Complexity of Equilibrium in Diffusion Games on Social Networks*

Seyed Rasoul Etesami, Tamer Başar
Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801
Email: (etesami1, basar1)@illinois.edu

Abstract—We revisit the competitive diffusion game on undirected connected graphs and study the complexity of the existence of pure-strategy Nash equilibrium for such games. We first characterize the utility of each player based on the location of its initial seed placements. Using this characterization, we show that the utility of each player is a sub-modular function of its initial seed set. Following this, a simple greedy algorithm provides an initial seed placement within a constant factor of the optimal solution. We show NP-completeness of the decision on the existence of pure-strategy Nash equilibrium for general networks. Finally we provide some necessary conditions for a given profile to be a Nash equilibrium and obtain a lower bound for the maximum social welfare of the game with two players.

Index Terms—Competitive diffusion game, pure Nash equilibrium, sub-modular function, greedy algorithm, NP-hardness, social welfare.

I. INTRODUCTION

The area of social networks is an important one which has attracted a lot of attention in recent years. Modeling the behavior of agents in a network using the possible interactions among them has been addressed in the literature [1]. One of the widely studied models in social networks is the diffusion model, where the goal is to propagate a certain type of product or behavior in a desired way through the network [2], [3], [4] and [5]. However, in many of the applications there is more than one party that want to spread information on his own products. This imposes a sort of competition among the providers who are competing for the same set of resources and their goal is to diffuse information on their own product in a desired way through the society. Such a competition can be modeled within a game theoretic framework and hence, a natural question one can ask is characterization of the set of equilibria of the game. In the literature there are some existing results which address this [6], [2], [7], [8] and [9]. However what enriches such models is their descriptive capabilities and traceable properties.

One of the models that describes such a competitive behavior in networks is known as competitive diffusion game. It was shown earlier [10] that in general such games do not admit pure-strategy Nash equilibria. It has been shown in [11] that such games may not even have a pure-strategy Nash equilibrium on graphs of small diameter. In fact, the authors in [11] have shown that for graphs of diameter 2 and under some additional assumptions on the topology of the network, the diffusion game admits a general potential function and hence an equilibrium. However checking these assumptions at the outset and for graphs of diameter at most 2 does not seem to be realistic.

In this paper, we consider the competitive diffusion game where each player chooses a set of initial seed nodes, and study the best response of each player. This can be viewed as the best seed placement problem which was extensively studied for different processes [12], [13], [14] and [15]. Since in most of the cases the best placement problem is NP-hard, providing an algorithm to approximate the optimal placement as close as possible is desired. In the current work we address this issue and propose a simple greedy algorithm.

The paper is organized as follows: in Section II we review the competitive diffusion game and some of its properties. In Section III we develop some results to characterize the utility of the agents based on the relative locations of the players’ initial seed placements. Using this fact, we show sub-modularity of players’ utilities. In Section IV we show that, in general, making a decision on the existence of Nash equilibrium is an NP-complete problem. In Section V we provide a necessary condition for a given profile to be a Nash equilibrium. Moreover, we give a lower bound on the maximum social welfare when there are two players in the game. We conclude the paper in Section VI.

Notations: For a vector $v \in \mathbb{R}^n$, we let $v_i$ be the $i$th entry of $v$. Similarly, for a matrix $A$, we let $A_{ij}$ be the $i$th entry of $A$ and we denote the $i$th row of $A$ by $A_i$. We denote the transpose of a matrix $A$ by $A^T$. Moreover, we let $I$ and $1$ be, respectively, the identity matrix and the column vector of all ones of proper dimensions. We let $G = (V, E)$ to be an undirected graph with the set of vertices $V$ and the set of edges $E$. Corresponding to $G$ we let $A_G$ be its adjacency matrix, i.e. $A_G(i, j) = 1$ if and only if $(i, j) \in E$ and $A_G(i, j) = 0$, otherwise. Given a graph $G = (V, E)$ and two vertices $x, y \in V$, we define $d_G(x, y)$ to be the length of the shortest graphical path between $x$ and $y$. Also, for a set of vertices $S \subseteq V$ and a vertex $x$, we let $d_G(x, S) = \min_{y \in S} \{d_G(x, y)\}$. For a vector $(s_1, s_2, \ldots, s_n)$ we sometimes write $(s_i, s_{-i})$, where $s_{-i}$ is the set of all entries except the $i$th one. For a real number $a$ we let $\lceil a \rceil$ to be the smallest integer greater than or equal to $a$.

---

*Research supported in part by the "Cognitive & Algorithmic Decision Making" project grant through the College of Engineering of the University of Illinois, and in parts by AFOSR MURI Grant FA 9550-10-1-0573 and NSF grant CCF 11-11342.
II. COMPETITIVE DIFFUSION GAME

In this section we introduce the competitive diffusion game as was introduced earlier in [10] and state some of its properties. For sake of simplicity, we state the model when there are only two players in the game, however, it can be naturally extended to the case when there are more than two players.

A. Game Model

We consider a network \(G\) of \(n\) nodes and two players (types) \(A\) and \(B\). Initially at time \(t = 0\), each player decides to choose a subset of nodes in the network and place his own seeds. After that, a discrete time diffusion process unfolds among uninfected nodes as follows:

- If at some time step \(t\) an uninfected node is neighbor to infected nodes of only one type (\(A\) or \(B\)), it will adopt that type at the next time step \(t + 1\).
- If an uninfected node is connected to nodes of both types at some time step \(t\), it will change to a gray node at the next time step \(t + 1\) and does not adopt any type afterward.

This process continues until no new adoption happens. Finally, the utility of each player will be the total number of infected nodes of its own type at the end of the process. This is in contrast to other types of processes where the state of the nodes does not change above process is a particular case of progressive diffusion streams of diffusion to other uninfected nodes.

Moreover, let \(N\) be the set of gray and uninfected (white) nodes by the end of the process. Then,

\[ N \subseteq \{i : d_G(i, A) = d_G(i, B)\}, \]

and

\[ \{i : d_G(i, A) < d_G(i, B)\} \subseteq N_A \subseteq \{i : d_G(i, A) \leq d_G(i, B)\} \]

\[ \{i : d_G(i, B) < d_G(i, A)\} \subseteq N_B \subseteq \{i : d_G(i, B) \leq d_G(i, A)\}. \]

Proof: The proof is by induction on the time step \(t = 0, 1, \ldots\). At time \(t = 0\) there is no gray node and the result follows. At time \(t = 1\), if there exists some gray nodes, it requires the gray nodes to be a neighbor of both \(i_A\) and \(i_B\), as otherwise, they will adopt \(B\) or \(A\), respectively. Therefore, for all gray nodes \(j\) which are born at time \(t = 1\) we must have \(d_G(i, j) = d_G(i, B) = 1\). Now, assume that the statement of the Lemma is true for all the gray nodes such that they are born at steps \(t \leq k\), and let \(T_k\) be the collection of all these nodes.

Consider a gray node \(\ell\) which is born at time \(t = k + 1\). Note that if there is no such node, there is nothing to show and we can go one step forward. Assume \(P_1\) and \(P_2\) are two different shortest paths from \(i_A\) and \(i_B\) to \(\ell\), respectively. We consider the following cases:

1. \(P_1\) and \(P_2\) do not have any other gray node except \(\ell\).
2. \(P_1\) (or similarly \(P_2\)) has at least one more gray node other than \(\ell\), but \(P_2\) (or similarly \(P_1\)) does not have any other gray node except \(\ell\).
3. \(P_1\) and \(P_2\) both have at least one gray node other than \(\ell\).

In the first case, we note that since \(\ell\) has changed to gray at time \(t = k + 1\) and since the shortest paths \(P_1\) and \(P_2\) have no gray nodes inside, we thus conclude that \(d_G(i, \ell) = d_G(i, B) = k + 1\), as otherwise \(\ell\) will adopt either \(A\) or \(B\). Moreover, after \(k\) time steps, all the internal nodes of \(P_1\) and \(P_2\) will adopt \(A\) and \(B\), respectively.

In the second case, let \(z\) be the first internal gray node in \(P_1\) which is born by running the process from time \(t = 0\) to time \(t = k\). Since \(P_1\) is the shortest path between \(i_A\) and \(\ell\), and also \(z\) is located on this path, we observe that \(P_1(i_A \rightarrow \ell)\) is also the shortest path between \(i_A\) and \(z\). Since by the definition of \(z\) there is no other gray node on \(P_1(i_A \rightarrow z)\), we thus conclude that \(z\) is born at step \(t_0 = |P_1(i_A \rightarrow z)| < |P_1(i_A \rightarrow \ell)|\) and hence \(z \in T_k\). Moreover, since \(P_1\) is the shortest path between the initial seed \(i_A\) and \(\ell\), and since \(\ell\) changes to gray at step \(k + 1\), it means that \(\ell\) does not change to gray for the first \(|P_1(i_A \rightarrow \ell)| - 1\) steps. Thus, \(|P_1(i_A \rightarrow \ell)| \leq k + 1\). Putting these inequalities together, we get \(t_0 \leq k\) and by the induction hypothesis we have \(d_G(i, z) = d_G(i, B, z)\). Now we can write

\[ d_G(i, \ell) = d_G(i, A, z) + d_G(z, \ell) = d_G(i, B, z) + d_G(z, \ell) \geq d_G(i, B, \ell). \]

We claim that equality must hold in (2). Otherwise, if \(d_G(i, \ell) > d_G(i, B, \ell)\), then since \(P_2\) does not have any gray node other than \(\ell\), it requires that after \(d_G(i, B, \ell) - 1\) steps all the nodes in \(P_2\) other than \(\ell\) adopt \(B\). Therefore, at time
reach. In other words, the marginal return given in the second case for the seed nodes \(i_A\) and \(i_B\), we get 
\[ d_G(i_A, \ell) \geq d_G(i_B, \ell) \]
and 
\[ d_G(i_B, \ell) \geq d_G(i_A, \ell), \]
which gives us 
\[ d_G(i_A, \ell) = d_G(i_B, \ell). \]

Finally, in the last case, by repeating the same argument given in the previous case, for the seed nodes \(i_A\) and \(i_B\), we get 
\[ d_G(i_A, \ell) \geq d_G(i_B, \ell), \]
and 
\[ d_G(i_B, \ell) \geq d_G(i_A, \ell), \]
which gives us 
\[ d_G(i_A, \ell) = d_G(i_B, \ell). \]

Now, let us assume that \(\ell'\) is a node which remains infected (white) at the end of the process. It means that there is no path of types either \(A\) or \(B\) which connects the seed nodes \(i_A\) or \(i_B\) to \(\ell'\). Without loss of generality assume that 
\[ d_G(i_A, \ell') < d_G(i_B, \ell') \]
denote one of the shortest paths between \(i_A\) and \(\ell'\) by \(P'_1\). This path must have at least one gray node which we denote its \(z\) by \(z'\). We can write 
\[ d_G(i_A, \ell') = d_G(i_A, z') + d_G(z', \ell') = d_G(i_B, z') + d_G(z', \ell') \geq d_G(i_B, \ell'). \]

This contradiction shows that 
\[ d_G(i_A, \ell') = d_G(i_B, \ell'). \]
Gathering all the above results, we get 
\[ N \subseteq \{i : d_G(i_A, i) = d_G(i_B, i)\}. \]

To prove the complete, we only need to show one of the relations in \([1]\), as the proof for the other one would be the same. Assume that for some \(i^*\) we have 
\[ d_G(i_A, i^*) < d_G(i_B, i^*). \]

Using \([2]\) we know that 
\[ d_G(i_A, z^*) = d_G(i_B, z^*). \]

Thus, 
\[ d_G(i_A, i^*) = d_G(i_A, z^*) + d_G(z^*, i^*) = d_G(i_B, z^*) + d_G(z^*, i^*) \geq d_G(i_B, i^*). \]

This contradiction shows that \(i^*\) will adopt \(A\) and hence 
\[ \{i : d_G(i_A, i) < d_G(i_B, i)\} \subseteq N. \]
This completes the proof. \(\blacksquare\)

Note that in Lemma \([1]\) we assumed that each player can only place one seed in the network as its initial placement. However, the result can be easily generalized to the case when each player (for example player \(A\)) is allowed to choose a set of nodes \(S_A \subseteq V\) as its initial seed placements. In this case we just need to replace 
\[ d_G(S_A, x) \]
instead of 
\[ d_G(i_A, x) \]
in the statement of Lemma \([1]\) and all the other results carry naturally. Moreover, a similar result can be proven when there are more than two players in the game.

**Definition 1:** Given a set \(\Omega\), a set function 
\[ f : 2^\Omega \to \mathbb{R} \]
is called a sub-modular function if for any two subsets \(A, B \subseteq \Omega\) with \(A \subseteq B\) and any \(x \in \Omega \setminus B\), we have
\[ f(A \cup \{x\}) - f(B \cup \{x\}) \geq f(B) - f(A). \]

**Remark 2:** Sub-modular functions feature a natural diminishing returns property. In other words, the marginal return when a single element is added to an input set decreases as the size of the input set increases. An equivalent form of the above definition is to say that for any two sets \(A, B \subseteq \Omega\), we have 
\[ f(A) + f(B) \geq f(A \cup B) + f(A \cap B). \]

In the next lemma we show the sub-modularity of the players’ utility functions.

**Lemma 2:** The utility of players is a sub-modular function of their initial seed placement sets.

**Proof:** Given an initial strategy profile \((S_1, S_2, \ldots, S_m)\), let us denote the set of all the agents who adopt player \(i\)’th type at the end of the process by 
\[ U_i(S_i, S_{-i}), i \in \{1, 2, \ldots, m\}. \]
We show that the utility of the \(i\)th player \([U_i(S_i, S_{-i})]\) is a sub-modular function with respect to \(S_i\). In fact, for an arbitrary but fixed \(x \in V\) and every two sets \(S_i \subseteq S'_i\) of \(V\), we show that 
\[ |U_i(S_i \cup \{x\}, S_{-i})| - |U_i(S_i, S_{-i})| \geq |U_i(S'_i \cup \{x\}, S_{-i})| - |U_i(S'_i, S_{-i})| \]
Note that since 
\[ U_i(A, S_{-i}) \subseteq U_i(A \cup \{x\}, S_{-i}) \]
for all \(A \subseteq V(G)\), we have 
\[ |U_i(A, S_{-i})| = |U_i(A \cup \{x\}, S_{-i})| \]
for all \(A \subseteq V(G)\). Therefore, in order to prove \([4]\), we show that for every \(S_i \subseteq S'_i\) we have 
\[ U_i(S'_i \cup \{x\}, S_{-i}) \subseteq U_i(S_i \cup \{x\}, S_{-i}) \]
for every \(x \in V\). Since \(S_i \subseteq S'_i\) and 
\[ y \notin U_i(S'_i, S_{-i}), \]
we can write 
\[ y \in U_i(x, S_{-i}) \subseteq U_i(S'_i, S_{-i}) \]
and this completes the proof. \(\blacksquare\)

In fact, Lemma \([2]\) tells us that by applying a simple greedy algorithm, each player can find its optimal seed placement up to a constant factor of \(1 - \frac{1}{e}\). This simply follows by applying the result of greedy algorithm for sub-modular cost functions \([18]\).

**IV. NP-COMPLETENESS OF DECISION ABOUT NASH EQUILIBRIUM**

As we discussed before, competitive diffusion game may or may not accept pure-strategy Nash equilibria based on the topology of the network \(G\), number of the players, and their initial seed placements. In fact, by a closer look at Lemma \([1]\) one can see that there is some similarity between Voronoi games \([19]\) and competitive diffusion games. Note, however, that in the competitive diffusion game a diffusion process unfolds while there is no notion of diffusion in Voronoi games. However, since at the end of the process both games demonstrate behavior close to each other, it seems natural to compare the complexity of Nash equilibria in these two games. In fact, in the following we show that the decision on the existence of Nash equilibrium in a
diffusion game is NP-complete. Toward that goal, we first prove some relevant results and modify the configuration of the diffusion game to make a connection with Voronoi games. Borrowing some of the existing results from the Voronoi games literature, we prove the NP-completeness of verification of existence of Nash equilibrium in the diffusion game. We prove it by reduction from the 3-partitioning problem which is shown to be an NP-complete problem [20]. In the 3-partitioning problem we are given integers \(a_1, a_2, \ldots, a_{3m}\) and \(B\) such that \(B \neq \frac{m}{2}\) for every \(1 \leq i \leq 3m\). \(\sum_{i=1}^{3m} a_i = mB\) and have to partition them into disjoint sets \(P_1, \ldots, P_m \subseteq \{1, 2, \ldots, 3m\}\) such that for every \(1 \leq j \leq m\) we have \(\sum_{i \in P_j} a_i = B\). We begin by stating and proving the following lemma.

**Lemma 3:** Let \(U\) be a subset of \(V(G)\). Then one can extend \(G\) to \(\tilde{G}\) such that every Nash equilibrium in \(G\) which places the seed nodes in \(U\) is equivalent to an equilibrium in \(\tilde{G}\). In other words, there is an equivalence relationship between the set of Nash equilibria of \(G\) and the set of Nash equilibria of \(\tilde{G}\) with strategy set restricted to \(U\).

**Proof:** Consider the graph \(\tilde{G}\) depicted in Figure 1 which is constructed using \(G\) by adding \([U][U+1]\) new edges and \(n/2\) new edges. Note that \([U]\) denotes the number of the nodes in \(U\) and \(n \gg |V(G)|\) is a positive integer. With each node \(i \in U\) we associate a set of \(n\) new nodes \(C_i = \{c_{i1}, c_{i2}, \ldots, c_{in}\}\) and we connect all of them to node \(i\). Furthermore, for \(j = 1, \ldots, n\), we connect all the nodes \(c_{ij}, \ldots, c_{U[ij]}\) to each other. In other words, nodes \(\{c_{1j}, \ldots, c_{U[ij]}\}\) form a clique for each \(j = 1, \ldots, n\).

![Fig. 1. Extension of graph \(G\) to \(\tilde{G}\)](image)

Now assume that at least one player puts his node seed in \(k \in U\). We refer to this player as the first player and denote its type by \(A\). We claim that all the other players must play in \(U\) as well. To prove this, suppose that another player which we refer to as the second player with corresponding type \(B\) chooses node \(\ell \in U, \ell \neq k\). In this case he will earn at least \(n\) due to winning all the nodes in \(C_\ell\). Now let us assume that the second player plays in the bottom part of the graph (Figure 1), i.e. \(L = \cup_{i=1}^{|B|} C_i\). We consider two cases:

1. He plays in \(C_k\) and without any loss of generality and by symmetry, we may assume that he plays in \(c_{k1}\).
2. He plays in \(C_\ell\) for some \(\ell \neq k\) such as \(c_{\ell1}\).

In the first case and after the first step of diffusion, all the elements \(c_{11}, c_{21}, \ldots, c_{U[1]}\) will adopt type of the second player, i.e. \(B\) (because there is a direct link between them and \(c_{k1}\)), and all the elements \(c_{k2}, c_{k3}, \ldots, c_{kn}\) will adapt type of the first player, i.e. \(A\). At the second time step, the second player can adopt all the elements of \(U \cup \{k\}\) in the best case. On the other hand, all the elements of \(L \cup \{c_{11}, c_{21}, \ldots, c_{U[1]}\}\) will change to \(A\). Therefore, in this case the second player can gain at most \(|U| + |V(G)| \ll n\).

In the second case and after the first step of diffusion, the second player can adopt only \(\{c_{11}, c_{21}, \ldots, c_{U[1], \ell} \} \setminus \{c_{11}\}\) while the first player will adopt at least \(\{c_{k2}, \ldots, c_{kn}\}\), (note that node \(c_{k1}\) will change to gray). However, at the second time step all the nodes in \(C_\ell \setminus \{c_{\ell1}\}\) and also in \(L \cup C_\ell \cup \{c_{11}, c_{21}, \ldots, c_{U[1]}\}\) will change to gray. Therefore, in this case, the second player gains at most \(|U| + |V(G)| \ll n\).

Furthermore, if the second player places his seed at a node in \(V(G) \setminus U\), then in the best scenario it will take at least two steps for the seed to be diffused to nodes of \(L\). On the other hand, type \(A\) can be diffused through every node of \(L\) in no more than 2 steps. Thus all the nodes in \(L\) either adopt \(A\) or they change to gray and hence, in this case the second player can not earn more than \(|V(G)| - 1 \ll n\). From the above discussion it should be clear that in either of the above cases, if a player is playing in \(L \cup V(G) \setminus U\) he can always gain more by deviating to the set \(U\). Thus in each equilibrium players must put their seeds in \(U\).

Finally suppose that \(Q\) is a Nash equilibrium profile in \(\tilde{G}\). By the above argument we know that all the players must play in \(U\) and thus, each of these players gains exactly \(n\) from the set \(L\). Therefore, the utility of players is equal to the utility that they would gain by playing on \(G\) plus \(n\). This shows that \(Q\) must be an equilibrium for \(\tilde{G}\) when the strategies of players are restricted to \(U\). Similarly, if \(Q\) is an equilibrium of \(\tilde{G}\) when the strategies of players are restricted to \(U\), it is also an equilibrium for \(G\) as we know all the equilibria seeds of players (if there is any) must be in \(U\). This shows that the set of equilibria of \(\tilde{G}\) is equivalent to the set of equilibria of \(G\) with the restricted strategy set \(U\).

**Theorem 1:** Given a graph \(G\) and \(m \geq 2\) players, the decision process on the existence of Nash equilibrium for the diffusion game on \(G\) is NP-complete.

**Proof:** Assume that integers \(a_1, a_2, \ldots, a_{3m}\) and \(B > 3\) are given such that \(B < a_i < \frac{m}{2}\) for every \(1 \leq i \leq 3m\), \(\sum_{i=1}^{3m} a_i = mB\). Moreover, let \(c = \frac{3m}{B}\) and choose an integer \(d\) such that \((\frac{B-1}{2}c < d < \frac{B}{2}c\). Now we consider the graph depicted in Figure 2 where the set \(U\) is defined to be the subgraph induced in the dashed-line area.

It has been shown in [11] that the most right graph in Figure 2 does not admit any pure-strategy Nash equilibrium when there are two agents playing on it. In fact, if there is only one player on this graph, he will gain at least \(9d\) and if there are two players on it, one of them can always deviate to gain at least \(4d\). Calling the graph in Figure 2 \(G\) and applying Lemma 3 to construct \(\tilde{G}\), we can see that...
any Nash equilibrium of $\tilde{G}$ is an equilibrium of $G$ when the strategies of players are restricted to $U$. We claim that $Q$ is an equilibrium for $G$ with the restricted strategy set $U$ (and equivalently an equilibrium for $\tilde{G}$) if and only if there is a 3-partitioning of $\{a_i\}_{i=1}^{3m}$.

Assume that there is a solution $P_1, \ldots, P_m$ to the 3-partition. For every $1 \leq q \leq m$, if $P_q = \{i,j,k\}$, then player $q$ is assigned to $u_{i,j,k}$. Let us also assume that player $m+1$ is assigned to one of the nodes in the most right part of the graph, which makes his utility to be $9d$. If player $m+1$ moves to a vertex $u_{i,j,k}$, his utility will be $1 < 9d$, because all the other $m$ players already covered all the $\sum_{\ell=1}^{n} c\alpha_{\ell}$ nodes in the most left side of the graph and his movement will not result in any additional payoff for him except producing some gray nodes. Now if one of the players $1 \leq q \leq m$ moves from vertex $u_{i,j,k}$ to one of the nodes in the most right part of the graph, then his gain can be at most $4d$ by which the selection of $d$ would be less than what he was getting before (Bc). Finally, if player $q$ or equivalently node $u_{i,j,k}$ moves to another node $u_{i',j',k'}$ for some $\{i',j',k'\} \neq \{i,j,k\}$ then since $P_q$ was part of 3-partitioning before, it means that his payoff after deviating will be at most $c\max\{a_i + a_j, a_i + a_k, a_j + a_k\} < cB$. Moreover, by Lemma 3, no player at equilibrium will be out of $U$ and hence the proposed profile using the 3-partitioning is an equilibrium.

Now let us suppose that there exists a Nash equilibrium for $\tilde{G}$. We show that it corresponds to a solution of 3-partitioning. First we note that there cannot be two players in the most right part of the graph, otherwise it is not an equilibrium by [11]. Moreover, if there are 3 players or more, one of them can gain at most $3d$. Since in this case there are at most $m-2$ players in the middle part, we can find a set $\{i',j',k'\}$ such that the corresponding set of all the other players does not have any intersection with it. Therefore, if a player with the least gain (at most $3d$) from the right side deviates to $u_{i',j',k'}$ in the middle part, he will gain at least $\frac{3d}{4}$ which is greater than $3d$. Thus the most right part can have either one player or nothing. However since $9d > \frac{3d}{2}$, at least one player would want to move to the most right part of the graph if there is no other player there. Therefore, the most right part of the graph has exactly one player. Thus, the rest of the $m$ players must not only play in $U$ (because of Lemma 3) but also they must form a partition. Otherwise one of them can move to an appropriate vertex of the middle graph and increase his utility. Finally, in this partitioning, each player must gain exactly $Bc$, because if this is not true and one of the players namely $u_{i,j,k}$ gets less than $Bc$, he will gain at most $(B-1)c$ (note that $c$ is a rescaling factor) and thus, he can always move to the most right side of the graph and gain $4d > (B-1)c$. Thus this partitioning must be a 3-partitioning. This proves the equivalence of the existence of Nash equilibrium in $\tilde{G}$ and existence of 3-partitioning for the set $\{a_i\}_{i=1}^{3m}$.

V. NECESSARY CONDITIONS FOR NASH EQUILIBRIUM AND OPTIMUM SOCIAL WELFARE

In this section, we present some necessary conditions for Nash equilibrium and a lower bound on the average social welfare players can achieve. We will obtain sharp results for certain type of graphs. In the following proposition we provide a necessary condition for a particular profile to be a Nash equilibrium. This result is for the case of two players; however, it can be extended quite naturally to an arbitrary number of players.

Proposition 2: Suppose that $(a,b) \in V \times V$ is an equilibrium profile for the diffusion game. Then,

\[
\left[ \frac{n-1}{d_G(a)} \right] \leq u(b), \quad \left[ \frac{n-1}{d_G(b)} \right] \leq u(a),
\]

\[
\left[ \frac{n-1}{d_G(a)} \right] + \left[ \frac{n-1}{d_G(b)} \right] \leq n - |\text{gray nodes in allocation } (a,b)|,
\]

where $u(a)$ and $u(b)$ denote the utilities of players $A$ and $B$ given the initial seed placement at $(a,b)$.

Proof: Assume that player $A$ and $B$ place their seeds at nodes $a$ and $b$ and receive payoffs of $u(a)$ and $u(b)$, respectively. We claim that there exists a neighbor of $a$ where player $b$ can gain at least $\left[ \frac{n-1}{d_G(b)} \right]$ by deviating to it. Toward showing this, let us also denote all the neighbors of $a$ by $i_1,i_2,\ldots,i_d(a)$. Let us denote the nodes that adopt $B$ for the initial seed allocation $(a,i_j)$ by $S_j$, $j = 1,\ldots,a(d)$. Then, we have $\cup_{j=1}^{d(a)} S_j = V \setminus \{a\}$. In fact, for every $x \in V \setminus \{a\}$, the shortest path from $x$ to $a$ must pass through at least one of the neighbors of $a$ such as $i_t$. This means that $d_G(x,i_t) < d_G(x,a)$ and using Lemma 1 we can see that $x \in S_t$. Therefore we have $n-1 = |\cup_{j=1}^{d(a)} S_j| \leq \sum_{j=1}^{d(a)} |S_j|$ and this means that there exists at least one $i_{j_t}$ such that $|S_{j_t}| \geq \left[ \frac{n-1}{d_G(a)} \right] = \left[ \frac{n-1}{d(a)} \right]$. Since we assumed that $(a,b)$ is an equilibrium, player $b$ can not gain more by deviating to $i_{j_t}$. This means that $\left[ \frac{n-1}{d_G(b)} \right] \leq u(b)$ and using the same argument for the other player we get $\left[ \frac{n-1}{d_G(a)} \right] \leq u(a)$. Using these relations and noting that $u(a) + u(b) = n - |\text{gray nodes in allocation } (a,b)|$, we arrive at the desired result.

Note that the results in Proposition 2 can be improved by noting that the inequality $|\cup_{j=1}^{d(a)} S_j| \leq \sum_{j=1}^{d(a)} |S_j|$ can be
strict and there are nodes which might be counted in different sets of $S_j$. In fact, it is not hard to see that if a node $x$ belongs to two of these sets such as $S_i$ and $S_{i_2}$, $x$ must be in an even cycle emanating $a$ and including the nodes $i_1$ and $i_2$. In such a case, to every cycle of even length which includes $a$ and does not contain another smaller cycle, one can associate a node which is counted twice in different sets. We call such cycles simple even cycles emanating from $a$. Therefore, we can write

$$n-1 = \left| \bigcup_{d(a)} S_j \right| + \left| \{ \text{Simple even cycles emanating from } a \} \right| \leq \sum_{j=1}^{d(a)} |S_j|,$$

and therefore the bound in the Proposition 2 will change to

$$\left| \frac{n-1 - \left| \{ \text{Simple even cycles emanating } a \} \right|}{\sigma(a)} \right| \leq u(b).$$

We next define the zero pattern of a nonnegative matrix, which is followed by a theorem which provides a lower bound on the optimal social welfare using the adjacency matrix of $G$.

**Definition 2:** Given an arbitrary nonnegative matrix $X$, define $\sigma(X)$ to be

$$\sigma(X)_{ij} = \begin{cases} 1, & X_{ij} > 0 \\ 0, & X_{ij} = 0. \end{cases}$$

**Theorem 3:** Given a graph $G = (V, E)$ of $n$ nodes, diameter $D$ and two players $A$ and $B$, there exists an initial seed placement $(a, b)$ for players such that the social utility $u(a) + u(b)$ is at least

$$n + 1 - \frac{\sum_{k=1}^{D} \| (\sigma(A^k) - \sigma(A^{k-1}))1 \|^2}{n(n-1)},$$

where $\| \cdot \|$ denotes the standard Euclidean norm.

**Proof:** Let us define $G^{(k)} = (V, E^{(k)})$, where $E^{(k)} = \{(i, j) | d_G(i, j) = k\}$. We consider all the initial placements over different pairs of nodes, and then compute the average utility gained by players. To do that, we consider an array $Q$ of $\binom{n}{2}$ rows and $n$ different columns. For $i \neq j, k \in \{1, 2, \ldots, n\}$, we let $Q(i, j, k) = 1$ if and only if node $k$ adopts either $A$ or $B$ during the diffusion process for the initial placement $(i, j)$, and $Q(i, j, k) = 0$, otherwise. We count the maximum number of zeros in column $x$. For an arbitrary but fixed node $x \in V = \{1, 2, \ldots, n\}$, we count the number of different initial seed placements which result in node $x$ turning to gray. In other words, we count the maximum number of zeros in column $x$ of $Q$. For this purpose we note that, using Lemma 1, if node $x$ turns to gray, it must be at equal distance from seed nodes. On the other hand, for a given $k = 1, 2, \ldots, D$, the number of choosing two nodes at distance $k$ from node $x$ (as possible seed placements which may turn $x$ to gray) is the same as the number of choosing two nodes among neighbors of $x$ in $G^{(k)}$, i.e. $(d^{(k)}_G(x))^2$. Thus, the maximum number of zeros in column $x$ of $Q$ is upper bounded by $\sum_{k=1}^{D} (d^{(k)}_G(x))^2$ and hence the number of zeros in $Q$ is bounded from above by

$$\sum_{x \in V} \sum_{k=1}^{D} \left( d^{(k)}_G(x) \right)^2.$$ 

This means that the average number of ones in each row of $Q$ is at least

$$n \left( \frac{1}{n} \right) - \sum_{x \in V} \sum_{k=1}^{D} \left( d^{(k)}_G(x) \right)^2 = n - \sum_{k=1}^{D} \sum_{x \in V} \left( d^{(k)}_G(x) \right)^2 = n - \frac{n(n-1)}{2} \sum_{k=1}^{D} \sum_{x \in V} \left( d^{(k)}_G(x) \right)^2 \frac{1}{n(n-1)} \sum_{k=1}^{D} \sum_{x \in V} d^{(k)}_G(x) \frac{1}{n(n-1)} = n - \frac{n}{n-1} \sum_{k=1}^{D} \sum_{x \in V} d^{(k)}_G(x) \frac{1}{n(n-1)},$$

where the last equality follows because $\{E^{(k)} \}_{k=1}^{D}$ partitions all the edges of a complete graph with $n$ nodes. This yields a lower bound on the maximum social welfare for the case of two players on the graph.

Finally, Using the zero pattern definition of a matrix, we can compute the last quantity in (5) in the following way. Let us take $A_G$ to be the adjacency matrix of graph $G$ of diameter $D$ and $A = I + A_G$, where $I$ denotes the identity matrix of appropriate dimension. It is not hard to see that $\sigma(A^k) - \sigma(A^{k-1})$ is the adjacency matrix of $G^{(k)}$. In other words, $A_G^{(k)} = \sigma(A^k) - \sigma(A^{k-1})$. Therefore, if we let $1$ be the column vector of all ones, the degree of each node $x$ in $G^{(k)}$ can be found easily by looking at the $x$ coordinate of vector $[\sigma(A^k) - \sigma(A^{k-1})]1$. Thus we have the following equality

$$\sum_{x \in V} \sum_{k=1}^{D} \left( d^{(k)}_G(x) \right)^2 = \sum_{k=1}^{D} \sum_{x \in V} \left( d^{(k)}_G(x) \right)^2 = \left[ \sigma(A^k) - \sigma(A^{k-1}) \right]1$$

Note that the total number of operations needed to compute the expression given in Theorem 3 is at most polynomial in terms of the number of nodes in the graph.

**VI. Conclusion**

In this paper, we have studied the competitive diffusion game on an undirected connected network. We showed sub-modularity of the players’ utilities based on their initial seed placements. Also, we proved NP-completeness of the decision process on the existence of pure-strategy Nash equilibrium in general networks. Moreover, we have presented some necessary conditions for a given profile to be an equilibrium in general graphs, and we gave a lower bound
on the maximum social welfare of the game in the case of two players.

As a future direction of research, an interesting problem would be to identify the class of networks which admit Nash equilibria for the case of two players. It is not hard to see that a tree construction which is a special case of bipartite graphs always leads to a pure-strategy equilibrium [17]. Also, one can think of a probabilistic model which can approximate the deterministic version of the diffusion game and it is easier to be traced as long as the diffusion process proceeds.

REFERENCES

[1] S. R. Etesami, T. Başar, A. Nedic, and B. Touri, “Termination time of multidimensional Hegselmann-Krause opinion dynamics,” in American Control Conference (ACC), 2013. IEEE, 2013, pp. 1255–1260.

[2] S. Goyal and M. Kearns, “Competitive contagion in networks,” in Proceedings of the 44th Symposium on Theory of Computing. ACM, 2012, pp. 759–774.

[3] H. P. Young, “The diffusion of innovations in social networks,” Economy as an Evolving Complex System. Proceedings Volume in the Santa Fe Institute Studies in the Sciences of Complexity, vol. 3, pp. 267–282, 2002.

[4] D. Acemoglu, A. Ozdaglar, and E. Yildiz, “Diffusion of innovations in social networks,” in Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on. IEEE, 2011, pp. 2329–2334.

[5] D. Kempe, J. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” in Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2003, pp. 137–146.

[6] M. Richardson and P. Domingos, “Mining knowledge-sharing sites for viral marketing,” in Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2002, pp. 61–70.

[7] S. Bharathi, D. Kempe, and M. Salek, “Competitive influence maximization in social networks,” in Internet and Network Economics. Springer, 2007, pp. 306–311.

[8] S. Brânzei and K. Larson, “Social distance games,” in Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence-Volume One. AAAI Press, 2011, pp. 91–96.

[9] Y. Singer, “How to win friends and influence people, truthfully: influence maximization mechanisms for social networks,” in Proceedings of the Fifth ACM International Conference on Web Search and Data Mining. ACM, 2012, pp. 733–742.

[10] N. Alon, M. Feldman, A. D. Procaccia, and M. Tennenholtz, “A note on competitive diffusion through social networks,” Information Processing Letters, vol. 110, no. 6, pp. 221–225, 2010.

[11] R. Takehara, M. Hachimori, and M. Shigeno, “A comment on pure-strategy Nash equilibria in competitive diffusion games,” Information Processing Letters, vol. 112, no. 3, pp. 59–60, 2012.

[12] E. Yildiz, D. Acemoglu, A. Ozdaglar, A. Saberi, and A. Scaglione, “Discrete opinion dynamics with stubborn agents,” Available at SSRN 1744113, 2011.

[13] M. Fazli, M. Ghodsi, J. Habibi, P. J. Khalilabadi, V. Mirokni, and S. S. Sadjehabab, “On the non-progressive spread of influence through social networks,” in LATIN 2012: Theoretical Informatics. Springer, 2012, pp. 315–326.

[14] E. Ackerman, O. Ben-Zwi, and G. Wolfowitz, “Combinatorial model and bounds for target set selection,” Theoretical Computer Science, vol. 411, no. 44, pp. 4017–4022, 2010.

[15] A. Gionis, E. Terzi, and P. Tsaparas, “Opinion maximization in social networks,” arXiv preprint arXiv:1301.7455, 2013.

[16] L. Small and O. Mason, “Information diffusion on the iterated local transitivity model of online social networks,” Discrete Applied Mathematics, 2012.

[17] C. Dürr and N. K. Thang, “Nash equilibria in voronoi games on graphs,” in Algorithms–ESA 2007. Springer, 2007, pp. 17–28.

[18] M. R. Garey and D. S. Johnson, “Complexity results for multiprocessor scheduling under resource constraints,” SIAM Journal on Computing, vol. 4, no. 4, pp. 397–411, 1975.