

THE DYNAMIC COMPETITION HYPOTHESIS

by

Rehan Taj Malik

Bachelor in Science, University of Sindh, Pakistan, 1991

Master of Science in Mathematics, QAU, Pakistan, 1995

Master of Arts in Economics, University of Sindh, Pakistan, 1996

Bachelor in Education, Preston University, Pakistan, 2015

A thesis presented to Ryerson University

in partial fulfillment of the
requirements for the degree of

Master of Science

in the program of

Applied Mathematics

Toronto, Ontario, Canada, 2018

© Rehan Malik, 2018

**AUTHOR'S DECLARATION FOR
ELECTRONIC SUBMISSION OF A THESIS**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

The Dynamic Competition Hypothesis

Master of Science, 2018

Rehan Taj Malik

Applied Mathematics, Ryerson University

Abstract

We study competitive and negative interactions in real world social network in which nodes represent agents and edges appear over discrete time step. We consider directed social network of competing agents that evolve dynamically over time, where directed edges represent some kind of negative relationships between the agents in the social network.

We present a novel hypothesis to identify the alliances and leaders within the the dynamic competition networks. We verify our hypothesis by using historical voting data of the social game shows Survivor and Big Brother.

Acknowledgements

I would first like to thank my supervisor Dr. Anthony Bonato for giving me the opportunity to write a fulfilling thesis. Throughout my thesis-writing period, he provided encouragement, sound advice, excellent teaching, and lots of good ideas.

It is also a pleasure to thank Dr. Dejan Delić and Dr. Peter Danziger for serving as my committee members and for their valuable comments and suggestions.

Lastly and most importantly, I would like to thank my wife for her support and understanding throughout my graduate studies. I dedicate this thesis to my parents, Malik Taj Muhammad and Shamim Taj Malik. They raised me, supported me, taught me, and loved me. I could not be successful without their prayers.

TABLE OF CONTENTS

Declaration	ii
Abstract	iii
Acknowledgements	iv
List of Figures	vii
List of Tables	x
Chapter 1. Introduction	1
1.1. Motivation	1
1.2. Introduction to Graph Theory	3
1.3. Summary of the Thesis	13
Chapter 2. The Dynamic Competition Hypothesis	14
2.1. Structural Balance Theory	14
2.2. Centrality metrics in networks	20
2.3. Dynamic Competition Networks	28
2.4. The Hypothesis	30
Chapter 3. Data and Methods	33
3.1. Survivor	33

3.2.	Big Brother	35
3.3.	Social Network Analysis tools	36
3.4.	Borneo	37
3.5.	China	40
3.6.	Game Changers	42
3.7.	Heroes vs Healers vs Hustlers (HHH)	44
3.8.	Big Brother	47
3.9.	Alliances	48
Chapter 4.	Conclusion and Future Work	51
4.1.	Conclusion	51
4.2.	Future Work	52
	Bibliography	54

List of Figures

1.1	G has multiple edges and a loop and H is a simple graph.	3
1.2	An undirected graph $G = (V, E)$.	4
1.3	A path P_5 of order 5 and length 4.	5
1.4	A connected graph G and a disconnected graph H .	6
1.5	A connected graph G on five nodes.	7
1.6	A graph G , a subgraph H , an induced subgraph Q , and a spanning subgraph R .	8
1.7	Complete graphs K_4 and K_6 .	8
1.8	A cycle C_6 .	9
1.9	Isomorphic graphs.	9
1.10	Examples of Trees.	10
1.11	A directed graph G .	11
1.12	Independent set on 5 nodes.	12
1.13	$ED(P) = 0.4$, $ED(Q) = 1$, $ED(R) = 1.3$.	12
2.1	A signed graph.	15
2.2	X , Y , Z are mutual friends: balanced.	16

2.3	X and Y are friends with Z as a mutual enemy: balanced.	16
2.4	X is friends with Y and Z , but Y and Z are enemy to each other : unbalanced.	17
2.5	X , Y and Z are mutual enemies: unbalanced.	17
2.6	The labeled four-node complete graph on the left is balanced and on the right is unbalanced.	19
2.7	A directed graph.	21
2.8	A directed graph with nodes labeled with their in-degree.	22
2.9	A directed graph with nodes labeled with their out-degree.	23
2.10	Two graphs with nodes labeled with their betweenness centralities.	24
2.11	A tree with nodes labeled with their closeness centralities.	25
2.12	The node w is the common out-neighbor of u and v .	26
2.13	The number of common out-neighbor of u and v is given by $\text{CON}(u, v)=3$.	26
2.14	CON score of $u=\text{CON}(u)=5$.	27
2.15	In these networks we have that $\text{ED}(\text{R}) = 0.6$ and $\text{ED}(\text{S}) = 1.2$.	28
2.16	A heat map representation of dynamic competition networks according to the DCH, where nodes closer to the center have higher closeness and CON scores. Larger nodes have higher CON scores, lower in-degree, and higher out-degree.	

The induced subgraphs (which are the circles) correspond to alliances.	32
3.1 Survivor: Borneo.	39
3.2 Survivor: China.	42
3.3 Survivor: Game Changer.	44
3.4 Survivor: HHH.	46
3.5 Big Brother, Season 12.	48

List of Tables

1	Survivor: Borneo, Season 1.	39
2	Survivor: China, Season 15.	41
3	Survivor: Game Changer, Season 34.	43
4	Survivor: Heroes vs Healors vs Hustlers (HHH), Season 35.	46
5	Big Brother, Season 12.	47
6	Edge densities of the Alliances.	49

CHAPTER 1

Introduction

1.1. Motivation

We are connected to each other via our social interactions, either as individuals or in large groups. A *social network* is a collection of agents that are connected by some particular types of social interactions such as friendship, enemies, likes or dislikes. We can observe, analyse and measure social networks with the help of observable data. Social Network Analysis (SNA) considers networks, where nodes are individual agents in the network and edges are the relationship or links between the agents. For example, we denote edges as following in Twitter, while edges represent friendship in Facebook.

We may apply SNA [17] to analyse the networks in the field of economics, computer science, political science, communication studies, geography, social psychology and also in reality shows in television. In social networks, we often consider social interaction as a positive relationship such as friendship, likeness or some type of common interest between the individuals or groups. We may also see the negative social relationship in network, which is also play an important role to find the hidden interactions between the connections.

Bonato et al. [5], introduced the *Iterated Local Transitivity* (ILT) model for a social networks. The central idea of ILT model is *transitivity*: if x is a friend of y , and y is a friend of z , then x is a friend of z . On the other hand *Structural Balance Theory* [9, 13] gives the inverse idea: if x is an enemy of y , and y is an enemy of z , then x is a friend of z .

As evidence of negative social interactions, we may consider the enmity between the nations, rival gangs or competition between players in games. Another example is the stock market graph [4], where the nodes are stocks and two stocks are adjacent if they are negatively correlated.

Based on the competitive and negative interactions in a social network, we consider directed social networks of competing agents that evolve dynamically over time, where directed edges represent some kind of negative relationships between the agents in the network. As an example relevant to this thesis, a directed edge may be a vote of one contestant to other in the game show *Survivor*. One of our main contributions, as first described in Chapter 2, is the introduction of the Dynamic Competition Hypothesis (DCH) to identify the alliances and top players within the dynamic competition network. We verify our hypothesis by using the voting history data of social game shows *Survivor* and *Big Brother*.

1.2. Introduction to Graph Theory

In this section, we discuss some basic definitions from graph theory which we use throughout our thesis. A *graph* $G = (V, E)$ is a pair consisting of a *node set* $V = V(G)$, an *edge set* $E = E(G)$ consisting of pairs of nodes. Note that E is taken as a multiset, as its elements may occur more than once. We write uv if u and v form an edge, and say that u and v are *adjacent*.

A *loop* is an edge whose endpoints are equal. *Multiple edges* are edges having the same pair of endpoints. We restrict our attention to *simple graphs*; that is, graphs without loops and multiple edges. Further, we only consider finite graphs.



FIGURE 1.1. G has multiple edges and a loop and H is a simple graph.

The *degree* of a node v in G , written $\deg_G(v)$, is the number of neighbors of v in G . We will drop the subscript G if the graph is clear from context.

The set of all nodes adjacent to a node u , is called the *neighborhood* of a node $u \in V(G)$, denoted by $N(u)$. We denote $N[u]$ be the *closed neighborhood* of u , written as $N[u] = N(u) \cup \{u\}$. We denote by $|V(G)|$

the order of nodes V in G ; the total number of nodes in V in graph G . We denote by $|E(G)|$ the order of edges E in G ; the total number of edges in E in graph G . We note that

$$|E| \leq \binom{|V|}{2}.$$

The degree of the node $u \in V(G)$, denoted by $\deg(u)$ is equal to the number of edges incident on u . The node u is *leaf or pendant* if $\deg(u) = 1$. The node u is *isolated* if $\deg(u) = 0$.

Now we give an example of an undirected graph $G = (V, E)$, where the nodes $V = \{a, b, c, d\}$ and edges $E = \{ab, bc, bd, da\}$.

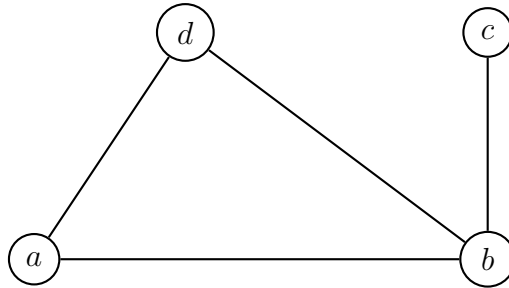


FIGURE 1.2. An undirected graph $G = (V, E)$.

In Figure 1.2, node a is adjacent to nodes d and b but it is not adjacent to c . The neighborhood of node a is the graph with two nodes d and b and one edge connecting b and d . The degree of nodes are $\deg(a)=\deg(d)=2$, $\deg(b)=3$, $\deg(c)=1$. We note that c is leaf.

The following is often referred to as the *First Theorem of Graph Theory* (which is included for completeness).

THEOREM 1. [5] *If G is a graph, then we have that*

$$\sum_{v \in V(G)} \deg(v) = 2|E(G)|.$$

PROOF. Let $S = \sum_{v \in V(G)} \deg(v)$. Notice that in S , we count each edge exactly twice. Thus, $S = \sum_{v \in V(G)} 2|E(G)|$. \square

A sequence of n nodes such that each node is adjacent to the next node in the sequence, is called the *path* of order n . We denote this graph by P_n and the *length* of the path is $n - 1$.

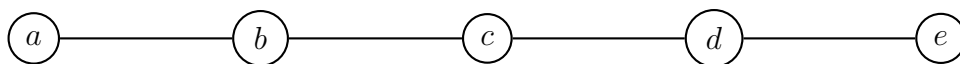


FIGURE 1.3. A path P_5 of order 5 and length 4.

A *walk* in a graph $G = (V, E)$ is an alternating sequence

$$W = \langle v_o, e_1, v_1, e_2, \dots, v_{n-1}, e_n, v_n \rangle$$

between the nodes and edges such that $e_i = (v_{i-1}, v_i)$, for all $v_i \in V(G)$ and $e_i \in E(G)$, $i = 1, 2, \dots, n$. The number of edge-steps in the walk called the *length* of the walk. A walk that begins and ends at the same node is said to be *closed walk*. A walk that begins and ends at different node called an *open walk*. A walk that does not pass through the same edges twice is called a *trail*.

A graph G is said to be *connected* if for each distinct pair of nodes $u, v \in V(G)$ there is a path from u to v in graph G ; otherwise the graph

G is disconnected. The *distance* from the nodes u to v is the length of the shortest path from u to v in a connected graph G , denoted by $d(u, v)$. If u and v are in different components, then G is said to be disconnected such that $d(u, v) = \infty$.

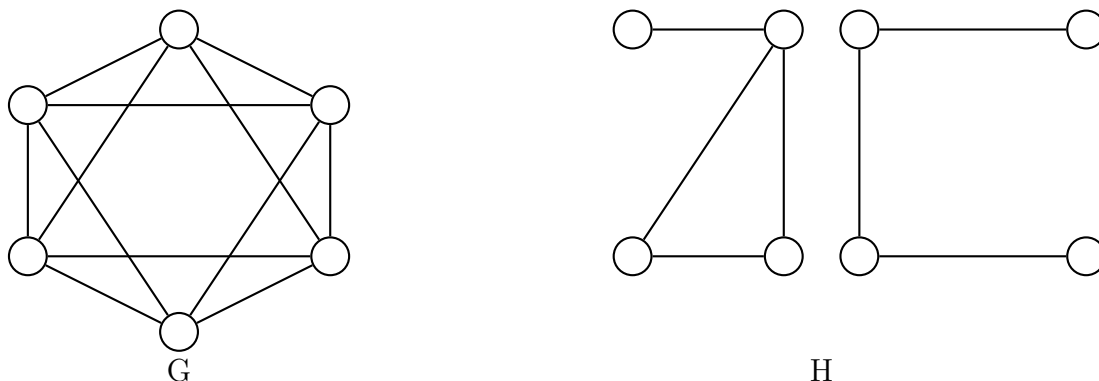


FIGURE 1.4. A connected graph G and a disconnected graph H .

The eccentricity of a node v , denoted by $e(v)$, of a connected graph $G = (V, E)$ is the maximum distance from node v to any other node $u \in V(G)$; that is,

$$e(v) = \max_{u \in V(G)} \{d(u, v)\}.$$

The maximum eccentricity in a graph G is the *diameter* of a graph G , denoted by $\text{diam}(G)$; that is,

$$\text{diam}(G) = \max_{v \in V(G)} \{e(v)\}.$$

The minimum eccentricity in the graph G is the *radius* of a graph G , denoted by $\text{rad}(G)$; that is,

$$\text{rad}(G) = \min_{v \in V(G)} \{e(v)\}.$$

A node with minimum eccentricity in a graph G is called a *central node* and we write

$$e(v) = \text{rad}(G).$$

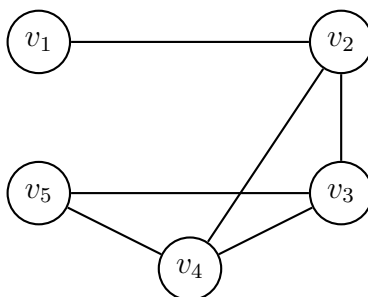


FIGURE 1.5. A connected graph G on five nodes.

In Figure 1.5, the eccentricities of the nodes of graph G are $e(v_1) = e(v_5) = 3$, and $e(v_2) = e(v_3) = e(v_4) = 2$. The diameter and radius of G are $\text{diam}(G) = 3$ and $\text{rad}(G) = 2$, respectively.

A graph H is a *subgraph* of a graph G when all of its nodes and edges are contained in G . That is $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, which we represent by $H \subseteq G$. A *spanning* subgraph H of a graph G is a graph in which $V(H) = V(G)$. A subgraph $H = (V(H), E(H))$ of graph $G = (V(G), E(G))$ is said to be an *induced* subgraph of G , if for all

$u, v \in V(H)$ joined by an edge $(u, v) \in E(G)$ then $(u, v) \in V(H)$, we represent *induced* subgraph of G by $H[G]$.

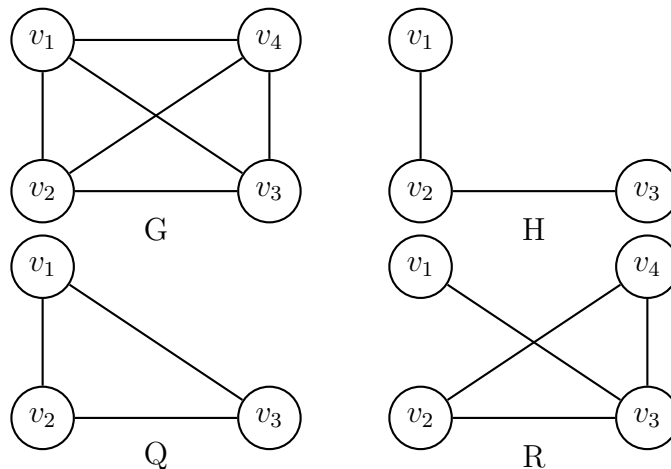


FIGURE 1.6. A graph G , a subgraph H , an induced subgraph Q , and a spanning subgraph R .

A graph of n nodes is said to be *complete* or *clique* if all distinct nodes are joined to each other by an edge. We represent a complete graph by K_n .

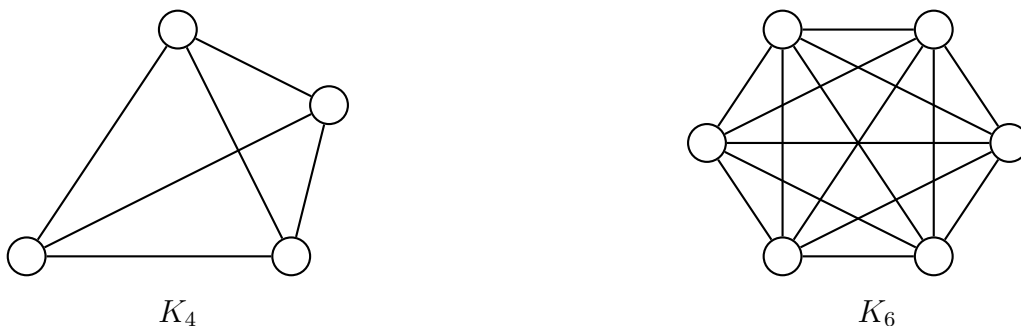


FIGURE 1.7. Complete graphs K_4 and K_6 .

A connected graph with n nodes is said to be a *cycle* in which the number of nodes equal to the number of edges, we represent a cycle by C_n . A graph with no cycle is said to be *acyclic*.

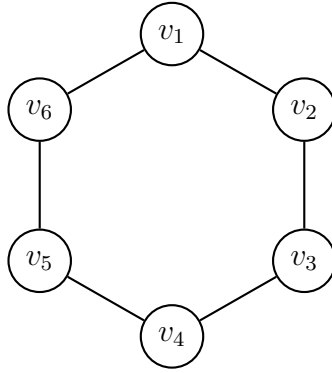


FIGURE 1.8. A cycle C_6 .

Let $G = (V(G), E(G))$ and $H = (V(H), E(H))$ be two graphs. A *homomorphism* of G to H is a function $f : V(G) \rightarrow V(H)$ such that for all pair of nodes $x, y \in V(G)$, $xy \in E(G)$ implies that $f(x)f(y) \in E(H)$. We denote a *homomorphism* from G to H as $G \rightarrow H$. An *isomorphism* is a *bijective function* $f : G \rightarrow H$ such that xy is an edge in G if and only if $f(x)f(y)$ is an edge in H . Note that two isomorphic graphs share all the same properties.

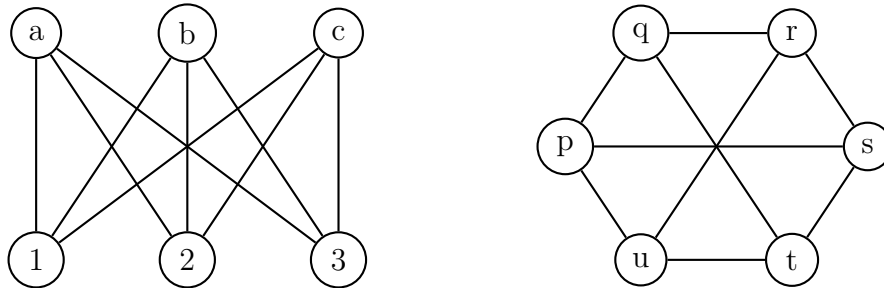


FIGURE 1.9. Isomorphic graphs.

A connected graph $G = (V, E)$ that has no cycle is called a *tree*. A node has degree one is called *leaf* in a tree, whereas *branches* are called the edges in tree. In a graph in which all of whose *connected components*

are trees said to be *forest*; the *connected components* of a graph is a subgraph in which any two nodes are connected to each other by *path*.

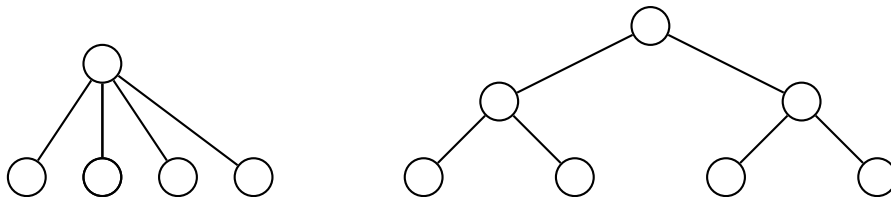


FIGURE 1.10. Examples of Trees.

Here we include below some equivalent definitions of a tree (which we include for completeness).

THEOREM 2. [18] *Given a graph $G = (V, E)$ with $|V(G)| = n$, the following are equivalent.*

- (a) *G is a tree.*
- (b) *G has no cycles.*
- (c) *The number of edges of G is one less than the number of nodes of G .*
- (d) *Every pair of nodes $u, v \in G$ is connected by unique path.*
- (e) *Removing any edge from G gives a graph which is not connected.*

We now define directed graphs. A graph $G = (V, E)$ is said to be *directed graph* or *digraph* if the edge set $e_i \in E(G)$ consists of ordered node pairs (u, v) , for all $u, v \in V(G)$. The edge e is said to be *directed edge* from u to v if it is associated with the pair (u, v) of nodes. The number of edges incoming to a node called *in-degree*, denoted by $\deg^-(u)$,

where u is a node. The number of edges outgoing from a node called *out-degree*, denoted by $\deg^+(u)$. A node has zero *in-degree* called *source* node, where a node has zero *out-degree* called *sink* node. We have the following theorem, which is part of folklore (which is included for completeness).

THEOREM 3. *For every digraph G ,*

$$\sum_{v \in V(G)} \deg^+(v) = \sum_{v \in V(G)} \deg^-(v),$$

PROOF. For every digraph $G = (V(G), E(G))$, an edge has two ends with one end of the edge adding 1 to the out-degree of some $v \in V(G)$, while the other end of the edge adding 1 to the in-degree of the some $v \in V(G)$. \square

In Figure 1.11, we find an example of a directed graph G with nodes $v_1, v_2, v_3 \in V(G)$, $\deg^-(v_1) = 1$, $\deg^-(v_3) = 2$ and $\deg^-(v_2) = 0$, where $\deg^+(v_1) = 1$, $\deg^+(v_2) = 2$ and $\deg^+(v_3) = 0$. We see that node v_2 is a *source* node and node v_3 is *sink* node.

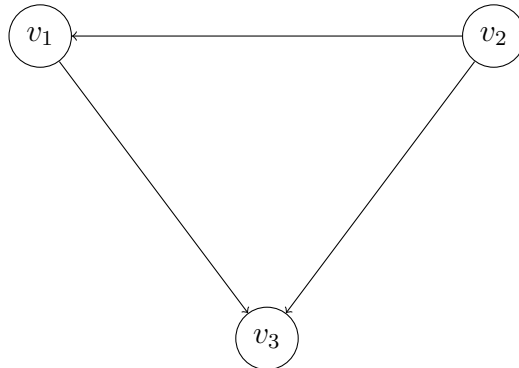


FIGURE 1.11. A directed graph G .

A digraph G is said to be *weighted graph* if all of its nodes or edges assigned a weight. An *edge-weighted graph* has weight on its edges and *node weighted graph* has weight on its nodes.

A set of nodes $W \subseteq V(G)$ with no directed edges within W , then the set W is an *independent set*.

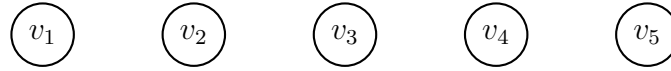


FIGURE 1.12. Independent set on 5 nodes.

An *independent set* W of a digraph G is said to be *maximum* if G has no independent set W' such that $|W'| > |W|$. The *independence number* of G , denoted by $\alpha(G)$, is the number of nodes in a maximum independent set of G . The *edge density* of the set of nodes S is the ratio

$$ED(S) = \frac{|E(S)|}{\binom{|S|}{2}}.$$

See Figure 1.13 for an example.

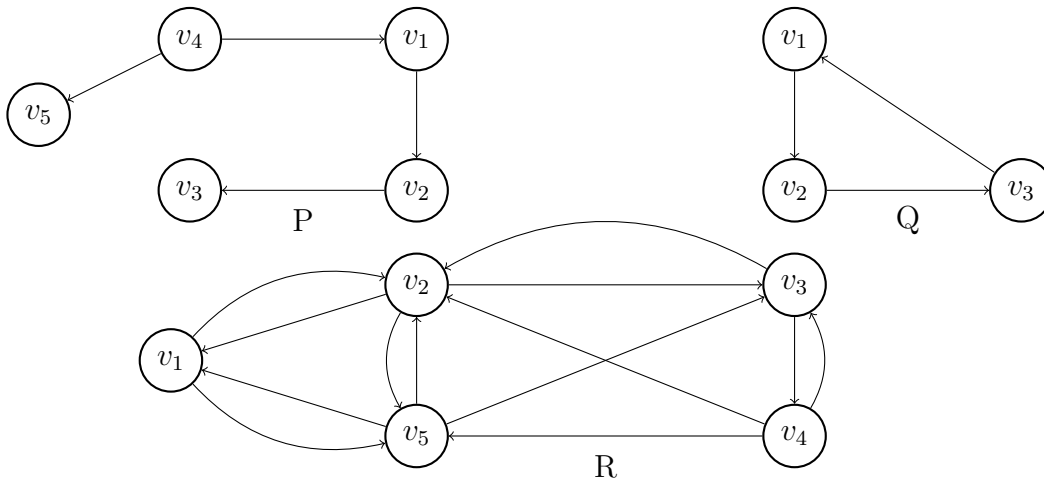


FIGURE 1.13. $ED(P) = 0.4$, $ED(Q) = 1$, $ED(R) = 1.3$.

1.3. Summary of the Thesis

In this chapter, we discussed the basic definitions and notation from graph theory, citing various properties and theorems. In Chapter 2, we will define centrality metrics on graphs, with the aim of determining important nodes in social networks. We will discuss structural balance theory and alliances in social networks. We will introduce and study the properties of dynamic competition networks. We formulate the Dynamic Competition Hypothesis that detects the leaders and measures the relative strength of the alliances in the social network.

In Chapter 3, we will present voting history data of four seasons from *U.S. Survivor* and one data set from *Big Brother*. We will analyse and verify these data sets by applying network metrics to validate the Dynamic Competition Hypothesis. We will find that the Dynamic Competition Hypothesis exactly predicts the alliances, top players and finalists in the game shows.

In the last chapter, we summarize and discuss our results. We will present open problems derived from our analysis and describe how the Dynamic Competition Hypothesis can be applied in other types of complex networks.

CHAPTER 2

The Dynamic Competition Hypothesis

In this chapter, we introduce and discuss structural balance theory. We define a model of positive and negative relationships in a complete graph. We discuss balanced and unbalanced graphs in signed social networks with examples. We define standard centrality metrics in network science that will be used to state the Dynamic Competition Hypothesis in Chapter 3. We introduce dynamic competition networks in real-world social networks. We formulate the Dynamic Competition Hypothesis with the help of graph theoretic tools to identify the alliances and leaders in a social network.

2.1. Structural Balance Theory

We discuss here the basic model of positive and negative relationships modeled by a complete graph. Suppose we have a social network representing a set of people, in which everyone knows everyone; that is, there is an edge joining each pair of nodes. We label each edge with either a positive sign (+) or a negative sign (−). A positive edge between two nodes denotes a positive relationship (such as friendship or an alliance) and a negative edge between two nodes denotes a negative relationship

(such as hatred or competition). Such a labeled graph is called a *signed network*.

In signed social networks, there is the concept of *balanced* and *unbalanced* cycles. A *balanced cycle* is defined as a cycle where the product of all the signs are positive. Balanced graphs represent a group of people who are unlikely to change their opinions of the other people in their group. In contrast, the unbalanced graphs represent a group of people who are very likely to change their opinions of the other people in their group.

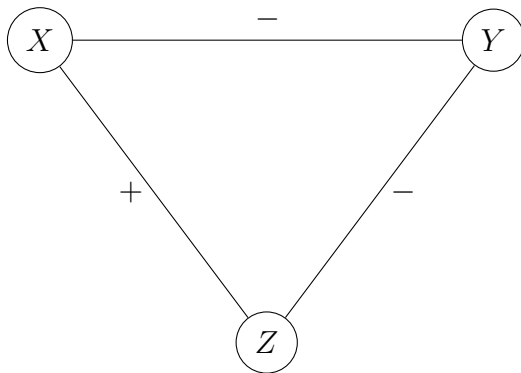


FIGURE 2.1. A signed graph.

The principles underlying structural balance are based on theories in social psychology from the work of Heider in the 1940s [14], and the theory was generalized and extended to the language of graphs with the work of Cartwright and Harary in the 1950s; see [7, 8, 12]. The critical idea of structural balance theory is that if we consider any *two* people in isolation, then the edge between them can be $+$ or $-$; that is, they are either friends or enemies. But when we look at sets of *three* people at a

time, there are four distinct ways to label the three edges among three people with $+$'s and $-$'s. We can differentiate these four possibilities as follows.

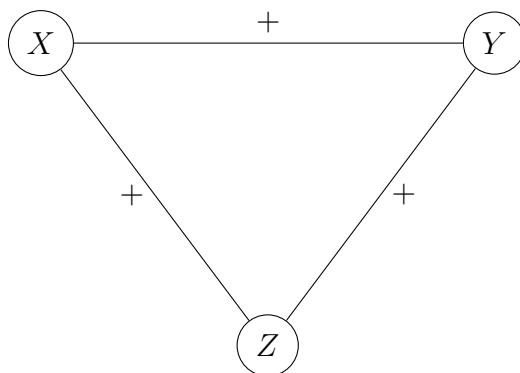


FIGURE 2.2. X , Y , Z are mutual friends: balanced.

In Figure 2.2, it is a natural situation if in a given set of three people X , Y and Z , they have three positive signs among them: it corresponds to three people who are mutual friends.

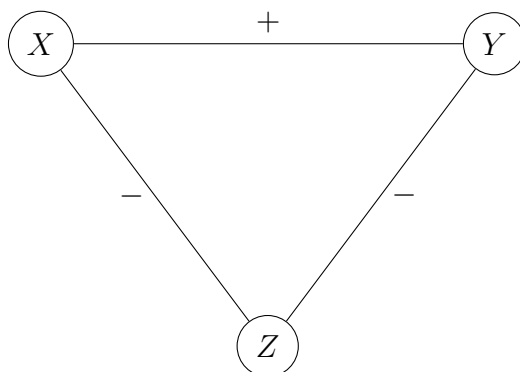


FIGURE 2.3. X and Y are friends with Z as a mutual enemy: balanced.

In Figure 2.3, it is also natural if a single positive sign and two negative signs in the relations among the three people. It means that two of the three are friends, and they have a mutual enemy in the third.

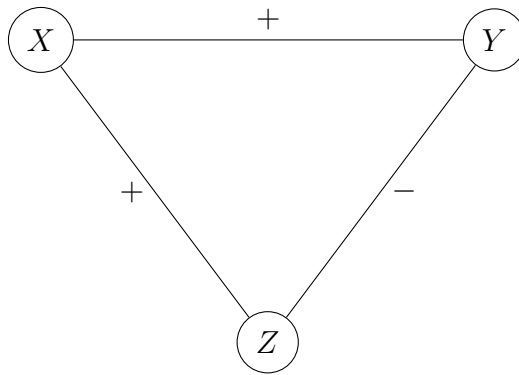


FIGURE 2.4. X is friends with Y and Z , but Y and Z are enemy to each other : unbalanced.

In Figure 2.4, a triangle with two positive signs and one negative sign corresponds to a person X who is a friends with each of Y and Z , but Y and Z are not friend with each other. This type of labeling shows some kind of psychological *stress* or *instability* between the relationships. In this situation, there would be implicit forces pushing X to try to get Y and Z to become friends (changing the YZ edge label to $+$); or else for X to side with one of Y or Z against the other (changing one of the edge labels incident with X to a $-$).

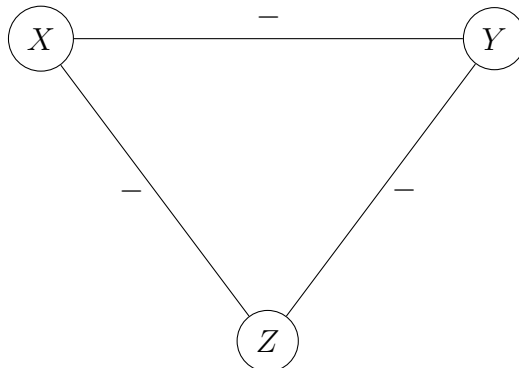


FIGURE 2.5. X , Y and Z are mutual enemies: unbalanced.

In Figure 2.5, there are also sources of instability in an arrangement where X , Y and Z are mutual enemies. In this situation, there would be forces motivating two of the three people to become partners against the third (changing one of the three edge labels to a $+$).

Due to the above reasons in the labeled triangles, we indicate the triangles who have one or three $+$ labels as *balanced* and triangles with zero or two $+$ labels as *unbalanced*. In structural balance theory, each labeled triangle must have one or three positive edges.

We may view unbalanced triangles as sources of stress or psychological dissonance, which strive to minimize in their personal relationships. Hence, they will be less abundant in real social settings than balanced triangles.

We next generalize our definition of structural balance of complete graphs for groups of three nodes into an arbitrary number of nodes with edges labeled by $+$'s and $-$'s. In particular, the labeled complete graph is *balanced* if it obeys the following property:

Structural Balance Property: For every set of three nodes, if we consider the three edges connecting them, either all three of these edges are labeled $+$, or else exactly one of them is labeled $+$.

We consider an example of two labeled networks in Figure 2.6. We see that the labeled four-node complete graph on the left consists of two

groups of friends X , Y and Z , R , with negative relations between people in different groups. It is balanced because each set of three nodes satisfies the Structural Balance Property. The labeled four-node complete graph on the right is not balanced because there are exactly two edges labeled $+$ between three nodes X , Y , Z and the triangle Y , Z , R violates the Structural Balance Property.

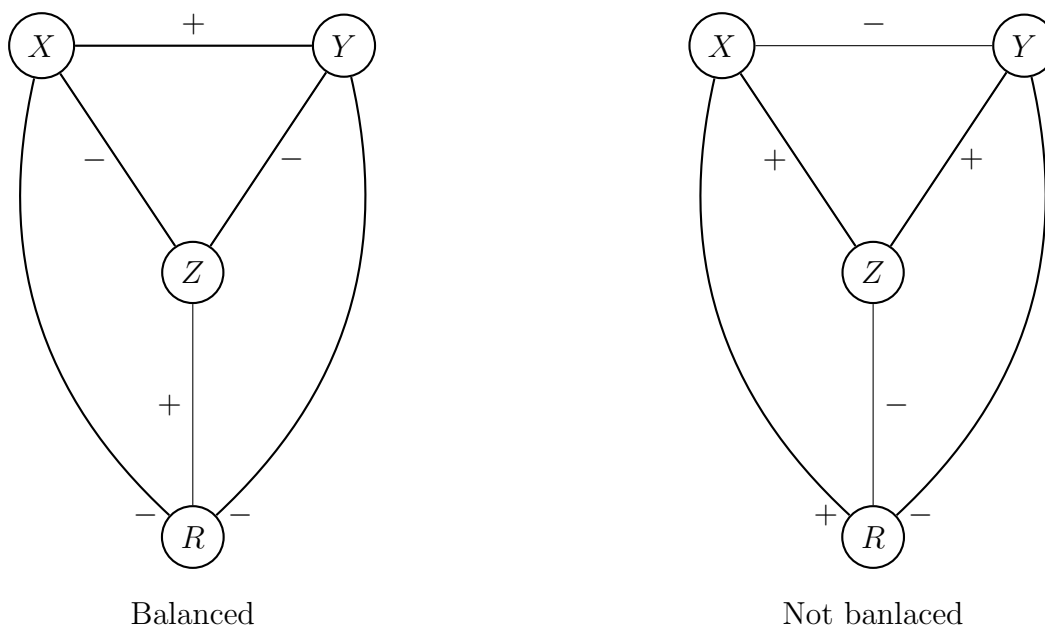


FIGURE 2.6. The labeled four-node complete graph on the left is balanced and on the right is unbalanced.

We see that there are only two basic ways for structural balance network to occur: everyone likes each other (see Figure 2.2), or the network consists of two groups of mutual friends with complete antagonism between the groups. This fact was originally discovered by Drown Cartwright and Frank Harary in the 1950s.

THEOREM 4. [8, 12] *If a labeled complete graph is balanced, then either all pairs of nodes are friends, or else the nodes can be divided into two groups, X and Y , such that every pair of people in X like each other, every pair of people in Y like other, and everyone in X is the enemy of everyone in Y .*

2.2. Centrality metrics in networks

In this section, we present some graph-theoretic terminology, and we will define standard centrality metrics in network science that we will use later. For example, by using various centrality metrics, we can find the most influential nodes.

A network can be represented by its matrix called the adjacency matrix. We define the *adjacency matrix* of a graph $G = (V(G), E(G))$ with n nodes by the $n \times n$ matrix with entries $A_{i,j}$ given by:

$$A_{i,j} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E(G) \\ 0 & \text{otherwise,} \end{cases}$$

for all $v_i, v_j \in V(G)$.

We consider an example to find the adjacency matrix of the following graph, whose nodes are v_1, v_2, v_3, v_4, v_5 and v_6 , respectively.

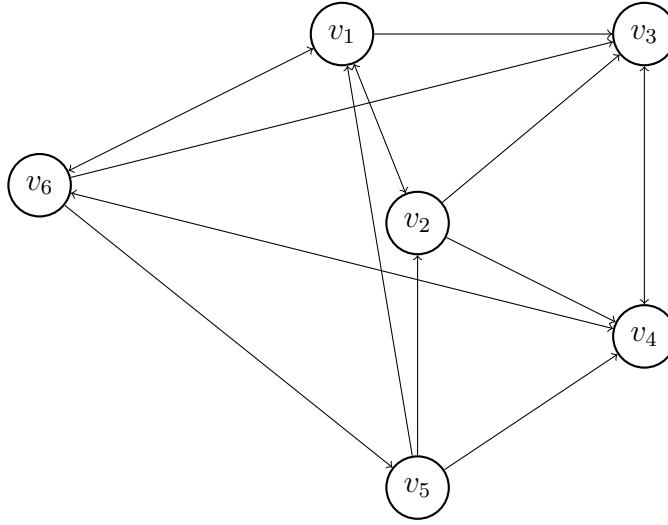


FIGURE 2.7. A directed graph.

The adjacency matrix of the directed graph in Figure 2.7 is the following:

$$A_{i,j} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Degree centrality simply measures the degrees of nodes. Degree may be interpreted in terms of the immediate risk of node catching whatever is flowing through the network (such as a virus, influence, or information). For a directed network, we define two types of degree centrality: in-degree and out-degree.

Let $A_{i,j}$ be the adjacency matrix of a directed graph. The in-degree centrality x_i of node i is equal to:

$$x_i = \sum_k A_{k,i}.$$

See Figure 2.8 for an example.

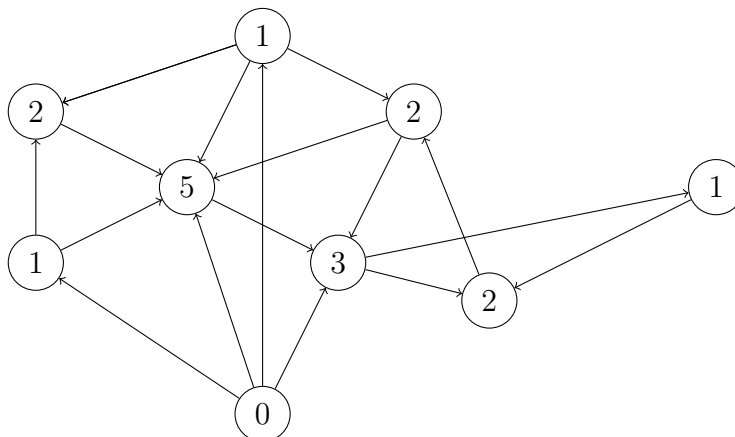


FIGURE 2.8. A directed graph with nodes labeled with their in-degree.

The out-degree centrality y_i of node i is equal to:

$$y_i = \sum_k A_{i,k}.$$

See Figure 2.9 for an example.

We consider degree centrality as a *local metric* since it only considers neighbors of nodes. For example, in an on-line social network such as Facebook, degrees measure the relative popularity of users.

Betweenness centrality is an important statistical property of a network. This is applied in many real-world problems, such as finding influential people in a social network, finding crucial hubs in a computer

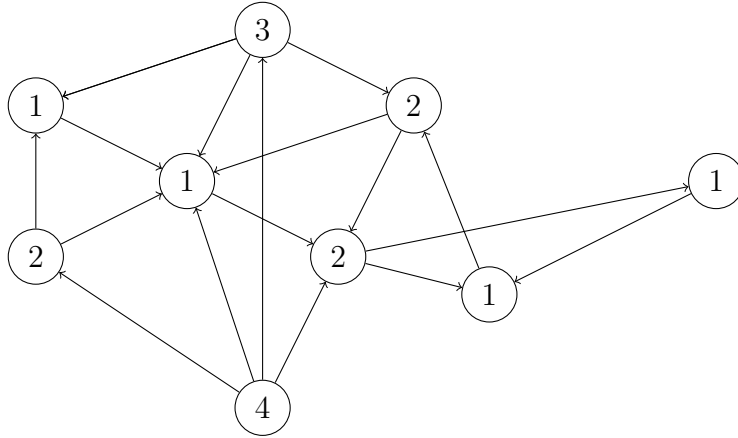


FIGURE 2.9. A directed graph with nodes labeled with their out-degree.

network, or finding border crossing points which have the most trade flow. See [10, 11].

The *betweenness* of a node u in a graph $G = (V, E)$ is computed as follows:

- (1) For each pair of nodes (v, w) , compute the shortest paths between them.
- (2) For each pair of nodes (v, w) , determine the fraction of shortest paths that pass through the node u .
- (3) Sum this fraction over all pairs of nodes (v, w) .

More formally, we define the *betweenness centrality* of u by

$$B(u) = \sum_{v, w \in V(G) \setminus \{u\}} \sigma_{vw}(u) / \sigma_{vw},$$

where σ_{vw} denote the total number of shortest one-way, directed paths between v , w and $\sigma_{vw}(u)$ denote the number of shortest one way, directed

paths between v and w that pass through the node u . If there does not exist any shortest path between v and w , then we have that $B(u) = 0$.

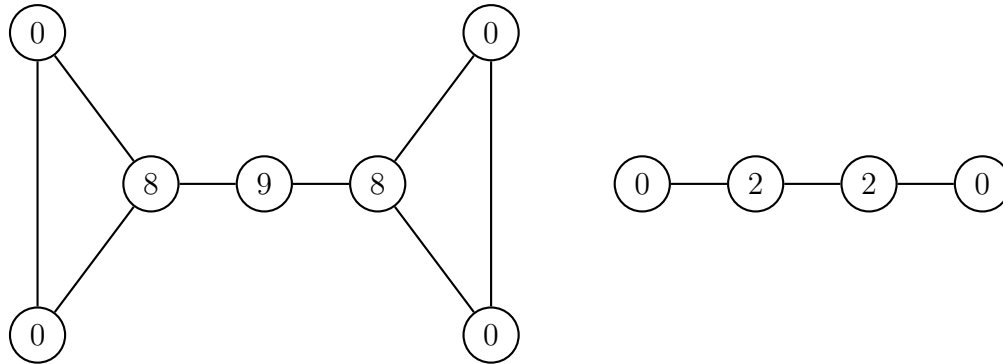


FIGURE 2.10. Two graphs with nodes labeled with their betweenness centralities.

In connected graphs there is a natural distance metric between all pairs of nodes, defined by the length of their shortest paths. The *farness* of a node u is defined as the sum of its distances from all other nodes in a graph. The *farness* of node u is defined by

$$F(u) = \sum_{v \in V(G)} d(u, v).$$

Closeness centrality indicates how long it will take for information from a given node to reach other nodes in the network. We describe the *closeness centrality* as the reciprocal of the *farness*. For a connected digraph $G = (V, E)$ and a node $v \in G$, we define *closeness centrality* of u by

$$C(u) = \left(\sum_{v \in V(G) \setminus \{u\}} d(u, v) \right)^{-1},$$

where $d(u, v)$ is the distance measured by one-way, directed paths from u to v . Thus, the more central a node is the lower its total distance from all other nodes. Note that taking distances *from* or *to* all other nodes is irrelevant in undirected graphs, whereas in directed graphs distances *to* a node are considered a more meaningful measure of centrality; see [2].

Both closeness and betweenness are well-studied centrality measures for complex networks; see [6]. For example, centrality of sports networks is often used to rank teams; see [15].

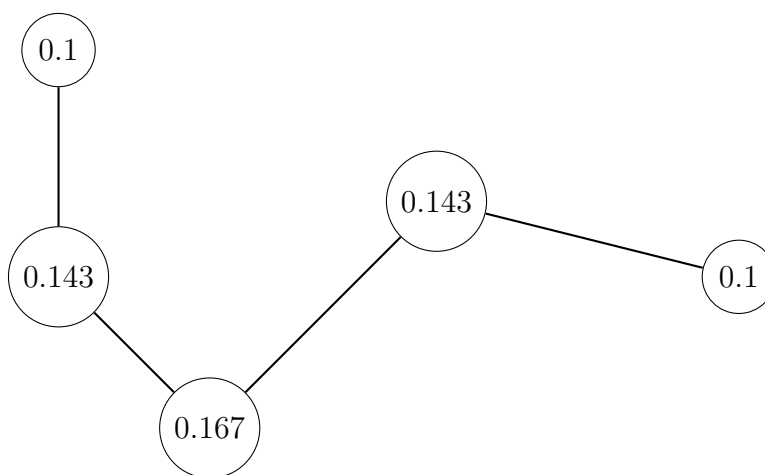


FIGURE 2.11. A tree with nodes labeled with their closeness centralities.

In Chapter 3, we will apply network science in the social game show *Survivor*. Given the nature of the voting network in *Survivor*, we also consider the number of common out-neighbors as a key metric which define as follows.

For nodes u , v , and w , we say that w is a *common out-neighbor* of u and v if (u, w) and (v, w) are directed edges.

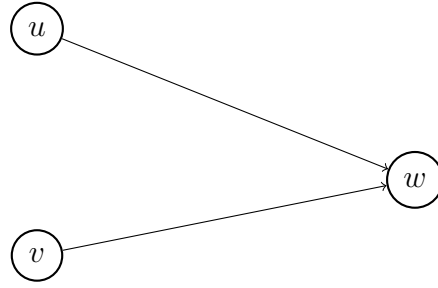


FIGURE 2.12. The node w is the common out-neighbor of u and v .

For a pair of distinct nodes u, v , we define $\text{CON}(u, v)$ to be the number of common out-neighbors of u and v .

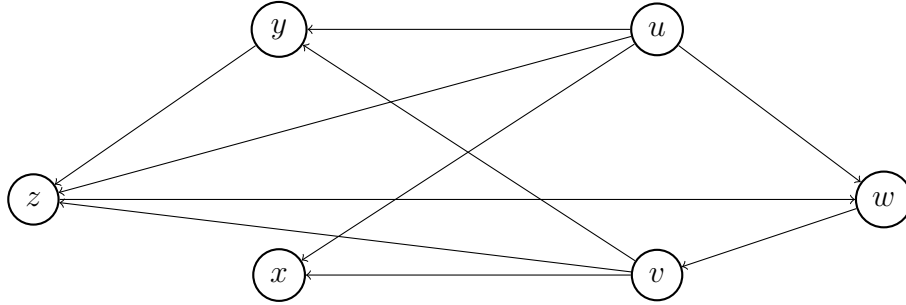


FIGURE 2.13. The number of common out-neighbor of u and v is given by $\text{CON}(u, v)=3$.

For a fixed node u , define

$$\text{CON}(u) = \sum_{v \in V(G)} \text{CON}(u, v).$$

We call $\text{CON}(u)$ the *CON score* of u .

For a set of nodes S with at least two nodes, we define

$$\text{CON}(S) = \sum_{u, v \in S} \text{CON}(u, v).$$

We note that $\text{CON}(S)$ is a non-negative integer.

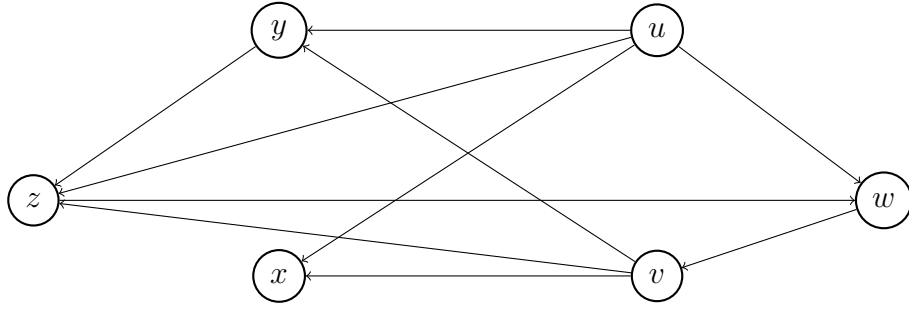


FIGURE 2.14. CON score of $u = \text{CON}(u) = 5$.

For a set S of nodes, we denote its *edge density* by $ED(S)$, and compare this to that of an *independent set*. We observe that $ED(S)$ may be greater than 1 as there may be multiple edges in the digraphs as we defined in Chapter 1 (see Figure 1.13).

If $ED(S) \leq \epsilon$, then a set S is said to be ϵ -near independent, where ϵ is a non-negative real number. We measure the relative density of sets of nodes by the parameter ϵ . If the set S is ϵ -near independent for some positive value of ϵ then S is called near independent. Generally, in applications, we take ϵ to be small. The value of ϵ will often be heuristically determined in a real-world networks by considering a ranking of subsets by their edge density (we highlight that every set of nodes is ϵ -near independent for a suitably chosen ϵ). We note that independent sets are near independent by choosing the value of $\epsilon = 0$.

In Figure 2.15, we consider two sets of nodes R and S where we have that $ED(R) = 0.6$ and $ED(S) = 1.2$, respectively. We choose ϵ_1 and ϵ_2 such that $\epsilon_1 = 0.6$ and $\epsilon_2 = 1.3$. We see that $ED(R) \leq \epsilon_1$ and

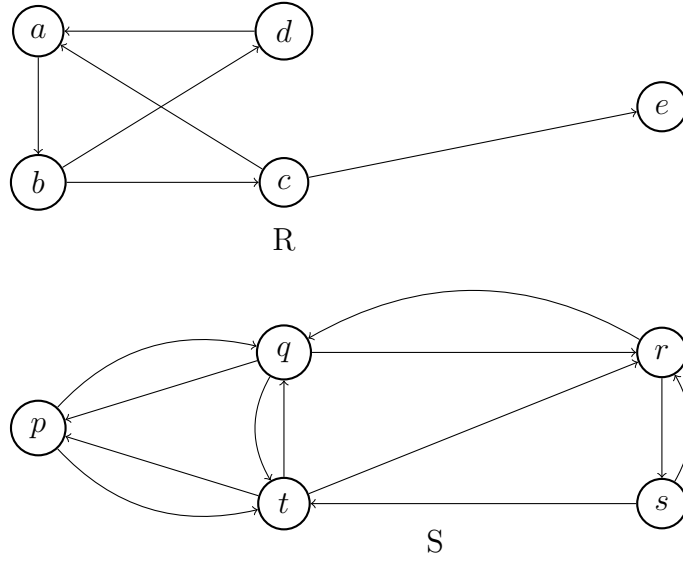


FIGURE 2.15. In these networks we have that $ED(R) = 0.6$ and $ED(S) = 1.2$.

If $ED(S) \leq \epsilon_2$, then we say that set R is ϵ_1 -near independent and set S is ϵ_2 -near independent.

2.3. Dynamic Competition Networks

We consider a network G in which nodes represent agents, and there may be directed edges from nodes u and v . We say G is a *competition network* if agent u is in competition with agent v . A directed edge may correspond to some kind of negative social interaction from nodes u to v ; that is, a directed edge (u, v) may represent a vote against v . Note that competition networks need not be *tournaments*, which are directed graphs isomorphic to those where every pair of nodes has exactly one directed edge between them.

A competition network is said to be a *dynamic competition network* if directed edges are added over discrete time-steps. For example, in the

game show Survivor, contestants cast votes against each other, and a directed edge may be a vote of one contestant to other; see Chapter 3. We consider another example from real-world networks, where nodes represents nation states and edges correspond to conflicts between them. Dynamic competition networks may have multiple edges. In our thesis, we focus on the underlying structure of social networks of competitors which we call a dynamic competition network. Specifically, we focus on the networks arising on the American television series Survivor and Big Brother. Before we describe our hypothesis, we discuss some important definitions which we use in our hypothesis and whole thesis.

Alliance means a bond or connection between families, states, individuals, or parties. In the real-world network, we find alliances in many ways, for example:

- (1) Companies with common economic interests.
- (2) People who unite by relationship or friendship.
- (3) Political parties.
- (4) Friendly nation states.

We define alliances in the social game show Survivor, as groups of contestants who work together to vote off contestants outside the alliance. In particular, alliances are induced subgraphs of dynamic competition networks. In the case of strong alliances, members of an alliance are less likely to vote for each other.

Persons that hold a dominant or superior position in the network are called *leaders* and are able to exercise a high degree of control or influence over others; that is, edges emanating from leaders may influence edge creation in other agents. Leaders may be the winner in the game show Survivor or may be the non-winning players with a strong hold on the results of the game. Our main goal is to apply network science to help determine alliances and leaders in dynamic competition networks arising in social networks.

2.4. The Hypothesis

Structural Balance Theory posits relationships as either positive or negative, and in our approach, we focus on negative relationships. We apply network science to identify the alliance and top players within a dynamic competition network such as Survivor. Players in Survivor are represented by nodes and there is a directed edge from player x to player y if x votes for y . The directed edges in Survivor correspond to competition or rivalry between the players, and so correspond to negative relationships. Directed edges are added over discrete time steps in dynamic competition network.

We indicate the top players in Survivor with their in-degree. Players who have low in-degree to be considered as top players. Alliances have low edge densities and are near independent. We note that the top

players have high CON scores and we consider them as a leader within the alliance.

An important measure of top players is the closeness. Players who are centrally located in the network have highest closeness. We discover that the closeness appears as a reliable predictor in Survivor.

The *Dynamic Competition Hypothesis* (or DCH) asserts that dynamic competition networks arising from a social network satisfy the the following four properties.

- (1) Alliances are near independent sets.
- (2) Strong alliances have low edge density.
- (3) Members of an alliance with high CON scores are more likely leaders.
- (4) Leaders exhibit high closeness, high CON scores, low in-degree, and high out-degree.

We see the visualization of DCH in Figure 2.16. Now we observe that the above four properties jointly support each other. When we detect an alliance from property (1) then we are able to compute the strength of an alliance relative to other alliances by using property (2). We can separate leaders within alliances by applying property (3). Specifically, we anticipate leaders to be in alliances because of having their eminent local control in the alliances. We consider property (4) is independent of

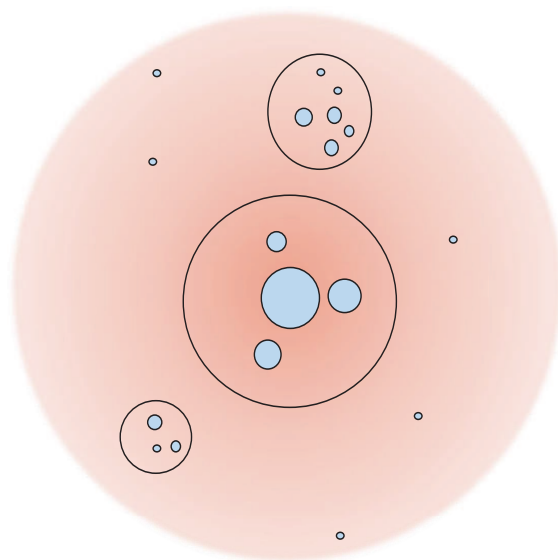


FIGURE 2.16. A heat map representation of dynamic competition networks according to the DCH, where nodes closer to the center have higher closeness and CON scores. Larger nodes have higher CON scores, lower in-degree, and higher out-degree. The induced subgraphs (which are the circles) correspond to alliances.

alliances and we determine leaders through global metrics of the network; that is, we measure closeness centrality.

In Chapter 3, we will test the Dynamic Competition Hypothesis (DCH) with historical voting data from four seasons of the U.S. television social game shows *Survivor* and one season from *Big Brother*.

CHAPTER 3

Data and Methods

In this chapter, we provide background on the U.S. based social game shows Survivor and Big Brother. We concentrate on four seasons from Survivor and one season from Big Brother, in detail. We examine the voting data of these seasons by using tools from network science. We observe that the Dynamic Competition Hypothesis (DCH) predicts the formation of alliances and finalists with a high degree of accuracy. As indicated in Chapter 2, the *Dynamic Competition Hypothesis* (or DCH) asserts that the dynamic competition networks arising from a social network satisfy the following four properties (which we restate for convenience):

- (1) Alliances are near independent sets.
- (2) Strong alliances have low edge density.
- (3) Alliance members who have high CON scores are leaders.
- (4) Leaders reveal high closeness, high CON scores, low in-degree, and high out-degree.

3.1. Survivor

Survivor is a reality-based competition television show. Survivor is an example of a social game, where social interactions establish the direction

of the play and winners. We observe that there are versions of Survivor in many countries. In our thesis, we consider the U.S. version of Survivor.

In Survivor, a group of contestants called *survivors* are marooned in an isolated location, usually with a tropical climate. Survivors are forced to provide food, water, and shelter for themselves with insufficient help from the outside world. Survivors are split into two or more *tribes*, which cohabitate and work together. As a tribe, the survivors must endure the elements, build shelter, look for water, and scrounge for food and other requirements for the whole period of the game, which is generally 39 days in the U.S. version.

In particular, in the first half of the game, the tribes compete in challenges, some for rewards of food, shelter, or luxury items, while others are for immunity, preventing the winning tribe from having to go to the *tribal council*. The losing tribe goes to tribal council, where they vote out one of their own tribesman to eliminate them from the game.

In the second half of the game, the tribes are merged into a single tribe, and the remaining survivors compete for individual rewards and immunity. At subsequent tribal councils, eliminated players start to form the *jury*, who sit in on all subsequent tribal councils but they do not participate in the voting. When there are a small number of remaining survivors who are *finalists*; that is, only two or three, they attend the *final* tribal council, where the jury is given the opportunity to ask them

questions. After this, the jury members vote in favor of one of them to become the *Sole Survivor* who receives a cash prize of one million dollars.

3.2. Big Brother

Big Brother is a reality-based competition television game show in which a group of contestants are called *HouseGuests*. HouseGuests co-habitate in a custom-built *house* under constant video surveillance. While in the house, the contestants are isolated from the outside world; that is, no phone, television, Internet, magazines, newspapers, or have no communication with those not in the house. This rule could be broken in the case of medical injury and any family emergency. The HouseGuests are free to leave from the game but they are not allowed to entry back into the house. HouseGuests could be expelled from the house if they break the rules of the game.

Each week, the HouseGuests compete for the title of *Head of Household*. Head of Household nominates two HouseGuests for eviction. The HouseGuests vote to evict one of them. The HouseGuests who has the most votes is evicted from the game. Once only two HouseGuests remain, the members of the Jury cast their votes for who should win the series. The winner receives a cash prize of half a million dollars.

Various twists have been introduced during the seasons in both Survivor and Big Brother to keep the players vigilant and to avoid players relying a similar game plan used in preceding seasons. For example, in

Survivor, these twists include tribal switches, seasons starting with more than two tribes, the ability to exile a player from a tribe for a short time, hidden immunity idols that players can use to save themselves or others at tribal council or special voting powers which can be used to influence the result at tribal council. As a disclaimer, our use of network analysis is insensitive to these twists.

3.3. Social Network Analysis tools

Social network analysis tools are used to identify, analyze, visualize or simulate nodes (agents) and edges (relationships) from various types of input data including mathematical models of social networks. There are several tools available for analysis of social networks. We selected the software Gephi [1], which is freely available for use and can handle large graph sizes. Gephi is a tool to explore and understands graphs. We used Gephi in our thesis for computing centrality metrics and the visualizations of dynamic competition graphs.

We have gathered data from Survivor Wiki [16] and Big Brother Wiki [3]. Data includes the information on contestants, their voting history and tribes, and catalogues of alliances.

We now present the visualizations of dynamic competition networks for Survivor and Big Brother. We discuss four seasons from Survivor; that is, Borneo, China, Game Changer, and Heroes vs Healers vs Hustlers (HHH), respectively. The reason for choosing these seasons is because

Borneo and HHH introduce the new players and Game Changer includes the returning contestants from the preceding seasons. In addition, several new twists and changes have been made in these seasons. We also discuss a data from Season 12 of Big Brother. We note that in both Survivor and Big Brother, we choose data when all the players cast their votes against each other. We also provide the tables which contains the summary of the relevant network statistics. In the tables below, we list the contestants in the order which they were voted out from the game. The winner of the game is the first entry of the table and the others are placed by when they were voted out from the game. We analyse the voting history data from Survivor and Big Brother by using Gephi in which we check whether the DCH predicts the top players (which may be the finalist) and the emergence of the alliances. We use the abbreviations ID, OD, C, CON, and B which stand for in-degree, out-degree, closeness centrality, CON-score, and betweenness centrality, respectively.

3.4. Borneo

We discuss the first season of Survivor set in Borneo. Borneo consisted of thirty-nine days of gameplay with sixteen competitors. The sixteen contestants were split into two tribes of eight: Tagi and Pagong. Both Tagi and Pagong coped equally in challenges but diverged in organisational structure. Three contestants from Pagong: Colleen, Jenna, Gervase formed an alliance, called the *Barbeque* alliance, whereas the four

contestants Richard, Rudy, Susan, and Kelly from other tribe formed an alliance, called the *Tagi* alliance. The *Tagi* alliance voted as a block to ensure their safety. The *Tagi* alliance took advantage of the lack of voting strategy of other contestants. There was also a dispute in the *Tagi* alliance as Kelly was considered untrustworthy. When ten players remained, the contestants merged into one tribe. After merging in one tribe, the *Tagi* alliance successfully used their votes to eliminate the other contestants.

When only the *Tagi* alliance members remained in the game, we see that the close allies Richard and Rudy voted for Susan, and Susan and Kelly voted for Richard, so the vote ended with tie. After the revote, Susan was eliminated because Kelly switched her vote from Richard to Susan. Kelly also voted for Rudy after winning her last immunity challenge. Richard and Kelly were the finalists and Richard won the game by 4 – 3.

In Table 1, we see that Richard is the the Sole Survivor of the season. We note that Richard has one of the highest closeness and CON scores. Rudy and Susan have higher scores and closeness but high in-degree as compared to Richard. We see that Kelly won four consecutive immunity challenges near the end of the game and was ineligible for elimination. Since Kelly voted for both Rudy and Susan, we note that this was the decisive factor for her not winning the finale.

TABLE 1. Survivor: Borneo, Season 1.

Name	Tribe	ID	OD	C	CON	B
Richard	Tagi	6	10	0.737	42	28.7
Kelly	Tagi	0	12	0.682	34	0
Rudy	Tagi	8	11	0.778	45	36.483
Susan	Tagi	7	10	0.778	44	16.467
Sean	Tagi	9	9	0.7	38	17.917
Colleen	Pagong	7	8	0.636	29	33.067
Gervaise	Pagong	6	7	0.636	31	8.583
Jenna	Pagong	11	6	0.583	27	27.85
Greg	Pagong	6	5	0.412	15	4.833
Gretchen	Pagong	4	4	0.56	17	7.233
Joel	Pagong	4	3	0.412	17	1
Dirk	Tagi	4	3	0.5	12	1.317
Ramona	Pagong	6	2	0.412	10	17.733
Stacey	Tagi	6	2	0.452	4	1.733
B.B.	Pagong	6	1	0.298	5	0.333
Sonja	Tagi	4	1	0.452	4	0.75

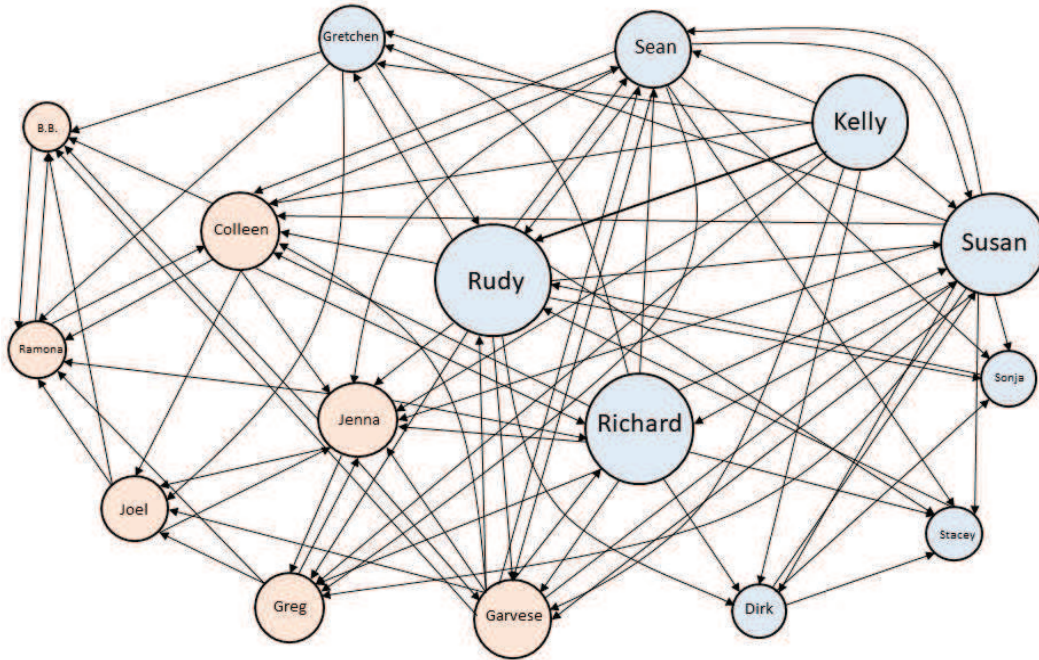


FIGURE 3.1. Survivor: Borneo.

We note that comparing betweenness of the players is vague as a predictor of leaders. For example, the betweenness of Richard, Kelly, and Rudy are 28.7, 0, and 36.5, respectively. One of the reason behind this is that the leaders tend to have lower in-degree, which may reduce the number of paths traversing through them. In the other seasons of Survivor, we do not consider, therefore, betweenness scores.

3.5. China

We next move to Survivor: China, which is the fifteenth season of U.S. reality television show. The reason to choose China is because it represents a sample after the game became more familiar to audiences, where contestants better understood the mechanics of the game.

Two tribes, Fei Long and Zhan Hu, were predetermined prior to the start of the game. Two alliances were formed in the tribes. Seven contestants formed the *Fie Long* alliance: Todd, Courtney, Amanda, Aaron, Denise, James, and Frost and three contestants Pie-Gee, Erik, and Jaime formed the *Zhan Hu* alliance.

The tribes merged in one tribe when ten contestants remained in the game, with the advantage of given to the Fei Long alliance (which comprised six contestants in the new tribe). After the elimination of two of the Fie Long's contestants, the remaining Fei Long members held on to their lead, and Todd, Amanda, and Courtney's alliance held strong.

Todd, Amanda, and Courtney went to the final tribal council. Todd won the game with four votes to Courtney’s two and Amanda’s one.

In Survivor: China, we note that in Table 2, Todd is the Sole Survivor. From the table, we note that Todd has highest closeness and high CON scores. Courtney and Amanda appear also as leaders based on their CON scores.

TABLE 2. Survivor: China, Season 15.

Name	Tribe	ID	OD	C	CON
Todd	Fei Long	5	9	0.765	49
Courtney	Fei Long	0	9	0.667	39
Amanda	Fei Long	0	9	0.737	49
Denise	Fei Long	3	9	0.722	40
Peih-Gee	Zhan Hu	8	10	0.722	41
Erik	Zhan Hu	5	9	0.722	41
James	Fei Long	9	6	0.591	31
Frosti	Zhan Hu	7	7	0.65	39
Jean-Robert	Fei Long	12	4	0.5	23
Jaime	Zhan Hu	7	5	0.481	26
Sherea	Zhan Hu	6	4	0.448	24
Aaron	Fei Long	3	2	0.406	12
Dave	Zhan Hu	6	3	0.382	11
Leslie	Fei Long	6	1	0.342	9
Ashley	Zhan Hu	8	2	0.464	10
Chicken	Zhan Hu	5	1	0.333	6

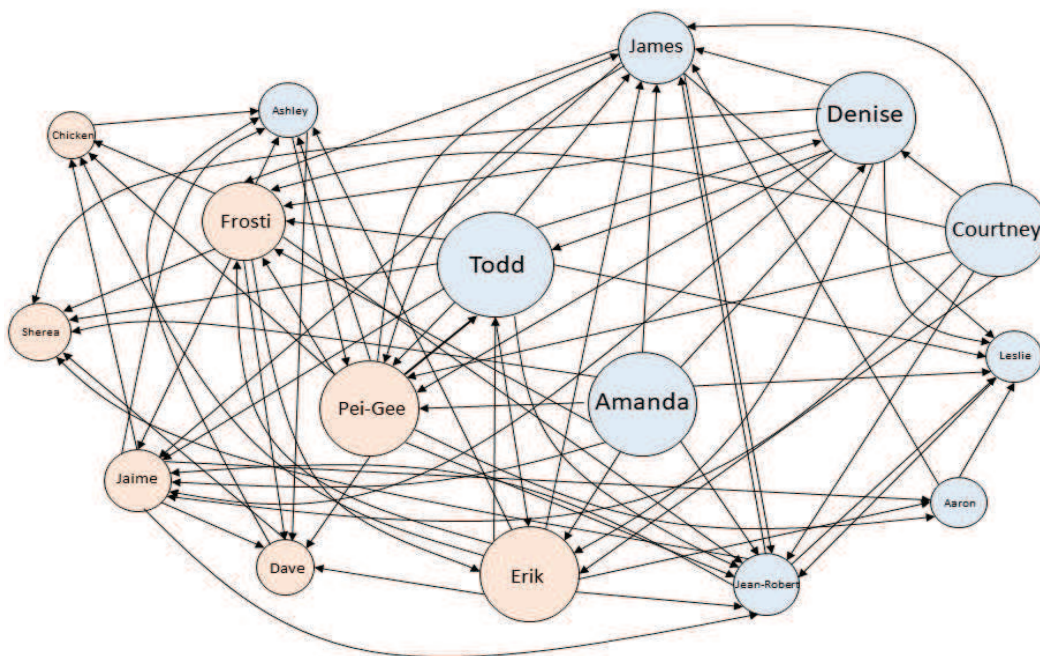


FIGURE 3.2. Survivor: China.

3.6. Game Changers

We next turn to Survivor: Game Changers, as the third-to-last season of the show. The twenty players were initially divided into two tribes: Mana and Nuku. The Nuku fared much better in challenges and the Nuku members were able to eliminate most of the members from the Mana tribe. When the tribes merged, two main factions appeared: *Power Six* alliance led by Sierra and Brad and *Tavva* alliance led by Andrea and Cirie, with Sarah navigating between the two groups to eliminate members from each. However, Sarah settled with Tavva to eliminate the other alliance members. Sarah betrayed Tavva and then aligned with Brad and Troyzan. Sarah, Brad, and Troyzan used a series of immunities and advantages to secure the majority and reach the end of

the game together. There were ten members in the jury and Sarah won the game by taking seven votes with Brad's three.

In Table 3, We see that the Sole Survivor Sarah had higher CON scores and closeness than Brad and Troyzan. We note that Tai and Aubry had higher closeness than Sarah, Brad and Troyzan and they also had the high CON scores than both Brad and Troyzan. We note that Tai and Aubry had high in-degrees than Sarah, Brad and Troyzan which likely disadvantaged them to be in the final three.

TABLE 3. Survivor: Game Changer, Season 34.

Name	Tribe	ID	OD	C	CON
Sarah	Nuku	3	13	0.692	64
Brad	Nuku	2	12	0.643	49
Troyzan	Mana	2	12	0.643	55
Tai	Nuku	12	13	0.72	56
Aubry	Mana	9	13	0.72	61
Cirie	Nuku	0	8	0.613	45
Michaela	Mana	11	11	0.643	51
Andrea	Nuku	14	8	0.581	39
Sierra	Nuku	15	7	0.581	34
Zeke	Nuku	11	6	0.6	39
Debbie	Nuku	6	7	0.545	32
Ozzy	Nuku	7	4	0.5	22
Hali	Mana	8	5	0.474	28
Jeff	Mana	6	5	0.529	33
Sandra	Mana	5	5	0.581	34
JT	Nuku	3	2	0.45	18
Malcom	Mana	5	3	0.439	24
Caleb	Mana	5	3	0.4	21
Tony	Mana	7	2	0.439	15
Ciera	Mana	9	1	0.4	8

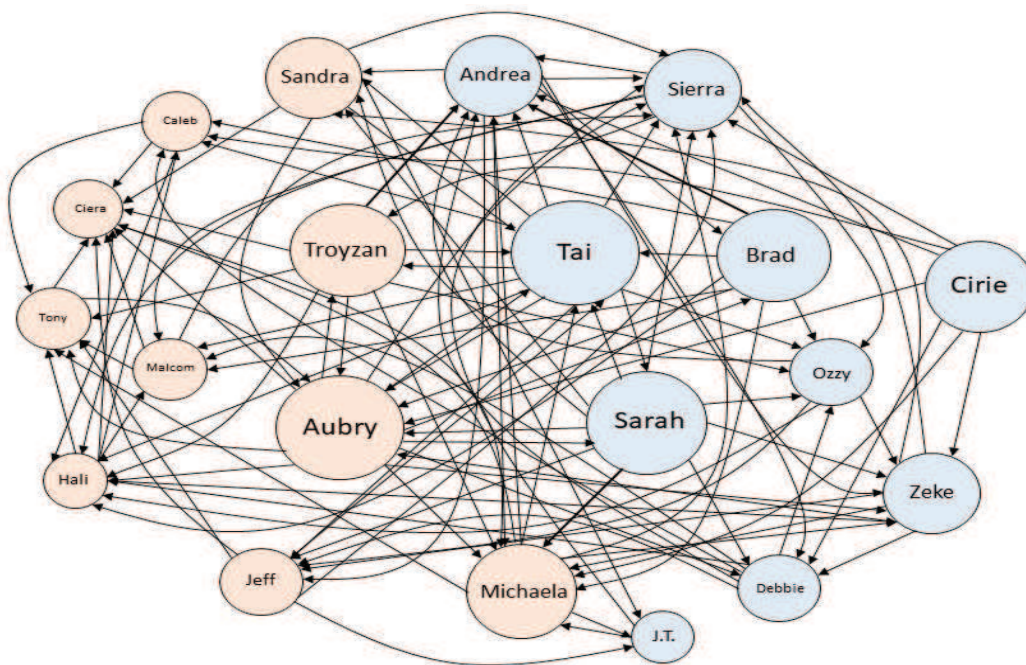


FIGURE 3.3. Survivor: Game Changer.

3.7. Heroes vs Healers vs Hustlers (HHH)

We now turn to the second-last season of Survivor: Heroes vs Healers vs Hustlers (HHH) which is the 35th season. This season featured 18 new players divided into three tribes based on a dominant perceived trait: with *Heroes* representing courage, *Healers* representing compassion, and *Hustlers* representing tenacity. The 18 players were initially divided into three tribes: Levu (Heroes), Soko (Healers) and Yawa (Hustlers). Ryan from the Hustlers tribe found a Super Immunity Idol that was only valid at the first Tribal Council. As his tribe won immunity, he sent the Idol to Chrissy who was a member of the losing Heroes tribe. Ryan and Chrissy both became dominant players on their tribes, forming strong alliances

with Devon and Ben, respectively. The Healers tribe went undefeated, while the Heroes lost one member and the Hustlers lost two.

When the tribes were switched, Ryan and Chrissy aligned on the new Soko tribe, working to eliminate the Healers and their allies. Three alliances were formed after the tribes merged in one tribe: that is, *healers*, *The Round Table*, and *Final Four alliance*. Ryan and Chrissy brought their former Heroes and Hustlers tribemates together against the Healers. When Chrissy won the final immunity challenge, she had to grant one castaway additional immunity, which would force the remaining two to compete against each other in a fire-making challenge to determine the third finalist. She chose to save Ryan. Ben defeated Devon in the firemaking challenge to join Chrissy and Ryan in the finals. At the final tribal council, Ben won five jury votes, with Chrissy winning two votes, and Ryan winning one. Ben was awarded the Sole Survivor of the season.

In Table 4, we see that Ryan and Devon had the highest overall closeness of 0.708 and highest overall CON scores of 47 and 55, respectively followed by Chrissy whose closeness and CON score of 0.68 and 44, respectively. However, Ben, the Sole Survivor, had lower CON score, closeness, and high in-degree than the other finalists; that is, 41, 0.63, and 11, respectively. Ben secured his place in the final three by successfully playing three consecutive hidden immunity idols.

TABLE 4. Survivor: Heroes vs Healers vs Hustlers (HHH), Season 35.

Name	Tribe	ID	OD	C	CON
Ben	Levu	11	11	0.63	41
Chrissy	Levu	7	13	0.68	44
Ryan	Yawa	2	14	0.708	47
Devon	Yawa	2	11	0.708	55
Mike	Soko	9	9	0.63	37
Ashley	Levu	8	10	0.607	46
Lauren	Yawa	3	7	0.63	39
Joe	Soko	12	6	0.607	26
JP	Levu	6	8	0.586	25
Cole	Soko	7	4	0.531	26
Desi	Soko	11	3	0.515	9
Jessica	Soko	7	1	0.415	6
Ali	Yawa	3	4	0.5	19
Roark	Soko	3	1	0.415	6
Alan	Levu	2	2	0.415	11
Patrick	Yawa	5	2	0.405	6
Simone	Yawa	5	1	0.293	4
Katrina	Levu	5	1	0.386	5

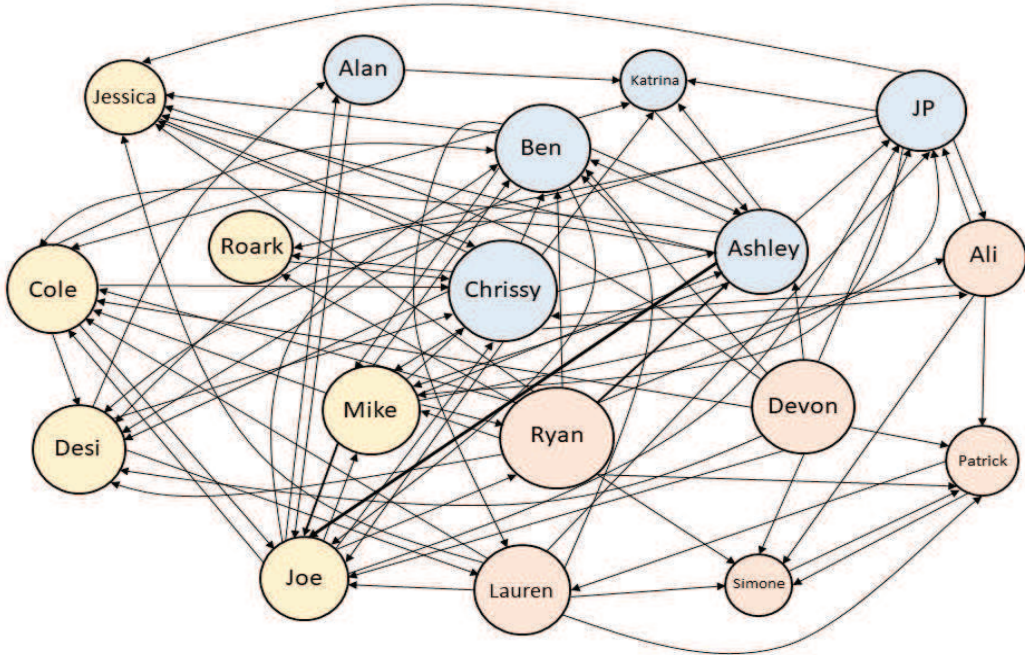


FIGURE 3.4. Survivor: HHH.

3.8. Big Brother

We now move to data from Survivor to another social game show: Big Brother. We only focus on Big Brother 12 which is the twelfth season of the U.S. version of the reality show. There were thirteen HouseGuests in this season and they spent seventy five days. Hayden, Enzo, Lane, and Matt formed an alliance called *The Brigade*. Hayden and Lane secured their place in the final two. The jury voted to give Hayden the grand prize by a vote of 4 to 3. Hayden received half a million dollars for spending seventy five days in the Big Brother House.

TABLE 5. Big Brother, Season 12.

Name	ID	OD	C	CON
Hayden	3	16	0.923	44
Lane	3	10	0.857	46
Enzo	4	9	0.8	48
Britney	4	10	0.8	43
Regan	5	8	0.706	49
Brendon	7	9	0.706	40
Matt	9	7	0.632	35
Kathy	7	4	0.6	20
Rachel	8	6	0.667	24
Kristen	7	3	1	25
Andrew	9	2	1	17
Monet	8	1	1	10
Annie	11	0	0	0

In Table 5, we note that Hayden, the winner of the season, had highest closeness and also the one of the high CON scores, with HouseGuests Lane and Enzo rounding out the top three. Lane and Enzo emerged also as leaders based on their scores.

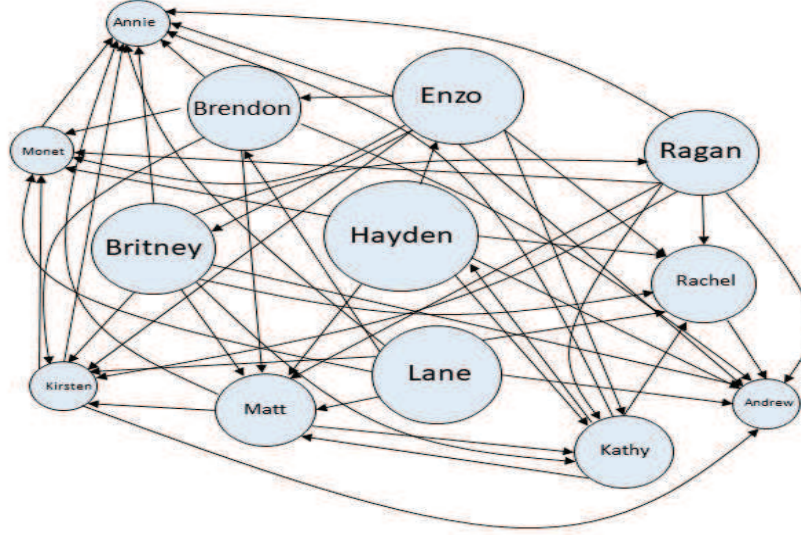


FIGURE 3.5. Big Brother, Season 12.

We discussed and analyzed the five seasons of two reality different shows, it is clear that the data conforms to the predictions of the Dynamic Competition Hypothesis (DCH) with respect to leaders which may be the winner of the game. We also found that the closeness is more predictive centrality measure than the betweenness in dynamic competition networks.

3.9. Alliances

In this section, we analyzed the alliances of four seasons from Survivor and one season from Big Brother. We computed their edge densities for predicting the winner in the game. We see that all alliances are near independent sets (with stronger alliances showing lower edge density), and they conform to the Dynamic Competition Hypothesis.

TABLE 6. Edge densities of the Alliances.

Season	Winner	Finalists	Alliances	ED
Borneo	Richard	Kelly	<i>Barbecue</i> : Colleen, Jenna, Ger-vase	1.667
			<i>Tagi</i> : Richard, Rudy, Susan, Kelly	1.5
China	Todd	Courtney	<i>Fei Long</i> : Todd, Courtney, Amanda, Aaron, Denise, James, Frosti	0.667
		Amanda	<i>Zhan Hu</i> : Peih-Gee, Erik, Jaime	0.0
Game Changers	Sarah	Brad	<i>Power Six</i> : Sarah, Brad, Troyzan, Sierra, Debbie, Tai	0.933
		Troyzan	<i>Tavua</i> : Aubry, Cirie, Michaela, Ozzy, Andrea, Zeke, Sarah	1.238
HHH	Ben	Chrissy	<i>Healers</i> : Joe, Desi, Jessica, Cole, Mike	0.6
		Ryan	<i>The Round Table</i> : Chrissy, Ryan, Devon, JP, Ben, Ashley, Lauren	0.905
			<i>Final Four</i> : Ashley, Lauren, Ben, Devon	1.333
BB 12	Hayden	Lane	<i>The Brigade</i> : Enzo, Hayden, Lane, Matt	0.5

In Table 6, we provide the edge densities of each alliance in the aforementioned five seasons. We note that some alliances have high edge density as compared to other alliances. Although, when we remove the edges of those contestants who play against their alliance, then the alliance has lower edge density. For Example, in *Survivor: Borneo*, the edge density of the *Tagi* alliance is 1.5 but when we narrow down the alliances to subsets of finalist, then the edge density decreases: that is, the edge density of the subset {Kelly,Richard} is 1 and {Richard,Rudy} is 0. In the *Fie-Long* alliance in *Survivor: China*, the subset {Amanda, Courtney, Todd} has edge density 0. We also see that in *Round Table*

alliance in Survivor: HHH, the subset $\{\text{Chrissy, Ryan, Devon}\}$ has edge density 0.

CHAPTER 4

Conclusion and Future Work

4.1. Conclusion

We introduced the notion of dynamic competition networks and studied their properties. In Chapter 2, we introduced structural balance theory. We defined a model of positive and negative relationships in a complete graph. We discussed, with examples, the balanced and unbalanced graphs in signed social networks. We also generalized the definition of signed complete graphs for groups of three nodes into an arbitrary numbers of nodes. We defined standard centrality metrics in network science and some graph theocratic tools that were used to state of Dynamic Competition Hypothesis (DCH). We formulated and presented the Dynamic Competition Hypothesis (DCH) which resolves dynamic competition networks arising from social networks into alliances, detects leaders, and measures the relative strength of alliances.

In Chapter 3, we discussed the background of U.S. reality television social game shows Survivor and Big Brother. We focused and analysed four seasons from Survivor and one season from Big Brother. We assessed the Dynamic Competition Hypothesis (DCH) with voting data for these seasons by using tools from network science. We observed that the DCH

predicted the alliances corresponding to near independent sets, CON scores accurately determine leaders of alliances, and leaders were detected via their CON scores and closeness.

4.2. Future Work

In our thesis, we only analysed four seasons from U.S. Survivor and one season from Big Brother. In future work, we will extract data from all U.S. and international seasons of Survivor and Big Brother. We will also verify the DCH more extensively for different data sets of Survivor and Big Brother within the lens of structural balance theory and social network analysis. A weakness with the hypothesis is that contestants who remain longer stay in a season acquire more prestige simply because of their survival. In particular, we note that all those players in Survivor and Big Brother who remain long in the game, have a greater chance to increase their CON-scores and other metrics.

In future work, we will analyse data at earlier stages of the formation of the network. We can examine the DCH in the other fields; for examples, we will consider food webs, geo political networks and subgraphs (with negative ties) of the signed networks.

It would be intriguing to invert the DCH to determine the low ranked members of dynamic competition networks. Additionally, it would be convenient to develop a mathematical model anticipating the evolution of

dynamic competition networks, which demonstrably simulate properties anticipate by the DCH.

Bibliography

- [1] M. Bastian, S. Heymann, M. Jacomy, Gephi: an open source software for exploring and manipulating networks, In: *Proceedings of the International AAAI Conference on Weblogs and Social Media*, 2009.
- [2] A. Bevelas, Communication patterns in task-oriented groups, *The Journal of the Acoustical Society of America* **22**(6): 725–730, 1950.
- [3] Big Brother Wiki http://bigbrother.wikia.com/wiki/Big_Brother_Wiki.
- [4] V. Boginski, S. Butenko, P.M. Pardalos, On structural properties of the market graph, In: A. Nagurney, editor, *Innovation in Financial and Economic Networks*, Edward Elgar Publishers, pp. 29–45.
- [5] A. Bonato, *A Course on the Web Graph*, American Mathematical Society Graduate Studies Series in Mathematics, Providence, Rhode Island, 2008.
- [6] U. Brandes, T. Erlebach (Eds.) *Network Analysis: Methodological Foundations*, LNCS 3418, Springer, 2005.
- [7] D. Cartwright, F. Harary, Structure balance: A generalization of Heider’s theory, *Psychological Review* **63**(5): 277–293, 1956.
- [8] J.A. Davis, Structural balance, mechanical solidarity, and interpersonal relations, *American Journal of Sociology* **68** (1963) 444–462.
- [9] D. Easley, J. Kleinberg, *Networks, Crowds, and Markets Reasoning about a Highly Connected World*, Cambridge University Press, 2010.
- [10] L.C. Freeman, S.P. Borgatti, D.R. White, Centrality in valued graphs: A measure of betweenness based on network flow, *Social Networks* **13**(2): 141–154, 1991.
- [11] M.M. Hakeem, K. Suzuki, Centrality Measures of Trade and Investment Networks, *Australian Academy of Accounting and Finance Reveiw* **1**(2): 103–118, 2015.
- [12] F. Harary, On the notion of balance of a signed graph, *Michigan Mathematical Journal* **2**(2): 143–146, 1953.
- [13] F. Heider, *The Psychology of Interpersonal Relations*, John Wiley & Sons, 1958.

- [14] F. Heider, Attitudes and cognitive organization, *Journal of Psychology* **21** (1946) 107—112.
- [15] A.N. Langville, C.D. Meyer, *Who's #1? The Science of Rating and Ranking*. Princeton University Press, 2012.
- [16] Survivor Wiki http://survivor.wikia.com/wiki/Main_Page.
- [17] S. Wasserman, , K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1995.
- [18] D.B. West, *Introduction to Graph Theory, 2nd edition*, Prentice Hall, 2001.