Developments in Data Analysis
A. Ferligoj and A. Kramberger (Editors)
Metodološki zvezki, 12, Ljubljana: FDV, 1996

Structural Balance and Partitioning Signed Graphs

Patrick Doreian and Andrej Mrvar*

Abstract

Signed graphs have been used frequently to model affect ties for social actors. In this context, structural balance theory has been used as a theory for the organization and form of these relations. Based on a graph theoretical formalization of this theory, for signed graphs, there are theorems stating that balanced structures have clear partitioned forms. We discuss the use of local optimization procedures as a method of obtaining these partitioned structures – or partitioned structures that are as close to a balanced partition as possible for structures that are not balanced. In addition, this method provides a measure of the extent to which a signed graph is imbalanced. For a time series of signed graphs, this permits a test of the balance hypothesis that human groups tend towards balance over time. Examples are provided of interpersonal ties in networks and institutionalized relations among collective units.

Since the pioneering work of Heider (1946) there has been an enduring concern with theories of structural balance. See, for example, Taylor (1970) and Forsythe (1990). Central to this formulation is the idea of triads made up of two actors, p and q, and some object, x. There is a *direct social relation* between the two actors and there are *unit formation* links between the actors and the object. The set of such triads is displayed in Figure 1.

The relation between the actors can be positive (for example, liking) or negative (disliking) and the unit formation relations can be positive (for example, approve) or negative (disapprove).

Definition 1 *A triad is balanced if the product of the signs that it contains is positive and is unbalanced when the product of its signs is negative.*

The structures in the top row of Figure 1 are all balanced while those in the second row are imbalanced.[1] Heider's key substantive insight is that people prefer balanced configurations to those that are imbalanced. For Heider, x can be another person, an event, an idea or a thing. Of interest here is the formalization of Heider's version of this theory by Cartwright and Harary (1956). In their formalization, the distinction between direct affect ties and unit formation ties was deleted – a

*Faculty of Social Sciences, University of Ljubljana, P.O. Box 47, 1109 Ljubljana, Slovenia
[1] As we use them, the terms imbalanced and unbalanced are interchangeable.

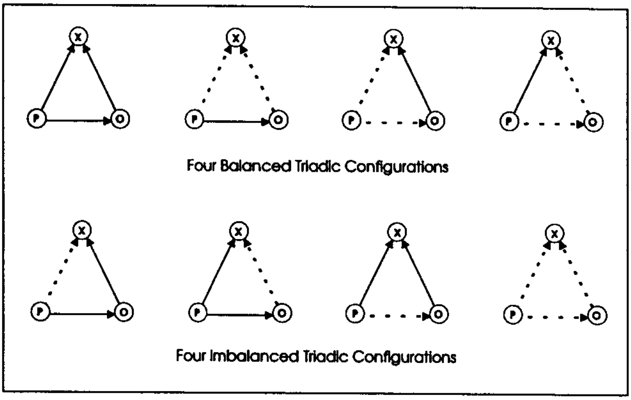


Figure 1: Triadic Configurations.

reasonable step if the 'object' can include other people. Indeed, the most frequent use of structural balance ideas is when p, q and x are all social actors – in which case, o is used as a label for the third social actor.

1 Signed graphs and structural balance

Cartwright and Harary (1956) used graphs and digraphs to formalize structural balance ideas.

Definition 2 *A signed digraph is an ordered pair, (G, σ), where (Batagelj, 1994:56):*

- $G = (V, R)$ is a digraph, without loops, having a set of vertices, V, and a set of arcs, $R \sqsubseteq V \times V$;
  and
- $\sigma : R \to \{+1, -1\}$ is a sign function. The positive arcs are assigned to $+1$ while the negative arcs are assigned to -1.

Clearly, such an assignment can be made for each of the triads shown in Figure 1. However, Cartwright and Harary extended the consideration of triads to include cycles and semicycles of any length.

Definition 3 *A cycle (semicycle) is positive if the product of the signs of its edges (arcs) is positive and negative if the product of the signs of its edges (arcs) is negative.*

Definition 4 *A signed graph (digraph) is balanced if all of its cycles (semicyles) are positive.*

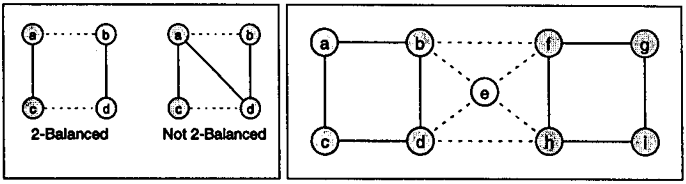


Figure 2: Two Undirected Signed Graphs (left) and A 3-Balanced Undirected Signed Graph (right).

In terms of these definitions, Cartwright and Harary (1956: 286) proved the, so-called, structure theorem:

Theorem 1 *If a signed graph (digraph) is balanced, then the set of vertices, V, can be partitioned into two subsets, called plus-sets, such that all of the positive edges (arcs) are between vertices within a plus-set and all negative edges (arcs) are between vertices in different plus-sets.*

This result is illustrated on the left side of Figure 2 where the graph on the left is balanced while the graph on the right is not. On the left, the plus sets are $\{a, c\}$ and $\{b, d\}$. The positive edge between a and c links vertices in $\{a, c\}$ and the other positive edge, between b and d, links vertices within $\{b, d\}$. The negative edges, between a and b and between c and d, link vertices in different plus sets. For the graph on right, the presence of the positive edge between a and d, means that the graph is not balanced. The triads $\{a, b, d\}$ and $\{a, c, d\}$ are both negative and there is no partition consistent with the structure theorem.

Consider next the graph shown on the right side of Figure 2, one that is not balanced in the above sense. In terms of triads (or 3-cycles) it is clear that $\{b, e, f\}$ and $\{d, e, h\}$ are both negative. Thus there is no partition into two plus-sets that is consistent with Theorem 1. Yet, viewing the structure, there is a clear partition. There is one subset, $\{a, b, c, d\}$, and another subset, $\{f, g, h, i\}$, that are plus-sets with only negative edges between them. And both of these subsets have negative edges with e. It seems reasonable, then, to think in terms of partitioning the vertices of a graph into more than two plus-sets. (In the imagery of human groups, the graph on the right side of Figure 2 describes two subgroups with internal positive ties and external negative ties plus another individual that is disliked by some members of each of the subgroups.) For the classical concept of structural balance there are two plus sets and we can label this as 2-balance. For the graph on the right side of Figure 2 we can talk of 3-balance where there are 3 plus-sets where all of the positive lines are within plus-sets and all of the negative lines are between plus-sets. This can be extended to a concept of k-balance in an obvious way with k the number of plus-sets in the partition of nodes. In effect, Davis (1967: 183) pursued this line of thought and established:

Theorem 2 *A signed graph (digraph) is k-balanced if and only if it contains no semicycles with exactly one negative edge (arc).*

For structural balance, the all negative triad is defined as imbalanced. Davis argued that it is just as reasonable to declare this as balanced. When this done, the structure Theorem 1 becomes the more general structure theorem 2. We label 'balance' for Theorem 2 as 'generalized balance' (i.e. for $k > 2$) and reserve 'structural balance' for $k = 2$.

The psychological balance theories claim that the (network) structures are either balanced or move towards a state of balance. For the latter case, measures of imbalance are needed if we want to trace movement towards balance. Two types of measures exist. One is couched in terms of cycles while the other is defined in terms of edges (or arcs). For cycles, if $C(G)$ is the total number of cycles in G and $C_+(G)$ is the number of positive cycles, then imbalance is measured as $\left(1 - \frac{C_+(G)}{C(G)}\right)$ (Cartwright and Harary, 1956: 288). Variations of this basic definition consider also the length of the cycles. Cycles beyond a certain length can be ignored or cycles can have (inverse) weights depending on their lengths. If we focus attention on edges (arcs), another type of measure of imbalance is the number of links that are inconsistent with the balance partition. For structural balance (Harary et al., 1965: 348), it is the number of links, *imb*, that have to change sign in order to create a balanced graph (and sign changes for all proper subsets of the links contributing to *imb* do not do so.)[2] We will use *imb* as a measure of imbalance for both structural and generalized balance.

2 Partitioning signed graphs

The structure theorems are stated for signed graphs that are balanced (for either structural or generalized balance). They are existance theorems and are silent about the composition of the plus-sets. For imbalanced graphs they have no relevance. Yet most empirical structures are not balanced. Doreian and Mrvar (1995) have proposed a method for obtaining clusters as close to balance (in either sense) as possible together with a minimized value of criterion function stated in terms of *imb*.

Let N be the number of *negative* ties (edges or arcs) that link directly vertices in the *same* plus-set. Similarly, let P be the number of *positive* ties that link vertices in *different* plus-sets. The obvious criterion function is $N + P$ and we seek partitions that minimize this value. The clustering problem is this: determine the clustering C^* for which

$$P(C^*) = \min_{C \in \Phi} P(C)$$

where C is a clustering of a given set of actors, V, Φ is the set of all possible clusterings and $P : \Phi \to \Re$ is a criterion function[3] where $\Re = N + P$. A slightly

[2] Another definition of imbalance is the number of edges that must be removed in order to create a balanced graph and is termed an edge-deletion measure. Harary et al. (1965) prove that the edge-deletion measure is the same as the sign change measure.

[3] This criterion function is simply *imb*.

Table 1: Partitioned matrices of the graph on the right side of Figure 2.

	a	b	c	d	e	f	g	h	i
a	.	1	1	.	.	.	.	.	.
b	1	.	.	1	-1	-1	.	.	.
c	1	.	.	1	.	.	.	.	.
d	.	1	1	.	-1	.	.	-1	.
e	.	-1	.	-1	.	-1	.	-1	.
f	.	-1	.	.	-1	.	1	1	.
g	.	.	.	.	.	1	.	.	1
h	.	.	.	-1	-1	1	.	.	1
i	.	.	.	.	.	.	1	1	.

	a	b	c	d	e	f	g	h	i
a	.	1	1	.	.	.	.	.	.
b	1	.	.	1	-1	-1	.	.	.
c	1	.	.	1	.	.	.	.	.
d	.	1	1	.	-1	.	.	-1	.
e	.	-1	.	-1	.	-1	.	-1	.
f	.	-1	.	.	-1	.	1	1	.
g	.	.	.	.	.	1	.	.	1
h	.	.	.	-1	-1	1	.	.	1
i	.	.	.	.	.	.	1	1	.

more general criterion function is

$$\Re = \alpha N + (1 - \alpha) P$$

where $0 \leq \alpha \leq 1$. For $\alpha = 0.5$, both types of error are weighted equally. With $0 \leq \alpha < 0.5$, positive errors are more important while, for $0.5 < \alpha \leq 1$, negative errors are more important. Here, we focus on the case where $\alpha = 0.5$.

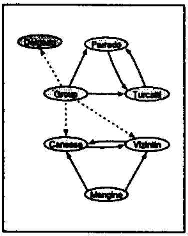
Consider the graph shown on the right side of Figure 2. If we consider partitions into 2 plus-sets, there are three such partitions, all with the minimum error of 2 (with $\alpha = 0.5$). These partitions are

$\{a, b, c, d, e\}$, $\{f, g, h, i\}$
$\{a, b, c, d\}$, $\{f, g, h, i, e\}$
$\{a, b, c, d, f, g, h, i\}$, $\{e\}$

The permuted and partitioned matrix corresponding to the first of these partitions is shown on the left side of Table 1.

The ties that contribute to the error are the negative links between b and e and between d and e. If the second partition is considered, the ties contributing to the error are the negative ties between e and f and between e and h. For the third partition, the ties contributing to the error do not involve e. Instead, they are the negative ties between b and f and between d and h. In general, we can expect that there will be more than one partition with a minimum number of errors and that, depending of the partitions identified, different sets of links can be identified as being inconsistent with a form of k-balance. However, there are cases where a partition is unique. Such a case is the partition of the graph shown on the right side of Figure 2 into three plus-sets where there are no ties inconsistent with 3-balance (the minimized value of the criterion is 0). The partition is shown on the right side of Table 1.

The local optimization procedure is a relocation procedure. It starts with the specification of the number of clusters for the partition. A random partition into the given number of clusters is then generated. Following this, vertices can be moved from one cluster to another or pairs of vertices can be interchanged between clusters.



		1	2	3	4	5	6	7
Parrado	1	0	1	0	0	0	0	0
Turcatti	2	1	0	0	0	0	0	0
Delgado	3	0	0	0	0	0	0	0
Canessa	4	0	0	0	0	1	0	0
Vizintin	5	0	0	0	1	0	0	0
Mangino	6	0	0	0	1	1	0	0
Group	7	1	1	-1	-1	-1	0	0

Figure 3: Directed Signed Graph for Andes Survivors.

For each such change, the criterion function is evaluated. If a relocating change leads to a decrease of the criterion function, the new partition is retained and the process continues until the criterion function cannot be lowered. This is repeated for many random starts and the partitions with the lowest value of $\Re$ are retained. We select the smallest value of the criterion function across all values of k. As noted above, there may be multiple partitions with the same minimized value of the criterion function. See Doreian and Mrvar (1995) for further details.

We now illustrate this partitioning procedure with some empirical examples.

3 Empirical examples

3.1 Human groups

3.1.1 Andes survival group

Read (1974) describes a sequence of events following a aeroplane crash high in the Andes mountains. A rugby team was traveling on the plane when it went down and, with the flight crew killed, the survivors had to fend for themselves. The pattern of relationships among the survivors changed during the ordeal. Forsyth (1990:126) presents some sociometric information in the form a signed graph. This is displayed in Figure 3 together with the data matrix. There are six named actors plus the "other group members". For the purpose of this demonstration, we consider the remaining actors as 'group' in a single node.

The structure of the graph is simple – indeed, one could do the partitoning analysis by hand. There is one plus-set of {Parrado, Turcatti, Group}, another with {Canessa, Vizintin, Mangino} and a singleton {Delgado}. Stated in this form, the example has an exact 3-balance partition. Clearly, it is possible to group {Delgago} with {Canessa, Vizintin, Mangino} for an exact 2-balance partition. These are shown in Table 2. Both partitions are exactly balanced and, of the two, the 3-balance solution seems the best as Delgado has no positive ties with any of the other actors.

We next consider a more complicated example where there are no exactly k-balanced partitions and there are enough actors so that a visual examination is of

Table 2: Exact 2-Balanced and 3-Balanced Partitions of the Andes Survival Group.

Actor	id	1	2	7	3	4	5	6
Parrado	1	.	1	.	.	.	.	.
Turcatti	2	1	.	.	.	.	.	.
Group	7	1	1	.	-1	-1	-1	.
Delgado	3	.	.	.	.	.	.	.
Canessa	4	.	.	.	.	.	1	.
Vizintin	5	.	.	.	.	1	.	.
Mangino	6	.	.	.	.	1	1	.

Actor	id	1	2	7	3	4	5	6
Parrado	1	.	1	.	.	.	.	.
Turcatti	2	1	.	.	.	.	.	.
Group	7	1	1	.	-1	-1	-1	.
Delgado	3	.	.	.	.	.	.	.
Canessa	4	.	.	.	.	.	1	.
Vizintin	5	.	.	.	.	1	.	.
Mangino	6	.	.	.	.	1	1	.

little value. Here, some algorithm is needed to get partitions as close to k-balance and a measure of imbalance.

3.1.2 Sampson monastery data

The Sampson (1968) monastery data have been used frequently as a test bed for network analytic procedures, a usage that we continue. Sampson obtained network data for a variety of ties (likes, esteems, influences and sanctions) through time for a group of young men undergoing training in a monastery. While the full study had four time periods, only the last three, labeled T_2, T_3 and T_4, had the same set of actors. Doreian and Mrvar (1995) considered these time points. Here, we consider only the liking relation and T_2. For both the positive and negative citations, Sampson asked the trainee monks to rank order these like and dislike ties. A citation of $+3$ goes to the most liked actor while a citation of -3 goes to the most disliked other actor.[4] These data are shown in Table 3.

If these ties are binarized to put the data in a form compatible with Definition 2, we get the data as shown in Table 4. (We report the analysis of the valued data shortly.) When the partitioning procedure discussed in Section 2 is used for these data, the partition with the lowest value of the criterion function is the one shown in Table 5.

The same procedure can be applied to the valued data of Table 3 if we assume that the rank scores (-3, -2, -1, 1, 2, 3) can be summed. The criterion function $\Re$ can be computed by summing ranks. When this is done, the partition shown in Table 6 results. A comparison of the two partitions, in terms of which actors belong to the identified plus-sets, shows that they are identical.

Doreian and Mrvar (1995) repeated the analysis, using the valued data, for the other time points. The values obtained for each time point for partitions with up to five clusters are shown in Table 7. It is clear that the smallest value of the criterion function is for the 3-cluster partitions and that this value declines through time. We note also that the measure for structural balance (the 2-cluster partitions) declines

[4]As is clear from inspecting the table, not all of the monks in training conformed with this request. Bonaventure and Romuald, for example, have no negative ties among their nominations.

Table 3: Valued Affect Ties in the Sampson Monastery Data at T_2.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	JohnBosco	0	0	2	0	3	-2	-1	0	0	-3	0	0	0	1	0	0	0	0
2	Gregory	3	0	0	0	0	0	2	0	0	-1	0	0	-3	1	0	0	-2	0
3	Basil	2	3	0	-1	0	0	0	-3	-2	0	0	0	0	0	0	0	1	0
4	Peter	0	0	-2	0	3	1	-3	0	0	2	0	0	0	0	0	0	0	-1
5	Bonaventure	0	0	0	3	0	0	0	0	0	0	2	0	1	0	0	0	0	0
6	Berthold	1	0	0	3	0	0	-3	-1	2	0	0	-2	0	0	0	0	0	0
7	Mark	0	2	0	-3	-1	-2	0	1	0	0	0	0	0	0	0	3	0	0
8	Victor	3	2	-3	0	0	0	0	0	1	0	0	0	0	-2	0	0	-1	0
9	Ambrose	0	0	-3	0	2	0	0	3	0	0	0	0	0	0	0	1	-2	-1
10	Romuald	0	0	0	3	0	0	0	1	0	0	0	0	0	2	0	0	0	0
11	Louis	0	0	-1	0	3	0	0	1	0	0	0	0	-3	2	0	0	-2	0
12	Winfrid	3	2	-1	-3	0	0	0	0	0	0	-2	0	0	1	0	0	0	0
13	Amand	0	-3	0	0	2	-2	1	0	0	0	0	-1	0	0	0	0	0	3
14	Hugh	3	0	0	0	0	0	0	-2	0	0	1	2	-3	0	2	0	-1	0
15	Boniface	3	2	-2	0	0	0	0	0	0	0	0	0	-3	1	0	0	-1	-1
16	Albert	1	2	0	0	0	0	3	0	0	0	0	0	-1	0	0	0	-3	-2
17	Elias	0	0	3	-3	-2	0	0	0	0	0	-1	0	2	0	0	0	0	1
18	Simplicius	2	3	0	-3	0	-2	1	0	0	0	0	0	0	0	0	-1	0	0

Table 4: Binary Affect Ties in the Sampson Monastery Data at T_2.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	JohnBosco	0	0	1	0	1	-1	-1	0	0	-1	0	0	0	1	0	0	0	0
2	Gregory	1	0	0	0	0	0	1	0	0	-1	0	0	-1	1	0	0	-1	0
3	Basil	1	1	0	-1	0	0	0	-1	-1	0	0	0	0	0	0	0	1	0
4	Peter	0	0	-1	0	1	1	-1	0	0	1	0	0	0	0	0	0	0	-1
5	Bonaventure	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0
6	Berthold	1	0	0	1	0	0	-1	-1	1	0	0	-1	0	0	0	0	0	0
7	Mark	0	1	0	-1	-1	-1	0	1	0	0	0	0	0	0	0	1	0	0
8	Victor	1	1	-1	0	0	0	0	0	1	0	0	0	0	-1	0	0	-1	0
9	Ambrose	0	0	-1	0	1	0	0	1	0	0	0	0	0	0	0	1	-1	-1
10	Romuald	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
11	Louis	0	0	-1	0	1	0	0	1	0	0	0	0	-1	1	0	0	-1	0
12	Winfrid	1	1	-1	-1	0	0	0	0	0	0	-1	0	0	1	0	0	0	0
13	Amand	0	-1	0	0	1	-1	1	0	0	0	0	-1	0	0	0	0	0	1
14	Hugh	1	0	0	0	0	0	0	-1	0	0	1	1	-1	0	1	0	-1	0
15	Boniface	1	1	-1	0	0	0	0	0	0	0	0	0	-1	1	0	0	-1	-1
16	Albert	1	1	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	-1	-1
17	Elias	0	0	1	-1	-1	0	0	0	0	0	-1	0	1	0	0	0	0	1
18	Simplicius	1	1	0	-1	0	-1	1	0	0	0	0	0	0	0	0	-1	0	0

Table 5: Plus-set Partition (for $k = 3$) of the Actors at T_2 for Binary Ties.

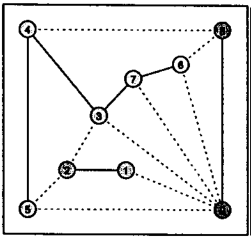
Actor	id	1	2	7	12	14	15	16	3	13	17	18	4	5	6	8	9	10	11
John Bosco	1	.	.	-1	.	1	.	.	1	.	.	.	.	1	-1	.	.	-1	.
Gregory	2	1	.	1	.	1	.	.	.	-1	-1	.	.	.	.	.	.	-1	.
Mark	7	.	1	.	.	.	.	1	.	.	.	.	-1	-1	-1	1	.	.	.
Winfrid	12	1	1	.	.	1	.	.	-1	.	.	.	-1	.	.	.	.	.	-1
Hugh	14	1	.	.	1	.	1	.	.	-1	-1	.	.	.	.	-1	.	.	1
Boniface	15	1	1	.	.	1	.	.	-1	-1	-1	-1	.	.	.	.	.	.	.
Albert	16	1	1	1	.	.	.	.	.	-1	-1	-1	.	.	.	.	.	.	.
Basil	3	1	1	.	.	.	.	.	.	.	1	.	-1	.	.	-1	-1	.	.
Amand	13	.	-1	1	-1	.	.	.	.	.	.	1	.	1	-1	.	.	.	.
Elias	17	.	.	.	.	.	.	.	1	1	.	1	-1	-1	.	.	.	.	-1
Simplicius	18	1	1	1	.	.	.	-1	.	.	.	.	-1	.	-1	.	.	.	.
Peter	4	.	.	-1	.	.	.	.	-1	.	.	-1	.	1	1	.	.	1	.
Bonaventure	5	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	1
Berthold	6	1	.	-1	-1	.	.	.	.	.	.	.	1	.	.	-1	1	.	.
Victor	8	1	1	.	.	-1	.	.	-1	.	-1	.	.	.	.	.	1	.	.
Ambrose	9	.	.	.	.	.	.	1	-1	.	-1	-1	.	1	.	1	.	.	.
Romuald	10	.	.	.	.	1	.	.	.	.	.	.	1	.	.	1	.	.	.
Louis	11	.	.	.	.	1	.	.	-1	-1	-1	.	.	1	.	1	.	.	.

Table 6: Plus-set Partition (for k=3) of the Actors at T_2 for Valued Ties.

Actor	id	1	2	7	12	14	15	16	3	13	17	18	4	5	6	8	9	10	11
John Bosco	1	.	.	-1	.	1	.	.	2	.	.	.	.	3	-2	.	.	-3	.
Gregory	2	3	.	2	.	1	.	.	.	-3	-2	.	.	.	.	.	.	-1	.
Mark	7	.	2	.	.	.	.	3	.	.	.	.	-3	-1	-2	1	.	.	.
Winfrid	12	3	2	.	.	1	.	.	-1	.	.	.	-3	.	.	.	.	.	-2
Hugh	14	3	.	.	2	.	2	.	.	-3	-1	.	.	.	.	-2	.	.	1
Boniface	15	3	2	.	.	1	.	.	-2	-3	-1	-1	.	.	.	.	.	.	.
Albert	16	1	2	3	.	.	.	.	.	-1	-3	-2	.	.	.	.	.	.	.
Basil	3	2	3	.	.	.	.	.	.	.	1	.	-1	.	.	-3	-2	.	.
Amand	13	.	-3	1	-1	.	.	.	.	.	.	3	.	2	-2	.	.	.	.
Elias	17	.	.	.	.	.	.	.	3	2	.	1	-3	-2	.	.	.	.	-1
Simplicius	18	2	3	1	.	.	.	-1	.	.	.	.	-3	.	-2	.	.	.	.
Peter	4	.	.	-3	.	.	.	.	-2	.	.	-1	.	3	1	.	.	2	.
Bonaventure	5	.	.	.	.	.	.	.	.	1	.	.	3	.	.	.	.	.	2
Berthold	6	1	.	-3	-2	.	.	.	.	.	.	.	3	.	.	-1	2	.	.
Victor	8	3	2	.	.	-2	.	.	-3	.	-1	.	.	.	.	.	1	.	.
Ambrose	9	.	.	.	.	.	.	1	-3	.	-2	-1	.	2	.	3	.	.	.
Romuald	10	.	.	.	.	2	.	.	.	.	.	.	3	.	.	1	.	.	.
Louis	11	.	.	.	.	2	.	.	-1	-3	-2	.	.	3	.	1	.	.	.

Table 7: Values of Criterion Function for Number of Clusters and Time.

Clusters	T_2	T_3	T_4
1	48.5	48.0	47.0
2	21.5	16.0	12.5
3	17.5	11.0	10.5
4	19.0	13.5	12.5
5	20.5	16.0	15.0



Clan	1	2	3	4	5	6	7	8	9
1 Mulina I	0	1	0	0	0	0	0	0	-1
2 Mulina II	1	0	-1	0	-1	0	0	0	0
3 Ainaona	0	-1	0	1	0	0	1	0	-1
4 Malabuabua	0	0	1	0	1	0	0	-1	0
5 Foloyai	0	-1	0	1	0	0	0	0	-1
6 Lulavile I	0	0	0	0	0	0	1	-1	-1
7 Lulavile II	0	0	1	0	0	1	0	0	-1
8 Nouneya	0	0	0	-1	0	-1	0	0	1
9 Iwaoyana	-1	0	-1	0	-1	-1	-1	1	0

Figure 4: Kuluana example.

in the same fashion. As such, and concluded by Doreian and Mrvar, this provides strong support for the *substantive* balance hypothesis.[5]

The partition is the same for each time point and is largely consistent with the partition identified by Sampson and obtained via blockmodeling all of the (stacked) ties (Breiger et al., 1975). John Bosco, Gregory, Mark, Winfrid, Hugh, Boniface and Albert are the Young Turks. Sampson identified Basil, Simplicius and Elias as the Outcasts. The generalized balance partition includes Amand with them. Part of the ethnography provided by Sampson suggests that Amand belongs here as he publically supported Basil in an election and cast the only vote for Basil in the subsequent voting. The pattern of ties in the partitioned matrix support this location. The remaining actors belong to the Loyal Opposition identified by Sampson. When the partition identified by Sampson is used as the start point, it yields a higher *imb* score than for the partition shown in Table 7, suggesting that this partition does *not* have the best fit to the data.

3.2 Institutionalized signed structures

3.2.1 Clans and subclans of Kaluana

We turn now to examine two examples drawn from anthropological sources. Both concern arrangements found in the Highlands of New Guinea. Young (1971) provides an example of exchange relations where there are positive and negative ties among clans or major clan segments for the Kaluana. The data come from the discussion of Hage and Harary (1991: 53). There are nine units as shown in Figure 4 with the matrix of ties on the right. While there are no balanced partitions for 2-balance, there is one such partition for 3-balance. This partition is obvious from Figure 4 and is described in Table 8.

[5]This does need to be qualified as the data were collected retrospectively. Also, for T_4 there are two partitions having the minimized value of the criterion function.

Table 8: Plus-set Partition for the Kaluana Example.

		Clans								
Clan	id	1	2	3	4	5	6	7	8	9
Mulina I	1	.	1	.	.	.	.	.	.	-1
Mulilna II	2	1	.	-1	.	-1	.	.	.	.
Ainaona	3	.	-1	.	1	.	.	1	.	-1
Malabuabua	4	.	.	1	.	1	.	.	-1	.
Foloyai	5	.	-1	.	1	.	.	.	.	-1
Lulavile I	6	.	.	.	.	.	.	1	-1	-1
Lulavile II	7	.	.	1	.	.	1	.	.	-1
Nouneya	8	.	.	.	-1	.	-1	.	.	1
Iwaoyana	9	-1	.	-1	.	-1	-1	-1	1	.

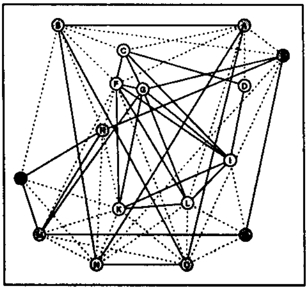


Figure 5: Gahuku-Gama Subtribes of Highland New Guinea.

3.2.2 Gahuku-Gama subtribes of highland New Guinea

Next, we consider a more complicated example. Read (1954) provides data and an extensive discussion of some of the cultures of highland New Guinea. A focus of his discussion was a set of political alliances and oppositions among the Gahuku-Gama subtribes. The alliances and oppositions depicted by Read are shown in Figure 5. These tribes are distributed in a particular area and engage in prolonged warfare. "Warfare ... is that activity which characterizes the tribes of the Gahuku-Gama as a whole and which differentiates them from groups in other socio-geographic regions" (1954:39). "Traditional enemies were separated by only a few miles. Scattered raids and organized, concerted attacks were constantly expected" (1954:22). This is an institutionalized form of warfare that is expected to continue indefinately. As such, it is a quite stable arrangement.

Read's account points out that there is an element of expediency in some of the political arrangements and some positive ties are broken (and presumeably some new negative ties are created). The whole system is stable with only minor changes at any time point. These minor changes do not change the basic structure of the

Table 9: Plus-set Partition for 3-Balance for Gahuku-Gama Example.

Subtribe	id	1	2	14	15	3	4	6	7	9	11	12	5	8	10	13	16
A	1	.	1	1	1	-1	-1	-1	.	.	.	-1	-1	.	.	.	.
B	2	1	.	1	1	-1	.	-1	.	.	.	.	-1	-1	-1	.	.
N	14	1	1	.	1	.	.	.	.	.	-1	-1	-1	-1	-1	-1	.
O	15	1	1	1	.	.	.	-1	.	.	-1	-1	-1	.	.	-1	-1
C	3	-1	-1	.	.	.	1	1	1	1	.	.	.	.	.	.	.
D	4	-1	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.
F	6	-1	-1	.	-1	1	.	.	1	1	1	1	.	-1	.	-1	.
G	7	.	.	.	.	1	.	1	.	1	1	1	1	.	.	1	.
I	9	.	.	.	.	1	1	1	1	.	1	1	.	.	.	.	-1
K	11	.	.	-1	-1	.	.	1	1	1	.	1	.	-1	-1	-1	.
L	12	-1	.	-1	-1	.	.	1	1	1	1	.	.	.	.	.	-1
E	5	-1	-1	-1	-1	.	.	.	1	.	.	.	.	1	.	.	1
H	8	.	-1	-1	.	.	.	-1	.	.	-1	.	1	.	1	1	.
J	10	.	-1	-1	.	.	.	.	.	.	-1	.	.	1	.	1	.
M	13	.	.	-1	-1	.	.	-1	1	.	-1	.	.	1	1	.	1
P	16	.	.	.	-1	.	.	.	.	-1	.	-1	1	.	.	1	.

system. Consistent with balance theory, such stability should be close to balance – but not necessarily in terms of structural balance. When these ties were examined, the partition with the lowest value of the criterion function is for 3-balance and is shown in Table 9. The values of the criterion function is 2. A closer examination of the distribution of ties reveals that all of the ties contributing to the criterion function are associated with subtribe G. Read's account emphasises a condition of universal warfare and that all of the tribes are warlike. Yet subtribe G has only positive links (to some of the other tribes). This suggests that G is an unusual subtribe worthy of further attention or that there may be some errors in the data. In either interpretation, the partitioning method has located something of interest. See also the discussion of Hage and Harary (1983).

3.3 Discussion

We have provided a general method for partitioning signed graphs in terms of k-balance. These partitions are those as close as is possible to exact balance and the method provides a measure of imbalance, for any value of k, for a signed graph (network). The set of ties inconsistent with a particular k-balance are identified also[6]. For $k = 2$ we have traditional structural balance and for $k > 2$ we have a generalized balance. This method can be used for all signed graphs and is general. For empirical analyses of social networks (Section 3.1) or institutionalized networks (Section 3.2) it is likely that the values of k will be quite small. In the examples considered, k was never greater than 3. Other examples are considered in Doreian, Batagelj and Ferligoj (1995). For the empirical examples considered thus far, some ideas meriting further attention include:

[6]Of course, if there are multiple partitions with the same minimized value of the criterion function, the identified links are not unique.

- Institutionalized signed networks are closer to a balanced state compared to networks of individuals.

- For human groups and systems the relevant values of k are small.

- When the ties that are inconsistent with balance, for any value of k, by far, the positive ties between plus-sets are far more numerous than negative ties within plus-sets. This suggests that the negative ties within plus-sets are far more costly and are unlikely to be tolerated. In contrast, positive ties between plus sets seem like idiosyncratic events that can be tolerated. (They may even prove useful in conflict reduction efforts.)

- The characterization of a graph (network) that is made with regard to balance using this method is of a macro-level property of the graph (network). Doreian and Mrvar (1995) argue that the links identified as inconsistent with balance need not be the ones that change over time. Characterizing the mechanisms for such changes is an open problem.

References

[1] Batagelj V. (1994): Semirings for Social Network Analysis. *Journal of Mathematical Sociology*, **19**, 53-68.

[2] Breiger R.L., Boorman S.A., and Arabie P. (1975): An Algorithm for Clustering Relational Data with Applications to Social Network Analysis and Comparison with Multidimensional Scaling. *Journal of Mathematical Psychology*, **12**, 328-383.

[3] Cartwright D. and Harary F. (1956): Structural balance: A generalization of Heider's theory. *Psychological Review*, **63**, 277-292.

[4] Davis J. A. (1967): Clustering and structural balance in graphs. *Human Relations*, **20**, 181-187.

[5] Doreian P., Batagelj V., and Ferligoj A.: Partitioning networks based on generalized concepts of equivalence. *Journal of Mathematical Sociology*, **19**, 1-27.

[6] Doreian P. and Mrvar A. (1996): A partitioning approach to structural balance. *Social Networks*, **18**, 149-168.

[7] Forsyth D.R. (1990): *Group Dynamics*, Pacific Grove: Brooks/Cole.

[8] Hage P. and Harary F. (1983): *Structural Models in Anthropology*, Cambridge: Cambridge University Press.

[9] Hage P. and Harary F. (1991): *Exchange in Oceania: A Graph Theoretic Analysis*, Oxford: Clarendon Press.

[10] Harary F., Cartwright D., and Norman R. (1965): *Structural Models: An Introduction to Directed Graphs*, New York: John Wiley and Sons.

[11] Heider F. (1946): Attitudes and Cognitive Organization. *Journal of Psychology*, **21**, 107-112.

[12] Read K.E. (1954): Cultures of the Central Highlands, New Guinea. *Southwestern Journal of Anthropology*, **10**, 1-43.

[13] Read P.P. (1974): *Alive*, New York: Avon.

[14] Sampson S.F. (1968): *A Novitiate in a Period of Change: An Experimental and Case Study of Social Relationships*, Cornell University, PhD thesis.

[15] Taylor H.F. (1970): *Balance in Small Groups*, New York: Van Nostrand Reinhold.

[16] Young M.W. (1971): *Fighting with Food*, Cambridge: Cambridge University Press.