Nash Stability in Social Distance Games

Alkida Balliu
Gran Sasso Science
Institute, Italy.
alkida.balliu@gssi.infn.it

Michele Flammini
DISIM - University of
L'Aquila & Gran Sasso
Science Institute, Italy.
michele.flammini@univaq.it

Giovanna Melideo
DISIM - University
of L'Aquila, Italy.
giovanna.melideo@univaq.it

Dennis Olivetti
Gran Sasso Science
Institute, Italy.
dennis.olivetti@gssi.infn.it

Abstract

We consider Social Distance Games (SDGs), that is cluster formation games in which agent utilities are proportional to their harmonic centralities in the respective coalitions, i.e., to the average inverse distance from the other agents. We adopt Nash stable outcomes, that is states in which no agent can improve her utility by unilaterally changing her coalition, as the target solution concept. Although SDGs always admit a Nash equilibrium, we prove that it is NP-hard to find a social welfare maximizing one and obtain a negative result concerning the game convergence. We then focus on the performance of Nash equilibria and provide matching upper bound and lower bounds on the price of anarchy of $\Theta(n)$, where n is the number of nodes of the underlying graph, and a lower bound on the price of stability of $6/5 - \epsilon$. Finally, we characterize the price of stability of SDGs for graphs with girth 4 and girth at least 5.

1 Introduction

Coalition formation is a pervasive aspect of social life and it has been studied extensively in algorithmic game theory using the natural model of *Hedonic Games* (HGs), introduced in (Dreze and Greenberg 1980) and further explored in (Aziz, Brandt, and Harrenstein 2011; Aziz, Brandt, and Seedig 2013; Banerjee, Konishi, and Sönmez 2001; Bogomolnaia and Jackson 2002; Elkind and Wooldridge 2009; Elkind, Fanelli, and Flammini 2016; Gairing and Savani 2010). A HG consists of a set of selfish agents (humans, robots, software agents, etc.) having preferences over coalitions that might include them, regardless of which other coalitions may or may not be present. The outcome is a partition of the agent set into disjoint coalitions (or clusters), referred to as a *clustering* or *coalition structure*.

Stability is the main criterion that has been used to analyze which coalition structures will arise: an outcome should be resistant to individual/group deviations, with different types of deviations giving rise to different notions of stability (such as core stability, individual stability, Nash stability). See (Aziz and Savani 2016) for a survey from a more computational point of view.

A clustering is *Nash stable*, or it is a Nash equilibrium, if no agent can improve her utility by unilaterally changing her own cluster. In this setting, one of the main tools for evaluating the degradation of the system performance induced by the selfish behavior of its agents is the price of anarchy (PoA) (Koutsoupias and Papadimitriou 1999; Papadimitriou 2001), defined as the worse-case ratio between the social welfare (i.e., the sum of the agents utilities) of a best clustering and the social welfare of a Nash stable clustering. A related optimistic measure for evaluating the welfare of the best possible equilibrium is the price of stability (PoS) (Anshelevich et al. 2004), that is the best-case ratio between the social welfare of a best clustering and the social welfare of a Nash stable clustering. A growing interest in HGs has concerned subclasses in which agents entertain preferences over the other agents, which are then naturally lifted to preferences over coalitions, additively-separable HGs being a noticeable example (Olsen 2009).

In this paper we focus on *Social Distance Games* (SDGs), an important subclass of HGs introduced in (Brânzei and Larson 2011) where agent utilities are based on the concept of social distance (i.e., the number of hops required to reach one node from another), which has become famous since Milgram's study on six degrees of separation. In SDGs the utility of an agent is given by the average inverse distance from all the other nodes in her coalition, that is by her harmonic centrality (Boldi and Vigna 2014) divided by the size of the coalition. The basic idea is that the agents prefer to maintain ties with other agents who are close to them. The utility formulation is a variant of the closeness centrality and reflects the principle of homophily, that similarity breeds connection and people tend to form communities with similar others (McPherson, Lovin, and Cook 2001). Moreover, it is related to several other classical measures from network analysis, such as degree, closeness, betweenness, and eigenvector centrality (Gomez et al. 2003), all of which are used to determine how a node is embedded in the network.

Related Work In the literature, an extensive research considered HGs from a strategic cooperative point of view (Bogomolnaia and Jackson 2002; Banerjee, Konishi, and

Sönmez 2001; Elkind and Wooldridge 2009; Gairing and Savani 2010) with the aim of characterizing the existence and the properties of coalition structures such as, for instance, the core. A clustering is core stable if there is no group of agents who can all be better off by forming a new deviating coalition. Furthermore, examples of non-cooperative studies on HGs in which self-organized clusterings are obtained from the decisions taken by independent and selfish agents can be found in (Bloch and Diamantoudi 2011; Feldman, Lewin-Eytan, and Naor 2015; Gairing and Savani 2010). Namely, (Banerjee, Konishi, and Sönmez 2001) study properties guaranteeing the existence of core allocations for HGs games with additively separable utility, while (Bogomolnaia and Jackson 2002) consider several forms of clustering stability like the core and the Nash. (Ballester 2004; Aziz, Brandt, and Seedig 2011; Olsen 2009) deal with computational complexity issues related to HGs, also considering additively separable utilities. In particular, (Olsen 2009) shows that the problem of deciding whether a Nash stable clustering exists in an additively separable HG is NP-complete, as well as the one of deciding whether a Nash stable clustering different from the grand coalition exists in an additively separable HG with non-negative and symmetric preferences. (Bloch and Diamantoudi 2011) study non-cooperative games of coalition formation and identify conditions for stable outcomes. In a similar way, (Apt and Witzel 2009) study how certain proposed rules can transform clusterings into other ones with specific properties. (Feldman, Lewin-Eytan, and Naor 2015) investigate some interesting subclasses of HGs from a non-cooperative point of view, by characterizing Nash equilibria and providing upper and lower bounds on both the price of stability and the price of anarchy.

Fractional Hedonic Games (FHGs) have been traditionally investigated under the additively-separable property. Namely, the utility of an agent is given by the sum of her preferences for each single member of her coalition, i.e. by her degree centrality, divided by the size of the coalition. In a sense, SDGs are FHGs in which in the definition of agents utilities the degree centrality measure is substituted by the harmonic centrality. FHGs have been investigated in (Aziz, Brandt, and Harrenstein 2014; Aziz et al. 2015; Brandl, Brandt, and Strobel 2015) from a cooperative perspective and in (Bilò et al. 2014; 2015; Olsen 2012) from a non-cooperative viewpoint.

To the best of our knowledge *Social Distance Games* have been considered only in the cooperative context of core stability in (Brânzei and Larson 2011). In particular, the authors show that finding the best clustering is NP-hard and provide an algorithm to approximate the optimal welfare within a factor of two. They adopt core stable solutions as the target solution concept to analyze its welfare and stability properties. Finally they show that core stable structures have small world characteristics.

Studying strategic solutions under a non-cooperative scenario such as Nash equilibria becomes of fundamental importance when considering huge environments characterized by decentralization, autonomy, and general lack of coordination among the entities or where the cost of coordination is tremendously high.

Our Contribution In this paper we approach SDGs from the viewpoint of non-cooperative game theory with the aim of understanding the existence, computability and performances of Nash stable clusterings. We first focus on the existence of Nash stable clusterings: even if SDGs always admit a Nash equilibrium, we first show that they may not converge to Nash equilibria and then give a polynomial reduction from the NP-Complete RESTRICTED EXACT COVER *by 3*-SETS (RXC3) problem (Gonzalez 1985) to prove that it is NP-hard to find a best Nash equilibrium. We then study the performances of the Nash equilibria and provide matching upper and lower bounds on the price of anarchy of $\Theta(n)$, where n is the number of nodes of the underlying graph, and a lower bound on the price of stability of $6/5-\epsilon$. Finally, we characterize the price of stability of SDGs for graphs with girth 4 and girth at least 5, the *girth* being the length of the shortest cycle in the graph.

Due to space limitations, sometimes only proof sketches are provided and some details are omitted.

2 Model and Preliminaries

Consider an undirected graph $G = (V, E)$. A *coalition* or *cluster* is a non-empty subset of V. The set of all nodes V is called the *grand coalition*, whereas a coalition of size 1 is said a *singleton coalition* and its node *singleton* or *isolated*. For any integer $n > 0$ denote with $[n]$ the set of integers $\{1, \ldots, n\}$ and with $[k, n]$, for $k \leq n$ the subset $\{k, \ldots, n\} \subseteq [n]$. A *clustering* or *coalition structure* is a partition of V into $k > 0$ coalitions $\mathcal{C} = \{C_1, \ldots, C_k\}$ such that $C_i \subseteq V$ for each $i \in [k]$, $\bigcup_{i\in[k]} C_i = V$ and $C_i \cap C_j = \emptyset$ for any $i, j \in [k]$ with $i \neq j$. Given a coalition C, we denote by $G(C)$ the subgraph induced by C. For brevity, we will often identify $G(C)$ directly with the corresponding coalition C.

Given a set of n selfish agents and an undirected graph $G = (V, E)$ with n nodes, we consider coalition forming games in which each node $x \in V$ is associated with an agent. In the following, for the sake of simplicity, we will often identify an agent with its node $x \in V$.

Let the *harmonic centrality* of a node $x \in V$ be defined as $\mu_x(G) = \sum_{y\in V\setminus\{x\}} \frac{1}{d_G(x,y)}$, where for any pair of nodes $x, y \in V$, $d_G(x, y)$ denotes the length of a shortest path connecting x and y in G. The sum of the inverse of the social distances can be viewed as the similarity of an agent with the other agents of the coalition, and it indicates the centrality of the agent in that coalition.

Definition 2.1 *(Brânzei and Larson 2011) A* Social Distance Game *SDG(G), or SDG in short, is represented as an undirected graph $G = (V, E)$ where (i) V is the set of n agents and (ii) the* utility *of an agent $x \in V$ in a given coalition C is a suitable function of her harmonic-centrality in the subgraph induced by C, that is:*

$$u_x(C) = \frac{\mu_x(G(C))}{|C|} = \frac{1}{|C|} \sum_{y\in C\setminus\{x\}} \frac{1}{d_C(x, y)}.$$

If x and y are disconnected in C, then $d_C(x, y) = \infty$.

We denote by $\mathcal{C}(x)$ the coalition of $\mathcal{C}$ including node x. Let us use $u_x(\mathcal{C})$ as a shorthand for the utility $u_x(\mathcal{C}(x))$ of x in a given solution $\mathcal{C}$. Each agent chooses the coalition to belong to with the aim of maximizing her utility. A *solution* or *outcome* of a SDG is a clustering $\mathcal{C}$. The *social welfare* $SW(\mathcal{C})$ of a clustering $\mathcal{C}$ is the sum of the agents' utilities, i.e., $SW(\mathcal{C}) = \sum_{x \in V} u_x(\mathcal{C})$.

For any pair of agents x and y, we denote with $(\mathcal{C}, x, y)$ the clustering obtained from $\mathcal{C}$ by moving x from $\mathcal{C}(x)$ to $\mathcal{C}(y)$. A clustering $\mathcal{C}' = (\mathcal{C}, x, y)$ is an improving deviation for agent x in $\mathcal{C}$ if $u_x(\mathcal{C}) < u_x(\mathcal{C}')$. In this case we say that agent x makes an improving move. An improving dynamics is a sequence of improving deviations. We assume that the input graph is connected, since disconnected graphs can be analyzed componentwise.

We say that an agent x is *Nash stable* in a clustering $\mathcal{C}$ if she cannot perform any improving deviation, that is if $\forall y \in V$, $u_x(\mathcal{C}) \geq u_x(\mathcal{C}')$, where $\mathcal{C}' = (\mathcal{C}, x, y)$. A clustering $\mathcal{C}$ is *Nash stable*, or is a *Nash equilibrium*, if every agent i is Nash stable in $\mathcal{C}$.

We are interested in bounding the performances of Nash equilibria with respect to the social optimum OPT, i.e. the social welfare of a best clustering $\mathcal{C}^*$ that maximizes $SW(\mathcal{C})$, that is such that $\text{OPT} = SW(\mathcal{C}^*) = \max_{\mathcal{C}} SW(\mathcal{C})$. Notice that $\mathcal{C}^*$ is not necessarily an equilibrium. If $\mathcal{N}$ denotes the set of the Nash equilibria, a best (worst) Nash equilibrium is a Nash equilibrium $\mathcal{C} \in \mathcal{N}$ that maximizes (minimizes) $SW(\mathcal{C})$. The best (worst) social cost is the social welfare of a best (worst) Nash equilibrium. The *price of anarchy* (PoA) is the ratio between the social optimum and the worst social cost in a game i.e., $\text{PoA} = \max_{\mathcal{C} \in \mathcal{N}} \frac{\text{OPT}}{SW(\mathcal{C})}$; the *price of stability* (PoS) is the ratio between the social optimum and the best social cost, i.e., $\text{PoS} = \min_{\mathcal{C} \in \mathcal{N}} \frac{\text{OPT}}{SW(\mathcal{C})}$.

3 Nash equilibria: existence, convergence and complexity of their finding

It is easy to see that a SDG always admits a Nash equilibrium. In fact, the grand coalition is Nash stable as no agent can have any improving deviation. Nevertheless, the following negative result on the convergence holds.

Theorem 1 *SDGs may not converge to Nash equilibria.*

Proof. We exhibit an instance where there is an infinite sequence of improving deviations. Consider the instance SDG(G) where G is the bipartite graph depicted in Figure 1. Let $X = \{x_i | i \in [24]\}$ and $Y = \{y_i | i \in [18]\}$. In the following, for $i \leq j$, we write $X_{i,j}$ and $Y_{i,j}$ as a shorthand for $\{x_i, \dots, x_j\} \subseteq X$ and $\{y_i, \dots, y_j\} \subseteq Y$, respectively.

Let $\mathcal{C} = \{X_{1,12} \cup \{z_1\}, Y, X_{13,24} \cup \{z_2\}\}$. Starting from the clustering $\mathcal{C}$, we prove the existence of a cycle, that is a sequence of improving deviations such that $\mathcal{C}$ is again reached.

We consider the following improving dynamics. Let $\mathcal{C}^{(1)} = (\mathcal{C}, z_1, y_1) = \{X_{1,12}, Y \cup \{z_1\}, X_{13,24} \cup \{z_2\}\}$ and $\mathcal{C}^{(2)} = (\mathcal{C}^{(1)}, z_2, y_1) = \{X_{1,12}, Y \cup \{z_1, z_2\}, X_{13,24}\}$. It

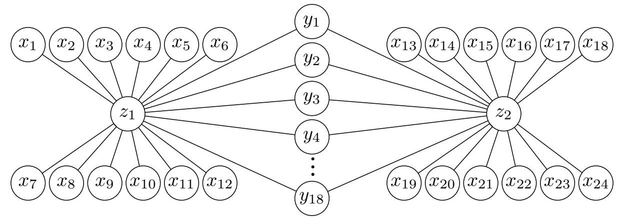

Figure 1: A SDG not converging to a Nash equilibrium.

is easy to see that $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ are two improving deviations for z_1 and z_2, respectively. In fact, since $u_{z_1}(X_{1,12} \cup \{z_1\}) = u_{z_2}(X_{13,24} \cup \{z_2\}) = \frac{12}{13}$, agent z_1 can indeed improve her utility by moving to Y and then z_2 can move to $Y \cup \{z_1\}$, achieving a utility of $u_{z_1}(\mathcal{C}^{(1)}) = u_{z_1}(Y \cup \{z_1\}) = \frac{18}{19} > \frac{12}{13}$ and $u_{z_2}(\mathcal{C}^{(2)}) = u_{z_2}(Y \cup \{z_1, z_2\}) = \frac{18 + \frac{1}{2}}{20} = \frac{37}{40} > \frac{12}{13}$.

Since $\forall x \in X_{1,12}$, $\forall x' \in X_{13,24}$, $u_x(X_{1,12}) = u_{x'}(X_{13,24}) = 0$, agents $x_1, \dots, x_6$ and $x_{13}, \dots, x_{18}$ can increase their utility by moving one after another to $Y \cup \{z_1, z_2\}$. Thus starting from $\mathcal{C}^{(2)}$, there are 12 improving moves inducing the clustering $\mathcal{C}^{(14)} = \{X_{7,12}, Y \cup \{z_1, z_2\} \cup X_{1,6} \cup X_{13,18}, X_{19,24}\}$.

Notice that $u_{z_1}(\mathcal{C}^{(14)}) = \frac{24 + \frac{1}{2} + \frac{6}{3}}{32} = \frac{53}{64}$, and thus agent z_1 can increase her utility by moving back to $X_{7,12}$, achieving a utility of $\frac{6}{7} > \frac{53}{64}$. Thus $\mathcal{C}^{(15)} = (\mathcal{C}^{(14)}, z_1, x_7) = \{X_{7,12} \cup \{z_1\}, Y \cup \{z_2\} \cup X_{1,6} \cup X_{13,18}, X_{19,24}\}$.

Since $u_{z_2}(\mathcal{C}^{(15)}) = \frac{24}{31}$, agent z_2 induces an improving deviation by moving back to $X_{19,24}$, that is $\mathcal{C}^{(16)} = (\mathcal{C}^{(15)}, z_2, x_{19}) = \{X_{7,12} \cup \{z_1\}, Y \cup X_{1,6} \cup X_{13,18}, X_{19,24} \cup \{z_2\}\}$.

Agents $x_1, \dots, x_6$ and $x_{13}, \dots, x_{18}$ have now utility zero and thus they can increase the utility by moving back one after another to the cluster containing z_1 and z_2, respectively. Then $\mathcal{C}$ is again reached and therefore the claim follows. $\square$

While determining the existence of a Nash equilibrium can be trivially done in polynomial time, as the grand coalition is stable, in the sequel of this section we show that unfortunately computing a best Nash equilibrium for SDGs is NP-hard. To this aim, we provide a polynomial time reduction from the NP-Complete RESTRICTED EXACT COVER *by 3-*SETS *(RXC3)* problem (Gonzalez 1985). An instance of RXC3 consists of a universe set U of $3p$ elements and a collection $\mathcal{B} = \{B_1, \dots, B_m\}$ of 3-elements subsets of U (triples) such that $\bigcup_{j=1}^{m} B_j = U$ and each element of U appears exactly in three triples. The problem is to decide if there exists an exact cover of U, that is a subcollection $\mathcal{F} \subseteq \mathcal{B}$ such that every element in U appears in exactly one triple of $\mathcal{F}$.

We reduce $RXC3$ to SDG. In particular, given a generic instance $(U, \mathcal{B})$ of $RXC3$, we build an instance of SDG by specifying the underlying undirected graph $G = (V, E)$ as follows:

- for each triple $B_i \in \mathcal{B}$, for $i \in [m]$, we associate a set of

5 nodes $X_i = \{a_i, b_i, c_i, d_i, e_i\}$ such that every two distinct nodes in X_j are adjacent, except a_j and e_j. Roughly speaking, G contains m cliques $K_5(X_j)$ without edge (a_j, e_j) for every node set X_j.

- for each element $u_j \in U$, for $j \in [3p]$, we consider a node y_j and a set of 3 edges $E_j = \{(y_j, e_i) | u_j \in B_i\}$.

Therefore, $|V| = 3p + 5m$ and $E = 9(p + m)$. Clearly such a reduction can be done in polynomial time.

In order to prove the hardness, we now show that the constructed SDG has a Nash stable clustering of social welfare at least $\frac{21}{4}p + \frac{19}{5}(m - p)$ if and only if the RXC3 instance has an exact cover. Let us first give some useful lemmas.

Lemma 1 *Let $\mathcal{C}$ be a Nash stable clustering for SDG. If agents e_i and e_j belong to the same coalition $C = \mathcal{C}(e_i) = \mathcal{C}(e_j)$, then $X_i \cup X_j \subseteq C$, that is agents in X_i and agents in X_j must belong to the same coalition as well.*

Proof. (Sketch) A case analysis shows that, in order to guarantee stability, agents a_i, b_i, c_i, d_i must be in the same coalition. Moreover, it is possible to check that if e_i and e_j are in the same coalition and a_i, b_i, c_i, d_i belong to another coalition, then e_i can improve her utility by connecting to a_i, b_i, c_i, d_i. A symmetric argument applies to e_j, thus proving the claim. □

Lemma 2 *Let $\mathcal{C}$ be a Nash stable clustering for SDG. If agents e_i and e_j belong to the same coalition, then the utility of all the agents in the coalition is strictly less than* 0.635.

Proof. Let $C = \mathcal{C}(e_i) = \mathcal{C}(e_j)$. By Lemma 1 we know that $X_i \cup X_j \subseteq C$. Notice that if there is not a path between e_i and e_j, we can repartition the coalition increasing the social welfare. Thus we can give an upper bound to the utility of all agents as follows (see Figure 2).

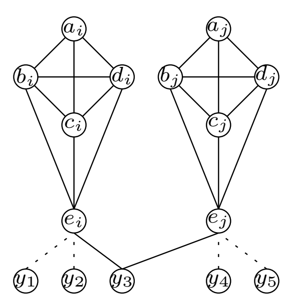

Figure 2: Nash stable solution where e_i and e_j belong to the same coalition.

- Each a_i has at least 3 agents at distance 1 (b_i, c_i, d_i), one agent at distance 2 (e_i), one agent at distance 4 (e_j), 3 agents at distance 5 (b_j, c_j, d_j) and one agent at distance 6 (a_j). All the other x nodes (including $y_1, \ldots, y_5$) have distance at least 3, thus the utility of each a_i is at most $\frac{3+\frac{1}{2}+\frac{x}{3}+\frac{1}{4}+\frac{3}{5}+\frac{1}{6}}{10+x} \leq 0.44$.
- Each b_i (and symmetrically c_i and d_i) has at least 4 agents at distance 1 (a_i, c_i, d_i, e_i), one agent at distance 3 (e_j), 3 agents at distance 4 (b_j, c_j, d_j) and one agent at distance 5 (a_j). All the other x nodes (including $y_1, \ldots, y_5$) have distance at least 2, thus, the utility of each b_i is at most $\frac{4+\frac{x}{2}+\frac{1}{3}+\frac{3}{4}+\frac{1}{5}}{10+x} \leq \frac{317}{600} \approx 0.53$.
- Each e_i can have at most 6 agents at distance 1 $(b_i, c_i, d_i, y_{i_1}, y_{i_2}, y_{i_3})$, 2 agents at distance 2 $(a_i,\ e_j)$, 3 agents at distance 3 (b_j, c_j, d_j) and one agent at distance 4 (a_j). All the other x nodes (including y_4, y_5) have distance at least 2, so that the utility of each e_i is at most $\frac{6+\frac{x+2}{2}+\frac{3}{3}+\frac{1}{4}}{13+x} \leq \frac{33}{52} \approx 0.634$.
- Each y_t has at most 3 agents at distance 1 (e_i, e_j, e_k), 6 agents at distance 2 $(b_i, c_i, d_i, b_j, c_j, d_j)$, 2 agents at distance 3 (a_i, a_j). This gives an upper bound of the utility of each y_t agent of $\frac{3+\frac{6+x}{2}+\frac{2}{3}}{12+x} \leq \frac{5}{9} \approx 0.55$

□

Lemma 3 *Let $\mathcal{C}$ be a Nash stable clustering for SDG(G). If agents e_i and e_j belong to the same coalition C, then C can be split in subcoalitions obtaining a new stable claustering $\mathcal{C}'$ with a strictly higher social welfare.*

Proof. We can divide C in many subcoalitions of the forms depicted in Figure 3, where each subcoalition contains a single node of the type e_k and the associated nodes a_k, b_k, c_k, d_k.

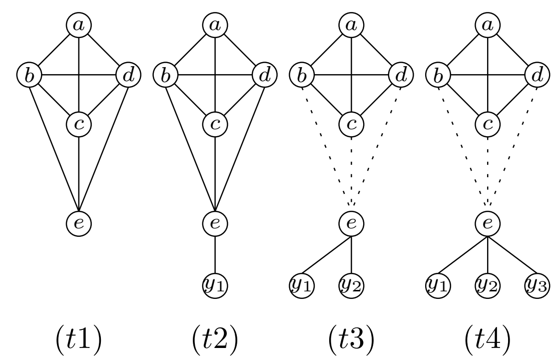

Figure 3: Nash stable solutions.

The social welfare and the average utility of the coalition members in every case are:

- $SW(t1) = \frac{3*4+2(3+1/2)}{5} = \frac{19}{5}$. $\bar{u}(t1) = \frac{19/5}{5} = \frac{19}{25} \approx 0.76$
- $SW(t2) = \frac{37}{9}$. $\bar{u}(t2) = \frac{37/9}{6} = \frac{37}{54} \approx 0.68$
- $SW(t3) = \frac{94}{21}$. $\bar{u}(t3) = \frac{94/21}{7} = \frac{94}{147} \approx 0.639$
- $SW(t4) = \frac{21}{4}$. $\bar{u}(t4) = \frac{21/4}{8} = \frac{21}{32} \approx 0.656$

Since the worst average utility 0.639 is greater than the best utility in the case where e_i and e_j are in the same coalition (0.635), we have that the social welfare strictly increases. Moreover, a case analysis shows that the obtained clustering is Nash stable and that no other Nash stable solution exists achieving a strictly higher social welfare. □

Lemma 4 *If there is an exact cover for the input instance of $RXC3$, then there exists a Nash equilibrium in the reduced instance of the SDG game with social welfare at least $\frac{21}{4}p + \frac{19}{5}(m-p)$.*

Proof. If there is an exact cover then there is a clustering composed by exactly p copies of $t1$ and $(m-p)$ copies of $t4$. Again, the stability can be shown by a case analysis, and the claim follows simply by summing up the utilities of all the agents. $\square$

Lemma 5 *If there is not an exact cover for the input instance of $RXC3$, then every Nash equilibrium in the reduced instance of the SDG has social welfare strictly less than $\frac{21}{4}p + \frac{19}{5}(m-p)$.*

Proof. Assume that there is not an exact cover. By the previous lemmas, the stable clusterings achieving the best social welfare are composed only by coalitions of the 4 types depicted in Figure 3. Consider any such a clustering $\mathcal{C}$. For each $i \in [m]$, we rearrange the social welfare of each coalition of $\mathcal{C}$ in such a way that $\frac{19}{5}$ is always accounted to the nodes a_i, b_i, c_i, d_i, e_i and the surplus is equally shared among the y agents. We have three cases for charging y agents, depending on the number of the them contained in each cluster: in the case $t2$ there is a single y agent and we account all the remaining social welfare to it, that is $\frac{14}{45} \approx 0.31$; in the case $t3$ there are two y agents, and we account $\frac{71}{210} \approx 0.338$ to each of them; in the case $t4$ there are three y agents, and we account to each $\frac{29}{60} \approx 0.483$. Notice that, since $\frac{29}{60} > \frac{71}{210} > \frac{14}{45}$, a clustering with social welfare equal to $3p \cdot \frac{29}{60} + \frac{19}{5}m = \frac{21}{4}p + \frac{19}{5}(m-p)$ can be obtained if and only if $\frac{29}{60}$ is accounted to all the y nodes, which would imply the existence of an exact cover: a contradiction. $\square$

We are now ready to claim the following theorem, whose proof comes directly from the previous 2 lemmas.

Theorem 2 *Computing a best Nash equilibrium for $SDGs$ is NP-hard.*

4 Price of Anarchy and Price of Stability

In this section we first provide matching upper and lower bounds on the price of anarchy of SDGs. Then we focus on the price of stability and prove that for general instances the lower bound of the PoS is $6/5 - \epsilon$. Finally we study the PoS for graphs with girth 4 and girth at least 5, the *girth* being the length of the shortest cycle in the graph. A *star* is a tree consisting of one vertex (the center) adjacent to all the other vertices (the leaves).

Theorem 3 *The price of anarchy of Social Distance Games is $\Theta(n)$.*

Proof. The definition of the game directly implies that the social welfare of any clustering is upper bounded by $n-1$ (and therefore $\text{OPT} \le n-1$). Such an upper bound can be obtained only by the grand coalition on complete graphs. On the other hand, since in any equilibrium every agent has at least one neighbor in her coalition, the utility of each node is at least $\frac{1}{n}$, and thus $SW(\mathcal{C}) \ge 1$.

It remains to show that there exists an SDG having price of anarchy $\Omega(n)$. To this end, consider the graph depicted in Figure 4. In the Nash stable solution $\mathcal{C}$ illustrated in Figure 5, $SW(\mathcal{C}) = \frac{n}{4} \cdot \frac{2(1+\frac{1}{2}+\frac{1}{3})+2(2+\frac{1}{2})}{4} = \frac{13n}{24}$. Another Nash stable solution $\mathcal{C}'$ is shown in Figure 6. In this case, the social welfare is $SW(\mathcal{C}') = \frac{(2(1+\frac{1}{2}+\frac{1}{3})+2(2+\frac{1}{2}))\cdot\frac{\frac{n}{2}}{4}}{n/2} = \frac{13}{3}$ and therefore $\text{PoA} = \frac{\frac{13n}{24}}{\frac{13}{3}} = \frac{n}{8} = \Omega(n)$. Thus, the claim follows. $\square$

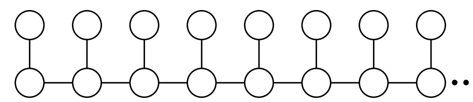

Figure 4: A SDG with n agents having $\text{PoA} = \Omega(n)$.

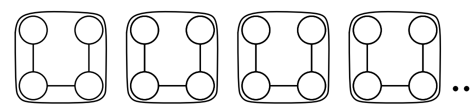

Figure 5: A Nash stable solution with $SW = \frac{13n}{24}$ for the SDG depicted in Figure 4.

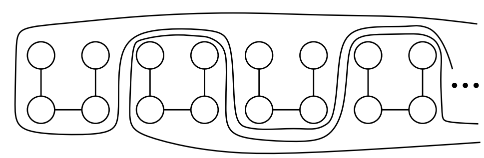

Figure 6: A Nash stable solution with $SW = \frac{13}{3}$ for the SDG depicted in Figure 4.

Theorem 4 *The PoS of SDGs is at least $6/5 - \epsilon$, for every $\epsilon \ge 0$.*

Proof. For any positive integer t, consider a graph of $2t+2$ nodes, composed by a clique of $t+1$ agents $K_{t+1} = \{c_i | i \in [t+1]\}$, each connected to the center x of a star of t leaves. Let $S = \{x\} \cup \{s_i | i \in [t]\}$ the set of the agents in the star. In Figure 7 is depicted the graph for the case $t = 3$.

Consider the clustering $\mathcal{C} = \{K_{t+1}, S\}$ (see Figure 7b), that gives a social welfare $SW(\mathcal{C}) = \frac{t(3t+5)}{2(t+1)}$. $\mathcal{C}$ is not Nash stable, since agent x can increase her utility from $\frac{t}{t+1}$ to $\frac{t+1}{t+2}$ by moving in K_{t+1}.

Due to space limitation, we leave the reader to verify that the grand coalition (Figure 7a) is the best Nash stable solution with $SW(V) = \frac{t(5t+11)+4}{4t+4}$. We briefly point out that the subgraph induced by $K_{t+1} \cup \{x\}$ is a clique K_{t+2} and provides utility at least $\frac{t+1}{t+2}$ to each agent. We notice also that, in any stable solution, each node s_i among $s_1, \ldots, s_t$ must be connected to x, otherwise it would be $u_{s_i}(\mathcal{C}(s_i)) = 0$. Thus, the ratio $\frac{SW(\mathcal{C})}{SW(V)} = \frac{2t(3t+5)}{t(5t+11)+4}$ tends to $6/5$ as t tends to infinity, proving the theorem. $\square$

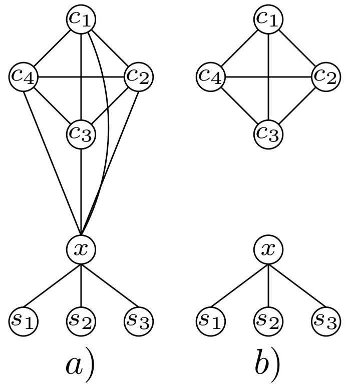

Figure 7: A SDG (a) and a corresponding clustering which is not Nash stable (b).

Theorem 5 *The PoS of SDGs in which the underlying graph has girth* 4 *is at least* $169/160 = 1.05625$.

Proof. (Sketch) Consider the graph in Figure 8. The best solution is given by the following clustering: $\{\{a_1, \ldots, a_5, b_1, \ldots, b_5\}, \{t, c_1, c_2\}\}$ that achieves a SW of $\frac{26}{3}$. This is not a Nash stable solution, since t can deviate and increase its utility from $\frac{2}{3}$ to $\frac{15}{22}$. The best stable solution is the grand coalition that achieves a SW of $\frac{320}{39}$. The result derives by the ratio of such SWs. □

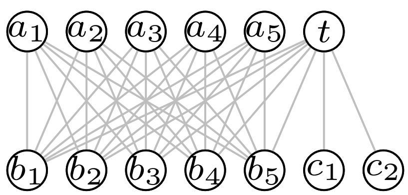

Figure 8: A SDG with PoS $= \frac{169}{160}$.

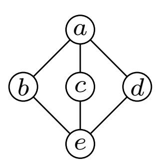

Figure 9: A graph with girth 4.

The following two lemmas show that the girth of the underlying graph is a crucial parameter for determining the structure of equilibria.

Lemma 6 *There exists an instance of SDGs in which the underlying graph has girth* 4 *and the only stable solution is the grand coalition.*

Proof. Consider the graph in Figure 9 and let us analyze all possible ways for partitioning the nodes. First of all, notice that an agent cannot be isolated and that, if a cluster is unconnected, an agent will deviate because in that case her utility would be zero. So, besides the grand coalition, we can partition V in two connected clusters of size 2 and 3. We only look at the case $A = \{\{a, b\}, \{c, d, e\}\}$ since all other cases are symmetric. A is not stable since a can increase its utility from $\frac{1}{2}$ to $\frac{5}{8}$. This implies the theorem. □

For a greater girth the following result holds.

Lemma 7 *Given any graph with girth* > 4*, we can obtain a stable solution with a* $SW \geq \frac{n}{2}$ *by partitioning the graph into stars.*

Proof. A simple procedure allows partitioning the graph into stars, each of at least 2 nodes. It is easy to check that in a star all the leaves have utility $\frac{1}{2}$, while the center has utility $\geq \frac{1}{2}$. For what concerns the stability, consider an agent in a star. Since the girth is > 4, she can have only one connection to an agent of a different star, hence she can not achieve an utility $> \frac{1}{2}$ by deviating. □

Theorem 6 *The upper bound of the PoS of SDGs in which the underlying graph has girth* > 4 *is* $\frac{1}{2} + \frac{1}{\sqrt{2}} \approx 1.2$.

Proof. Let δ_i be the degree of node i. We can give an upper bound to the utility of i by considering its neighbors at distance 1 and all the other nodes as if they were at distance 2. Hence, the social welfare would be: $SW \leq \sum_{i\in V} \frac{\delta_i + \frac{n-\delta_i-1}{2}}{n} = \sum_{i\in V} \frac{\frac{n-1}{2} + \frac{\delta_i}{2}}{n} = \frac{n-1}{2} + \sum_{i\in V} \frac{\delta_i}{2n} = \frac{n-1}{2} + \frac{2|E|}{2n} = \frac{n-1}{2} + \frac{|E|}{n}$.

From (Dutton and Brigham 1991), we know that in a graph with girth > 4, $|E| \leq \frac{n\sqrt{n-1}}{2}$. Thus,

$$SW \leq \frac{n-1}{2} + \frac{\frac{n\sqrt{n-1}}{2}}{n} = \frac{n-1}{2} + \frac{\sqrt{n-1}}{2}.$$

From Lemma 7, we can guarantee a SWof at least $\frac{n}{2}$, hence

$$PoS \leq \frac{\frac{n-1}{2} + \frac{\sqrt{n-1}}{2}}{\frac{n}{2}} = \frac{n-1+\sqrt{n-1}}{n} = 1 + \frac{\sqrt{n-1}}{n} - \frac{1}{n}.$$

If we maximize the above formula, we obtain $\frac{1}{2} + \frac{1}{\sqrt{2}}$, which proves the theorem. □

5 Conclusions

We investigated Nash stability in SDGs. Several issues remain open. First of all, for general graphs, while a 6/5 lower bound on the price of stability holds, it would be nice to provide a corresponding upper bound. Another relevant question that naturally emerges is whether there exists a polynomial time algorithm for determining the existence of a Nash stable clustering for SDGs different from the grand coalition. As we have seen, this is not guaranteed for a girth less or equal to 4. We notice that, besides this computational issue, a deeper understanding of the conditions for the existence of such non trivial equilibria would be particularly important for providing better bounds on the price of stability. A related open question is that of identifying special graphs in which a best equilibrium or a best non stable clustering can be computed in polynomial time. Finally, it would be interesting to generalize our results to weighted graphs and to consider classes of hedonic and fractional hedonic games induced by other classical centrality measures, like the ones presented in (Gomez et al. 2003). On this respect, it would be particularly worth to consider models in which being a singleton is not the worst choice.

References

Anshelevich, E.; Dasgupta, A.; Kleinberg, J.; Tardos, E.; Wexler, T.; and Roughgarden, T. 2004. The price of stability for network design with fair cost allocation. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '04, 295–304. Washington, DC, USA: IEEE Computer Society.

Apt, K. R., and Witzel, A. 2009. A generic approach to coalition formation. *International Game Theory Review* 11(03):347–367.

Aziz, H., and Savani, R. 2016. *Hedonic Games*. Cambridge University Press. chapter 15.

Aziz, H.; Gaspers, S.; Gudmundsson, J.; Mestre, J.; and Täubig, H. 2015. Welfare maximization in fractional hedonic games. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, 461–467.

Aziz, H.; Brandt, F.; and Harrenstein, P. 2011. Pareto optimality in coalition formation. In *Proceedings of the 4th International Conference on Algorithmic Game Theory*, SAGT, 93–104. Springer-Verlag.

Aziz, H.; Brandt, F.; and Harrenstein, P. 2014. Fractional hedonic games. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, 5–12. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Aziz, H.; Brandt, F.; and Seedig, H. G. 2011. Stable partitions in additively separable hedonic games. In *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '11, 183–190. International Foundation for Autonomous Agents and Multiagent Systems.

Aziz, H.; Brandt, F.; and Seedig, H. G. 2013. Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence* 195:316–334.

Ballester, C. 2004. Np-completeness in hedonic games. *Games and Economic Behavior* 49(1):1–30.

Banerjee, S.; Konishi, H.; and Sönmez, T. 2001. Core in a simple coalition formation game. *Social Choice and Welfare* 18(1):135–153.

Bilò, V.; Fanelli, A.; Flammini, M.; Monaco, G.; and Moscardelli, L. 2014. Nash stability in fractional hedonic games. In *Web and Internet Economics - 10th International Conference, WINE 2014, Beijing, China, December 14-17, 2014. Proceedings*, 486–491.

Bilò, V.; Fanelli, A.; Flammini, M.; Monaco, G.; and Moscardelli, L. 2015. On the price of stability of fractional hedonic games. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May 4-8, 2015*, 1239–1247.

Bloch, F., and Diamantoudi, E. 2011. Noncooperative formation of coalitions in hedonic games. *International Journal of Game Theory* 40(2):263–280.

Bogomolnaia, A., and Jackson, M. O. 2002. The stability of hedonic coalition structures. *Games and Economic Behavior* 38(2):201–230.

Boldi, P., and Vigna, S. 2014. Axioms for centrality. *Internet Mathematics* 10(3-4):222–262.

Brandl, F.; Brandt, F.; and Strobel, M. 2015. Fractional hedonic games: Individual and group stability. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, 1219–1227. International Foundation for Autonomous Agents and Multiagent Systems.

Brânzei, S., and Larson, K. 2011. Social distance games. In *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, 91–96.

Dreze, J., and Greenberg, J. 1980. Hedonic coalitions: Optimality and stability. *Econometrica* 48(4):987–1003.

Dutton, R. D., and Brigham, R. C. 1991. Edges in graphs with large girth. *Graphs and Combinatorics* 7(4):315–321.

Elkind, E., and Wooldridge, M. 2009. Hedonic coalition nets. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '09, 417–424. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Elkind, E.; Fanelli, A.; and Flammini, M. 2016. Price of pareto optimality in hedonic games. In *AAAI*, 475–481. AAAI Press.

Feldman, M.; Lewin-Eytan, L.; and Naor, J. S. 2015. Hedonic clustering games. *ACM Trans. Parallel Comput.* 2(1):4:1–4:48.

Gairing, M., and Savani, R. 2010. Computing stable outcomes in hedonic games. In *Algorithmic Game Theory - Third International Symposium, SAGT 2010, Athens, Greece, October 18-20, 2010. Proceedings*, 174–185.

Gomez, D.; Gonzlez-Arangena, E.; Manuel, C.; Owen, G.; del Pozo, M.; and Tejada, J. 2003. Centrality and power in social networks: a game theoretic approach. *Mathematical Social Sciences* 46(1):27 – 54.

Gonzalez, T. F. 1985. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science* 38:293–306.

Koutsoupias, E., and Papadimitriou, C. 1999. Worst-case equilibria. In *Proceedings of the 16th Annual Symposium On Theoretical Aspects of Computer science*, 404–413.

McPherson, M.; Lovin, L. S.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27(1):415–444.

Olsen, M. 2009. Nash stability in additively separable hedonic games and community structures. *Theory of Computing Systems* 45(4):917–925.

Olsen, M. 2012. On defining and computing communities. In *Proceedings of the Eighteenth Computing: The Australasian Theory Symposium - Volume 128*, CATS '12, 97–102. Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

Papadimitriou, C. H. 2001. Algorithms, games, and the internet. In *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings*, 1–3.