

Signed networks

Milka Gruber^a, Dominik Hrovat^a, Simon Goričar^a, and Konstantin Bojchevski^a

^aUniversity of Ljubljana, Faculty of Computer and Information Science, Večna pot 113, SI-1000 Ljubljana, Slovenia

The manuscript was compiled on June 8, 2025

Social media has provided new ways of analyzing our interpersonal relationships. While traditional social networks represent connections between individuals, they overlook a critical dimension - the nature of these relationships. The distinction between friendly and antagonistic interactions provides nuances crucial to our understanding of social dynamics. Based on decades of sociological and psychological research, signed networks have emerged as the standard for modeling relationships with opposing polarities. In this paper, we discuss the motivation for using signed networks and provide a mathematical background for understanding their structure, which we discuss through concepts of balance and status theory. We present the state-of-the-art methods for analyzing them and experiment on a few practical examples.

Problem definition In many social networks, relationships are not simply present or absent, but they carry a positive or negative sentiment. For example, in a small community, every individual has a positive or a negative opinion about any other individual. We cannot analyse such networks as normal undirected networks because we cannot store that information in such a graph. The idea is instead to expand the definition of networks by assigning a positive or negative sign to every edge. We will call such networks *signed networks* (1).

Note

This is not a weighted network, because in a weighted network, we do not allow negative weights.

Let $G = (V, E, \sigma)$ be an undirected signed network, where V is the set of vertices, E a set of edges and σ is the sign function $\sigma : E \rightarrow \{-1, +1\}$. The set of edges is represented by E . In E , we have negative edges, denoted by m^- , and positive edges, denoted by m^+ , adding up to a total of $m = m^+ + m^-$ edges. We denote the graph density by $\delta = \frac{2m}{n(n-1)}$. The symmetric signed adjacency matrix and the unsigned adjacency matrix are denoted by $A = [a_{uv}]_{uv \in E}$ and $|A| = [|a_{uv}|]_{uv \in E}$ respectively. Their entries are defined as follows:

$$a_{uv} = \begin{cases} \sigma(u, v) & \text{if } (u, v) \in E, \\ 0 & \text{if } (u, v) \notin E. \end{cases} \quad [1]$$

$$|a_{uv}| = \begin{cases} 1 & \text{if } (u, v) \in E, \\ 0 & \text{if } (u, v) \notin E. \end{cases} \quad [2]$$

The positive degree and negative degree of node i are denoted by k_i^+ and k_i^- respectively, representing the number of positive and negative edges incident on node i . They are calculated as:

$$k_i^+ = \frac{1}{2} \left(\sum_j |a_{ij}| + \sum_j a_{ij} \right), \quad k_i^- = \frac{1}{2} \left(\sum_j |a_{ij}| - \sum_j a_{ij} \right).$$

The degree of node i , represented by k_i , equals the number of edges incident on node i and is calculated by:

$$k_i = k_i^+ + k_i^- = \sum_j |a_{ij}|.$$

A walk of length ℓ in G is a sequence of nodes $v_0, v_1, \dots, v_{\ell-1}, v_\ell$ such that for each $i = 1, 2, \dots, \ell$ there is an edge from v_{i-1} to v_i . If $v_0 = v_\ell$, the sequence is called a *closed walk* of length ℓ . If all the nodes in a closed walk are distinct except the endpoints, it is called a *cycle* of length ℓ .

Definition

The sign of a cycle is the product of the signs of its edges.

This means that a triangle (i.e. a cycle with 3 nodes) is positive, if and only if, it has even amount of negative signs.

Definition

A cycle is *balanced* if its sign is positive, and *unbalanced* otherwise.

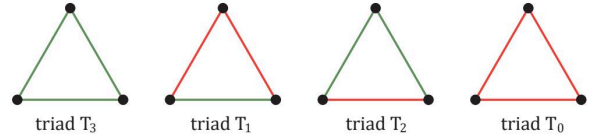


Fig. 1. Triangles T_1 and T_3 are balanced, while T_0 and T_2 are unbalanced. Red represents a negative sign and green represents a positive sign.

The total number of balanced cycles of length ℓ is denoted by O_ℓ^+ . Similarly, O_ℓ^- denotes the total number of unbalanced cycles (closed walks) of length ℓ . The total number of cycles of length ℓ is represented by:

$$O_\ell = O_\ell^+ + O_\ell^-.$$

We use $G_r = (V, E, \sigma_r)$ to denote a reshuffled graph in which the sign function σ_r is a random mapping of E to $\{-1, +1\}$ that preserves the number of negative edges. The reshuffling process preserves the underlying graph structure.

Balance theory

Structural balance theory examines how relationships among three individuals can be signed (positive for friends, negative for foes), and which configurations are more commonly observed in real networks. According to the theory formulated

by Heider in the 1940s (2) in the field of social psychology, balanced triangles tend to occur more regularly than unbalanced triangles. In a social network we can reason that a friend of my friend is usually also my friend (Figure 1, T_3), or a friend of my enemy is usually my enemy (Figure 1, T_1). These two situations are more plausible than two enemies that have a common friend (Figure 1, T_2) or that three people are mutual enemies (Figure 1, T_0). Balance theory states that a signed social network tends to have a higher proportion of balanced triads.

Weak structural balance. Davis (3) proposed that this balance theory is too strict. He proposed that only triads with one negative edge (Figure 1, T_2) are unbalanced. In this way we get the notion of weak structural balance.

Mathematical background of balance theory

In this section, we briefly examine the first mathematical results obtained by Harary (4). Harary defined a signed graph as balanced if all of its cycles are positive. Based on this he developed a characterization of balance for signed graphs. For signed graphs that have all possible edges the following theorem holds.

Theorem 1. Say that $G = (V, E, \sigma)$ is a complete signed graph. Then G is balanced if and only if there exists such a partition of vertices V into two disjoint sets V_1 and V_2 (where one can be empty), such that all edges that link vertices inside the sets are positive and all edges between V_1 and V_2 are negative.

Proof of this theorem can be found in the article (4). We observe that every subgraph of a balanced graph is balanced. This follows from the fact that every cycle in a subgraph is also a cycle in the original graph and the sign of the cycle is inherited from the original graph. Hence all cycles in the subgraph will be positive. The following theorem is an important result that states the previous characterization can be extended to arbitrary signed graphs.

Theorem 2 (Structure theorem). A signed graph $G = (V, E, \sigma)$ is balanced if and only if there exists such a partition of vertices V into two disjoint sets V_1 and V_2 (where one can be empty), such that every negative edge has one point in V_1 and another in V_2 and every positive edge lies inside one of the sets V_1 or V_2 .

Proof. Let's start with an arbitrary signed graph, which has a partition like the one described in the theorem. Let's construct a complete signed graph in the following way. Take two vertices u and $v \in V$, such that $uv \notin E$. If u and v both lie in the same subset V_i , add a positive edge uv to the new signed graph G' . If they are in different subsets, add a negative edge uv . By theorem 1 the obtained complete signed graph is balanced. And because G is a subgraph of a complete balanced signed graph, G is also balanced.

We prove the other direction with induction on the number of edges. We show that it is always possible (until the graph is complete) to add one additional unique edge to a balanced signed graph such that the graph remains balanced. In this way we obtain a complete balanced graph, which by theorem 1 implies the existence of the appropriate sets V_1 and V_2 . Let u and v be any non-adjacent vertices in the graph G . All

paths joining vertices u and v have the same sign. We now add an edge uv with the same sign determined by the paths. Every cycle that was formed in this way will be positive, so the resulting graph is balanced. $\square$

We can interpret this theorem in a social setting in the following way: a social network is balanced if it consists entirely of two groups such that if two people have a relationship inside the group, it is a friendly one, and any relationship between people in different groups is antagonistic. If this holds in a signed network, it is obvious that every triangle will be balanced. However, this will rarely happen in real signed networks. And even if it does, there is no reason for the network to be partitioned only into two opposing factions. The following real-life example is depicted in figure 2. Reaserchers (5) have proposed that such a structure consisting of multiple parts does not destabilize balance.

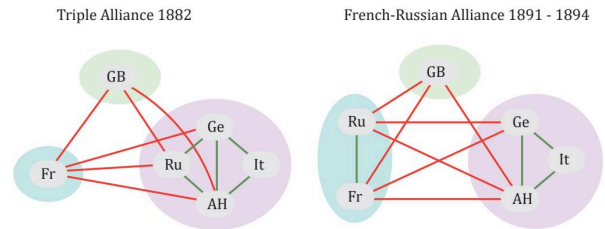


Fig. 2. Illustration of the tripartite structure formed by European powers in the 19th century. Green lines represent friendships and red lines indicate a negative relationship.

Analyzing balance based only on triads

For analyzing balance in a signed network as described in the article (1), we use the definition of a signed network, where each present edge has to have either a positive or negative sign. We denote the fraction of positive edges in the network with $p = \frac{m^+}{m}$. We begin by counting the occurrences of triads T_0, T_1, T_2 and T_3 . Let us denote $|T_i|$ as the number of occurrences of the triad T_i in the signed network. We denote the fraction of a triad as $p(T_i) = \frac{|T_i|}{\Delta}$, where Δ is the number of all triangles in the network.

To see if the values of $p(T_i)$ are statistically significant, we have to compare our empirical data to a signed network of the same structure, where the values of negative and positive signs are randomly reshuffled. The signed network G_r should preserve the fractions p and $p(T_i)$ for all triads. After performing such reshuffling, we denote the fraction of triads of type T_i in G_r with $p_0(T_i)$. We say that a triad T_i is overrepresented in a signed network if $p(T_i) > p_0(T_i)$ and underrepresented if $p(T_i) < p_0(T_i)$. The surprise $s(T_i)$ is the number of standard deviations by which the actual quantity of the T_i triad differs from the ones in the reshuffling model.

Based on this analysis of triads, we define the following measure of balance (6).

Definition

The *balance index* β of the signed network is the proportion of triads that are balanced. We use the previous notation $\beta = \frac{|T_1|+|T_2|}{\Delta} = p(T_1) + p(T_3)$ for strong structural balance, and $\beta^* = \frac{|T_1|+|T_2|+|T_3|}{\Delta} = p(T_1) + p(T_2) + p(T_3)$ for weak structural balance.

However, it is important to note that some researchers have concluded that analysis solely based on this measure is incomplete (5). They propose alternative approaches like considering the smallest number of edges whose reversal of their sign leads to a balanced network. This measure makes sense in a social setting, where we can ask ourselves how many actors are creating tension and preventing an otherwise balanced state. The following example in figure 3 shows how just one edge can create tension.

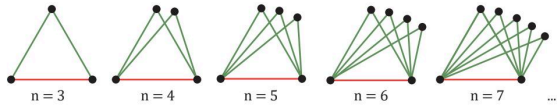


Fig. 3. Illustration of a simple signed network, where the number of unbalanced triads in a network with n vertices is $n - 2$. The balance index of this network is 0 for every $n \geq 3$. If we reversed the sign of just one edge the balance index would be 1 for every $n \geq 3$.

Status theory

Balance theory can be viewed as a model of likes and dislikes. However, there are many more possible interpretations of a sign in a signed network (1). Consider two friends, A and B , in a directed social network. A positive link from A to B could mean that A is friends with B . However, it could also mean that A has a lower social status than B . This is the idea of status theory, which provides a different organizing principle for directed networks. In this theory of status, we consider a positive directed link to indicate that the creator of the link views the recipient as having a higher status, and a negative directed link indicates that the recipient is viewed as having a lower status.

Note

Both balance and status theory are based on patterns observed from social interactions. Other non-social networks might not follow these principles.

Differences between status theory and balance theory

Let us consider a simple example in which a user A links positively to a user B , and B links positively to a user C . If C forms a link to A , what kind of sign should we expect this link to have? Balance theory predicts that since C is a friend of A 's friend, we should see a positive link from C to A . Status theory, on the other hand, predicts that A regards B as having higher status, and B regards C as having higher status. This means that C should regard A as having low status and hence be inclined to link negatively to A . In this example we can see that the two theories would suggest assigning different signs.

Mathematical background of status theory

Authors of the article (7) have proposed a similar connection from local to global structure for status theory as did Harary for balance theory (4). First, we define a type of directed graph used in their formulation.

Definition

A *tournament* is a directed graph, where each pair of nodes is connected in one direction or another.

If we consider a directed triad that follows status theory, it must be possible to assign three distinct status values to the vertices in such a way that the positive edges among them (if any) point from nodes of lower status to nodes of higher status, and the negative edges (if any) point from nodes with higher status to nodes of lower status. Intuitively, the following situations are impossible to occur in social situations and are thus not status-consistent, since the relation of having a higher or lower status is transitive:

- A has lower status than B , B has lower status than C and C has lower status than A ,
- A has higher status than B , B has higher status than C and C has higher status than A ,
- A has higher status than B , B has higher status than C , but A has lower status than C and
- A has lower status than B , B has lower status than C , but A has higher status than C .

If a triad consisting of nodes A , B , and C follows the rule described above, we say that it is *status-consistent*. We can see all status consistent configurations on picture 4. The following proposition 1 gives a simple method of recognizing status-consistent triads.

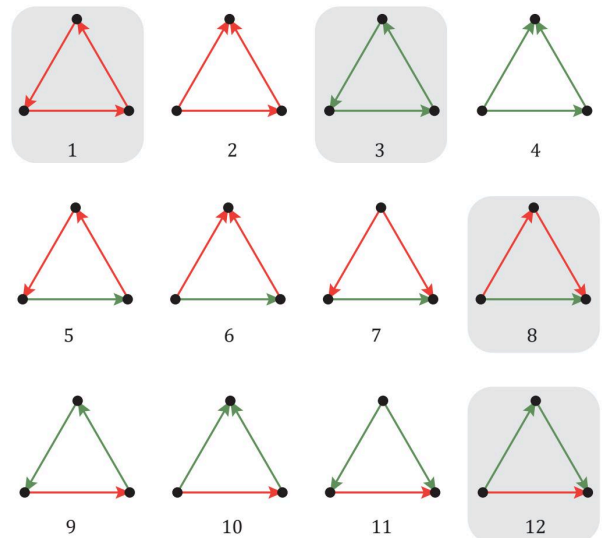


Fig. 4. There are 12 possible configurations of directed triads on three vertices. Configurations 1, 3, 8, and 12, which are not status consistent, are marked with grey.

Proposition 1. A directed triad t in a signed network is status-consistent if the triad t' constructed from t is acyclical. We construct t' by reversing the direction of negative edges in t and flipping their sign.

Proof. We prove this by proving the equivalent statement: if the triad t isn't status-consistent, then the triad t' must be cyclical. If we examine all of the possible triads (1, 3, 8, 12) that are not status-consistent as depicted in 4, we can see that the constructed triad t' is cyclical for each case, which concludes the proof. $\square$

We can now formulate the following result, which says that if all triads are status-consistent, then the whole graph obeys status-consistency.

Theorem 3. Say that $G = (V, E, \sigma)$ is a signed directed tournament and suppose all of its triads are status-consistent. Then it is possible to order the nodes of G as $v_1, v_2, \dots, v_n$ in such a way that each positive edge (v_i, v_j) satisfies $i < j$, and each negative edge (v_i, v_j) satisfies $j < i$.

Proof. We begin by applying the same procedure to the entire graph G as we did to the triad in proposition 1. The graph G' obtained in this manner will only have positive signs. Notice that all of the triads in the graph that were status-consistent before the conversion remain such in G' . If any of the triads in G' were to form cycles, this would violate status consistency. Hence all triads in G' are acyclic. Applying a well-known theorem on tournaments (8), it follows that G' itself is acyclic. Thus we can find a topological ordering $v_1, v_2, \dots, v_n$ of G' . This ordering satisfies the conditions of the theorem. Suppose that an edge (v_i, v_j) is positive in G . This means that it is also an edge in G' and by the property of the topological ordering it follows $i < j$. If the edge (v_i, v_j) is negative in G , then (v_j, v_i) is an edge in G' , and by the property of topological ordering it follows $j < i$. $\square$

Status prediction

We are interested in how the links and their corresponding signs form in time. Instead of just looking at a signed network at one given point in time, we also want to consider the order in which the links are added to the network.

Say we are interested in the social dynamics of a group of people attending a Standard and Latin American dance course, where everybody has a strong opinion about everybody else. Let's say that the opinions of the participants are roughly consistent with reality. We interview each participant and ask them to rank other participants in comparison to their own skills. If dancer A ranks dancer B as having higher skill than oneself, then we add a +1 sign to the edge (A, B) . If they rank the dancer B as having worse skills, we add a -1 sign to the edge (A, B) . Suppose that we then come across a participant who doesn't want to give their opinion on another participant. We could make a prediction of their answer from the signs of the links we do know.

Say we have a situation where dancer X rated both dancers A and B positively. We can use the following reasoning for making a prediction for the status between A and B.

- Firstly we can argue that, because dancer B already received a positive sign, they are more likely to have

above average skills. This means it is more likely that A will evaluate B positively than some other random dancer.

- Since A has also received a positive evaluation, A is also more likely to have above average skills. Therefore, B is less likely to receive a positive evaluation from A than from some other random dancer.

We now introduce some definitions to formalize this reasoning. First, we need a way to describe the context in which a link is made.

Definition

A *contextualized link* or a *c-link* is a triple $(A, B; X)$ with the property that a link forms from A to B after each of the nodes A and B already has a link either to or from X.

If we consider three fixed vertices and all possible ways of how the links are formed, we can see that there are 16 possible c links, as depicted in figure 5.

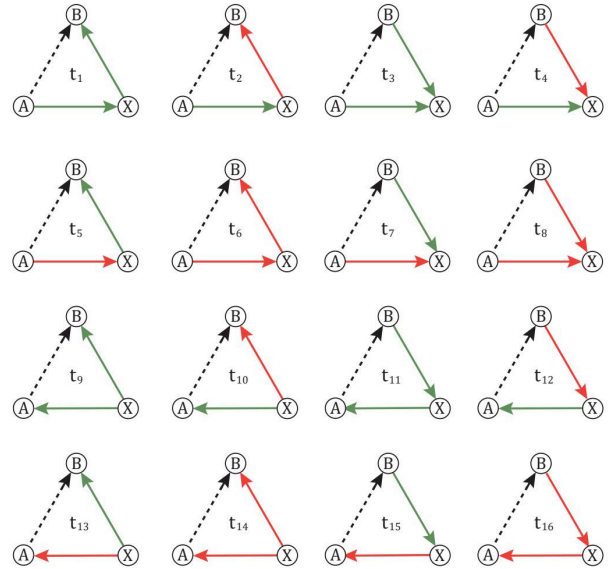


Fig. 5. All possible c-links formed on three vertices. The dashed edge is the edge that is being made in the existing context.

For a fixed type of c-link, we look at the set of all such c-links (A, B, X) and compute the fraction of the links from A to B that are positive in this set.

Definition

Let $(A_1, B_1; X_1), \dots, (A_k, B_k; X_k)$ be all the instances of c-links of type t in a signed network. The *generative baseline* for type t is the sum of the generative baselines of all nodes A_i . The *generative surprise* $s_g(t)$ for type t is the signed number of standard deviations by which the actual number of positive $A_i - B_i$ edges in the data differs from this expectation.

In other words, if the context provided by the node X and its links with A and B had no effect on the sign of the $A - B$ link being formed — in which case each node A_i would make a link with a sign according to its generative baseline — then we would expect the generative surprise of 0 for type t .

Definition

Let $(A_1, B_1; X_1), \dots, (A_k, B_k; X_k)$ be all the instances of c-links of type t in a signed network. The receptive baseline for type t is the sum of the receptive baselines of all nodes B_i . The receptive surprise $s_r(t)$ for type t is the signed number of standard deviations by which the actual number of positive $A_i - B_i$ edges in the data differs from this expectation.

We now assign status values to vertices A and B in any type of c-link in the following way. Firstly we assign the vertex X a status of 0.

- If X links to A with a positive sign or A links to X with a negative sign, we conclude that A has higher status than X , hence we assign A the status of 1.
- If X links to A with a negative sign or A links to X with a positive sign, we conclude that A has lower status than X , hence we assign A the status of -1.
- If X links to B with a positive sign or B links to X with a negative sign, we conclude that B has higher status than X , hence we assign B the status of 1.
- If X links to B with a negative sign or B links to X with a positive sign, we conclude that B has lower status than X , hence we assign B the status of -1.

We say that generative surprise for type t is consistent with status if B has a status that has the same sign as its generative surprise. Note that if this is true, B receives more positive evaluations than we expect from the generative baseline of the node A who is producing the link.

We say that the receptive surprise for type t is consistent with status if A 's status has the opposite sign from the receptive surprise. This means that high status generators of links A produce fewer positive evaluations than would be expected from the receptive baseline of the node B receiving the link.

Link reciprocity in directed signed networks

An interesting observation that was made in the article (1) is that, in an undirected signed network in the portion of the graph where the presence of links is symmetrical, the principles of balance are more pronounced. This makes sense as people who have an established relationship will treat the counterpart differently in the context of back-and-forth interaction, than in the case where the interaction is only in a single direction.

Definition

Let A and B be nodes in a signed directed network, where A links to B and B links to A . We call such edges *reciprocal*. If the edge $A - B$ was present before $B - A$, we say B *reciprocates* the link to A .

The results of the article show that the reciprocation of positive $A - B$ edges is closely consistent with balance theory rather than status theory. This means that, in most cases, if A links positively to B , B will also link positively to A . The reciprocation of negative $A - B$ edges follows a hybrid of the two principles.

Now let's consider the context of a third vertex X . Suppose there exists one directed link between A and B and also two directed links from or to the vertex X and vertices A and B . If the triad we get by ignoring the direction of the links is structurally balanced, it is highly likely that the reciprocal edge between A and B will have the same sign as the original edge between A and B .

Community detection in signed networks

Most community detection algorithms assume that the links between any pair of nodes have the same meaning everywhere in the network. However, this is not true in signed networks, so those algorithms cannot be applied directly to them. One of the widely used techniques for detecting communities in an unsigned network is *modularity* proposed by Newman (9). The idea behind this approach is to divide the network into communities, such that the number of edges within those communities is higher compared to the number of edges we would have gotten in a network where we reshuffled the edges. However, in signed networks, the goal of community detection is to find such sets of nodes that are densely connected with positive edges within sets and negatively connected between the sets. The approach developed for finding the best partition into communities is based on optimizing modularity and minimizing frustration. Modularity in unsigned networks measures the quality of the partition of the network. Here we give the modified modularity which was designed to take into account the contribution of links with binary values as given in the article (10).

Definition

Let $G = (V, E, \sigma)$ be a signed network with network partition $C = \{C_1, \dots, C_k\}$ into k communities. The *signed modularity* is defined as

$$Q_s = \frac{1}{2m} \sum_{i,j \in V} (A_{ij} + \frac{k_i^- k_j^-}{2m} - \frac{k_i^+ k_j^+}{2m}) \delta(c_i, c_j).$$

The frustration of the network partition is the sum of the number of negative edges between nodes inside the community and the number of positive edges between nodes belonging to different communities.

Definition

Let $G = (V, E, \sigma)$ be a signed network with network partition $C = \{C_1, \dots, C_k\}$ into k communities. The *frustration* is defined as

$$F(C) = \frac{1}{2} \sum_{i,j \in V} (|a_{ij}| - a_{ij}) \delta(c_i, c_j) + (|a_{ij}| + a_{ij})(1 - \delta(c_i, c_j))$$

Given a signed network, the objective of maximizing the sign modularity and minimizing frustration is to find a par-

tioning into k clusters, such that, internal connections are dense and most edges in the clusters are positive and connections between clusters are sparse and most of them are negative. A multi-objective algorithm SN-MOGA for solving this problem was recently proposed by researchers (11).

Empirical evaluation of signed networks

To put the theory we discussed to the test and obtain some empirical data, we evaluated 7 real-world networks (3 undirected and 4 directed):

- **Chess player network** (12) from a 2010 outcome prediction competition. This directed graph has 7301 nodes and 65053 edges. Nodes represent chess players and edges represent the outcomes of the matches between them.
- Two **Wikipedia adminship vote** networks, one from 2008 (13), and the other from 2013 (14). These two are directed graphs, as the edges represent votes and nodes represent the users (voters or admin candidates). The 2008 network contains 7118 nodes and 103675 edges, while the 2013 network contains 11381 nodes and 198275 edges. The 2013 network also contained some neutral edges, which we removed from the graph prior to our analysis.
- A **New Guinea Highlands tribes friendship** network (15). The graph is undirected, as the nodes represent tribes and edges represent alliances between them. It contains 16 nodes and 58 edges.
- A **US Congressional co-sponsorship** network (16), more specifically the House network from the 114th Congress. Edges indicate tendency to co-sponsor or not co-sponsor bills, while the nodes are representatives. The network contains 446 nodes and 75180 edges and is undirected.
- A **Wikipedia user interaction** network (17). The dataset was created in 2011 and contains interactions between users on the English Wikipedia, more specifically agreement and disagreement between users based on the edit histories of 563 Wikipedia pages related to politics. Nodes represent users and edges represent the positive or negative editorial relationship between them. The graph is directed and has 138592 nodes and 740397 edges.
- A **Wikipedia edit wars** network (18). The network represents editorial interactions among users on the English Wikipedia. Nodes represent users, while edges represent either agreement or conflict on the edit. The graph is undirected and has 118100 nodes and 2917785 edges.

For a more detailed look into the structure and statistics of our networks, see tables 1 and 2, for undirected and directed networks, respectively.

While several of the networks we listed are too large to meaningfully visualize, we can still look at some of the smaller ones (figures 6, 7, and 8). In the visualizations, red edges represent a negative sign and the green edges represent a positive sign. For example, in 6, we can very clearly see which tribes are allies and which are enemies. Even more: even with the naked eye we can see pockets of balance with, e.g., many allies of allies being allies (T_3) or tribes having a mutual enemy

(T_1). At the same time, as the node count grows, we can no longer meaningfully visualize the structure in such a simple way — for example, in figure 7, while perhaps hard to see, we can still somewhat observe the hierarchical structure of the games. In figure 8, however, no meaningful information can be observed anymore, as the node and edge count is simply too high. Therefore, we will need to analyse the networks in a different way.

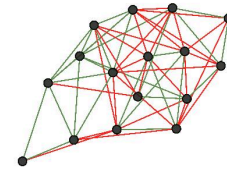


Fig. 6. Visualization of the New Guinea tribes network (15).

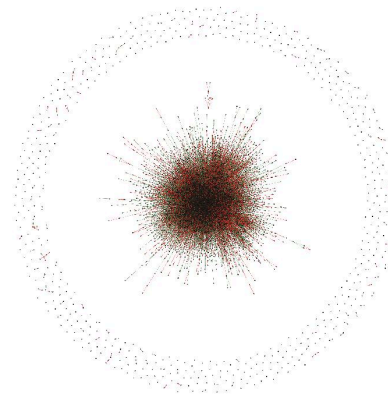


Fig. 7. Visualization of the Chess network (12).

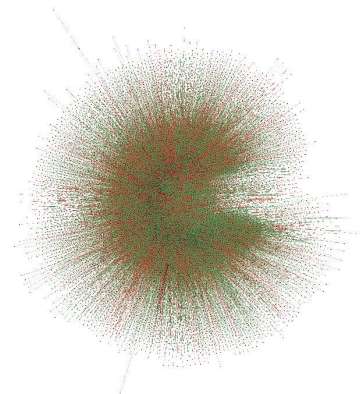


Fig. 8. Visualization of the Wikipedia 2013 adminship vote network (14).

Network	Nodes	Edges	Edge sign distribution		Triangles	Clustering coefficient	Connected components	
New Guinea tribes (15)	16	58	+	29 (50.0 %)	68	0.527	Total	1
			-	29 (50.0 %)			Largest	16 (100.0 %)
114th US Congress (16)	446	75180	+	26578 (35.35 %)	7026166	0.805	Total	1
			-	48602 (64.65 %)			Largest	446 (100.0 %)
Wikipedia edit wars (18)	118100	2915501	+	1242833 (42.63 %)	146159240	0.026	Total	3168
			-	1672668 (57.37 %)			Largest	112998 (95.68 %)

Table 1. Basic statistics of the undirected networks we analysed.

Network	Nodes	Edges	Edge sign distribution		Triangles	Clustering coefficient	Strongly connected components		Weakly connected components	
Chess (12)	7301	36387	+	21154 (58.14 %)	35594	0.092	Total	2536	Total	563
			-	15233 (41.86 %)			Largest	4747 (65.02 %)	Largest	6616 (90.62 %)
Wikipedia admin-ship (2008) (13)	7118	103675	+	81318 (78.44 %)	719769	0.136	Total	5819	Total	27
			-	22357 (21.56 %)			Largest	1300 (18.26 %)	Largest	7066 (99.27 %)
Wikipedia admin-ship (2013) (14)	11381	185627	+	145084 (78.16 %)	1826377	0.153	Total	8944	Total	123
			-	40543 (21.84 %)			Largest	2436 (21.40 %)	Largest	11259 (98.93 %)
Wikipedia inter-actions (17)	138592	740397	+	650653 (87.88 %)	3814040	0.042	Total	111863	Total	310
			-	89744 (12.12 %)			Largest	26699 (19.26 %)	Largest	137740 (99.39 %)

Table 2. Basic statistics of the directed networks we analysed.

Network	Strong Balance (%)		Weak Balance (%)		Status Consistency (%)	
	Original	Shuffled	Original	Shuffled	Original	Shuffled
New Guinea tribes	86.7	50.0	97.0	61.8	—	—
114th US Congress	15.0	48.8	26.2	75.7	—	—
Wikipedia edit wars	49.8	50.4	70.3	70.5	—	—
Chess	64.6	50.0	71.1	56.9	84.48	75.33
Wiki adminship (2008)	89.3	59.3	90.8	60.3	92.46	83.19
Wiki adminship (2013)	88.7	52.2	90.3	60.9	91.32	81.39
Wikipedia interactions	95.7	71.9	96.9	72.1	90.89	71.96

Table 3. Balance and status metrics of all analysed real-world networks.

As discussed in previous chapters, we can analyse the signs of its three edges for any given triangle in a graph. Considering strong structural balance, we say that the triangle is balanced only if the three edges are all positive, or two edges are negative and one is positive; otherwise the triangle is considered unbalanced. We analysed the real-world networks and obtained their triad sign distributions which we can see in table 4. Note that, for directed graphs (the second half of the table), we first converted them into undirected graphs before computing their triad sign distributions.

We also analysed the strong and weak balance, as well as the status consistency of the same networks. The results of those computations can be seen in table 3. The shuffled column represents the balance or status consistency ratio after

having reshuffled the entire network's edge signs, but leaving its structure alone.

For example, looking at the Wikipedia-related networks (except the political edit wars), which are in essence social networks, we can quickly deduce that the vast majority of interactions on the site are positive, either being positive adminship votes or agreement on page edits. While further balance and status consistency tests will give us a confirmation, we can already see that these three networks are somewhat balanced, as T_3 and T_2 prevail.

In the New Guinea tribes network, we can even at a glance see that balance is quite strong: there are only 9 unbalanced edges, i.e. only 2 occurrences where two enemies have a common ally and 7 occurrences where three tribes are all

enemies between each other. By weak balance, that means that only 2 out of 68 triangles are unbalanced.

Looking at the Congress and Wikipedia page edit wars networks, both of which are politically charged, we can observe an interesting contrast to e.g., the other Wikipedia social networks: both of these networks have a much higher rate of negative edges. Even more: they have a very high ratio of unbalanced triads (T_1 and T_0) compared to other networks, making them much less balanced.

Network		$ T_i $	$p(T_i)$
New Guinea tribes	T_3 + + +	19	0.279
	T_1 + - -	2	0.029
	T_2 + + -	40	0.588
	T_0 - - -	7	0.103
114th US Congress	T_3 + + +	1021473	0.145
	T_1 + - -	5186239	0.734
	T_2 + + -	34724	0.005
	T_0 - - -	783730	0.112
Wikipedia edit wars	T_3 + + +	4179371	0.302
	T_1 + - -	4101714	0.297
	T_2 + + -	2716664	0.196
	T_0 - - -	2833487	0.205
Chess	T_3 + + +	4865	0.216
	T_1 + - -	6511	0.289
	T_2 + + -	9681	0.430
	T_0 - - -	1453	0.065
Wikipedia adminship (2008)	T_3 + + +	408088	0.668
	T_1 + - -	57065	0.093
	T_2 + + -	137356	0.225
	T_0 - - -	8895	0.015
Wikipedia adminship (2013)	T_3 + + +	808882	0.648
	T_1 + - -	122004	0.098
	T_2 + + -	297980	0.239
	T_0 - - -	19861	0.016
Wikipedia interactions	T_3 + + +	2622454	0.869
	T_1 + - -	94513	0.031
	T_2 + + -	266005	0.088
	T_0 - - -	35998	0.012

Table 4. Triangle distributions of all analysed real-world networks.

Conclusion

Signed networks are a useful representation in situations where we want to assess connections as being positive or negative instead of just knowing about their existence. These types of networks can model a variety of opposed relationships, such as friendships and enemies, liking and disliking, agreement and disagreement, or trust and distrust.

With this addition, many interesting questions and new

metrics arise as well. For example, we can now detect mutually antagonistic groups in signed social networks, or attempt to predict the nature of a relationship between people before they meet. We've also used some of those methods that were developed specifically for analyzing such networks on real-world undirected and directed networks and analysed their triads, balance, and status consistency.

Although technology like mainstream social media platforms generate more data than ever about negative or positive interactions both between users and between users and published content, much of that data is not publicly available, making it hard for researchers to test their methods. Despite the fact that balance and status theory seem promising for improving recommendation systems, the lack of datasets, especially those that are dynamic in time or up-to-date, makes it hard for researchers to study these phenomena.

- Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Signed networks in social media, 2010. URL <https://arxiv.org/abs/1003.2424>.
- Fritz Heider. Attitudes and cognitive organization. *The Journal of Psychology: Interdisciplinary and Applied*, 21:107–112, 1946.
- James A. Davis. Clustering and structural balance in graphs. *Human Relations*, 20(2):181–187, 1967.
- Frank Harary. On the notion of balance of a signed graph. *Michigan Mathematical Journal*, 2: 143–146, 1953. URL <https://api.semanticscholar.org/CorpusID:120368252>.
- Ernesto Estrada. Rethinking structural balance in signed social networks. *Discrete Applied Mathematics*, 268:70–90, 2019. ISSN 0166-218X. URL <https://www.sciencedirect.com/science/article/pii/S0166218X1930229X>.
- Arti Patidar, Vinti Agarwal, and K.K. Bharadwaj. Predicting friends and foes in signed networks using inductive inference and social balance theory. pages 384–388, 08 2012. ISBN 978-1-4673-2497-7.
- Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Predicting positive and negative links in online social networks. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, page 641–650, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605587998. URL <https://doi.org/10.1145/1772690.1772756>.
- Maria Chudnovsky, Paul Seymour, and Blair D. Sullivan. Cycles in dense digraphs, 2007. URL <https://arxiv.org/abs/math/0702147>.
- Mark Newman. Newman mej.. modularity and community structure in networks. *proc natl acad sci usa* 103: 8577-8582. *Proceedings of the National Academy of Sciences of the United States of America*, 103:8577–82, 07 2006.
- Sergio Gomez, Pablo Jensen, and Alex Arenas. Analysis of community structure in networks of correlated data. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 80: 016114, 08 2009.
- Alessia Amelo and Clara Pizzuti. Community mining in signed networks: A multiobjective approach. pages 95–99, 08 2013.
- Jeff Sonas. Chess ratings - elo versus the rest of the world. <https://kaggle.com/competitions/chess>, 2010. Kaggle.
- Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Governance in social media: A case study of the wikipedia promotion process, 2010. URL <https://arxiv.org/abs/1004.3547>.
- Robert West, Hristo S. Paskov, Jure Leskovec, and Christopher Potts. Exploiting social network structure for person-to-person sentiment analysis. *Transactions of the Association for Computational Linguistics*, 2:297–310, 10 2014. ISSN 2307-387X. URL https://doi.org/10.1162/tacl_a_00184.
- K. E. Read. Cultures of the central highlands, new guinea. *Southwestern Journal of Anthropology*, 10(1):1–43, 1954. URL <https://doi.org/10.1086/soutjanth.10.1.3629074>.
- Zachary Neal. A Sign of the Times: Dataset of US Congress signed network backbones from co-sponsorship data, 1973-2016. 5 2019. URL https://figshare.com/articles/dataset/A_Sign_of_the_Times/8096429.
- Silviu Maniu, Talel Abdesslem, and Bogdan Cautis. Casting a web of trust over wikipedia: an interaction-based approach. In *Proceedings of the 20th International Conference Companion on World Wide Web, WWW '11*, page 87–88, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450306379. URL <https://doi.org/10.1145/1963192.1963237>.
- Ulrik Brandes, Patrick Kenis, Jürgen Lerner, and Denise van Raaij. Network analysis of collaboration structure in wikipedia. In *Proceedings of the 18th International Conference on World Wide Web, WWW '09*, page 731–740, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605584874. URL <https://doi.org/10.1145/1526709.1526808>.