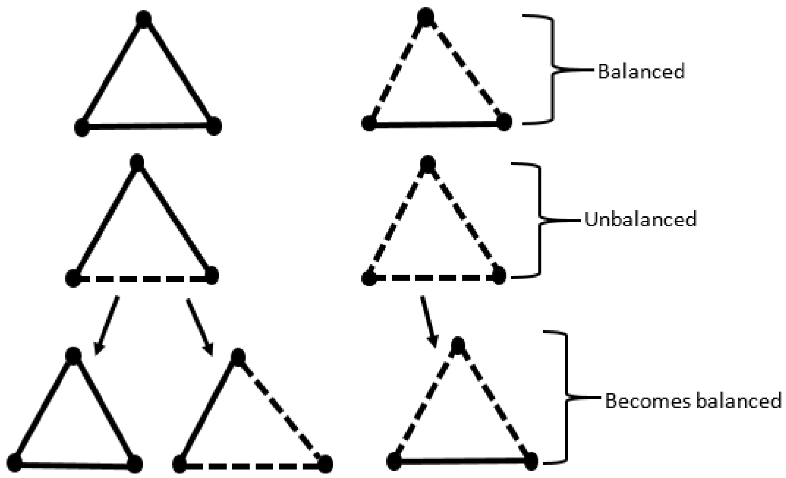

Figure 1.1 : Possible social networks consisting of three members.

In [3], Cartwright and Harary also determined the structure of all balanced networks.

Theorem 1 (Structure Theorem) *A signed graph is balanced if and only if its vertex set can be partitioned into two disjoint sets (one of which may possibly be empty) such that, edges inside each of the subsets (parts) are all positive and edges between vertices in different subsets are all negative.*

According to structural theorem all graphs in Fig. 1.2 and Fig. 1.3 are balanced.

In non-mathematical terms the theorem states that in all balanced social networks, either two groups exist where there are only positive feelings between members belonging to the same group and there are only negative feelings between members of different groups, or there is a single group in which there are only positive feelings between its members. The second case can be considered as a paradise case, but in

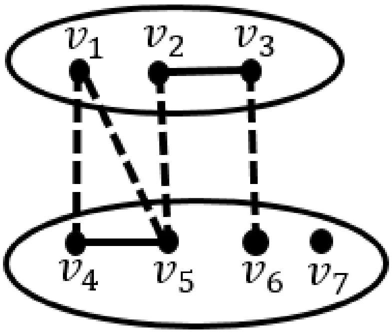

Figure 1.2 : The vertex set partitioned in to two disjoint sets and all the edges between the partites are negatives and the edges inside each partite are positive. Therefore, it is balanced.

general one of the implications of Heider's research is that for most societies given enough time polarization is unavoidable. Therefore, being able to determine the type of balanced network to which a given unbalanced network will transform, is of some importance. One could conjecture that large social networks, especially networks in which many members do not have an opinion about each other, may not transform into balanced social networks because individual members do not feel enough stress to change their relationships. Indeed, Leskovec et al. [9] seem to come this conclusion. One of the aims of this paper is to explain why even small complete networks, networks in which everyone has a positive or negative opinion of everyone else, may not reach balanced states.

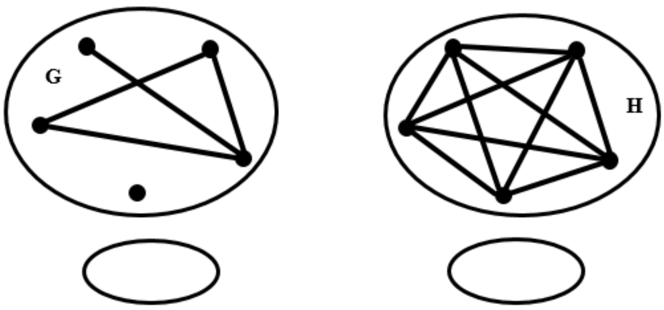

Figure 1.3 : For both graphs all the vertices lie in one set and the second set is empty. All the edges in the partites are positive. Therefore, both graphs are balanced.

When the balanced signed graph is a paradise case, that is for graphs of Fig. 1.3, we simply draw them by using the non - empty partite, which is Fig. 1.4.

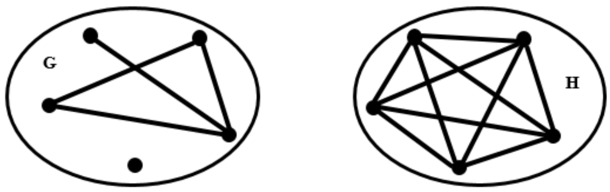

Figure 1.4 : The balanced signed graph like in Fig. 1.3, when one of the partite is empty set, we simply draw only the non - empty partite and omit the empty partite for simplicity.

1.3 Measure of Balance in Signed Graphs

Previous research in the field of structural balance has been concentrated on three main issues: Determining whether or not a given social network is balanced, developing reasonable metrics that quantify how balanced a given social network is, and developing algorithms that balance any social network and examining their correspondence to the transformations of real world social networks.

The first issue can easily be solved as has been done by Harary and Kabell in [10]. Theorem 1 implies that the set of negative edges forms a bipartite graph, which can be checked quickly. If the graph is bipartite, we can immediately identify the two parts during this process. What remains is answering whether there are any positive edges between the two parts, if not the network is balanced.

Since the decision problem of whether a signed graph is balanced or not is easy, and since most signed graphs are unbalanced, many different metrics have been developed to measure the amount of imbalance of a social network. Of particular concern to us are *degree of balance*, $\beta(G) = \frac{c^+(G)}{c(G)}$ (the ratio of balanced cycles to the total number of cycles) developed by Cartwright and Harary in [3], and the line index, $l(G)$, defined by Harary in [11]. Note that $l(G)$ is the minimum number of relationships in a social network which have to change from friendly to hostile or vice versa, so that the network becomes balanced.

In the signed graph F of Fig. 1.5 there are six cycles, of them three are positive cycles. Therefore, $\beta(F) = \frac{c^+(F)}{c(F)} = \frac{3}{6} = \frac{1}{2}$.

On the third issue of developing algorithms which balance social networks, one must be careful that a given algorithm is predictive of the transformations of real world

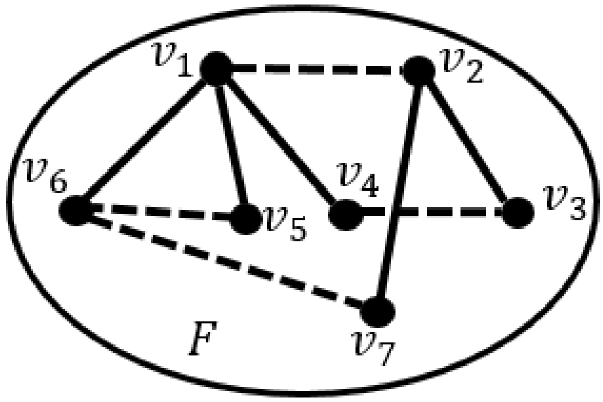

Figure 1.5 : F is a signed graph containing only the edges drawn.

social networks. After all for a given network, using Theorem 1 one could state a trivial algorithm that achieves balance by randomly dividing the vertex set into two disjoint sets, and changing each of the negative edges inside each set to positive, and each of the positive edges between the two sets to negative. However, it is clear that such a random algorithm would be incapable of explaining factual data. One of the reasons is that this trivial algorithm may impose more change on existing relationships than the network is capable of handling. In addition, algorithms which try to balance signed graphs by using either of the metrics above run into various problems.

The line index $l(G)$ is a reasonable metric to consider, since it measures the least number of relationships that need to change so that the network becomes balanced. It is only natural to expect that a given social network will transform into a balanced network which is as similar to the original as possible. However, Barahona [4] proved that calculating $l(G)$ is of NP-Complete complexity, meaning that most probably it is computationally infeasible even for relatively small networks. In fact even for networks in which all edges are negative, computing $l(G)$ is still of NP-Complete complexity, as was shown in [12]. This case is not difficult, since by Theorem 1, to preserve the largest number of the original negative edges, one needs to identify the largest bipartite subgraph of the original network. This is the well known NP-Complete problem of finding a MAX-CUT of a given graph. In fact, Xu [7] showed that for each signed graph G, there exists a signed graph H with only negative edges such that $l(G) = l(H)$.

As an example consider the signed graph G in Fig. 1.6 whose optimal cut that require the smallest number of edge sign change is the cut $(A, \overline{A})$ shown in the figure. In that

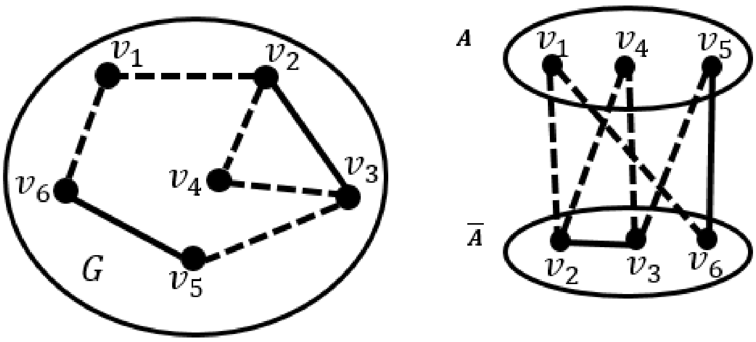

Figure 1.6 : A signed graph G containing the edges drawn with its optimal partition $(A,\overline{A})$.

case, $l(G) = l(A,\overline{A}) = 1$. Then, in graph G since there are two positive edges, v_2v_3 and v_5v_6, replacing them by two consecutive negative edges we get graph H, as shown in Fig. 1.7. Then, the cut that gives the minimum edge sign change in balancing H is $(B,\overline{B})$ and $l(H) = l(B,\overline{B}) = 1 = l(G)$, as required.

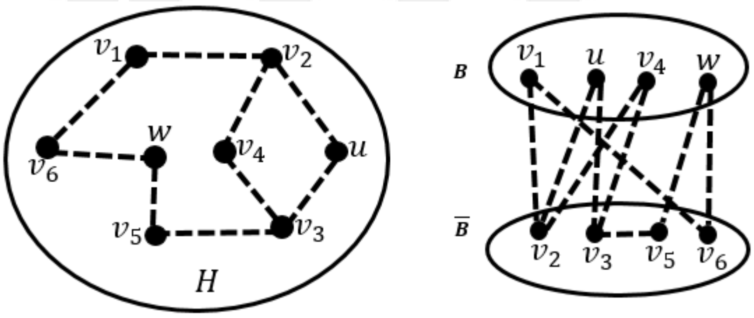

Figure 1.7 : A signed graph H obtained from graph G of Fig. 1.6 by replacing the positive edges by two consecutive negative edges with its optimal partition $(B,\overline{B})$.

1.4 Balance on Complete Signed Graphs

Antal et al. [5] have proposed two algorithms that for complete social networks negate some edges on triangles to decrease the number of negative triangles and thereby improve $\beta(G)$. For example, in their constrained triad dynamics algorithm, a random edge is negated if after the new sign assignment, the number of negative triangles decrease. It was shown in [5, 13] that both of the algorithms may not balance some networks because the algorithms may get stuck in ”jammed states”, which are local but not global minima of the energy functions used to evaluate the amount of imbalance in a network. It may be said that even for the class of signed graphs, where these algorithms achieve balance, due to the randomness involved in negating edges, the predictive power of these algorithms about real world networks could be questionable.

As a consequence of the perceived difficulties of the approaches mentioned above, new algorithms have been developed which only try to balance signed graphs without much consideration of their predictive power. In [6], Marvel et al. developed an algorithm based on a continuous model, that balances most complete signed graphs. However, their algorithm yields only two kinds of balanced states depending on the density of positive edges in the signed network. If the initial signed graph has more positive edges than negative ones, the algorithm yields a signed graph in which all edges are positive; if not, then the algorithm yields a balanced signed graph in which the two parts have equal size. The approach in [6] has been the inspiration of many new research articles that try to balance social networks using continuous models. Although from the perspective of being able to balance networks the work of Marvel et al. is a major achievement, we show in the last chapter that such an algorithm may fail to correspond with the transformations of actual social networks. We also show that a social network might not reach a balanced state where the partition of its members achieve $l(G)$.

In this thesis we will consider complete signed graphs from the perspective of their line indices. We will exhibit a discrete time greedy algorithm which balances any complete signed graph with n vertices by changing the signs of at most $\lceil \frac{n^2}{4} - \frac{n}{2} \rceil$ edges. This is best possible since $l(K_n^-) = \lceil \frac{n^2}{4} - \frac{n}{2} \rceil$. We then demonstrate that the problem of determining the line indexes of signed graphs where each vertex is incident to at least as many positive edges as negative edges, is still NP-complete. Furthermore, we will demonstrate an infinite family of complete signed graphs in which the ratio of positive edges to negative edges approaches 3, and for which once their line indices are calculated and are balanced accordingly, the resultant balanced networks have two parts with equal size. This example demonstrates that the solution in [6] may not have much predictive power.

2. BASIC RESULTS AND EXAMPLES

In this chapter we prove some results of the thesis which are important to prove some bigger results of the thesis in the next two chapters and we exemplify them with examples. We also generalize structure theorem stated in Theorem 1. For a a signed graph $G = (V, E, \sigma)$ and $(S, \overline{S})$ be a partition of V, we denote G_S for the balanced signed graph with vertex set V, edge set E, and parts S and $\overline{S}$, where the weights of the edges of G_S are obtained from the weights of the edges of G by negating the weight of each bad edge and leaving the remaining weights unchanged.

2.1 A Computational Example

In this section we demonstrate computations used in this thesis on a signed graph and the implications of these computations for actual social networks.

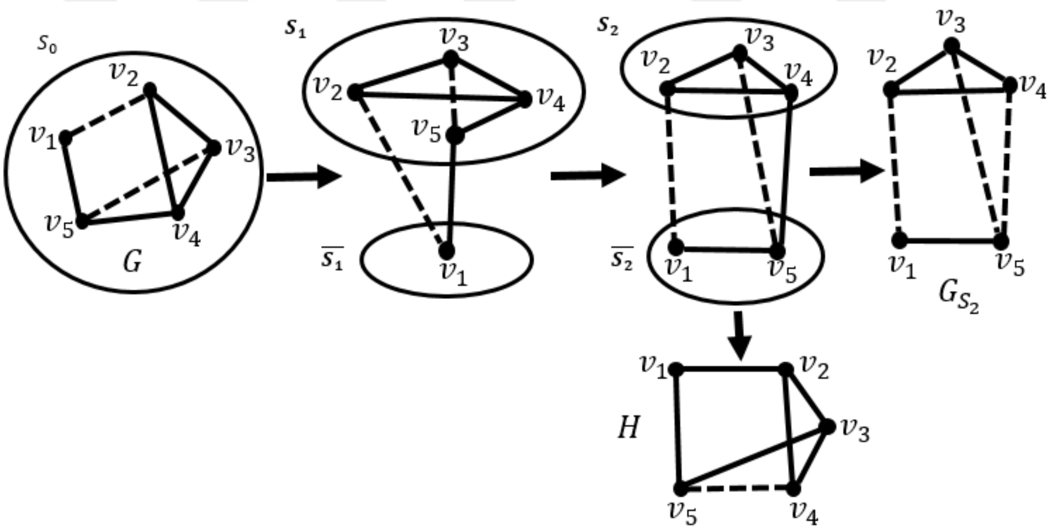

Figure 2.1 : G is a signed graph whose optimal partition is $(S_2, \overline{S_2})$ and H is the unified representation of the partition $(S_2, \overline{S_2})$. The balanced form of G then is G_{S_2}.

Consider the signed graph G in Fig. 2.1. We see this signed graph as the representation of a network with five members where, solid edges represents positive relationships and dashed edges represent negative relationships. Missing edges between vertices represent indifference. As a result solid edges have weight $+1$, negative edges have weight -1 and missing edges have weight 0. Observe that G contains the negative cycle $v_3v_4v_5$ and hence is not balanced. From the perspective of v_5, the current state of the social network is problematic, since v_3 is a friend of v_4 who is a friend of v_5, but v_3 and v_5 are enemies. According to Heider's observations we expect some relationships between members of the network to change until each member can consistently decide who their friends and enemies are. We would like to predict the final state of the network.

Originally, we partition the vertex set of G into $(S_0, \overline{S_0})$ where, $S_0 = V(G)$ and $\overline{S_0} = \emptyset$.

We compute the stability degree of each vertex of G as follows:

$$\partial(v_1) = \sum_{v_i \in S_0 = V(G)} \sigma(v_1 v_i) - \sum_{v_j \in \overline{S_0} = \emptyset} \sigma(v_1 v_j) = \sigma(v_1 v_2) + \sigma(v_1 v_5) = -1 + 1 = 0.$$

The term $\sum_{v_i \in S_0 = V(G)} \sigma(v_1 v_i)$ is the difference between the number of friends and enemies v_1 has in its group, while the term $\sum_{v_j \in \overline{S_0} = \emptyset} \sigma(v_1 v_j)$ is the difference between the number of friends and enemies v_1 has in the opposite group (in this case the opposite group is $\emptyset$). Consequently, a positive $\partial(v_1)$ value indicates that v_1 is happy to be a member of his group, while a negative $\partial(v_1)$ value indicates that v_1 would rather be a member of the opposite group.

Similarly, we find that $\partial(v_2) = 1$, $\partial(v_3) = 1$, $\partial(v_4) = 3$, $\partial(v_5) = 1$.

Thus, the stability degree of G is:

$$\mathscr{D}(G) = \sum_{v_i \in V} \partial(v_i) = 6.$$

Note that initially $\min_{v_i \in V} \partial(v_i) = \partial(v_1) = 0$, and consequently, it seems all members of the network are satisfied with the current partition. In this case, the bad edges of the partition are v_1v_2 and v_3v_5, and one would expect both of these negative relationships to transform to positive relationships to yield a balanced network. Notice that,

$$2 = l(S_0, \overline{S_0}) = \frac{2|E(G)| - \mathscr{D}(S_0, \overline{S_0})}{4} = \frac{14 - 6}{4}.$$

However, if we change the partition slightly by moving v_1, we see that it is not necessary to change both relationships v_1v_2 and v_3v_5. In other words, $l(S_0, \overline{S_0}) = 2 > l(G)$.

Consider the partition $(S_1, \overline{S_1})$ where $S_1 = \{v_2, v_3, v_4, s_5\}$ and $\overline{S_1} = \{v_1\}$. The stability degrees for this partition are:

$$\partial_{S_1}(v_1) = \sum_{v_i \in \overline{S_1}} \sigma(v_1 v_i) - \sum_{v_j \in S_1} \sigma(v_1 v_j) = -\sigma(v_1 v_2) + -\sigma(v_1 v_5) = 1 + -1 = 0.$$

$$\partial_{S_1}(v_2) = \sum_{v_i \in S_1} \sigma(v_2 v_i) - \sum_{v_j \in \overline{S_1}} \sigma(v_2 v_j) = \sigma(v_2 v_3) + \sigma(v_2 v_4) - \sigma(v_2 v_1) = 1 + 1 - (-1) = 3.$$

$$\partial_{S_1}(v_3) = \sum_{v_i \in S_1} \sigma(v_3 v_i) - \sum_{v_j \in \overline{S_1}} \sigma(v_3 v_j) = \sigma(v_3 v_2) + \sigma(v_3 v_4) + \sigma(v_3 v_5) = 1 + 1 + (-1) = 1.$$

$$\partial_{S_1}(v_4) = \sum_{v_i \in S_1} \sigma(v_4 v_i) - \sum_{v_j \in \overline{S_1}} \sigma(v_4 v_j) = \sigma(v_4 v_2) + \sigma(v_4 v_3) + \sigma(v_4 v_5) = 1 + 1 + 1 = 3.$$

$$\partial_{S_1}(v_5) = \sum_{v_i \in S_1} \sigma(v_5 v_i) - \sum_{v_j \in \overline{S_1}} \sigma(v_5 v_j) = \sigma(v_5 v_3) + \sigma(v_5 v_4) - \sigma(v_5 v_1) = 1 + (-1) - 1 = -1.$$

The bad edges of this partition are $v_1 v_5$ and $v_3 v_5$ which still implies $l(S_1, \overline{S_1}) = 2$, but we now observe that $\min_{v_i \in V} \partial_{S_1}(v_i) = \partial_{S_1}(v_5) = -1$. We deduce that v_5 is not happy with the current state of the network, and that the partitioning could be improved by moving v_5.

Let $(S_2, \overline{S_2})$ be the new partition where $S_2 = \{v_2, v_3, v_4\}$ and $\overline{S_2} = \{v_1, v_5\}$. The stability degrees for this partition are:

$\partial_{S_2}(v_1) = 2$, $\partial_{S_2}(v_2) = 3$, $\partial_{S_2}(v_3) = 3$, $\partial_{S_2}(v_4) = 1$, $\partial_{S_2}(v_1) = 1$, and $\mathscr{D}(S_2) = 10$.

The only bad edge of this partition is $v_4 v_5$ and, indeed, $1 = l(S_2, \overline{S_2}) = l(G)$. As a result we predict that the unbalanced social network represented by G will transform to the balanced network represented by G_{S_2} by flipping the sign of the edge $v_4 v_5$ as can be seen in Fig. 2.1. In terms of the actual real life network, we expect the relationship between persons represented by v_4 and v_5 to change from friendly to hostile.

The signed graph H given in Fig. 2.1 is the unified representation the partition $(S_2, \overline{S_2})$. We notice that for every $v_i \in V(H)$, $\partial_H(v_i) = \partial_{S_2}(v_i)$ and $l(H) = l_G(S_2, \overline{S_2}) = l(G) = 1$. As G and H are switching equivalent, the advantage of working with H instead of G is that we do not have to keep track of the partite sets S_2 and $\overline{S_2}$; the second partite in H is the $\emptyset$.

2.2 Basic Concepts

In this section, we restate the structure theorem, Theorem 1, using stability degrees. All the results are primarily stated for signed graphs, however, they can be generalized to weighted graphs after small modifications.

Claim 1 *H is a balanced signed graph if and only if $V(H)$ can be partitioned into sets S and $\overline{S}$ such that $\partial_S(v) = d(v)$ for all $v \in V(H)$ and $\mathscr{D}(S) = 2|E(H)|$.*

Proof. If H is a balanced signed graph, then by Theorem 1, $V(H)$ can be partitioned into sets S and $\overline{S}$ such that every edge inside each part is positive and every edge between the two parts is negative. Therefore, for the partition $(S,\overline{S})$ every edge is a good edge. Hence, for any vertex v, $\partial_S(v) = d(v)$. On the other hand if H is not a balanced signed graph, then any partition $(S,\overline{S})$ of $V(H)$ will contain a bad edge, and as a result, $\partial_S(v) < d(v)$ for some $v \in V(H)$. In this partition since every edge is counted twice and $\mathscr{D}(S) = \sum_{v_i \in V} \partial_S(v_i) = \sum_{v_i \in V} d(v_i) = 2|E(H)|$. □

We now describe how to balance a signed graph $G = (V,E,\sigma)$ given a partition $(S,\overline{S})$ of $V(G)$. Let $G = (V,E,\sigma)$ be a signed graph and $(S,\overline{S})$ be a partition of V. G_S is the balanced signed graph with vertex set V, edge set E, and parts S and $\overline{S}$, where the weights of the edges of G_S are obtained from the weights of the edges of G by negating the weight of each bad edge and leaving the remaining weights unchanged. The following lemma describes the relationship between $\mathscr{D}(S)$ and $l(S,\overline{S})$.

Lemma 1 *Let G be a signed graph and $(S,\overline{S})$ be a partition of $V(G)$. Then,*

$$l(S,\overline{S}) = \frac{2|E(G)| - \mathscr{D}(S)}{4}.$$

Proof. For each vertex $v \in V(G)$, $\partial_S(v)$ is the difference between the number of good edges and bad edges incident to v. Consequently, $d(v) - \partial_S(v)$ counts each bad edge incident to v twice. Summing over all $v \in V(G)$ we get $2|E(G)| - \mathscr{D}(S) = 4 \cdot l(S,\overline{S})$ because each bad edge gets counted twice more, once for each endpoint. □

Lemma 1 shows that as parameters $\mathscr{D}(S)$ and $l(S,\overline{S})$ are equivalent. If we know one we can easily calculate the other. Also as stated in [7], it is clear that $l(G) = \min_{(S,\overline{S})} l(S,\overline{S})$, which immediately yields that $l(G)$ is attained only by a partition which maximizes $\mathscr{D}(S)$. For a given signed graph G, we define a partition $(S,\overline{S})$ of $V(G)$ to be an *optimal partition* if $l(S,\overline{S}) = l(G)$.

The advantage of considering $\partial_S(v)$ is that unlike $l(G)$, which is a global parameter, $\partial_S(v)$ is a local parameter which measures the happiness of each individual with a given partition of the social network. If we see social networks as environments where individuals choose their alliances so as to maximize their own happiness, then the natural conclusion is that each person v makes choices only based on $\partial_S(v)$ because he/she is primarily concerned with his/her relationships with others and not with the relationships between other individuals. We can then calculate $\mathscr{D}(S)$ from $\partial_S(v)$. As $\mathscr{D}(S)$ and $l(S,\overline{S})$ are equivalent as parameters, $\mathscr{D}(S)$ becomes a stepping stone to tackle the global parameters $l(S,\overline{S})$ and $l(G)$. This insight forms the bases of our algorithm, where given a partition $(S,\overline{S})$ of $V(G)$, a person v decides to "move" to the opposite part if $\partial_S(v) < 0$. A signed graph G can be *split* if there exists a proper subset S of the vertex set such that $\mathscr{D}(G) < \mathscr{D}(S)$. In the next chapter we discuss this algorithm and show that even if each individual tries to maximize his/her happiness, then the network may not reach a balanced state.

3. THE DISCRETE TIME GREEDY ALGORITHM

In this chapter we develop our algorithm and apply to some basic examples.

3.1 The Algorithm

We now describe the discrete time greedy algorithm that can be used to balance any signed graph. The algorithm is inspired from the greedy Max-Cut algorithm. We move a vertex v from one partite set to the other, if $\partial_S(v) < 0$; in other words, the number of good edges incident to v is less than the number of bad edges incident to v. Once no such vertex is left, we output a balanced signed graph after minor modifications.

Algorithm.

Input: A signed graph G.

Output: A balanced signed graph G_{S_k}, where $(S_k, \overline{S_k})$ is a partition of $V(G)$ by negating the signs of every bad edge of $(S_k, \overline{S_k})$ in such a way $\partial_{S_k}(v) \geq 0$ for all $v \in V(G)$.

Initialization: Set $S_0 = V(G)$ and $\overline{S_0} = \emptyset$. (The algorithm could also take any bipartition of the vertex set as the initial condition.)

Iteration: At step i identify vertex v such that $\partial_{S_{i-1}}(v) = \min\limits_{w \in V(G)} \partial_{S_{i-1}}(w)$ and $\partial_{S_{i-1}}(v) \leq 0$. If no such vertex exists, construct $G_{S_{i-1}}$ and terminate. If such a vertex v exists and $\partial_{S_{i-1}}(v) < 0$, then construct S_i and $\overline{S_i}$ by removing v from the part containing it and adding v to the other part, and iterate. If $\min\limits_{w \in V(G)} \partial_{S_{i-1}}(w) = 0$, then let $A = \{x : 0 \leq \partial_{S_{i-1}}(x) \leq 1\}$. If there are vertices u and v such that $\partial_{S_{i-1}}(u) = 0$ and $v \in A$ and uv is a good edge, remove u from the partite containing it and add it to the other partite to construct S_i and $\overline{S_i}$, and iterate. If there are no such vertices, construct $G_{S_{i-1}}$ and terminate.

3.2 Examples

Example 3.1: Consider the signed graph G with vertex set $V(G) = S_0$ in Fig. 3.1.

The stability degree of the vertices in $V(G)$ are:

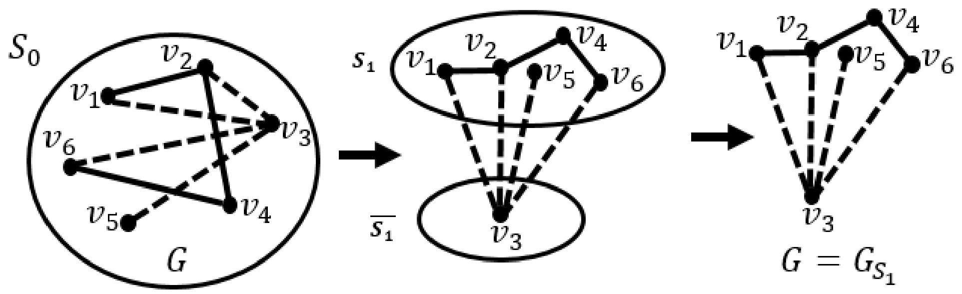

Figure 3.1 : G is a signed graph where all its edges are as drawn. And G_{S_1} is the balanced graph obtained after a single flip.

$\partial(v_1) = 0$, $\partial(v_2) = 1$, $\partial(v_3) = -4$, $\partial(v_4) = 2$, $\partial(v_5) = -1$, $\partial(v_6) = 0$.

Considering the stability degrees of the vertices, we have $\min_{w \in V(G)} \partial(w) = -4 = \partial(v_3)$. Thus, by the algorithm, we construct a new partition $(S_1, \overline{S_1})$ by removing v_3 from S_0.

In the partition $(S_1, \overline{S_1})$ the stability degree of the respective vertices are:

$\partial_{S_1}(v_1) = 2$, $\partial_{S_1}(v_2) = 3$, $\partial_{S_1}(v_3) = 4$, $\partial_{S_1}(v_4) = 2$, $\partial_{S_1}(v_5) = 1$, $\partial_{S_1}(v_6) = 2$.

In this partition every vertex has a positive stability degree, that is, there is no vertex with $\partial_{S_1}(v_i) \leq 0$. Hence, by our algorithm we terminate at the partition $(S_1, \overline{S_1})$. We also notice that, the stability degree of each vertex is equal to its corresponding degree. Therefore by Claim 1, we realize that G was already balanced. Our algorithm was still useful however, since using it we could easily identify the two mutually exclusive sets of alliances.

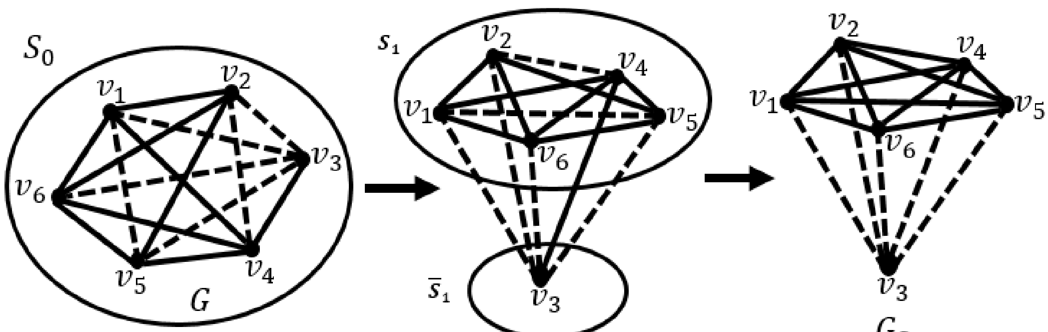

Figure 3.2 : G is a complete signed graph whose optimal partition is $(S_1, \overline{S}_1)$ obtained after one flip.

Example 3.2: Let $G = (V, E, \sigma)$ be the complete signed graph in Fig. 3.2.

We first calculate the stability degree of each vertex: $\partial(v_1)=1$, $\partial(v_2)=1$, $\partial(v_3)=-3$, $\partial(v_4)=3$, $\partial(v_5)=1$, $\partial(v_6)=3$. Clearly, the $\min_{w\in V(G)} \partial(w)=-3=\partial(v_3)$. Thus, we construct the partition $(S_1,\overline{S_1})$ by removing v_3 from $V(G)$ as shown in Fig. 3.2.

The stability degrees for the partition $(S_1,\overline{S_1})$ are:

$\partial_{S_1}(v_1)=3$, $\partial_{S_1}(v_2)=3$, $\partial_{S_1}(v_3)=3$, $\partial_{S_1}(v_4)=1$, $\partial_{S_1}(v_5)=3$, $\partial_{S_1}(v_6)=5$.

In this partition, the stability degree of every vertex is positive. Therefore, the process of balancing G terminates at the partition $(S_1,\overline{S_1})$. However, the stability degrees of all the vertices except v_6 are not equal to their corresponding degrees. This indicates that there are bad edges between the partite sets and inside S_1, which are clearly v_1v_5, v_2v_4 and v_3v_4. In other words, $l(S_1,\overline{S_1})=3$. We can verify this result by using Lemma 1 and replacing $2|E(G)|$ by $n(n-1)$. Note that the partition $(S_1,\overline{S_1})$ is optimal since $l(G)=3$.

By negating the signs of these edges we get the balanced graph G_{S_1} in Fig. 3.2.

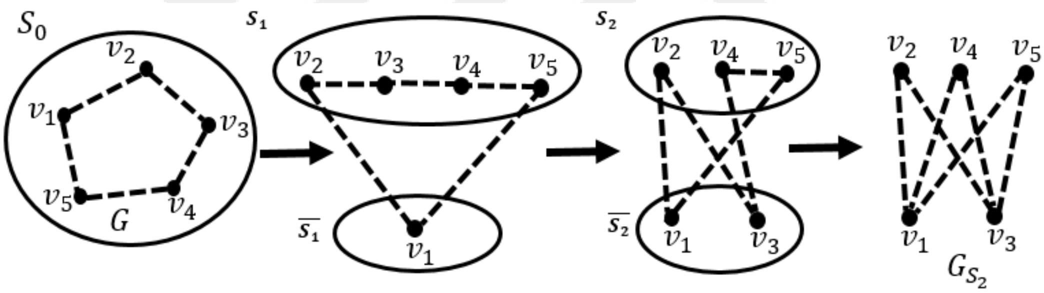

Figure 3.3 : G is a complete signed graph on 5 vertices where the missing edges are all positive edges. It shows the graphs before and after balancing.

Example 3.3: Let G be the complete signed graph in Fig. 3.3, where all the missing edges in the figure are positive.

Initially, $\partial(v)=0$ for all $v\in V$. We notice that the edge v_1v_3 is a good edge between two vertices of stability degree 0. Therefore according to the algorithm, we move v_1 to create the partition $(S_1,\overline{S_1})$, where:
$\partial_{S_1}(v_1)=0$, $\partial_{S_1}(v_2)=\partial_{S_1}(v_5)=2$, $\partial_{S_1}(v_3)=\partial_{S_1}(v_4)=-2$ and, therefore, the $\min_{w\in V(G)} \partial_{S_1}(w)=-2$ for v_3 and v_4. Thus, we move v_3 and create the partition $(S_2,\overline{S_2})$.
In this partition $\partial_{S_2}(v_1)=\partial_{S_2}(v_3)=2$, $\partial_{S_2}(v_2)=4$, $\partial_{S_2}(v_4)=\partial_{S_2}(v_5)=0$

Since $\min_{w\in V(G)} \partial_{S_2}(w)=0$ for the vertices v_4 and v_5 and because the edge v_4v_5 is a bad edge, the algorithm terminates by switching the signs of the bad edge, v_1v_4, v_3v_5 and

v_4v_5. This yields the balanced graph G_{S_2} as shown in Fig. 3.3. Again, the partition $(S_2, \overline{S_2})$ is optimal.

Unfortunately as we will see in Chapter 4, the algorithm does not always yield an optimal partition.

4. PROPERTIES OF THE ALGORITHM AND OPTIMIZATION

In this chapter, we examine some properties of the algorithm and use the algorithm to prove a generalization of the Structure Theorem.

4.1 Results

Lemma 2 *If v is the vertex that moves at the i^{th} iteration of the algorithm, then $\partial_{S_i}(v) = -\partial_{S_{i-1}}(v)$ and $\mathscr{D}(S_i) = \mathscr{D}(S_{i-1}) - 4\partial_{S_{i-1}}(v)$.*

Proof. Since we remove v from the part containing it and add it to the opposite part, each of its good edges become bad edges, and vice versa. As a result, $\partial_{S_i}(v) = -\partial_{S_{i-1}}(v)$. In addition, for each neighbor x of v, the contribution made by vx to $\partial_{S_{i-1}}(x)$, is negated when we calculate $\partial_{S_i}(x)$. Consequently, $\partial_{S_i}(x) = \partial_{S_{i-1}}(x) + 2$ for each bad neighbor x of v in the partition $(S_{i-1}, \overline{S_{i-1}})$, and $\partial_{S_i}(y) = \partial_{S_{i-1}}(y) - 2$ for each good neighbor y of v in the partition $(S_{i-1}, \overline{S_{i-1}})$. In summary, for each good edge incident to v, there is a contribution of -2, and for each bad edge there is a contribution of $+2$ to $\mathscr{D}(S_i)$. Since the difference between the number of good and bad neighbors of v is $\partial_{S_{i-1}}(v)$, the contribution of every vertex other than v to $\mathscr{D}(S_i)$ is $-2\partial_{S_{i-1}}(v)$. The contribution of v to $\mathscr{D}(S_i)$ is also $-2\partial_{S_{i-1}}(v)$. Consequently, $\mathscr{D}(S_i) = \mathscr{D}(S_{i-1}) - 4\partial_{S_{i-1}}(v)$. □

Note that, $\mathscr{D}(S_i)$ increases at each iteration except when $\min_{w \in V(G)} \partial_{S_{i-1}}(w) = 0$. In that case, either the algorithm terminates, or $\mathscr{D}(S_i)$ increases again in the next iteration. By Claim 1, $\mathscr{D}(S_i) \leq 2|E(G)|$, and therefore, the algorithm will terminate after a finite number of steps. Once the algorithm terminates, $\partial_{S_k}(v) \geq 0$ for all $v \in V(G)$, which implies that $\mathscr{D}(S_k) \geq 0$. As a result, using Lemma 1 we get:

$$l(G) \leq \frac{2|E(G)| - \mathscr{D}(S_k)}{4} \leq \frac{2|E(G)| - 0}{4} = \frac{|E(G)|}{2}.$$

Therefore to construct the balanced graph G_{S_k}, the algorithm would negate at most half the edges of $E(G)$. Presently we will slightly improve this bound in Theorem 2.

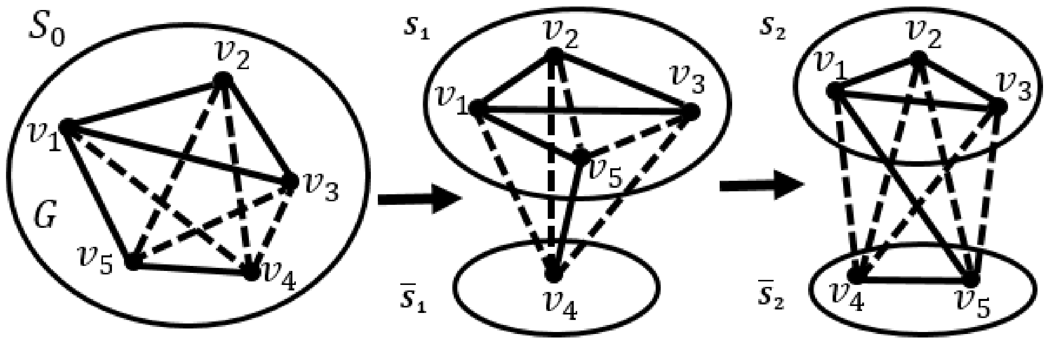

Figure 4.1 : G is a signed graph of 5 vertices. The partition $(S_2,\overline{S}_2)$ is the optimal partition for G.

For the partition $(S_1,\overline{S_1})$ of Fig. 4.1, we have $\partial_{S_1}(v_1) = 4$, $\partial_{S_1}(v_2) = 2$, $\partial_{S_1}(v_3) = 2$, $\partial_{S_1}(v_4) = 2$, $\partial_{S_1}(v_5) = -2$, and $\mathscr{D}(S_1) = 8$. After we move v_5 to create the partition $(S_2,\overline{S_2})$, we obtain $\partial_{S_2}(v_1) = 2$, $\partial_{S_2}(v_2) = 4$, $\partial_{S_2}(v_3) = 4$, $\partial_{S_2}(v_4) = 4$, $\partial_{S_2}(v_5) = 2$, and $\mathscr{D}(S_2) = 16$.

If we compare the stability degree of v_5 before and after the move, we notice that $\partial_{S_2}(v_5) = 2 = -\partial_{S_1}(v_5)$. After the move, the stability degree of v_1 which was the only good neighbor of v_5 in $(S_1,\overline{S_1})$ decreased by 2, and the stability degrees of the vertices which were bad neighbors of v_5 in $(S_1,\overline{S_1})$ increased by 2. Note specifically that although the edge v_4v_5 is positive, v_4 is a bad neighbor of v_5 in $(S_1,\overline{S_1})$. In addition, $\mathscr{D}(S_2) = 16 = \mathscr{D}(S_1) - 4\partial(v_5) = 8 - 4(-2)$.

We next examine the case of complete signed graphs. We will show that the maximum value of $l(G)$ is attained for a complete signed graphs K_n^- of n vertices, whose underlying graph is K_n, and whose edges all have negative signs.

Lemma 3 *$l(K_n^-) = \lceil \frac{n^2-2n}{4} \rceil$. Furthermore, if $(S,\overline{S})$ is an optimal partition of K_n^-, then $\mathscr{D}(S) = n$ if n is even and $\mathscr{D}(S) = n-1$ if n is odd.*

Proof. A partition $(S,\overline{S})$ of K_n^- will be optimal if it contains the maximum number of edges between the two parts. As a result $|S| = \lceil n/2 \rceil$ and $|\overline{S}| = \lfloor n/2 \rfloor$, or vice versa. Since only the edges inside each part have to be negated we have, $l(K_n^-) = 2\binom{n/2}{2}$ if n is even and $l(K_n^-) = \binom{\lceil n/2 \rceil}{2} + \binom{\lfloor n/2 \rfloor}{2}$ if n is odd, as desired.

To proof the second statement we use Lemma 1, If n is even, $l(K_n^-) = \lceil \frac{n^2-2n}{4} \rceil = \frac{n^2-2n}{4}$. Therefore, for the optimal partition $(S,\overline{S})$, since $l(S,\overline{S}) = l(K_n^-) = \frac{2|E(K_n^-)|-\mathscr{D}(S)}{4}$. Equivalently we have, $\mathscr{D}(S) = 2|E(K_n^-)| - 4.l(S,\overline{S})$, where $|E(K_n^-)| = \frac{n(n-1)}{2}$. Hence,

$$\mathscr{D}(S) = 2|E(K_n^-)| - 4.l(S,\overline{S}) = 2.\frac{n(n-1)}{2} - 4.\frac{n^2-2n}{4} = n^2 - n - n^2 + 2n = n.$$

If n is odd, $l(K_n^-) = \lceil \frac{n^2-2n}{4} \rceil = \frac{n^2-2n}{4} + \frac{1}{4}$. Therefore, for the optimal partition $(S,\overline{S})$, $\mathscr{D}(S) = 2|E(K_n^-)| - 4.l(S,\overline{S})$, where $|E(K_n^-)| = \frac{n(n-1)}{2}$. Hence,
$\mathscr{D}(S) = 2|E(K_n^-)| - 4.l(S,\overline{S}) = 2.\frac{n(n-1)}{2} - 4.\frac{n^2-2n+1}{4} = n^2 - n - n^2 + 2n - 1 = n - 1$.
□

One might suspect that among all signed graphs min $D(G)$ is achieved by K_n^-. This is true for complete signed graphs as we prove in Theorem 2, but false in general as can be seen in Fig. 4.2.

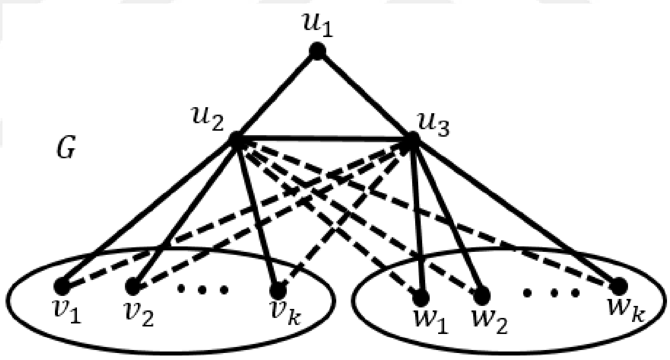

Figure 4.2 : G is a signed graph with $n = 2k+3$ vertices where $\mathscr{D}(G) = 6 < 2k+2 = n-1$, for $k \geq 3$.

Theorem 2 *For any signed graph G without isolated vertices $l(G) < \frac{|E|}{2}$. Additionally, if G is a complete signed graph on n vertices and $(S,\overline{S})$ is the partition of $V(G)$ identified by our algorithm, then $l(G) \leq l(S,\overline{S}) \leq l(K_n^-)$.*

Proof. We can assume G is connected because the line index of a graph is the sum of the line indices of its components. Let $(S,\overline{S})$ be the partition of $V(G)$ identified by our algorithm. If G does not have a vertex v with $\partial_S(v) = 0$, then $\mathscr{D}(S) \geq n$ and by Lemma 1, $l(G) \leq \frac{|E|}{2} - \frac{n}{4}$. So assume there is a vertex v such that $\partial_S(v) = 0$. Note that $\partial_S(v) = 0$ if and only if v has an equal number of good and bad neighbors.

Since G is connected and not an isolated vertex, $d(v) \geq 2$ and each of its $d(v)/2$ good neighbors must have stability degree at least 2, otherwise the algorithm would not have terminated. Consequently, $\mathscr{D}(S) \geq d(v)$ and $l(G) \leq \frac{|E|}{2} - \frac{d(v)}{4}$.

If G is a complete signed graph and n is even, then $d(v)$ is odd for all vertices in $V(G)$. Therefore, $\partial_S(v) \geq 1$ for all $v \in V(G)$ and $\mathscr{D}(S) \geq n$, which yields $l(S,\overline{S}) \leq l(K_n^-)$. On the other hand if n is odd and there exist a vertex v with $\partial_S(v) = 0$, then each of its $\frac{n-1}{2}$ good neighbors must have stability degree at least 2 so, $\mathscr{D}(S) \geq n-1$ and $l(S,\overline{S}) \leq l(K_n^-)$. □

A bipartition $(S,\overline{S})$ of a signed graph G is called *semi-optimal* if $\partial_S(v) \geq 0$ for each vertex v, and if there is no good edge xy such that $\partial_S(x) = 0$ and $\partial_S(y) \leq 1$.

Our algorithm finds a semi-optimal partition of any signed graph. In terms of the actual social network represented by the signed graph, the semi-optimal partition corresponds to the two antagonistic coalitions that will be formed if each individual tries to maximize his/her happiness or comfort. Note that as seen in Fig. 4.3, the algorithm does not necessarily find the optimal partition of a signed graph even if it is complete. To find an optimal partition, or in other words, the coalitions that will impose the least amount of changes to the relationships between members of a social network, one needs to consider the happiness of larger groups not just individuals.

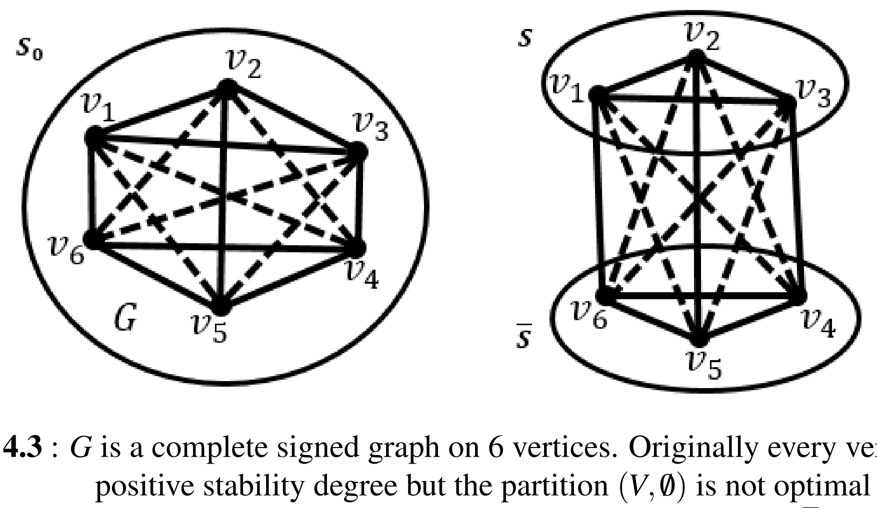

Figure 4.3 : G is a complete signed graph on 6 vertices. Originally every vertex has positive stability degree but the partition $(V,\emptyset)$ is not optimal but is semi-optimal. The optimal partition is attained by $(S,\overline{S})$, where $S = \{v_1, v_2, v_3\}$.

However, by using the algorithm we can show that the problem of determining the line index of signed graphs, in which each vertex is incident to at least as many positive edges as negative edges, is still NP-Complete. We make use of the notion of switching for signed graphs developed by Zaslavsky [14].

Theorem 3 *Determining the line index of signed graphs, in which each vertex is incident to at least as many positive edges as negative edges, is NP-Complete.*

Proof. Let G be a signed graph and $(S,\overline{S})$ be the partition of $V(G)$ identified by our algorithm. Since $\partial_S(v) \geq 0$ for all $v \in V(G)$, each vertex is incident to at least as many good edges as bad edges. Therefore, in the unified semi-optimal representation H of $(S,\overline{S})$ each vertex is incident to at least as many positive edges as negative edges. If $l(H)$ could be computed in polynomial time, so could $l(G)$ since $l(G) = l(H)$. □

We can also investigate the optimality of a bipartition of a signed graph G by considering its unified representation. Given a signed graph G and a subset S of $V(G)$, the *stability degree of* S, $\partial(S)$, is the sum of the weights of the edges which have one endpoint in S and one endpoint in $\overline{S}$. Clearly, $\partial(S) = \partial(\overline{S})$. Recall that given a graph G and a subset S of its vertex set, the induced graph on S is $G[S]$. This leads us to our next results:

Lemma 4 *Let $G = (V,E,\sigma)$ be a signed graph and $S \subseteq V$. Then,*

$$\mathscr{D}(G) = \mathscr{D}(G[S]) + 2\partial(S) + \mathscr{D}(G[\overline{S}]).$$

Proof.

$$\mathscr{D}(G) = \sum_{v\in V} \partial(v) = \sum_{v\in S} \partial(v) + \sum_{w\in \overline{S}} \partial(w)$$
$$= \mathscr{D}(G[S]) + \partial(S) + \mathscr{D}(G[\overline{S}]) + \partial(\overline{S})$$
$$= \mathscr{D}(G[S]) + 2\partial(S) + \mathscr{D}(G[\overline{S}]).$$ □

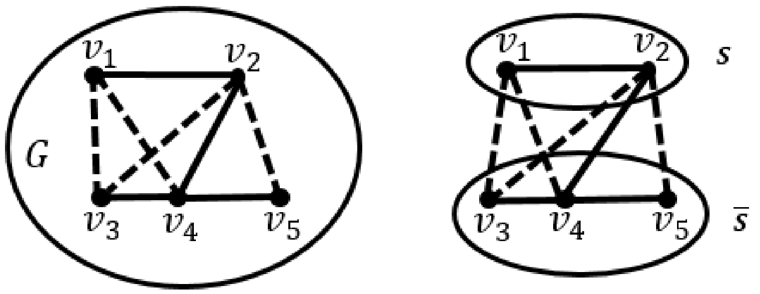

Figure 4.4 : *G* is a signed graph of 5 vertices and 8 edges.

For the signed graph G in Fig. 4.4, $\mathscr{D}(G) = 0$. Let $S = \{v_1, v_2\}$. Then, $\mathscr{D}(G[S]) = 2$ and $\partial(S) = -3$ and $\mathscr{D}(G[\overline{S}]) = 4$. Therefore by Lemma 4,

$$0 = \mathscr{D}(G) = \mathscr{D}(G[S]) + 2\partial(S) + \mathscr{D}(G[\overline{S}]) = 2 + 2(-3) + 4.$$

Lemma 5 *If $G = (V, E, \sigma)$ is a signed graph and $(S, \overline{S})$ is a partition of its vertex set, then $\mathscr{D}(S) = \mathscr{D}(G) - 4\partial(S)$.*

Proof. Since the good/bad edges between S and $\overline{S}$ in G (i.e in the partition $(V(G), \emptyset)$) are bad/good edges in the partition $(S, \overline{S})$, the contribution of the edges between S and $\overline{S}$ is negated when we compute $\mathscr{D}(S)$. The contribution of the edges in $G[S]$ and $G[\overline{S}]$ is the same in both $\mathscr{D}(G)$ and $\mathscr{D}(S)$. As a result, $\mathscr{D}(S) = \mathscr{D}(G[S]) - 2\partial(S) + \mathscr{D}(G[\overline{S}]) = \mathscr{D}(G) - 4\partial(S)$. □

Note that Lemma 5 yields a generalization of Lemma 2. For the example in Fig. 4.4, we had $\mathscr{D}(G) = 0$ and $\partial(S) = -3$. It can be easily verified that $\mathscr{D}(S) = 0 - 4(-3) = 12$.

Theorem 4 *Let $G = (V, E, \sigma)$ be a signed graph and $(S, \overline{S})$ be a partition of its vertex set and H be the unified representation of $(S, \overline{S})$. The following conditions are equivalent:*

1. $(S, \overline{S})$ is an optimal partition of G.

2. H cannot be split.

3. For any $A \subseteq V(H)$, $\partial(A) \geq 0$.

4. For any $A \subseteq V(H)$, $\mathscr{D}(H[A]) \leq \sum_{v \in A} \partial(v)$.

Proof. ($1 \Leftrightarrow 2$) H splits if and only if $(V,\emptyset)$ is not an optimal partition of V. As a result, $l(G) = l(H) < l_H(V,\emptyset) = l_G(S,\overline{S})$ if and only if $(S,\overline{S})$ is not an optimal partition of G.

($2 \Leftrightarrow 3$) Given $A \subseteq V(H)$, by Lemma 5, we know that $\mathscr{D}(A) = \mathscr{D}(H) - 4\partial(A)$. Therefore, H cannot be split if and only if $\partial(A) \geq 0$.

($3 \Leftrightarrow 4$) By Lemma 4, for any $A \subseteq V(H)$, $\sum_{v \in A} \partial(v) = \mathscr{D}(H[A]) + \partial(A)$. Consequently, $\partial(A) \geq 0$ if and only if $\mathscr{D}(H[A]) \leq \sum_{v \in A} \partial(v)$. □

In the complete signed graph G in Fig. 4.3, since the stability degree of each vertex in the partition $(V(G),\emptyset)$ is initially positive, our algorithm will immediately terminate. However as mentioned before, the partition $(V,\emptyset)$ is not optimal. If we consider $S = \{v_1, v_2, v_3\}$, we see that $\partial(S) = -3$. Therefore by Theorem 4, the partition $(S,\overline{S})$ is better, and indeed it is actually optimal.

In terms of actual social networks, Theorem 4 predicts that if members of a group are more loyal to their group than to the society as a whole, then polarization of the society may be unavoidable. Notice that although our algorithm may not be able to find an optimal partition of the vertex set of a given signed graph, Theorem 4 can be used to show that a partition is not optimal. If a subset A failing conditions 3 or 4 is identified, then $(A,\overline{A})$ is a better semi-optimal partition.

5. APPLICATION OF THE ALGORITHM TO REAL SOCIAL NETWORKS

In this chapter we consider two real life social networks and compare the predictions of our algorithm with the final state of those social networks. The first social network is the network formed by the major state participants of the Syrian conflict. Güner and Koç in [15] describe in detail the evolution of the social network which can be summarized in Fig. 5.1 below. The second social network consists of members of a karate club as discussed in [1] by Zachary.

5.1 The Syrian Conflict

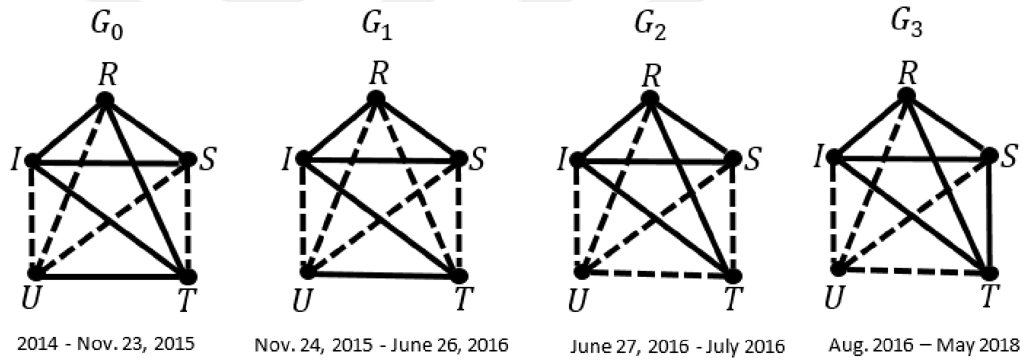

Figure 5.1 : These four graphs are complete signed graphs showing the evolution of the relationships between the five countries Turkey, USA, Russia, Iran and Syria, represented by T, U, R, I and S respectively, from 2014 to May 2018.

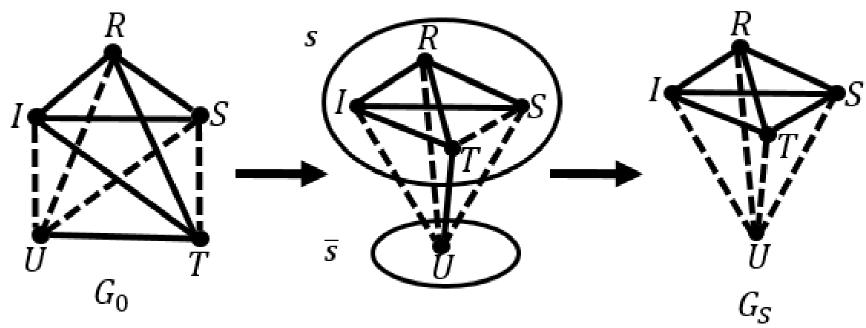

Figure 5.2 : The original graph G_0 of Fig. 5.1, and the partition $(S,\overline{S})$ identified by the algorithm, and the balanced graph G_S predicted by the algorithm.

As seen in Fig. 5.2, starting from the original network G_0, the algorithm of this thesis would yield the partitions $(S,\overline{S})$. Hence, the prediction of the algorithm would be that the relationship between US and Turkey would change from friendly to hostile, while the relationship between Turkey and Syria would change from hostile to friendly. These predictions correspond exactly to the final state G_3 of the network in Fig. 5.1. One could make the observation that the algorithm fails to predict the evolution of the network from state G_0 to G_1. However, since $\partial_S(T) = 0$ in Fig. 5.2, the easiest way for any party wishing to externally influence the evolution of the network would be to try to change the relationships between Turkey and the other countries. Indeed, Turkey officially blames members of the Fethullah Gulen movement for the shooting down of the Russian warplane. This incident caused the network to change to state G_1. It is worthwhile to note that, since in the original graph G_0 the number of positive edges is more than the number of negative edges, the algorithm in Marvel et al. [6] would predict the final state of the network to only have positive edges. Clearly, this is not the case in the Syrian Conflict.

5.2 The Karate Club

In [1], Zachary analyses the social network formed by members of a university-based Karate club. To protect the anonymity of the members of the club, neither the club nor the members are identified by name in [1]. Zachary collected data about the network from 1970 to 1972. During this period the Karate Club split into two clubs. The split originated from a disagreement between the club president and the instructor over the cost of the karate lessons. Zachary recorded which members consistently socialized outside of the classes and the club meetings before the split. The graph in Fig. 5.3 summarizes his data. Zachary used the Ford-Fulkerson algorithm to predict the members of each of the two karate clubs after the split, and compared them with the actual members of the two clubs. The Ford-Fulkerson algorithm predicted the correct memberships with 97% accuracy.

In this section we compare the prediction of our algorithm about the memberships of the clubs with the actual memberships after the split. This case study also demonstrates one of the strengths of our algorithm compared to the algorithms of Antal et al. [5] and Marvel et al. [6], which is the fact that our algorithm still works with any real

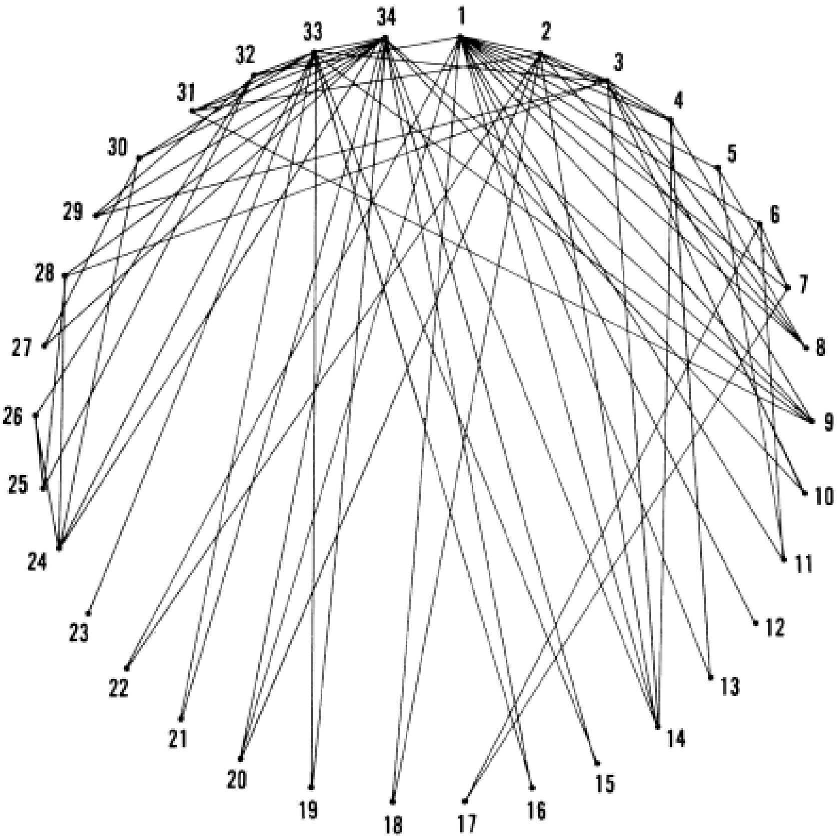

Figure 5.3 : This signed graph shows the relationships between the 34 members in the Zachary Karate Club given in [1] where the edges drawn represents the positive relation existing between them. Vertex 1 represents the instructor, while vertex 34 represents the president.

number weights assigned to the edges of an underlying signed graph signifying the intensity of the relationships. We modify the graph in Fig. 5.3 slightly in order to run our algorithm. The modification is necessitated by the fact that Zachary only recorded positive relationships, but not negative ones. We make two assumptions about the data:

1. The relationship between the president and the instructor is very negative.
2. The relationships between the president and his friends are very positive. Similarly, the relationships between the instructor and his friends are also very positive.

Assumption 1. is reasonable since Zachary goes into sufficient detail to describe the conflict between the instructor and the president. As they are the two polarizing figures, it is also natural to assume that the relationships involving the president or the instructor and other members are more intense than the relationships not involving the two, which justifies assumption 2. Fig. 5.4 contains the adjacency matrix of the modified graph.

```
0 8 8 8 8 8 8 8 8 0 8 8 8 8 0 0 0 8 0 8 0 8 0 0 0 0 0 0 0 0 0 8 0 -130
8 0 1 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0
8 1 0 1 0 0 0 1 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0
8 1 1 0 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 8
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 8
8 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 8
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 8
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 8
0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 8
8 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 8
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 8
8 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 1 8
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 8
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 8
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 8
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 8
0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 8
8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 1 8
0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 1 0 1 0 1 1 0 0 0 0 0 1 1 1 0 8
-130 0 0 0 0 0 0 0 8 8 0 0 0 8 8 8 0 0 8 8 8 0 0 8 0 0 8 8 8 8 8 8 8 0
```

Figure 5.4 : The adjacency matrix of the modified signed graph whose original graph is given in Fig. 5.3.

After running our algorithm we get the partition $(S, \overline{S})$ where,

$S = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_{11}, v_{12}, v_{13}, v_{14}, v_{17}, v_{18}, v_{20}, v_{22}\}$, and v_i corresponds to person i. Our prediction corresponds with the actual memberships with also 97% accuracy. In fact Zachary's and our predictions are exactly the same. The only misplaced member in both predictions is person 9. Zachary explained that person 9 switched alliances because unless he was a member of the club formed by the instructor, he risked loosing his black belt.

It is worthwhile noting that for both social networks in this chapter our algorithm mis-predicted the actual outcome when there was an extraordinary event/hidden information not captured by the initial data.

6. CONCLUSIONS AND RECOMMENDATIONS

The simple greedy algorithm discussed in this thesis has several advantages compared to other recent algorithms that balance signed graphs. Although much like Antal et al. [5] and other graph theoretic algorithms, the algorithm cannot find an optimal partition of the signed network, it nevertheless does balance any signed graph. Compared to the algorithm of Marvel et al. [6], our algorithm also forces much fewer relationship changes on the social network because a signed graph with mostly positive edges might still contain vertices that are incident to many negative edges. Therefore, the negation of all negative edges for such graphs might not be reasonable. In fact, the ratio of the number of positive to negative edges can be arbitrarily large and the the signed graph could still split to form coalitions predicted by our algorithm. See graph G from Fig. 6.1 for an example.

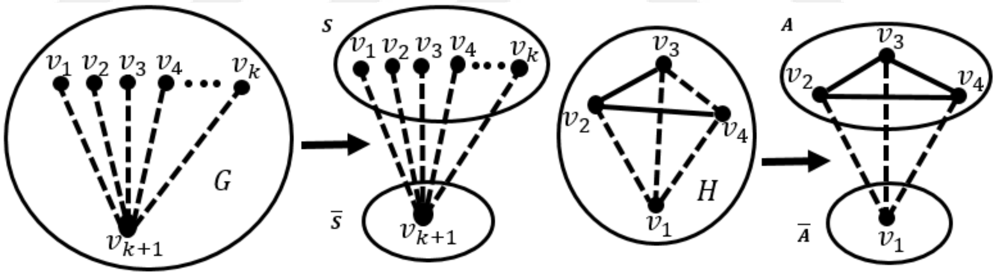

Figure 6.1 : G is a complete signed graph on $k+1$ vertices. The missing edges have positive sign. Note that G has $\binom{k}{2} = \frac{k^2-k}{2}$ positive edges and k negative edges, but the graph can be split by moving v_{k+1}. The signed graph H has optimal partition as shown; where the two partite sets do not have the same size.

In the initial state G of the network in Fig. 6.1, most of the relationships are friendly. Consequently, the algorithm of Marvel et al. [6] would predict a final network state where all parties are friendly to each other. It is clear that this prediction is not reflective of the evolution of an actual social network. One could easily provide an infinite family of signed graphs in which the density of positive edges is higher than the density of negative edges, but whose optimal partitions are clearly not $(V,\emptyset)$ as would be predicted by [6]. The same statement is also true of graphs where the density of negative edges is higher than the density of positive edges. Such a graph might have an optimal partition where the two partite sets do not have the same size, as would be predicted by [6].

We also wonder whether the assumption that all networks evolve towards a balanced state that achieves $l(G)$ is correct. As seen in Fig. 6.2, there is an infinite family of signed graphs, where each signed graph can be split. For each signed graph belonging to the family, every vertex is incident to $3k$ positive edges and $k+1$ negative edges. Therefore, the ratio of the number of positive edges to negative edges approaches 3 as k increases. Note also that initially $\partial(v) = 2k-1$ for each vertex. As a result, members of such a network are relatively happy with its current state and have no individual reasons to split the network and form opposing coalitions.

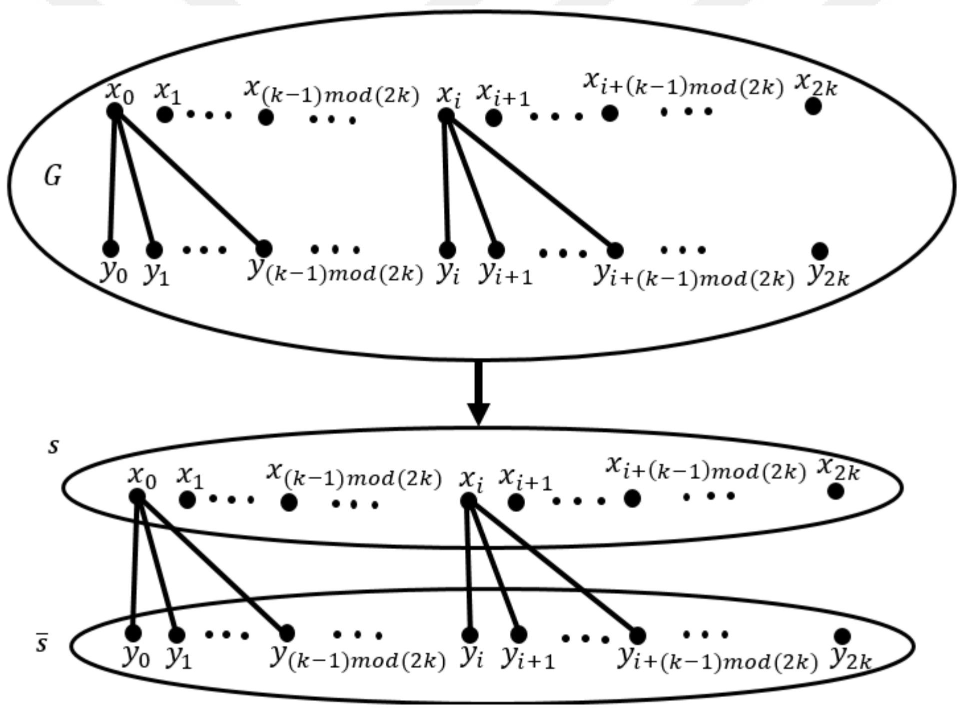

Figure 6.2 : G is a complete signed graph on $4k+2$ vertices. Solid edges have positive sign, while the missing edges have negative sign. Note that in G, $\partial(v) = 2k-1$ for all $v \in V$. However, the optimal partition is $(S,\overline{S})$, where $\partial_S(v) = 2k+1$ for all $v \in V$.

In fact research on large social networks support the idea that Heider's balance theory does not hold for such networks [16], [9]. Long unbalanced cycles do not seem to be as strong a factor as short unbalanced cycles on large networks. As a result, new approaches have been considered by weighing cycles according to their lengths. This and similar issues plague all algorithms that have a global approach to balance. It is not surprising that large social networks do not achieve balance in Heider's sense because, as computing $l(G)$ is very hard in general, it would be unreasonable to expect individual members to make choices that would be equivalent to calculating $l(G)$, especially when it is known that in real life networks individuals do not have information about the structure of the whole network. Since the algorithm in this thesis is a local one, it is largely unaffected by such problems.

Also for large networks, the assumption that all relationships are symmetric and that all vertices behave the same way does not hold. Powerful members of a network have a much greater effect on the evolution of a social network as was the case with the Karate Club network in the previous chapter. These concerns necessitate the need to consider directed signed graphs and differentiate between members when dealing with large networks. Unfortunately, lacking data on large social networks, we have not developed an algorithm suitable for predictions about large networks. This is an area which we hope to develop in future research.

REFERENCES

[1] **Zachary, W.** (1977). An information flow model for conflict and fission in small groups, *Journal of Anthropological Research*, *33*, 452–473.

[2] **Heider, F.** (1946). Attitudes and Cognitive Organization, *The Journal of Psychology*, *21*(1), 107–112, `http://dx.doi.org/10.1080/00223980.1946.9917275`, pMID: 21010780, `http://dx.doi.org/10.1080/00223980.1946.9917275`.

[3] **Cartwright, D. and Harary, F.** (1956). Structural balance: a generalization of Heider's theory., *Psychological Review*, *63*(5), 277.

[4] **Barahona, F.** (1982). On the computational complexity of Ising spin glass models, *Journal of Physics A: Mathematical and General*, *15*(10), 3241.

[5] **Antal, T.**, **Krapivsky, P.L. and Redner, S.** (2005). Dynamics of social balance on networks, *Phys. Rev. E (3)*, *72*(3), 036121, 10, `http://dx.doi.org/10.1103/PhysRevE.72.036121`.

[6] **Marvel, S.A.**, **Kleinberg, J.**, **Kleinberg, R.D. and Strogatz, S.H.** (2011). Continuous-time model of structural balance, *Proceedings of the National Academy of Sciences*, *108*(5), 1771–1776.

[7] **Xu, S.** (1998). The line index and minimum cut of weighted graphs, *European Journal of Operational Research*, *109*(3), 672–685.

[8] **West, D.B.** (2001). *Introduction to Graph Theory*, Prentice Hall, Upper Saddle River, N.J., second edition.

[9] **Leskovec, J.**, **Huttenlocher, D. and Kleinberg, J.** (2010). Signed networks in social media, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp.1361–1370.

[10] **Harary, F. and Kabell, J.A.** (1980). A simple algorithm to detect balance in signed graphs, *Mathematical Social Sciences*, *1*(1), 131–136.

[11] **Harary, F.** (1959). On the measurement of structural balance, *Behavioral Science*, *4*(4), 316–323, `http://dx.doi.org/10.1002/bs.3830040405`.

[12] **Loukakis, E.** (2003). A dynamic programming algorithm to test a signed graph for balance, *International Journal of Computer Mathematics*, *80*(4), 499–507.

[13] **Marvel, S.A.**, **Strogatz, S.H. and Kleinberg, J.M.** (2009). Energy landscape of social balance, *Physical Review Letters*, *103*(19), 198701.

[14] **Zaslavsky, T.** (1982). Signed graphs, *Discrete Applied Mathematics*, *4*(1), 47–74.

[15] **Güner, S.Ş. and Koç, D.E.** (2017). The Syrian Conflict: Driving Forces of Balances and Imbalances., *MERIA Journal*, *33*(3).

[16] **Estrada, E. and Benzi, M.** (2014). Walk-based measure of balance in signed networks: Detecting lack of balance in social networks, *Physical Review E*, *90*(4), 042802.

CURRICULUM VITAE

Name Surname: Araniyos Terefe WELDEGEBRIEL

Place and Date of Birth: Addis Ababa (Ethiopia), 29.11.1979

E-Mail: araniye@yahoo.com

EDUCATION:

- **B.Sc.:** October 1997-July 2001, Debub University in Ethiopia, Teacher Education Faculty, Department of Mathematics.
- **M.Sc.:** October 2004-August 2006, Addis Ababa University in Ethiopia, Science Faculty, Mathematics Department.

PROFESSIONAL EXPERIENCE AND REWARDS:

- August 2001-July 2003: Instructed the subject Mathematics and heading the department of Mathematics at Aysaeita Preparatory and Technical School for two years in Afar region, Ethiopia.
- 2003: Earned certificate for giving training about ”Alternate Basic Education” at Aysaeita, Ethiopia.
- September 2003-September 2004: Instructed the subject Mathematics at Alpha University College in the secondary school for one year and one month in Addis Ababa, Ethiopia.
- 2005: Earned certificate for preparing student text book and teacher guide on Mathematics for grade three students for Afar Region, Ethiopia.
- September 2006- October 2007: Served as a Branch Manager of Alpha University College at Debre Birhan Branch office for one year and two month, North Showa in Ethiopia.
- Sep.2007-August 2008: Earned certificate for the completion of Higher Diploma Training given for one year from Arba Mich University, South Ethiopia.
- October 2008: Earned certificate for the completion of Pedagogical Training from Arba Minch University, South Ethiopia.
- March 2009-July 2012: Lead Higher Diploma Program in Arba Minch University, South Ethiopia.

- October 2007- November 2011: Instructed different courses like Optimization Theory I, Advanced Mathematics for Engineers, Linear Optimization, Applied Mathematics I and II, Calculus I, Abstract Algebra, Number Theory, and etc at Arba Minch University, South Ethiopia.

PRESENTATIONS ON THE THESIS:

- **Weldegebriel, A. T.**, Stodolsky, B. Y., 2017. Optimization of complete social networks. *INTERNATIONAL CONFERENCE ON MATHEMATICS AND ENGINEERING (ICOME-2017),* May 10-12, 2017, Istanbul, Turkey

PUBLICATIONS ON THE THESIS:

- **Weldegebriel, A. T.**, Stodolsky, B. Y., accepted in Oct. 2018. On Optimization of Complete Social Networks. *Lobachevskii Journal of Mathematics,*.

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- **Araniyos Terefe Weldegebriel**, August 2011. Concepts of Linear Optimization with Application, *Lap Lambert Academic publishing in Germany,* ISBN-13 978-3-843323-802.