Community detection in networks with positive and negative links

V.A. Traag[1, 2, *] and Jeroen Bruggeman[1]

[1]*Department of Sociology and Anthropology, Faculty of Social and Behavioural Sciences, University of Amsterdam, Netherlands*

[2]*Data Quality and Integrity, Enterprise Risk Services, Deloitte Netherlands*

(Dated: November 14, 2008)

Detecting communities in complex networks accurately is a prime challenge, preceding further analyses of network characteristics and dynamics. Until now, community detection took into account only positively valued links, while many actual networks also feature negative links. We extend an existing (spin glass) approach to incorporate negative links as well, resulting in a method similar to the clustering of signed graphs, but more accurate and more general. To illustrate our method, we applied it to a network of international alliances and disputes. Using data from 1993–2001, it turns out that the world can be divided into six power blocs similar to Huntington's civilizations, with some notable exceptions.

Keywords: complex networks; community detection; modularity; negative links; social balance theory; international relations

I. INTRODUCTION

Many complex phenomena can be represented as networks, and subsequently be analyzed fruitfully [1, 2, 3]. One of the first targets of network analysis is the detection of communities on the basis of the links, i.e. the possibly valued, or weighted, arcs or edges that connect the nodes. Intuitively, an assignment of nodes to communities should be such that links within communities are relatively dense and between communities relatively sparse. This means that we should compare actual densities to expected densities of links within and between communities. Furthermore, since nodes, for example humans or proteins, can be members of different communities at the same time, e.g. organizations or protein complexes, respectively, the assignment should allow for the possibility that communities overlap.

In approaches to find appropriate community assignments, much progress has been made in recent years [4, 5, 6]. While current approaches take for granted that links are positively valued, scientists in numerous fields grapple with networks that also have *negative* links, for example neural networks, semantic webs, genetic regulatory networks, and last but certainly not least, social networks. In this paper, we generalize an existing spin glass model [7] for positive links to incorporate negative links as well. We will follow the intuition that the assignment of nodes related by negative links should be done the opposite way of positive links, with negative links sparse within and more dense between communities, generalizing an old idea from social balance theory [8]. We will then show our approach at work in a network wherein conflicts and alliances between national states are represented by negative and positive links, respectively.

Our approach is based on the concept of *modularity*, which quantifies the discrepancy between communities in actual networks and in their random counterparts [9]. Recently, it was shown that modularity might miss small communities embedded in larger ones [10], and is less accurate if the actual communities are highly different in size [11]. Our method, which is an extension of a Potts spin glass model [7], has two balancing parameters that address this problem to some extent [12]. Yet community detection through modularity remains a global rather than a local approach.

II. MODULARITY

We start by considering a directed binary graph G with n vertices and m links, which, as we will see, can be easily generalized to weighted graphs. We denote the total number of positive links in G as m^+ and the negative links as m^-, hence $m = m^+ + m^-$. We define the adjacency matrix of G as follows: if a positive link is present from node i to node j, $A_{ij} = 1$, if a negative link is present $A_{ij} = -1$, and $A_{ij} = 0$ otherwise. We separate the negative and positive links by setting $A^+_{ij} = A_{ij}$ if $A_{ij} > 0$ and zero otherwise, and $A^-_{ij} = -A_{ij}$ if $A_{ij} < 0$ and zero otherwise, such that $A_{ij} = A^+_{ij} - A^-_{ij}$. The positive and negative in- and outdegrees of i are defined as

$$\begin{aligned} {}^-k_i^{out} &= \textstyle\sum_j A^-_{ij} \qquad {}^-k_i^{in} &= \textstyle\sum_j A^-_{ji} \\ {}^+k_i^{out} &= \textstyle\sum_j A^+_{ij} \qquad {}^+k_i^{in} &= \textstyle\sum_j A^+_{ji}. \end{aligned} \tag{1}$$

For a weighted graph with weights ω_{ij}, we set $A_{ij} = \omega_{ij}$, and keep the above definitions as they are.

Our challenge is to assign each node i to one of c communities $\sigma_i \in \{1, \ldots, c\}$. The complete configuration of community assignments is denoted by $\{\sigma\}$, which assigns each node $i = 1, \ldots, n$ to a community $\sigma_1, \ldots, \sigma_n$. Following our intuition, we should reward (i) positive internal links; (ii) the absence of positive external links; (iii)

*Corresponding author: vtraag@f-m.fm

the absence of negative internal links; and, (iv) negative external links. Doing so results in a modularity[1] Q,

$$
\begin{aligned}
Q(\{\sigma\}) \;=\; & \frac{1}{m}\sum_{ij} a_{ij}A^{+}_{ij}B_{\sigma_i\sigma_j} \\
& + b_{ij}(1-A^{+}_{ij})(1-B_{\sigma_i\sigma_j}) \\
& + c_{ij}(1-A^{-}_{ij})B_{\sigma_i\sigma_j} \\
& + d_{ij}A^{-}_{ij}(1-B_{\sigma_i\sigma_j}).
\end{aligned}
\tag{2}
$$

If links between community σ_i and σ_j are allowed, $B_{\sigma_i\sigma_j} = 1$, else $B_{\sigma_i\sigma_j} = 0$. This is arguably the most general formulation of modularity, which can also be used to assign roles of nodes based on their pattern of relations [13]. Here we consider only $B_{\sigma_i\sigma_j} = \delta(\sigma_i, \sigma_j)$, where $\delta(\sigma_i, \sigma_j) = 1$ if $\sigma_i = \sigma_j$ and zero otherwise. Hence, only positive links within communities are rewarded, while links between communities are "punished," thereby gearing modularity towards optimizing community detection. By dropping all terms not depending on the specific configuration, we can simplify equation (2) to

$$
\begin{aligned}
Q(\{\sigma\}) = \frac{1}{m}\sum_{ij}\Big(& (a_{ij}+b_{ij})A^{+}_{ij} + \\
& (c_{ij}+d_{ij})A^{-}_{ij} - (b_{ij}-c_{ij})\Big)\delta(\sigma_i,\sigma_j).
\end{aligned}
\tag{3}
$$

We want to balance the rewarded and the punished links, and, writing $+/-$ as a shorthand for positive or negative, we do so by considering $p^{+/-}_{ij}$, such that $\sum_{ij} p^{+/-}_{ij} = \sum_{ij} A^{+/-}_{ij}$, and by setting

$$
\begin{aligned}
b_{ij} &= \gamma p^{+}_{ij} \\
a_{ij} &= 1-\gamma p^{+}_{ij} \\
c_{ij} &= \lambda p^{-}_{ij} \\
d_{ij} &= 1-\lambda p^{-}_{ij}.
\end{aligned}
$$

When $\gamma = \lambda = 1$, this leads to equal contributions of present and absent links, positive and negative. These two parameters thus balance the penalties for absent and present links in an assignment. If the researcher increases γ, the communities retrieved will exhibit higher density of positive links than for lower values of γ. If λ is increased, the density of negative internal links will be higher than for lower λ. Increasing γ and decreasing λ set higher thresholds for finding communities. These parameters may thus be used to find larger (overlapping) super communities, or to detect smaller subcommunities hierarchically nested in larger ones. The latter option helps to overcome a shortcoming of modularity, which for the default parameter setting tends to overlook small (sub) communities [10]. If we set $\lambda = \gamma = 0$, we can direct our search towards finding completely separate communities.

Using the above assumptions and the fact that $A^{+} - A^{-} = A$, we can simplify (3) to:

$$
Q(\{\sigma\}) = \frac{1}{m}\sum_{ij}(A_{ij} - (b_{ij} - c_{ij}))\delta(\sigma_i, \sigma_j). \tag{4}
$$

This definition resembles the original definition of modularity [14], with different weights in our case. If there are no negative links and $\gamma = 1$, our modularity reduces to the original one. Notice that $\sum_{ij} p^{+}_{ij} - p^{-}_{ij} = m^{+} - m^{-}$, which equals m in case $m = m^{+} - m^{-}$, not by counting the links but by summing their weights. The difference is that if we count the links, negative weights have a positive influence, while if we sum the weights, the negative weights have a negative influence.

For $p^{+/-}_{ij}$ there are various options available. In line with [14], we take the actual degree distribution into account, and set

$$
p^{+/-}_{ij} = \frac{{}^{+/-}k^{out}_i \, {}^{+/-}k^{in}_j}{m^{+/-}}. \tag{5}
$$

We can interpret $p^{+/-}_{ij}$ as the probability of a link from i to j. Furthermore, modularity can be rewritten as adhesion between, and cohesion within communities, making possible a formal definition of community (see Appendix); these notions are straightforward extensions of the earlier spin glass model for positive links [7].

Maximizing modularity is an NP-complete problem [15], and therefore any algorithm with a feasible running time will most likely produce suboptimal results in general. We opted for simulated annealing [16], which performs well [7, 17, 18, 19], although it's not the fastest algorithm [20]. While staying with the modularity approach, one might opt for a faster algorithm instead [21, 22]. We have implemented our algorithm in the `igraph` package [23] of the open source program `R` [24]. At the time of writing, our implementation was not yet available in the official release of the package, but it can be obtained from the first author upon request. We have conducted standard performance tests [18, 19] and it performs as expected, compared with the algorithm for positive links that had already been examined [18, 25].

III. SOCIAL BALANCE THEORY

A long standing interest of social scientists is *conflict* between people or other social entities, e.g countries, which social balance theorists represent by negative links between these entities [26]. The idea of balance originated in cognitive dissonance theory [27], which has it that if two people have a positive relationship, their attitudes should match.

[1] Our equation (2) is a spin-off from a spin glass Hamiltonian [7], from which the Girvan-Newman modularity [14] was inferred as a special case. An assignment for which modularity is maximal, then, is the ground state of the pertaining spin glass.

For example, if John (i) and Mark (j) are friends, both John and Mark should like or dislike the same things. On the other hand, if they are enemies, they should have opposite tastes, so if one of them likes something, the other should dislike it. If we take that "thing" to be a third person (k), the idea of "social balance" emerges. In a so called *signed* graph with link values -1 or 1 (or 0 if there is no link), the triad of i, j and k is said to be socially balanced if $A_{ij}A_{jk}A_{ki} = 1$. The entire graph is considered to be balanced if and only if all triads are balanced, and the notion was later generalized to cycles larger than three [26].

Along with investigating network dynamics, i.e. the global consequences of local balance restoring by flipping signs with given probabilities (of foes becoming friends or the other way around) [28], balance theorists asked if graphs can be partitioned into communities with only positive links within, and only negative links between them. It was proven that a graph can be partitioned into two communities if and only if it is balanced [8]. Empirically, however, populations of humans in conflict are often unbalanced to some extent, feature more than two communities, or both. Although balance theorists were aware of this complexity and attempted to accommodate the possibility of more than two communities [29], their assignment was somewhat arbitrary.

We therefore return to our modularity approach, of which the clustering of balanced signed graphs into opposing communities is a special case. After all, we can detect communities also in unbalanced and in weighted graphs. Our modularity approach relates to balance theory through the concept of *frustration* [30, 31], which increases in value if nodes that are related by negative links are assigned to the same community, and decreases with positive links within communities. A similar but less general approach was taken earlier [32]. We define the frustration of community s as

$$F_s = \sum_{ij}(A^-_{ij} - A^+_{ij})\delta(\sigma_i, s)\delta(\sigma_j, s).$$

If we set $\gamma = \lambda = 0$, we obtain

$$Q(\{\sigma\}) = -\frac{1}{m}\sum_s F_s.$$

To compare the frustration of an actual graph with the frustration in its random counterpart (with probabilities $p^{+/-}_{ij}$), we can write

$$\begin{aligned} F_s &= \sum_{ij}(A^-_{ij} - A^+_{ij})\delta(\sigma_i, s)\delta(\sigma_j, s) \\ &\quad - \sum_{ij}(c_{ij} - b_{ij})\delta(\sigma_i, s)\delta(\sigma_j, s), \end{aligned}$$

and thus recover our definition of modularity. In sum, minimizing frustration is the same as maximizing modularity.[2] In contrast to categorically disallowing negative links within communities, as balance theory has it, our approach is more appropriate for the messiness of real life.

IV. APPLICATION

To show how our method can be applied to an empirical network, we analyze international relations, where positive links are military alliances, and negative links are disputes. We use the Correlates of War [35, 36] data set for both alliances and disputes. The data set contains a wide variety of disputes, for example border tensions between Colombia and Venezuela, the deployment of Chinese submarines to Japanese islands, and Turkish groups entering Iraqi territory. In total 512 disputes that took place between 1993 and 2001 were recorded, in which 138 nations were involved. Disputes were assigned hostility levels, from "no militarized action" to "interstate war," and we chose the mean level of hostility between two countries over the given time interval as the weight of their negative link. The alliances we coded one of three values, for (1) entente, (2) non-aggression pact, or (3) defense pact. The disputes and alliances are normalized to values in the interval $[0, 1]$ which then bear equal weight in the overall link value ω_{ij}. The largest component consists of 161 nodes (countries) and 2517 links (conflicts and alliances).

The result of the analysis ($Q = 0.534$) is shown in Figure 1. Countries of the same color belong to the same community, which in this context is more appropriately labeled a *power bloc*. How strongly a country belongs to its power bloc is indicated by color intensity.[3] The power blocs can be identified as follows: (1) the West; (2) Latin America; (3) Russia & China; (4) West Africa; (5) North Africa & the Middle East; and, (6) a collection of "independents" not truly forming a bloc, lacking the cohesion that other blocs have. If we detect communities by using only positive links, there is an agreement of about 64% with the configuration in Fig. 1, while if using only negative links, there is an agreement of about 30%.

Our result resembles the configuration depicted in Huntington's renowned book *The Clash of Civilizations* [37], with a few notable exceptions. The West African power bloc is an additional insight that is absent in Huntington's configuration. Furthermore, the

---

[2] Notice that for a graph with only negative links, there will be no frustration if the number of communities is at least the number of colors necessary for coloring the graph—the chromatic number [33, p. 111]. Setting $\lambda = 0$ thus yields a graph coloring algorithm [34].

[3] The accounting of the number and strength of links between communities, and of a subcommunity's (e.g. a singleton's) links to its community, was done by using the notion of *adhesion*, see Appendix.

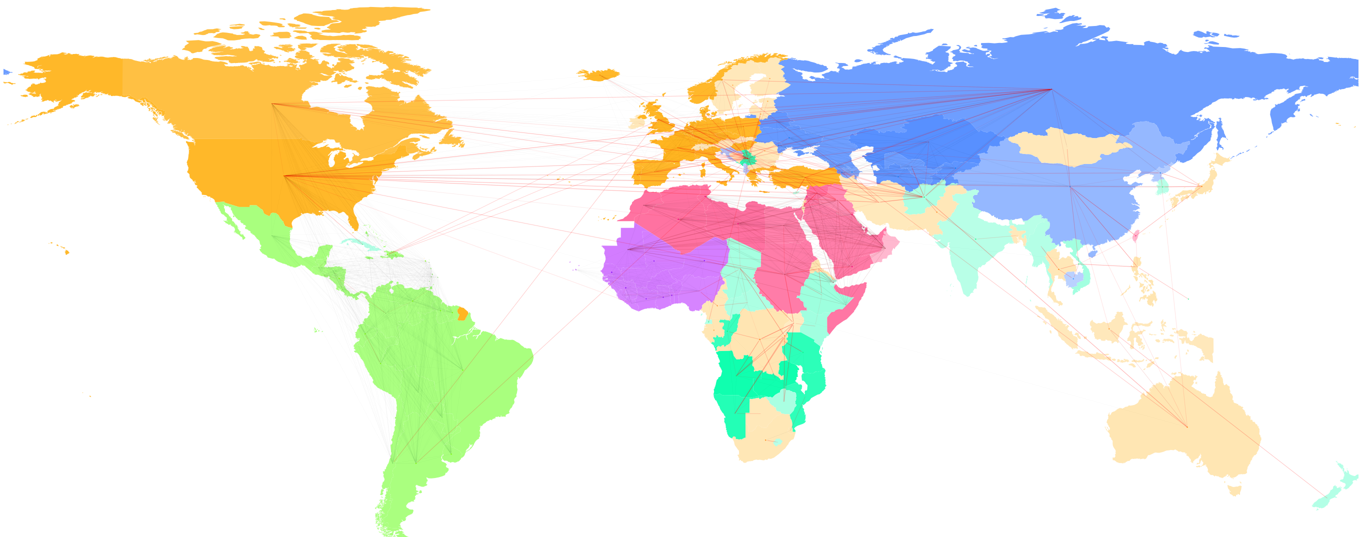

FIG. 1: Map of the communities in the conflict and alliance network ($Q = 0.534$). Negative links are red, and positive links are black.

collection of independents comprises militarily isolated countries like Afghanistan, Taiwan, Myanmar and Cuba, but also New Zealand, which according to Huntington belongs to the West, but has relatively few alliances. A major difference with Huntington is that China itself does not constitute a separate bloc, nor does Japan.

If we run the algorithm with $\gamma = 0.1$ and $\lambda = 1$, North America merges with Latin America, while Europe becomes an independent community, and North Africa and the Middle East align with Russia and China. When setting $\gamma = 1$ and $\lambda = 2$, in contrast, former Soviet countries separate from Russia and form an independent community. Using a range of values for γ and λ, one can detect various layers in the community structure, although it's theoretically unknown which parameter values are best to choose [12]. Setting a smaller γ usually results in larger communities, as does a larger λ, and may point to community overlaps. The opposite choices result in smaller communities, which need not be hierarchically nested, though, as our data show.

Our configuration does not imply that conflicts take place between power blocs only, as 24% of all conflicts actually take place *within* blocs. For example, Georgia and Russia had serious conflicts, and Zaire and Uganda had theirs, but each of these pairs is grouped together nevertheless. In these cases, the alliances overcame the conflicts in the grouping, confirming that a configuration of international relations is more than the sum of bilateral links.

Our political analysis here is limited, since we wish to demonstrate the method rather than present a complete coverage of international alliances and disputes. Other approaches that could be brought into play are the democratic peace theory [38, 39], which predicts few conflicts between democratic countries but fails to predict that in actuality, most conflicts occur between democratic and non-democratic countries; the realist school [40], which emphasizes geopolitical concerns; and, the trade-conflict theory [41], which argues that (strong) trade relations diminish the probability of a dispute, or lower its intensity. In sum, although Huntington's configuration of civilizations was questioned [42, 43], it seems to be fairly robust, and with some marked exceptions, is confirmed by our analysis.

V. CONCLUSION

We have extended the existing spin glass approach by adapting the concept of modularity to detect communities in complex networks where both positive and negative links are present. This approach solves a long standing problem in the theory of social balance, namely the clustering of signed graphs, which can now be done more accurately and more generally. Our formulation of modularity (2) can also be used to assign roles of nodes based on their patterns of relations [13].

As a case in point, we have analyzed a social network of international disputes and alliances. Other applications could be networks of references on the Web [44] or in blogs [45]. If in these data positive and negative references are distinguished, our method makes possible to detect not only thematic clusters, but also positional clusters with internal agreement and external disagreement.

For network data, the model's parameters (λ and γ) can be used to find smaller (sub) communities by trial and error, as there is currently no theoretical guidance to choose parameter values [12]. Even if there were such guidance, the modularity approach intrinsically aims at global rather than local optimization. Our implementation is based on simulated annealing, which performs quite well with standard tests, although for very large networks, faster algorithms will be necessary.

Whatever algorithms future researchers will use, or improvements of the concept of modularity they will develop, being able to detect communities in networks with both positive and negative links is important in numerous fields of science, and a stepping stone towards further analyses of complex networks.

APPENDIX: ADHESION AND COHESION

To facilitate different applications of modularity and comparison with other theories, the notion can be reformulated in terms of community *adhesion*. Taking negative ties into account, and following Reichardt and Bornholdt [7], we define the adhesion between community r and s as

$$a_{rs} = (m^+_{rs} - m^-_{rs}) - ([m^+_{rs}] - [m^-_{rs}]),$$

or alternatively as

$$a_{rs} = (m^+_{rs} - [m^+_{rs}]) + ([m^-_{rs}] - m^-_{rs}),$$

where

$$m^-_{rs} = \sum_{ij} (d_{ij} + c_{ij}) A^-_{ij} \delta(\sigma_i, r)\delta(\sigma_j, s)$$

and

$$[m^-_{rs}] = \sum_{ij} c_{ij} \delta(\sigma_i, r)\delta(\sigma_j, s),$$

with similar definitions for positive links. If we wish, we can express the expected links between communities, $[m^{+/-}_{rs}]$, analogously to the expected links between nodes, and modularity can then be written as

$$Q(\{\sigma\}) = -\frac{1}{m} \sum_{r \neq s}^{c} a_{rs},$$

where the sum runs over the communities pair-wise.

As a complement to adhesion between communities, Reichardt and Bornholdt defined the *cohesion*, c_s, of a community s as $c_s = a_{ss}$. This notion should be carefully distinguished from other conceptions of cohesion, though, wherein topological features, like the minimum number of independent paths connecting pairs of nodes [46], play a crucial role.

Some further considerations make it possible to define the notion of community. A configuration of assignments $\{\sigma\}^*$ is optimal if no re-assignment of nodes improves Q. Hence, $\Delta Q \leq 0$ for all nodes and alternative communities, and for any subset S of nodes in a community α, for every other community ϕ,

$$(a_{S\phi} + a_{\phi S}) - (a_{S\alpha} + a_{\alpha S}) \leq 0.$$

Therefore,

$$(a_{S\phi} + a_{\phi S}) \leq (a_{S\alpha} + a_{\alpha S}),$$

which implies that the total adhesion of subset S to the rest of its community α is always greater than the adhesion to any other community ϕ. Let's assume for now that $(a_{S\alpha} + a_{\alpha S})$ is negative, then it seems that we could increase Q by moving subset S to a yet unpopulated community. But since we examine a configuration in the ground state, we can't increase Q, hence $(a_{S\alpha} + a_{\alpha S}) \geq 0$, and the adhesion of S to α is always positive. This is also true for all nodes l in community α, hence c_α is always positive. Let's assume for the sake of argument that now the adhesion between two different communities, α and ϕ, is positive. Then we could increase Q by merging the two communities together, since then

$$(a_{\alpha\phi} + a_{\phi\alpha}) \geq 0.$$

This move, however, would contradict the fact that $\{\sigma\}^*$ is optimal. Therefore the adhesion between any two communities is always negative. It is interesting to notice that this implies that communities can exist only in relation to other communities, and a community can never be all-encompassing, analogously to social identities. We can use these arguments to *define* a community, generalizing Reichardt and Bornholdt's definition. A community is a cohesive set of nodes ($c_s \geq 0$) which is distinct from other sets ($a_{sr} + a_{rs} \leq 0$) and every member is positively associated with its community ($a_{is} + a_{si} \geq 0$). Finally, we can *normalize* cohesion and adhesion, $n(a)$, by dividing actual by maximal values such that $0 \leq n(c_s) \leq 1$ and $-1 \leq n(a_{rs}) \leq 0$.

[1] Boccaletti, S, V Latora, Y Moreno, M Chavez and DU Hwang (2006) Complex networks: structure and dynamics, *Phys Rep* 424:175-308.
[2] Dorogovtsev, SN AV Goltsev and JFF Mendes (2007) Critical phenomena in complex networks. *arXiv:0705.0010v6* [cond-mat.stat-mech].
[3] Bruggeman, J (2008) *Social networks: An introduction* (Routledge, London).
[4] Duch, J and A Arenas (2005) Community detection in complex networks using extremal optimization. *Phys Rev E* 72:027104.
[5] Newman, MEJ (2006) Finding community structure in networks using the eigenvectors of matrices. *Phys Rev E* 74:036104.
[6] Palla, G, I Derényi, I Farkas, and T Vicsek (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435:814-818.
[7] Reichardt, J and S Bornholdt (2006) Statistical mechanics of community detection. *Phys Rev E* 74:016110.
[8] Harary, F (1953) On the notion of balance of a signed graph. *Michigan Math J* 2:143–146.
[9] Girvan, M and MEJ Newman (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99:7821–7826.
[10] Fortunato, S and M Barthélemy (2007) Resolution limit in community detection. *Proc Natl Acad Sci USA* 104:36–41.
[11] Haifeng Du, DR White, Yike Ren and Shuzhuo Li (2008) A normalized and a hybrid modularity. *Draft paper.*
[12] Kumpala, JM, J Saramäki, K Kaski and J Kertész (2007) Limited resolution in complex network community detection with Potts model approach. *Eur Phys J B* 56:41–45.
[13] Reichardt, J and DR White (2007) Role models for complex networks. *Phys J B* 60:217–224.
[14] Newman, MEJ and Girvan, M (2004) Finding and evaluating community structure in networks. *Phys Rev E* 69:026113.
[15] Brandes, U, D Delling, M Gaertler, R Görke, M Hoefer, Z Nikoloski and D Wagner (2006) Maximizing modularity is hard. *arXiv:physics/0608255v2* [physics.data-an].
[16] Kirkpatrick, S, CD Gelatt and MP Vecchi (1983) Optimization by simulated annealing. *Science* 220:671–680.
[17] Lancichinetti, A, S Fortunato, and F Radicchi (2008) New benchmark in community detection. *arXiv:0805.4770v3* [physics.soc-ph].
[18] Danon, L, A Díaz-Guilera, J Duch and A Arenas (2005) Comparing community structure identification. *J Stat Mech* P09008.
[19] Guimerà, R, and LA Nunes Amaral (2005) Functional cartography of complex metabolic networks. *Nature* 433:895–900.
[20] Newman, MEJ (2006) Modularity and community structure in networks. *Proc Natl Acad Sci USA* 103:8577–8582.
[21] Ruan, Jianhua and Weixiong Zhang (2008) Identifying network communities with a high resolution. *Phys Rev E* 77:016104.
[22] Leung, Ian XY, Pan Hui, P Liò, and J Crowcroft (2008) Towards real time community detection in large networks *arXiv:0808.2633v2* [physics.soc-ph]
[23] Csardi, G and T Nepusz (2006) The igraph software package for complex network research, *InterJournal Complex Systems*, http://cneurocvs.rmki.kfki.hu/igraph/.
[24] R Development Core Team (2008) R: A language and environment for statistical computing, http://www.R-project.org.
[25] Reichardt, J and S Bornholdt (2006) When are networks truly modular? *Physica D* 74:016110.
[26] Cartwright, D and F Harary (1979) Balance and clusterability: An overview, in PW Holland and S Leinhardt (eds.) *Perspectives on social network research* (Academic Press, New York).
[27] Heider, F (1946) Attitudes and cognitive organization. *J Psych* 21:107–112.
[28] Antal, T, PL Krapivsky and S Redner (2005) Dynamics of social balance on networks. *Phys Rev E* 72:036121.
[29] Davis, JA (1967) Clustering and structural balance in graphs. *Human Relations* 20:181–187.
[30] Toulouse, G (1977) Theory of the frustration effect in spin glasses I. *Commun Phys* 2:115.
[31] Vannimenus, J and G Toulouse (1977) Theory of the frustration effect II. Ising spins on a square lattice. *J Phys C* 10:537–542.
[32] Doreian, P and A Mrvar (1996) A partitioning approach to structural balance. *Social Networks* 18:149–168.
[33] Diestel, R (2005) *Graph theory* (Springer-Verlag, New York).
[34] Cartwright, D and F Harary (1968) On the coloring of signed graphs. *Elemente der Mathematik* 23:85–89.
[35] Ghosn, F, G Palmer and S Bremer (2004) The MID3 data set, 1993–2001: Procedures, coding rules, and description. *Conflict Management and Peace Science* 21:133–154.
[36] Gibler, DM and M Sarkees (2004) Measuring Alliances: The Correlates of War Formal Interstate Alliance Data set, 1816-2000. *J Peace Res* 41:211–222.
[37] Huntington, SP (1996) *The clash of civilizations and the remaking of world order* (Simon & Schuster, New York).
[38] Hensel, PR, G Goertz and PF Diehl (2000) The democratic peace and rivalries. *J of Politics* 62:1173–1188.
[39] De Tocqueville, A (1835, trans 2000) *Democracy in America* (Perennial, New York).
[40] Kissinger, H (1994) *Diplomacy* (Simon & Schuster, New York).
[41] Polachek, SW, J Robst and Yuan-Ching Chang (1999) Liberalism and interdependence: extending the trade-conflict model. *J Peace Res* 36:405–422.
[42] Henderson, EA and R Tucker (2002) Clear and Present Strangers: The Clash of Civilizations and International Conflict. *Int Stud Quart* 45(2): 317–338.
[43] Russett, BM, JR Oneal and M Cox (2000) Clash of Civilizations, or Realism and Liberalism Déjà Vu? Some Evidence. *J Peace Res* 37:583–608.
[44] Flake, GW, S Lawrence, CL Giles and FM Coetzee (2002) Self-Organization of the Web and Identfication of Communities. *IEEE Comp* 35:66–71.
[45] Adamic, LA and N Glance (2005) The political blogosphere and the 2004 U.S. election: divided they blog. *LinkKDD '05: Proceedings of the 3rd international workshop on Link discovery.*
[46] White, DR and F Harary (2001) The cohesiveness of blocks in social networks *J Soc Method* 31:305–359.