Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)

Network, Popularity and Social Cohesion: A Game-Theoretic Approach

Jiamou Liu, Ziheng Wei
The University of Auckland
Department of Computer Science
Auckland, New Zealand
jiamou.liu@auckland.ac.nz z.wei@auckland.ac.nz

Abstract

In studies of social dynamics, cohesion refers to a group's tendency to stay in unity, which – as argued in sociometry – arises from the network topology of interpersonal ties. We follow this idea and propose a game-based model of cohesion that not only relies on the social network, but also reflects individuals' social needs. In particular, our model is a type of cooperative games where players may gain popularity by strategically forming groups. A group is socially cohesive if the grand coalition is core stable. We study social cohesion in some special types of graphs and draw a link between social cohesion and the classical notion of structural cohesion (White and Harary 2001). We then focus on the problem of deciding whether a given social network is socially cohesive and show that this problem is CoNP-complete. Nevertheless, we give two efficient heuristics for coalition structures where players enjoy high popularity and experimentally evaluate their performances.

1 Introduction

Human has a natural desire to bind with others and needs to belong to groups. By understanding the basic instruments that generate coherent social groups, one can explain important phenomena such as the emergence of norms, group conformity, self-identity and social classes (Festinger 1950; Huisman and Bruggeman 2012; Hogg 1992). For example, studies reveal that on arrival to Western countries, immigrants tend to form cohesive groups within their ethnic communities, which may hamper their acculturation into the new society (Nee and Sanders 2001). Another study identifies cohesive groups of inhabitants in an Austrian village that correspond to stratified classes defined by succession to farmland ownership (Brudner and White 1997).

Most theories of group dynamics rely on two fundamental drives: *cooperation* and *social needs*. Indeed, every group exists to accomplish certain tasks. Cooperation is desirable because combining skills and resources leads to better outcomes. The theory of *cooperative games* studies distribution of collective gains among rational agents (Peleg and Sudhölter 2010). Social need is another important drive of group dynamics. A society contains complex interpersonal relations. The theory of self-categorization asserts that individuals mentally associate themselves into groups based on relations such as friendship and trust (Hogg 1992).

Cohesion denotes a tendency for a social group to stay in unity, which traditionally consists of two views. Firstly, cohesion refers to a "pulling force" that draws members together (Festinger 1950); Secondly, cohesion also means a type of "resistance" of the group to disruption (Gross and Martin 1952). In (2001), White and Harary propose a notion of *structural cohesion*, which unifies these two views.

However, we identify insufficiencies in the existing models for social cohesion: 1) Cooperative game theory often misses the crucial social network dimension. 2) The structural cohesion of a network refers to the minimum number of nodes whose removal results in network disintegration (White and Harary 2001); this is a property of the network on the whole, and does not embody individual needs. Since cohesion embodies both the micro-focus of psychology (fulfillment of personal objectives and needs) and the macro-focus of sociology (emergence of social classes) (Carron and Brawley 2000), the main challenge is to build a general but rigorous model to bridge the micro- and the macro-foci.

In this paper, we define a type of cooperative games on networked agents. Outcomes of the game not only rely on the network topology but also reflect individuals' social needs. Our model is consistent with existing theories: Firstly, we follow the network approach to study social phenomena, which is initiated by early pioneers such as Simmel and Durkheim. Secondly, our game-theoretic formulation is in line with group dynamics theories that focus on the interdependence among members (Lewin 1943). Thirdly, we verify that networks with high structural cohesion also tend to be socially cohesive according to our definition.

People prefer to be in a group where they are seen as valuable and popular members. Hence the payoff should reflect in some sense players' *social positions*. *Popularity* – an important indicator of social position – arises from interpersonal ties such as liking or attraction (Lansu and Cillessen 2011). In particular, (Conti et al. 2013) uses the degrees of nodes as a measure of popularity and identify the economic benefits behind gaining popularity. Therefore, payoffs in our games are defined based on degrees of players.

We summarize our main contributions: (1) We propose *popularity games* on a social network and define *social co-*

hesion using core stability of the grand coalition. (2) We show that our notion is consistent with intuition for several standard classes of networks and connect structural cohesion with our notion of social cohesion (Thm. 6). (3) We prove that deciding socially cohesion of a network is computationally hard (Thm. 7). (4) Finally, we present two heuristics that decide social cohesion and compute group structures with high player payoffs and evaluate them by experiments.

Related works. The series of works (Narayanam et al. 2014; Michalak et al. 2013; Szczepański et al. 2015a; 2015b; Szczepański, Michalak, and Wooldridge 2014) investigates game-based network centrality. Their aim is to capture a player's centrality using various instances of semivalues, which are based on the player's expected payoff. In contrast, our study aims at games where the payoff of players are given a priori by degree centrality and focus on core stability. (Chen et al. 2011) uses non-cooperative games to explain community formation in a social network. Each player in their game decides among a fixed set of strategies (i.e. a given set coalitions); the payoff is defined based on *gain* and *loss* which depend on the local graph structures. (McSweeney, Mehrotra, and Oh 2014) studies community formation through cooperative games. The payoffs of players are given by modularity and modularity-maximizing partitions correspond to Nash equilibria. The focus is on community detection but not on social cohesion. (Moskvina and Liu 2016) studies strategies to build social networks by establishing interpersonal ties. Furthermore, our payoff function is not additively separable and hence does not extends from their model. Lastly, our work is different from community detection (Fortunato 2010). The notion of community structure originates from physics which focuses on a macro view of the network, while our work is motivated from group dynamics and focus on individual needs and preferences.

2 Popularity Games and Social Cohesion

A *social network* is an unweighted graph $G = (V, E)$ where V is a set of nodes and E is a set of (undirected) edges. An edge $\{u, v\} \subseteq V$ (where $u \neq v$), denoted by uv, represents certain social relation between u, v, such as attraction, interdependence and friendship. We do not allow loops of the form uu. The reader is referred to (Peleg and Sudhölter 2010) for more details on cooperative game theory.

Definition 1. *A* cooperative game (with non-transferrable utility) *is a pair $G = (V, \rho)$ where V is a set of players, and $\rho : V \times 2^V \to \mathbb{R}$ is a* payoff function.

A *coalition formation* of G is a partition of V $\mathcal{W} = \{V_1, \dots, V_k\}$, i.e., $\bigcup_{1 \leq i \leq k} V_i = V, \forall 1 \leq i < j \leq k \colon V_i \cap V_j = \varnothing$; each set V_i is called a *coalition*. The *grand coalition formation* is $\mathcal{W}_G = \{V\}$ where V is called the *grand coalition*. Cooperative games describe situations where players strategically build coalitions based on individual payoffs. Set $\rho_{\mathcal{W}}(u) \coloneqq \rho(u, S)$ where $u \in S$ and $S \in \mathcal{W}$.

Definition 2. *A non-empty set of players $H \subseteq V$ is* blocking for $\mathcal{W}$ *if $\forall u \in H \colon \rho(u, H) > \rho_{\mathcal{W}}(u)$; In this case we say that $\mathcal{W}$ is* blocked *by H. A coalition formation $\mathcal{W}$ of G is* core stable *w.r.t. (V, ρ) if it is not blocked by any set $H \subseteq V$.*

Social positions, as argued in sociometry, arise from the network topology (Cillessen and Mayeux 2004). A long line of research studies how different *centralities* (e.g. degree, closeness, betweenness, etc.) give rise to "positional advantage" of individuals. In particular, degree centrality refers to the number of edges attached to a node. Despite its conceptual simplicity, degree centrality naturally represents (sociometric) *popularity*, which plays a crucial role in a person's self-efficacy and social needs (Zhang 2010; Conti et al. 2013). Popularity depends on the underlying group: a person may be very popular in one group while being unknown to another. Hence individuals may gain popularity by forming groups strategically. We thus make our next definition. The *sub-network induced* on a set $S \subseteq V$ is $G{\restriction}S = (S, E{\restriction}S)$ where $E{\restriction}S = E \cap S^2$. $\deg_S(u)$ denotes $|\{v \colon uv \in E{\restriction}S\}|$ and we write $\deg(u)$ for $\deg_V(u)$.

Definition 3. *The* popularity *of a node u in a subset $S \subseteq V$ is $p_S(u) \coloneqq \deg_S(u)/|S|$.*

Note that $p_{\{u\}}(u) = 0$ for every node u. If $u \in S$ has an edge to all other nodes in the graph $G{\restriction}S$, then u is the most popular node in S with $p_S(u) = (|S| - 1)/|S|$.

Definition 4. *The* popularity game *on $G = (V, E)$ is a cooperative game $\Gamma(G) = (V, \rho)$ where $\rho \colon V \times 2^V \to [0, 1)$ is defined by $\rho(u, S) = p_S(u)$.*

An outcome of the popularity game $\Gamma(G)$ assigns any player u with a coalition $S \ni u$. The sum of popularity of members of S equals their *average degree* in S, i.e. $\sum_{u \in S} p_S(u) = \sum_{u \in S} \deg_S(u)/|S| = 2|E{\restriction}S|/|S|$. The average degree measures the *density* of the set S, which reflects the amount of interactions within S, and thus can be regarded as a collective utility. In this sense, the popularity game is *efficient* in distributing such collective utility among players according to their popularity.

Social cohesion represents a group's tendency to remain united (Cartwright and Zander 1953). We express cohesion through *core stability* w.r.t. the popularity game $\Gamma(G)$: Suppose a coalition formation $\mathcal{W}$ is not core stable. Then there is a set $S \subseteq V$ every member of which would gain a higher popularity in S than in their own coalitions in $\mathcal{W}$. Thus there is a latent incentive among members of S to disrupt $\mathcal{W}$ and form a new coalition S. On the contrary, a core stable $\mathcal{W}$ represents a state of the network that is resilience to such "disruptions". Thus, when the grand coalition formation $\mathcal{W}_G = \{V\}$ is core stable, all members bind naturally into a single group and would remain so as long as the network topology does not change.

Definition 5. *A network $G = (V, E)$ is* socially cohesive *(or simply* cohesive*) if the grand coalition formation $\mathcal{W}_G$ is core stable w.r.t. the popularity game $\Gamma(G)$.*

Example 1. Fig. 1(a) displays $G_1 = (V_1, E_1)$. The popularity $p_V(i)$ is $1/3$ for $i = b, f$, and is $1/2$ for $i = a, c, d, e$. The set $\{a, b, c\}$ blocks $\mathcal{W}_{G_1}$ as each member has popularity $2/3$. The only core stable formation is $\{\{a, b, c\}, \{d, e, f\}\}$. Adding the edge ad makes G_1 cohesive as the popularity of both a, d in V_1 reaches $2/3$. Fig. 1(b) displays $G_2 = (V_2, E_2)$ where $p_V(a) = 4/5$ and $p_V(i) = 1/5$ for all $i = b, \dots, e$. This graph is cohesive as $\mathcal{W}_{G_2}$ is not blocked.

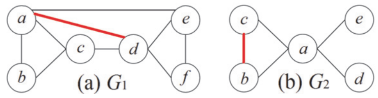

Figure 1: The graphs considered in Example 1 are in black. The added edges are highlighted in red.

However, adding the edge bc will destroy social cohesion as then $\{b, c\}$ blocks $\mathcal{W}_{G_2}$. Social psychological studies often presume that more ties leads to higher cohesion; this example displays a more complicated picture: Adding an edge may establish cohesion, but may also sabotage cohesion.

Theorem 1 (Connectivity)**.** *If a coalition formation $\mathcal{W}$ of G is core stable then any $S \in \mathcal{W}$ induces a connected subgraph. Thus any socially cohesive graphs are connected.*

Proof. Suppose $S = V_1 \cup V_2$ where V_1, V_2 are non-empty and no edge exists between any pair in $V_1 \times V_2$. Then for every $u \in V_1$, $p_{V_1}(u) = \frac{\deg_{V_1}(u)}{|V_1|} > \frac{\deg_{V_1}(u)}{|S|} = \frac{\deg_S(u)}{|S|} = p_S(u)$. Hence $\mathcal{W}$ does not contain S. $\square$

By Thm. 1, it is sufficient to consider only coalitions that induce connected sub-networks of a social network.

Definition 6. *A set $S \subseteq V$ is called a* social group *of G if S induces a connected sub-network. A* group structure *is a coalition formation containing only social groups.*

The next theorem shows that socially cohesive networks have bounded size.

Theorem 2. *Let $\delta(G)$ be the maximum degree of nodes in $G{=}(V, E)$. Then G is socially cohesive only when $|V| \leq 2\delta(G)$ unless $|V| = 1$.*

Proof. Suppose $|V| > 2\delta(G)$ and $|V| > 1$. If $E = \varnothing$, G is not socially cohesive by Thm. 1. Otherwise, pick an edge uv. Then $\max\{\deg(u), \deg(v)\} \leq \delta(G) < |V|/2$. This means that $\max\{p_V(u), p_V(v)\} < 1/2$, and the edge $\{u, v\}$ forms a blocking set. Thus G is not socially cohesive. $\square$

We now investigate our games on some standard classes of graphs and characterize core stable group structures.

Complete networks. A graph $G = (V, E)$ is *complete* if any pair of nodes is linked. Naturally, one would expect complete networks to be socially cohesive.

Theorem 3. *Let G be a complete network. The grand coalition is the only core stable group structure.* $\square$

Proof. Any induced sub-network $G{\restriction}S$ of a complete network $G = (V, E)$ is also complete. Thus $p_S(u) = \frac{|S|-1}{|S|} < \frac{|V|-1}{|V|} = p_V(u)$. Therefore any player's popularity is maximized in the grand coalition V. $\square$

Star networks. A *star network* contains a node c (*centre*), a number of other nodes $u_1, \ldots, u_m$ (*tails*) where $m > 1$, and edges $\{cu_1, \ldots, cu_m\}$. Intuitively, the centre c would like to be in a social group with as many players as possible, while a tail would like to be with as few others as possible.

Theorem 4. *A group structure $\mathcal{W}$ of a star network is core stable iff the centre is in the same social group with at least half of the tails. Thus, any star network is socially cohesive.*

Proof. Take any group structure $\mathcal{W}$ and suppose the centre c is in a social group S with ℓ tails. Then $p_S(c) = \ell/(\ell + 1)$ and for any tail $u_i \in S$, $p_S(u_i) = 1/(\ell + 1)$. All players not in S has popularity 0.

Suppose $\ell \geq m/2$. Take any set $S' \neq S$ that contains c. If $|S'| \leq |S|$, then $p_{S'}(c) \leq p_S(c)$. If $|S'| > |S|$, then $p_{S'}(v) < p_S(v)$ for some tail v. In either case S' does not block $\mathcal{W}$. Hence $\mathcal{W}$ is core stable. Suppose $\ell < m/2$. Then let N be the set of all tails not in S. Then $N \cup \{c\}$ blocks $\mathcal{W}$. Thus $\mathcal{W}$ is core stable iff $\ell \geq m/2$. $\square$

Complete Bipartite Graph. A *complete bipartite graph* $K_{n,n}$ consists of disjoint sets of nodes V_1, V_2 with n nodes each and $E = \{uv \mid u \in V_1, v \in V_2\}$. Let $\mathcal{W}$ be a group structure. For every $S \in \mathcal{W}$, we use $\ell(S)$ and $r(S)$ to denote $|\{v \mid v \in S \cap V_1\}|$ and $|\{v \mid v \in S \cap V_2\}|$, respectively.

Lemma 1. *$\mathcal{W}$ is core stable only if $\forall S \in \mathcal{W}\colon \ell(S) \geq r(S)$.*

Proof. Suppose there is $S \in \mathcal{W}$ with $\ell(S) < r(S)$. Since $m \geq n$, there is $H \in \mathcal{W}$ with $\ell(H) > r(H)$. Take any $u \in S \cap V_2$ and $v \in H \cap V_1$. Then we have $p_S(u) = \frac{r(S)}{\ell(S)+r(S)} < \frac{1}{2}$ and $p_H(v) = \frac{\ell(H)}{\ell(H)+r(H)} < \frac{1}{2}$. Hence, the set $\{u, v\}$ blocks $\mathcal{W}$ as $p_{\{u,v\}}(u) = p_{\{u,v\}}(v) = 1/2$. $\square$

We next characterize core stable group structure in $K_{n,n}$. In particular, *perfect matchings*, i.e., situations when every $v \in V_1$ is matched with a unique player in V_2, are core stable.

Theorem 5. *A group structure $\mathcal{W}$ of $K_{n,n}$ is core stable iff $\forall S \in \mathcal{W}\colon \ell(S) = r(S)$.*

Proof. By Lem. 1, if $\mathcal{W}$ is core stable then $\forall S \in \mathcal{W}\colon \ell(S) = r(S)$. Conversely, if $\forall S \in \mathcal{W}\colon \ell(S) = r(S)$, then any v has payoff $\frac{1}{2}$. Thus $\mathcal{W}$ is core stable as every $H \subseteq V$ contains some player with payoff at most $\frac{1}{2}$. $\square$

3 Structural Cohesion and Social Cohesion

Definition 7 (White and Harary 2001)**.** *The* structural cohesion *κ of a connected graph G is the minimal number of nodes upon removal of which G become disconnected.*

By Def. 7, a larger κ implies that G is more resilient to conflicts or the departure of group members, and is thus more cohesive. Moreover, Menger's theorem states that κ is the greatest lower bound on the number of paths between any pairs of nodes. Hence κ is a reasonable measure of cohesion. We next link κ with our notion of social cohesion. In (Granovetter 1973), a pair $uv \notin E$ is seen as a "structural hole" that forbids communication and is thus referred to as an *absent tie*. For each $S \subseteq V$ and $u \in S$, we define:

- $f_{\mathsf{in}}(u, S) := \deg_S(u)$ and $f_{\mathsf{out}}(u, S) := |\{v \notin S \mid uv \in E\}|$ are the numbers of actual ties of u within the group S and outside S, resp.
- $s_{\mathsf{in}}(u, S) := |S| - f_{\mathsf{in}}(u, S)$ and $s_{\mathsf{out}}(u, S) := |\{v \notin S \mid uv \notin E\}|$ are the number of absent ties in S (including u itself) and outside S, resp.

Intuitively, if $S \subseteq V$ is blocking, each member u tends to have many actual ties within S and absent ties outside S, i.e., high $f_{\mathsf{in}}(u,S)$ and $s_{\mathsf{out}}(u,S)$, and u tends to have few absent ties in S and actual ties outside S, i.e., low $f_{\mathsf{out}}(u,S)$ and $s_{\mathsf{in}}(u,S)$. Thus, we define for all $S \subseteq V$, $u \in S$,

$$\gamma(u,S) \coloneqq f_{\mathsf{in}}(u,S)s_{\mathsf{out}}(u,S) - f_{\mathsf{out}}(u,S)s_{\mathsf{in}}(u,S) \quad (1)$$

Lemma 2. *For all $S \subseteq V$, S blocks $\mathcal{W}_G = \{V\}$ iff $\forall u \in S\colon \gamma(u,S) > 0$.*

Proof. For each $u \in S$, $p_S(u) = \frac{f_{\mathsf{in}}(u,S)}{f_{\mathsf{in}}(u,S)+s_{\mathsf{in}}(u,S)}$ and

$$p_V(u) = \frac{f_{\mathsf{in}}(u,S) + f_{\mathsf{out}}(u,S)}{f_{\mathsf{in}}(u,S) + f_{\mathsf{out}}(u,S) + s_{\mathsf{in}}(u,S) + s_{\mathsf{out}}(u,S)}$$

The set S blocks $\mathcal{W}_G$ iff $\forall u \in S\colon p_S(u) > p_V(u)$, which can be shown to be equivalent to $f_{\mathsf{in}}(u,S) \cdot s_{\mathsf{out}}(u,S) > f_{\mathsf{out}}(u,S) \cdot s_{\mathsf{in}}(u,S)$ using the above equalities. □

A network G contains a *minimal cut* $A_0 \subseteq V$ of size κ, i.e., removing A_0 from G decomposes the graph into m distinct connected components $A_1, \ldots, A_m \subseteq V$ where $m \geq 2$. We further assume that $|A_1| \leq \cdots \leq |A_m|$ and A_0 is chosen in a way where $|A_1|$ is as small as possible. Let χ be the size $|A_1|$, and let μ be the largest possible length m of the sequence of A_i's. We first look at the case when $\kappa = 1$.

Lemma 3. *If κ=1 and G is socially cohesive, then χ<2.*

Proof. Suppose $\kappa = 1$ and $\chi \geq 2$. Let $(A_1, \ldots, A_m)$ be an optimal cut sequence. Take $u \in A_1$. As G contains a cut node, $f_{\mathsf{out}}(u,A_1) \leq 1$ and $s_{\mathsf{out}}(u,A_1) \geq |V| - \chi - 1$. Then $\gamma(u,A_1) \geq f_{\mathsf{in}}(u,A_1) \cdot (|V| - \chi - 1) - s_{\mathsf{in}}(u,A_1)$. Since $f_{\mathsf{in}}(u,A_1) + s_{\mathsf{in}}(u,A_1) = \chi$,

$$\begin{aligned}\gamma(u,A_1) &\geq f_{\mathsf{in}}(u,A_1)(|V| - \chi - 1) - (\chi - f_{\mathsf{in}}(u,A_1)) \\ &= f_{\mathsf{in}}(u,A_1)(|V| - \chi) - \chi.\end{aligned}$$

Since $|V| - \chi > \chi$, $\gamma(u,A_1) > 0$. Thus by Lem. 2, A_1 blocks the grand coalition $\mathcal{W}_G$. □

Lemma 4. *Suppose $\mu > 2$. Then any network G is socially cohesive only if $\chi < \frac{\kappa}{\mu-2}$.*

Proof. Suppose $\mu > 2$. Take an optimal cut sequence $(A_1, ..., A_\mu)$ and $u \in V_1$. Since $\deg(u) < \chi + \kappa$ and $|V| \geq \mu\chi + \kappa$, we have $p_V(u) < \frac{\chi+\kappa}{\mu\cdot\chi+\kappa}$. Suppose $\chi \geq \frac{\kappa}{\mu-2}$. Then $\mu\chi - 2\chi \geq \kappa$. One can then derive $p_V(u) < \frac{\chi+\kappa}{\mu\cdot\chi+\kappa} \leq \frac{1}{2}$. Thus any edge $\{u,v\}$ in $G{\restriction}V_1$ forms a blocking set of the grand coalition formation $\mathcal{W}_G$. □

Lem. 4 can be used as a (semi-)test for social cohesion when $\mu > 2$: whenever χ exceeds $\frac{\kappa}{\mu-2}$, G is not socially cohesive. Clearly, more graphs become socially cohesive as κ gets larger. Summarizing Lem. 3 and 4, we get:

Theorem 6. *Let G be a network. (1) If $\kappa = 1$, then G is not socially cohesive for all $\chi \geq 2$. (2) If $\kappa > 1$ and $\mu > 2$, then G is not socially cohesive for all $\chi \geq \frac{\kappa}{\mu-2}$.*

Remark. The only case left unexplained is when $\kappa > 1$ and $\mu = 2$. In this case there exist graphs with arbitrarily large χ but are socially cohesive.

4 The Computational Complexity of Deciding Social Cohesion

We are interested in the decision problem COH: Given a network $G = (V,E)$, decide if G is socially cohesive.

The *distance* between two nodes u and v, denoted by $\mathsf{dist}(u,v)$, is the length of a shortest path from u to v in G. The *eccentricity* of u is $\mathsf{ecc}(u) = \max_{v \in V} \mathsf{dist}(u,v)$. The *diameter* of the network G is $\mathsf{diam}(G) = \max_{u \in V} \mathsf{ecc}(u)$. A graph $G = (V,E)$ is *diametrically uniform* if all $v \in V$ have the same eccentricity; otherwise G is called *non-diametrically uniform*. We use NDU_2 to denote the set of all non-diametrically uniform connected graphs whose diameter is at most 2. Our goal is to show that the COH problem is computationally hard already on the class NDU_2.

Theorem 7. *The network G belongs to NDU_2 iff its nodes V can be partitioned into two non-empty set V_1 and V_2, where $V_1 = \{u \mid vu \in E \text{ for all } v \neq u\}$.* □

Let G be a graph in NDU_2. We call $\{V_1, V_2\}$ as described in Thm. 7 the *eccentricity partition* of G. We first present some simple properties of NDU_2.

Lemma 5. *The network G in NDU_2 is socially cohesive iff no set $S \subseteq V_2$ blocks $\mathcal{W}_G$.*

Proof. One direction (left to right) is clear. Conversely, suppose the network is not socially cohesive. Let $S \subset V$ be a blocking set of the grand coalition formation, i.e., $\forall u \in S\colon\ p_S(u) > p_V(u)$. If $S \cap V_1 \neq \varnothing$. Then $\forall u \in S \cap V_1\colon p_V(u) = \frac{|V|-1}{|V|}$. However, $p_V(u) \geq \frac{|S|-1}{|S|} \geq p_S(u)$ which contradicts that fact that S is a blocking set. □

By Lem. 5, the structure of $G{\restriction}V_2$ is crucial in determining social cohesion of G. For any $S \subseteq V_2$ and $u \in S$, we recall the notions $f_{\mathsf{in}}(u,S), f_{\mathsf{out}}(u,S), s_{\mathsf{in}}(u,S)$, and $s_{\mathsf{out}}(u,S)$ from Section 3, but re-interpret these values within the subnetwork $G{\restriction}V_2$. Hence, we now set $f_{\mathsf{out}}(u,S)$ as $|\{v \in V_2 \setminus S \mid uv \in E\}|$, i.e., the number of ties that u has within V_2 but not in S, the other variables remain as originally defined. Thus

$$|V_2| = f_{\mathsf{in}}(u,S) + f_{\mathsf{out}}(u,S) + s_{\mathsf{in}}(u,S) + s_{\mathsf{out}}(u,S) \quad (2)$$

We then define the value

$$\lambda(u,S) = \frac{f_{\mathsf{in}}(u,S) \cdot s_{\mathsf{out}}(u,S)}{s_{\mathsf{in}}(u,S)} - f_{\mathsf{out}}(u,S)$$

Theorem 8. *A network G in NDU_2 is socially cohesive iff for all $S \subseteq V_2$ there exists $v \in S$ s.t. $|V_1| \geq \lambda(v,S)$*

Proof. By Lem. 5, we only need to examine subsets $S \subseteq V_2$. Every $u \in S$ has $p_S(u) = \frac{f_{\mathsf{in}}(u,S)}{f_{\mathsf{in}}(u,S)+s_{\mathsf{in}}(u,S)}$ and

$$p_V(u) = \frac{f_{\mathsf{in}}(u,S) + f_{\mathsf{out}}(u,S) + |V_1|}{|V_2| + |V_1|}.$$

Applying (2), S blocks $\mathcal{W}_G$ iff $\forall u \in S : p_S(u) > p_V(u)$, iff $\forall u \in S : |V_1| < \frac{f_{\mathsf{in}}(u,S)s_{\mathsf{out}}(u,S)}{s_{\mathsf{in}}(u,S)} - f_{\mathsf{out}}(u,S) = \lambda(u,S)$. □

We now give a sufficient condition for social cohesion of NDU_2. The *size* of a network is its number of nodes. A *clique* is a complete subgraph. The *clique number* of G, denoted by $\omega(G)$, is the size of the largest clique. Turán's theorem relates $\omega(G)$ with the number of edges in G:

Theorem 9 (Turán 1941)**.** *For any $p{\geq}2$, if a graph G of size n has more than $\frac{p-2}{2(p-1)}n^2$ edges, then $\omega(G){\geq}p$.* □

Lemma 6. *For any social group $S \subseteq V$, there exists $u \in S$ with $\frac{f_{\mathsf{in}}(u,S)}{s_{\mathsf{in}}(u,S)}{\leq}\omega(G{\restriction}S){-}1$.*

Proof. Let $k = \omega(G{\restriction}S)$ and suppose for all $u \in S$, $\frac{f_{\mathsf{in}}(u,S)}{s_{\mathsf{in}}(u,S)} > k-1$. Since $|S| = f_{\mathsf{in}}(u,S) + s_{\mathsf{in}}(u,S)$,

$$\begin{aligned}&k-1 < f_{\mathsf{in}}(u,S)/(|S| - f_{\mathsf{in}}(u,S))\\ &f_{\mathsf{in}}(u,S) > k|S| - kf_{\mathsf{in}}(u,S) - |S| + f_{\mathsf{in}}(u,S)\\ &f_{\mathsf{in}}(u,S) > (k-1)|S|/k\end{aligned}$$

Thus $E{\restriction}S$ contains $> \frac{k-1}{2k}|S|^2$ edges. By Thm. 9, $G{\restriction}S$ contains a $(k+1)$-clique, contradicting k's definition. □

The set V_1 contains the most socially active members – those who interact with everyone else. Hence the larger V_1 gets, the more likely G will be socially cohesive. There is a bound such that once $|V_1|$ exceeds it, the network G is guaranteed to be socially cohesive.

Lemma 7. *Suppose $c = \omega(G{\restriction}V_2)$ and $|V_2| > c(c{-}1)$. Then G is socially cohesive if $|V_1| \geq (c-1)(|V_2| - c)$.*

Proof. Suppose $|V_2| > c(c-1)$, $|V_1| \geq (c-1)(|V_2| - c)$. Take any $S \subseteq V_2$. If S has a size-c clique, by Lem. 6 there exists $u{\in}V_2$ with $\frac{f_{\mathsf{in}}(u,S)}{s_{\mathsf{in}}(u,S)} \leq c{-}1$. Since $s_{\mathsf{out}}(u,S) \leq |V_2|{-}c$, $|V_1| \geq (c-1)(|V_2| - c) \geq \frac{f_{\mathsf{in}}(u,S)}{s_{\mathsf{in}}(u,S)}s_{\mathsf{out}}(u,S) > \lambda(u,S)$ and thus G is socially cohesive by Thm. 8.

Let $k{=}\omega(G{\restriction}S)$. If S contains no c-clique, $k{\leq}c{-}1$. By Lem. 6, $\frac{f_{\mathsf{in}}(u,S)}{s_{\mathsf{in}}(u,S)}{\leq}k{-}1{<}c{-}1$. Thus $c{-}\frac{|V_2|}{s_{\mathsf{in}}(v,S)}{\geq}1$. Since $s_{\mathsf{out}}(u,S){=}|V_2|{-}f_{\mathsf{in}}(u,V_2){-}s_{\mathsf{in}}(u,S)$, $\lambda(u,S) = \frac{f_{\mathsf{in}}(u,S)(|V_2|-f_{\mathsf{in}}(u,V_2))}{s_{\mathsf{in}}(u,S)} - f_{\mathsf{in}}(u,V_2)$. Since $f_{\mathsf{in}}(u,S){=}|S| - s_{\mathsf{in}}(u,S)$,

$$\lambda(u,S) = \frac{|S|\cdot|V_2|}{s_{\mathsf{in}}(u,S)} - \frac{|S|f_{\mathsf{in}}(u,V_2)}{s_{\mathsf{in}}(u,S)} - |V_2|$$

Hence, $|V_1| - \lambda(u,S)$ is at least

$$\begin{aligned}&(c-1)(|V_2|-c) - \left(\frac{|S|\cdot|V_2|}{s_{\mathsf{in}}(u,S)} - \frac{|S|f_{\mathsf{in}}(u,V_2)}{s_{\mathsf{in}}(u,S)} - |V_2|\right)\\ &= |V_2|\left(c - \frac{|S|}{s_{\mathsf{in}}(u,S)}\right) - c(c-1) + \frac{|S|f_{\mathsf{in}}(u,V_2)}{s_{\mathsf{in}}(u,S)} \geq 0\end{aligned}$$

The last step is by $c{-}\frac{|S|}{s_{\mathsf{in}}(v,S)}{\geq}1$ and $|V_2|{>}c(c{-}1)$. By Thm. 8, G is socially cohesive. □

Theorem 10. *The problem* COH *is* CoNP*-complete. Furthermore* CoNP*-hardness holds for the class* NDU_2.

Proof. The complement of COH, $\overline{\mathsf{COH}}$, asks whether a set S blocks the grand coalition $\mathcal{W}_G$ of a given network G; this problem is clearly in NP and thus COH is in CoNP. For hardness, we reduce $\mathsf{MaxClique}$ (asking whether a graph contains a clique of a given size k) to $\overline{\mathsf{COH}}$. CoNP-hardness of COH then follows from the NP-hardness of $\mathsf{MaxClique}$ (Garey and Johnson 1979).

Algorithm 1 Construction of H given $G{=}(V,E)$ and $k{>}2$

Set $d \coloneqq k \cdot \max\{\deg(u) \mid u \in V\}$
Create G' by adding $k(k{-}1){+}d$ isolated nodes to G
Let V_2 be the set of nodes in G'
Create a complete graph with $(k{-}1)(|V_2|{-}k){-}d$ nodes; Let V_1 be the set of these nodes
Create edges $\{uv \mid u \in V_1, v \in V_2\}$ to connect V_1, V_2. The resulting graph is H

To this end, we construct, for a given $G = (V,E)$ and $k > 2$, a graph $H \in \mathsf{NDU}_2$ as in Alg. 1. Our goal is to show that H is not socially cohesive iff G contains a clique of size k. It is clear that H is a NDU_2 network with eccentricity partition $\{V_1, V_2\}$. Let $c = \omega(G)$. Suppose $c < k$. By definition of V_1 and V_2, we have $|V_1| - (c-1)(|V_2| - c) = |V_2|(k-c) - k(k-1) + c(c-1) - d$. Since $k{>}c$ and $|V_2|{>}k(k{-}1){+}d$, $|V_1|{\geq}(c{-}1)(|V_2|{-}c)$. By Thm. 7, H is cohesive.

Conversely, suppose $\omega(G) \geq k$. Let $C \subseteq V_2$ be a clique of size k. Take $u \in C$. Since $f_{\mathsf{in}}(u,C){=}k{-}1$, $s_{\mathsf{out}}(u,C){=}|V_2|{-}k{-}f_{\mathsf{out}}(u,C)$ and $s_{\mathsf{in}}(u,C) = 1$,

$$\begin{aligned}\lambda(u,C) &= (k-1)(|V_2| - k - f_{\mathsf{out}}(u,C)) - f_{\mathsf{out}}(u,C)\\ &= (k-1)(|V_2| - k) - k \cdot f_{\mathsf{out}}(u,C)\end{aligned}$$

Hence, $\lambda(u,C){-}|V_1| = d - kf_{\mathsf{out}}(u,C)$. Since $d \geq k\deg(u)$, $\lambda(u,C) - |V_1| > 0$. By Lem. 8, G is not socially cohesive. Therefore, G contains a clique of size k iff H is not socially cohesive and the reduction is complete. □

5 Efficient Heuristics

We propose two heuristics that construct group structures of a given network where players enjoy high popularity. These heuristics (partially) solve COH despite COH's inherent complexity: Each heuristic builds a group structure $\mathcal{W}$ and checks if any set $S{\in}\mathcal{W}$ blocks $\mathcal{W}_G$. If G is socially cohesive, then no such S will be found; On the other hand, if a blocking set S is found, G is surely not socially cohesive.

Heuristic 1: Louvain's method (LM) We observe that blocking sets of W_G are usually tightly connected within, but are sparsely connected with nodes outside. This property corresponds to the well-studied notion of *communities* (Fortunato 2010). Therefore, the first heuristic uses a well-known community detection algorithm, Louvain's method (Blondel et al. 2008), to compute a group structure in G.

Heuristic 2: Average payoff (AP) The second heuristic aims to optimize the average payoffs of members of a coalition. Socially cohesive networks usually have small diameters (≤ 2). Thus we consider neighborhood $N(v) \coloneqq \{v\} \cup \{u \mid uv \in E\}$ of players $v \in V$. In Alg. 2, let $\nu(S)$ be the average payoff $\sum_{u\in S}\rho(u,S)/|S|$ in any set $S \subseteq V$.

Algorithm 2 AP: Given a network $G = (V, E)$

Initialize set $V' := V$ and a group structure $\mathcal{W} := \varnothing$
while $|V'| > 0$ **do**
  compute $\nu_w := \nu(N(w) \cap V')$ for every $w \in V'$
  $S := N(v) \cap V'$ such that $\forall u \in V'\colon \nu_u \leq \nu_v$
  $\mathcal{W} := \mathcal{W} \cup \{S\}$, $V' := V' \setminus S$
end while
return $\mathcal{W}$

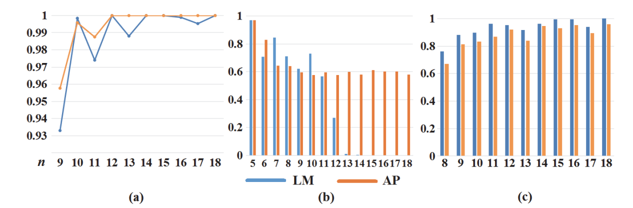

Figure 2: The results of running LM and AP on graphs with $n = 5, \ldots, 18$. (a) Accuracy (b) The percentage of graphs are found core stable by the heuristics (c) The percentage of nodes with an increased payoff in the found group structure than in the grand coalition.

Experiments. We evaluate the heuristics on graphs of size $n = 5, \ldots, 18$. For each size, let s be the total number of graphs, b be the number of graphs for which the heuristic finds a blocking set, and c be the number of socially cohesive graphs. The heuristic thus correctly solves COH for $b + c$ graphs. Hence the heuristic has *accuracy* $(b + c)/s$. For $n = 5, 6, 7$, we enumerate all connected graphs; LM has accuracy 96.8%, 97.7% and 83.1%, resp., while AP has 85.9%, 68.5% and 69.2%, resp. For each $n = 8, \ldots, 18$, we generate 10^5 Erdös-Renyí random graphs of size n. As shown in Fig. 2(a), both heuristics achieve high accuracy. As n increases, socially cohesive graphs become increasingly rare. The results show that the heuristics successfully find blocking sets in almost all cases, with AP performs slightly better (100% accuracy for $n \geq 12$). Note that the fluctuation is within a very small range and is due to the small graph sizes. We then consider the coalitions constructed by the heuristics. Fig. 2(b) shows that while LM fails to output core stable group structures, AP achieves core stability in 60% of the sampled cases when $n \geq 7$. Nevertheless, Fig. 2(c) shows that, compared to the payoffs in the grand coalition, more nodes get a higher payoff in the coalitions identified by LM.

Real world networks. We further evaluate the heuristics on 8 real-world networks: karate club ZA (Zachary 1977), dolphins DO (Lusseau et al. 2003), college football FT (Girvan and Newman 2002), Facebook FB, Enron email network EN (Leskovec et al. 2009), and three physics collaboration networks AS, CM and HE (Leskovec et al. 2007). We only use the largest components in each network; see details in Table. 1. Expectedly, none of these networks are socially cohesive. The box-and-whisker diagrams in Fig. 3 show the distribution of payoffs of players in the grand coalition as well as in the coalitions output by each heuristic (outliers

Table 1: N, E, C denote the number of nodes, edges, and communities, respectively.

	ZA	DO	FT	FB	EN	AS	CM	HE
N	34	62	115	3927	36696	17903	21363	11204
E	78	159	613	84210	180811	197031	91342	117649
C	4	5	10	27	248	36	58	37

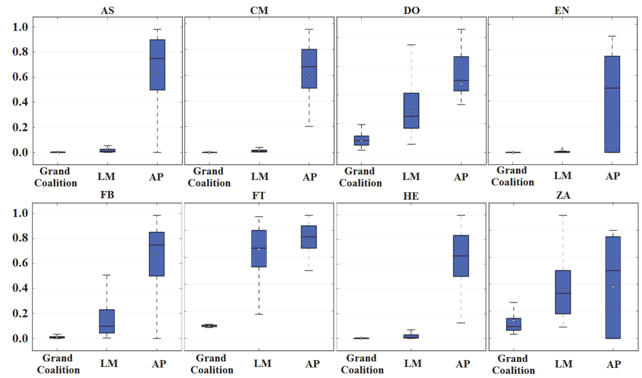

Figure 3: The distribution of payoffs of players in the grand coalition and in the coalitions computed by the heuristics in the real-world networks.

omitted). In all cases, the heuristics improve players' payoffs considerably compared to the grand coalition, while AP in particular achieves higher payoffs. Furthermore, Fig. 4 shows all nodes get higher payoffs through LM. In summary, both of the heuristics are useful in computing coalitions; while LM may benefit a larger portion of players, AP tends to obtain higher payoffs.

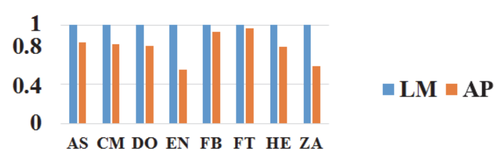

Figure 4: The percentage of nodes with a higher payoff (compared to the grand coalition) in coalitions generated by the heuristics for real-world networks.

6 Conclusion and Future Work

We aim to investigate natural game-theoretical and computational questions as future works: Does a core stable group structure exists for every network? What about other stability concepts? What would be strategies of players to improve popularity? The proposed games are instances of a general game-theoretical framework for networked agents, whose payoffs are given by various centrality indices. It will be interesting to extend the work by considering other centralities and different forms of social networks (e.g. directed, weighted, signed networks). Furthermore, one could also explore the evolution of social groups in a dynamic setting.

References

Blondel, V. D.; Guillaume, J. L.; Lambiotte, R.; and Lefebvre, E. 2008. Fast unfolding of communities in large networks. *J. Stat. Mech.-Theory E.* 10.

Brudner, L., and White, D. 1997. Class, property, and structural endogamy: Visualizing networked histories. *Theor. Soc.* 26:161–208.

Carron, A. V., and Brawley, L. R. 2000. Cohesion conceptual and measurement issues. *Small. Gr. Res.* 31(1).

Cartwright, D., and Zander, A. 1953. *Group Dynamics: Research and Theory*. Evanston, IL: Row, Peterson.

Chen, W.; Liu, Z.; Sun, X.; and Wang, Y. 2011. Community detection in social networks through community detection games. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'11)*, volume 22, No. 3.

Cillessen, A., and Mayeux, L. 2004. Sociometric status and peer group behavior: Previous findings and current directions. In Kupersmidt, J., and Dodge, K., eds., *Children's peer relations: From development to intervention. Decade of behavior*. American Psychological Association. 3–20.

Conti, G.; Galeotti, A.; Mueller, G.; and Pudney, S. 2013. Popularity. *J. Hum. Resour.* 48(4).

Festinger, L. 1950. The spatial ecology of group formation. (l. festinger, s. schachter, & k. back, eds.). *Social Pressure in Information Groups* 33–59.

Fortunato, S. 2010. Community detection in graphs. *Phys. Rep.* 486(3):75–174.

Garey, M. R., and Johnson, D. S. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H.Freeman.

Girvan, M., and Newman, M. E. 2002. Community structure in social and biological networks. In *P. Natl. Acad. Sci. USA*, volume 99(12).

Granovetter, M. S. 1973. The strength of weak ties. *Am J. Sociol.* 78(2):1360–1380.

Gross, N., and Martin, W. E. 1952. On group cohesiveness. *Am. J. Sociol.* 57:546–564.

Hogg, M. A. 1992. *The social psychology of group cohesiveness: From attraction to social identity*. Harvester Wheatsheaf.

Huisman, C., and Bruggeman, J. 2012. The social network, socioeconomic background, and school type of adolescent smokers. *Int. J. Behav. Dev.* 36(5).

Lansu, T. A., and Cillessen, A. H. 2011. Peer status in emerging adulthood: Associations of popularity and preference with social roles and behavior. *J. Adolescent. Res.*

Leskovec, J.; Kleinberg, J.; ; and Faloutsos, C. 2007. Graph evolution: Densification and shrinking diameters. *ACM TKDD* 1(1).

Leskovec, J.; Lang, K. J.; Dasgupta, A.; and Mahoney, M. W. 2009. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Math.* 6(1).

Lewin, K. 1943. Psychology and the process of group living. *J. Soc. Psychol.* 17:113–131.

Lusseau, D.; Schneider, K.; Boisseau, O. J.; Haase, P.; Slooten, E.; and Dawson, S. M. 2003. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behav. Ecol. Sociobiol.* 54(4).

McSweeney, P. J.; Mehrotra, K.; and Oh, J. C. 2014. Game-theoretic framework for community detection. In *Encyclopedia of Social Network Analysis and Mining*. Springer New York.

Michalak, T. P.; Aadithya, K. V.; Szczepanski, P. L.; Ravindran, B.; and Jennings, N. R. 2013. Efficient computation of the shapley value for game-theoretic network centrality. *J. Artif. Intell. Res.*

Moskvina, A., and Liu, J. 2016. How to build your network? A structural analysis. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16), New York, NY, USA, 9-15 July 2016*, 2597–2603.

Narayanam, R.; Skibski, O.; Lamba, H.; and Michalak, T. 2014. A shapley value-based approach to determine gatekeepers in social networks with applications. In *ECAI'14*, 651–656. IOS Press.

Nee, V., and Sanders, J. 2001. Trust in ethnic ties: Social capital and immigrants. *Trust in Society* 2.

Peleg, B., and Sudhölter, P. 2010. *Introduction to the Theory of Cooperative Games (2nd Ed.)*. Springer.

Szczepański, P. L.; Barcz, A.; Michalak, T. P.; and Rahwan, T. 2015a. The game-theoretic interaction index on social networks with applications to link prediction and community detection. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, volume 22, No. 3. AAAI Press.

Szczepański, P.; Tarkowski, M.; Michalak, T.; Harrenstein, P.; and Wooldridge, M. 2015b. Efficient computation of semivalues for game-theoretic network centrality. In *Proceedings of the 29th AAAI Conference on AI (AAAI'15)*. AAAI Press.

Szczepański, P.; Michalak, T.; and Wooldridge, M. 2014. A centrality measure for networks with community structure based on a gneralization of the owen value. In *Proceedings of the 21th European Conference on Artificial Intelligence (ECAI'14)*, 867–872.

Turán, P. 1941. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok (in Hungarian)* 48:436–452.

White, D. R., and Harary, F. 2001. The cohesiveness of blocks in social networks: Node connectivity and conditional density. *Social. Methodol.*

Zachary, W. 1977. An information flow model for conflict and fission in small groups. *J. Anthropol. Res.* 33.

Zhang, M. 2010. Social network analysis: History, concepts and research. *Handbook of Social Network Technologies and Applications*.