Theoretical Computer Science 412 (2011) 7147–7168

Contents lists available at SciVerse ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Bounded budget betweenness centrality game for strategic network formations☆

Xiaohui Bei [a,*], Wei Chen [b], Shang-Hua Teng [c], Jialin Zhang [a], Jiajie Zhu [d]

[a] Institute for Theoretical Computer Science, Tsinghua University, China
[b] Microsoft Research Asia, China
[c] University of Southern California, United States
[d] University of California at Los Angeles, United States

ARTICLE INFO

Article history:
Received 7 December 2009
Received in revised form 11 December 2010
Accepted 28 September 2011
Communicated by P. Spirakis

Keywords:
Algorithmic game theory
Network formation game
Nash equilibrium

ABSTRACT

In computer networks and social networks, the *betweenness centrality* of a node measures the amount of information passing through the node when all pairs are conducting shortest path exchanges. In this paper, we introduce a strategic network formation game in which nodes build connections subject to a budget constraint in order to maximize their betweenness in the network. To reflect real world scenarios where short paths are more important in information exchange in the network, we generalize the betweenness definition to only count shortest paths with a length limit ℓ in betweenness calculation. We refer to this game as the *bounded budget betweenness centrality* game and denote it as ℓ-B^{3}C game, where ℓ is the path length constraint parameter.

We present both complexity and constructive existence results about Nash equilibria of the game. For the nonuniform version of the game where node budgets, link costs, and pairwise communication weights may vary, we show that Nash equilibria may not exist and it is NP-hard to decide whether Nash equilibria exist in a game instance. For the uniform version of the game where link costs and pairwise communication weights are one and each node can build k links, we construct two families of Nash equilibria based on shift graphs, and study the properties of Nash equilibria. Moreover, we study the complexity of computing best responses and show that the task is polynomial for uniform 2-B^{3}C games and NP-hard for other games (i.e. uniform ℓ-B^{3}C games with $\ell \geq 3$ and nonuniform ℓ-B^{3}C games with $\ell \geq 2$).

1. Introduction

Many network structures in real life are not designed by central authorities. Instead, they are formed by autonomous agents who often have selfish motives [19]. Typical examples of such networks include the Internet where autonomous systems linked together to achieve global connection, peer-to-peer networks where peers connect to one another for online file sharing (e.g. [5,22]), and social networks where individuals connect to one another for information exchange and other social functions [21]. Since these autonomous agents have their selfish motives and are not under any centralized control, they often act strategically in deciding whom to connect to in order to improve their own benefits. This gives rise to the field

☆ An extended abstract of the paper appears in the proceedings of 17th Annual European Symposium (ESA2009).

* Corresponding author. Tel.: +86 1062777907.
E-mail addresses: bxh08@mails.tsinghua.edu.cn, beixiaohui@gmail.com (X. Bei), weic@microsoft.com (W. Chen), shanghua@usc.edu (S.-H. Teng), zhanggl02@mails.tsinghua.edu.cn (J. Zhang), jiajie@cs.ucla.edu (J. Zhu).

0304-3975/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2011.09.028

of *network formation games*, which studies the game-theoretic properties of the networks formed by these selfish agents as well as the process in which all agents dynamically adjust their strategies [1,10,15–17].

A key measure of importance of a node is its betweenness centrality. The *betweenness centrality* (or betweenness for short) is introduced originally from social network analysis as one of the measures on how central an individual is in a social network [12,18]. If we view a network as a graph $G = (V, E)$ (directed or undirected), the betweenness of a node (or vertex) i in G is

$$btw_i(G) = \sum_{u \neq v \neq i \in V,\, m(u,v)>0} w(u, v) \frac{m_i(u, v)}{m(u, v)} \tag{1}$$

where $m(u, v)$ is the number of shortest paths from u to v in G, $m_i(u, v)$ is the number of shortest paths from u to v that pass i in G (all shortest paths are with respect to the number of edges on the paths), and $w(u, v)$ is the weight on pair (u, v). Intuitively, if the amount of information from u to v is $w(u, v)$, and the information is passed along the shortest paths from u to v, and all shortest paths split the traffic equally, then the betweenness of node i measures the amount of information passing through i incurred by all pair-wise exchanges.

In this paper, we generalize the betweenness definition with a parameter ℓ such that only shortest paths with length at most ℓ are considered in betweenness calculation. Formally, we define

$$btw_i(G, \ell) = \sum_{u \neq v \neq i \in V,\, m(u,v,\ell)>0} w(u, v) \frac{m_i(u, v, \ell)}{m(u, v, \ell)}, \tag{2}$$

where $m(u, v, \ell)$ is the number of shortest paths from u to v in G with length at most ℓ, and $m_i(u, v, \ell)$ is the number of shortest paths from u to v that passes i in G with length at most ℓ. It is easy to see that $btw_i(G) = btw_i(G, n - 1)$, where n is the number of vertices in G.

Betweenness with path length constraint is reasonable in real-world scenarios. In social networks, researches (e.g. [2,3]) show that short connections are much more important than long-range connections. In fact, results of [2,3] motivate Kleinberg et al. to consider essentially $btw_i(G, 2)$ as part of the objective function in their game [15]. In other scenarios, shortest paths with length $l > 2$ are also considered in the analysis. For example in peer-to-peer networks such as Gnetella [22], query requests are searched only on nodes with a short graph distance (but greater than 2) away from the query initiator. In social networks, shortest paths of more than 2 could also be important. For example, consider the recent DARPA network challenge [7], where Defense Advanced Research Projects Agency (DARPA) asked the registered teams to identify the locations of 10 weather balloons in the continental US as fast as possible. The winning team from MIT media lab used incentives and social networking to recruit almost 5000 people in about 36 hours. The depths of their recruitment trees ranges from 2 to 14, showing that paths longer than 2 are also important for information dissemination [20]. Therefore, we believe it is reasonable to consider shortest paths with length greater than 2 in the betweenness computation, and our definition $btw_i(G, \ell)$ can be viewed as a generalization of [15] in this regard.

In a decentralized network with autonomous agents, each agent may have incentive to maximize its betweenness in the network. For example, in computer networks and peer-to-peer networks, a node in the network may be able to charge the traffic that it helps relaying, in which case the revenue of the node is proportional to its betweenness in the network. So the maximization of revenue is consistent with the maximization of the betweenness. In a social network, an individual may want to gain or control the most amount of information traveling in the network by maximizing her betweenness. Again considering the DARPA network challenge, we can see that individuals with high betweenness in the social network is more likely to collect reward money since more information diffusion paths pass through them. Therefore, individuals do have incentives to increase their betweenness in a social network.

In this paper, we introduce a network formation game in which every node in a network is a selfish agent who decides which other nodes in the network to build connection with in order to maximize its own betweenness. Building connections with other nodes incur costs. Each node has a budget such that the cost of building its connections cannot exceed its budget. We call this game the *bounded budget betweenness centrality* game or the B^3C game. When distinction is necessary, we use ℓ-B^3C to denote the games using generalized betweenness definition $btw_i(G, \ell)$.

Bounded budget assumption, first incorporated into a network formation game in [16], reflects real world scenarios where there are physical limits to the number of connections one can make. In computer and peer-to-peer networks, each node usually has a connection limit. In social networks, each individual only has a limited time and energy to create and maintain direct relationships with other individuals, which is referred to as the Dunbar limit in sociology literature [9]. An alternative treatment to connection costs appearing in more studies [1,10,15,17] is to subtract connection costs from the main objectives to be maximized, which means that as long as the benefit outweighs the cost, a node is allowed to build more connections without other physical constraints. This treatment, however, restricts the variety of Nash equilibria exhibited by the game. For example, Kleinberg et al. [15] show that all Nash equilibria in their network formation game (with a social network background) are dense graphs with at least $\Omega(n^2)$ number of edges for graphs with n vertices, because the game has no connection budget constraint. However, it is well known that social networks are sparse graphs, since individuals have physical constraints and thus can only build connections with a relatively small number of people. Therefore, in this paper we explicitly incorporate the bounded budget assumption to overcome the shortcomings existed in prior work such as [15], even though it makes the game model more complicated.

In this paper, we consider the directed graph variant of the game, in which links in the network are directed and nodes can establish outgoing links to other nodes. This is suitable for computer networks and peer-to-peer networks that relay traffics, in which downloading and uploading links are often not symmetric, as also modeled by [16]. In social networks, the relationship between a pair of individuals in terms of information exchange may not be symmetric, especially if the relationship was built by the unilateral effort of one individual with some selfish incentive. One concrete example is the Twitter follower network, in which users choose which other users to follow, and the follower–followee relationship is asymmetric.[1] Therefore, it is reasonable to use the directed graph model in our betweenness game.

Our results. We focus on the algorithmic aspect of computing Nash equilibria as well as their structures in the B^3C games. Since the game allows some trivial Nash equilibria (such as a network with no links at all), we study a stronger form called *maximal* Nash equilibria, in which no node can add more outgoing links without exceeding its budget constraint. Since adding outgoing links of a node can only help its betweenness, it is reasonable to study maximal Nash equilibria in the B^3C games.

We present both complexity results and existence results about this game. We first show that the general *nonuniform* ℓ-B^3C game may not have any maximal Nash equilibria for any $\ell \geq 2$. A nonuniform ℓ-B^3C game is specified by several parameters concerning the node budgets, link costs, and pairwise communication weights (see Section 2 for a formal definition). Moreover, given these parameters as input, we show that it is NP-hard to determine whether the game has a maximal Nash equilibrium. The result indicates that finding Nash equilibria in general ℓ-B^3C games is a difficult task.

We then address the complexity of computing best responses in ℓ-B^3C games. For uniform ℓ-B^3C games where all pair weights are one, all link costs are one, and all node budgets are given as an integer k, we show that with $\ell = 2$, computing a best response takes $O(n^3)$ time. For all other cases (uniform games with $\ell \geq 3$ or nonuniform games with $\ell \geq 2$), the task is NP-hard.

Next, we turn our attention to the construction of Nash equilibria in *uniform* ℓ-B^3C games and their properties. We introduce a type of multi-partite graphs that we call *shift graphs*, which are variants of better known De Bruijn graphs and Kautz graphs [8]. Based on these shift graphs, we construct two different families of Nash equilibria for *uniform* ℓ-B^3C games. One family gives a stronger form of Nash equilibria call strict Nash equilibria, while the other family belongs to what we call *ℓ-path-unique graphs* (ℓ-PUGs), which we show are always Nash equilibria for uniform ℓ-B^3C games.

Finally, we use ℓ-PUGs to study several properties of Nash equilibria. In particular, we show that (a) for any ℓ, k and large enough n ($n \geq (k+\ell)!/k!$), a maximal Nash equilibrium exists; (b) Nash equilibria may exhibit rich structures, e.g. they may be disconnected or unbalanced (some nodes have zero in-degree and zero betweenness while other nodes have very large in-degree and betweenness); and (c) for 2-B^3C games, when the in-degree are relatively balanced all maximal Nash equilibria must be 2-PUGs, a direct consequence of which is that Abelian Cayley graphs with sufficiently large n ($n \geq k^3 + k^2 + 2k$) cannot be Nash equilibria for 2-B^3C games.

Whenever applicable, we also state the results for B^3C games without the path length constraint.

To summarize, our contributions include: (a) we define the bounded budget betweenness centrality game to study the strategic network formation with maximizing betweenness as the goal, and we are the first to incorporate reasonable assumptions of both bounded budget and general path length constraint into betweenness related games; (b) we show that in the general version of the game where budgets, link costs and pairwise betweenness contribution may vary, Nash equilibria may not exist and it is NP-hard to decide if a game instance has a Nash equilibrium; (c) we show that computing best responses is polynomial-time solvable for uniform 2-B^3C games and is NP-hard for other variants; (d) for the uniform ℓ-B^3C games, we explicitly construct families of Nash equilibria and provide several features about Nash equilibria in these games. We hope that this research will motivate further studies on betweenness related network formation games.

Related work. There are a number of studies on network formation games with Nash equilibrium as the solution concept [1,10,15–17]. Most of the above work belong to a class of games in which nodes try to minimize their average shortest distances to other nodes in the network [1,10,16,17], which is called *closeness centrality* in social network analysis [12]. The game in [10] considers undirected edges and the cost of links in the network are part of the objective function to minimize. It focuses on the study of price of anarchy of the game and also presents results on the structure of Nash equilibria. In [1], Albers et al. extend the research of [10] by disproving a conjecture made in [10] that all Nash equilibria have a tree structure, and studying other variants of the game including the cost of an edge being shared by two end nodes. The game in [17] instead considers minimizing the average stretch of each node, where stretch is defined as the ratio between the shortest path distance of two nodes in the graph versus the geometric distance in the underlying space.

Our research is partly motivated by the work of [15], in which Kleinberg et al. study a different type of network formation games related to the concept of structural holes in organizational social network research. In this game, each node tries to bridge other pairs of nodes that are not directly connected. In a sense, this is a restricted type of betweenness where only length-two shortest paths are considered. Besides some difference in the game setup, such as they use undirected edges, there are two important differences between our work and theirs. First, we consider betweenness with a general path length constraint of ℓ as well as no path length constraints, while they only consider the bridging effect between two immediate

[1] In the Twitter follower network, users are building incoming links since they receive tweets (microblog posts) from people whom they follow. In Section 2 we will explain that building incoming links are equivalent as building outgoing links in terms of the game format.

neighbors of a node. For the case of $\ell = 2$, both studies include results showing that computing best response is polynomial-time solvable, but their work does not show in their non-uniform case if Nash equilibrium exists and whether deciding the existence of Nash equilibria is NP-hard, while we have results on both. Second, we incorporate budget constraints to restrict the number of links one node can build, while their work has not such constraint. As already discussed, without link budget constraints, they show that all Nash equilibria are dense graphs with $\Omega(n^2)$ edges where n is the number of vertices. This is what we want to avoid in our study. A couple of other studies [4,13] also address strategic network formations with structural holes, but they do not address the computation issue, and their game formats have their own limitations (e.g. star networks as the only type of equilibria [13] or limited to length-2 paths [4]).

Our game is also inspired by the BBC game of Laoutaris et al. [16]. This game considers directed links and bounded budgets on nodes, using minimization of average shortest distances to others as the objective for each node. It shows hardness results in determining the existence of Nash equilibria in general games, and provides tree-like structures as Nash equilibria for the uniform version of the game. It also shows that Abelian Cayley graphs cannot be Nash equilibria in large networks.

Solution concepts other than Nash equilibrium are also used in the study of network formation games. Authors in [6,14] consider games in which two end points of a link have to jointly agree on adding the link, and they use pairwise stability as an alternative to Nash equilibrium.

Paper organization. Section 2 provides the detailed definition of the ℓ-B^{3}C game and the related concepts. Section 3 provides the complexity result on determining the existence of Nash equilibria in nonuniform games, while Section 4 presents the results on the complexity of computing best responses. Section 5 presents the construction of Nash equilibria in uniform games via shift graphs and studies the properties of Nash equilibria. We conclude the paper and discuss future directions in Section 6.

2. Problem definition

We first define the *bounded-budget betweenness centrality game* (B^{3}C game) without path length constraint, and then extend it to the version with path constraint (ℓ-B^{3}C game). A *bounded-budget betweenness centrality* game with parameters (n, b, c, w) is a network formation game defined as follows. We consider a set of n players $V = \{1, 2, \ldots, n\}$, which are also nodes in a network. Function $b : V \to \mathbb{N}$ specifies the budget $b(i)$ for each node $i \in V$ ($\mathbb{N}$ is the set of natural numbers). Function $c : V \times V \to \mathbb{N}$ specifies the cost $c(i, j)$ for the node i to establish a link to node j, for $i, j \in V$. Function $w : V \times V \to \mathbb{N}$ specifies the weight $w(i, j)$ from node i to node j for $i, j \in V$, which can be interpreted as the amount of traffic i sends to j, or the importance of the communication from i to j.[2]

The strategy space of player i in B^{3}C game is $S_i = \{s_i \subseteq V \setminus \{i\} \mid \sum_{j\in s_i} c(i, j) \le b(i)\}$, i.e., all possible subsets of outgoing links of node i within i's budget. A strategy profile $s = (s_1, s_2, \ldots, s_n) \in S_1 \times S_2 \times \cdots \times S_n$ is referred to as a *configuration* in this paper. The graph induced by configuration s is denoted as $G_s = (V, E)$, where $E = \{(i, j) \mid i \in V, j \in s_i\}$. For convenience, we will also refer G_s as a configuration.

The utility of a node i in configuration s is defined by the *betweenness centrality* of i in the graph G_s as follows:

$$btw_i(s) = btw_i(G_s) = \sum_{u\neq v\neq i\in V,\, m(u,v)>0} w(u, v)\frac{m_i(u, v)}{m(u, v)}, \tag{3}$$

where $m(u, v)$ is the number of shortest paths from u to v in G_s and $m_i(u, v)$ is the number of shortest paths from u to v that passes i in G_s.

We now generalize the definition of betweenness, such that a shortest path from u to v contributes to the betweenness of a node i on the path only when the path length is at most ℓ, for some parameter ℓ. Formally, given a graph G_s (corresponding to a configuration s) and a parameter $\ell \in \mathbb{N}$, we define

$$btw_i(G_s, \ell) = \sum_{u\neq v\neq i\in V,\, m(u,v,\ell)>0} w(u, v)\frac{m_i(u, v, \ell)}{m(u, v, \ell)}, \tag{4}$$

where $m(u, v, \ell)$ is the number of shortest paths from u to v in G_s with length at most ℓ, and $m_i(u, v, \ell)$ is the number of shortest paths from u to v that passes i in G_s with length at most ℓ. Since the longest shortest path in G_s is at most $n - 1$, we know that $btw_i(G_s) = btw_i(G_s, n - 1)$. We use ℓ-B^{3}C game to denote the version of B^{3}C game with parameter ℓ and $btw_i(G_s, \ell)$ as the utility of node i.

Definition 1. A configuration s is a *(pure) Nash equilibrium*, if in s no node can increase its own utility by changing its own strategy unilaterally, and we also say that s is *stable*.

Definition 2. A configuration s is a *strict Nash equilibrium*, if in s any strategy change of any node strictly decreases the utility of the node.

2 We may also define a distance function specifying distances between every pair of nodes, but it is not needed throughout our paper.

The following Lemma shows the monotonicity of betweenness centrality when adding new edges to a node, which motivates our definition of maximal Nash equilibrium. It is stated for $btw_i(G)$ and B^{3}C games, but is also applicable to $btw_i(G, \ell)$ and ℓ-B^{3}C games. Its proof is straightforward and omitted.

Lemma 1. *Adding an outgoing edge to a node i does not decrease i's betweenness. That is, for any graph $G = (V, E)$ with $i \in V$ and $(i, j) \notin E$ for some $j \in V$. Let $G' = (V, E \cup \{(i, j)\})$. Then $btw_i(G) \leq btw_i(G')$.*

Given a nonuniform B^{3}C game with parameters (n, b, c, w), a *maximal strategy* of a node v is a strategy with which v cannot add any outgoing edges without exceeding its budget. We say that a graph (configuration) is *maximal* if all nodes use maximal strategies in the configuration. By the monotonicity of betweenness centrality, it makes sense to study maximal graphs where no node can add more edges within its budget limit. Moreover, some trivial non-maximal graphs are trivial Nash equilibria, e.g. empty graphs with no edges. However, when nodes add more edges into the graph allowed by their budgets, other nodes may have chance of improving their utilities by changing their strategies. Therefore, for the rest of the paper, we focus on Nash equilibria in maximal graphs. In particular, we say that a configuration is a *maximal Nash equilibrium* if it is a maximal graph and it is a Nash equilibrium.

The following lemma states the relationship between maximal Nash equilibria and strict Nash equilibria, a direct consequence of the monotonicity of betweenness centrality.

Lemma 2. *Given a B^{3}C game with parameters (n, b, c, w), any strict Nash equilibrium in the game is a maximal Nash equilibrium.*

Based on the above lemma, for positive existence of Nash equilibria, we sometimes study the existence of strict Nash equilibria to make our results stronger.

We also consider a special case of B^{3}C(or ℓ-B^{3}C) game which is defined as following.

Definition 3. A B^{3}C(or ℓ-B^{3}C) game with parameters $n, k \in \mathbb{N}$ is *uniform* if $b(i) = k$ for all $i \in V$, and $c(i, j) = w(i, j) = 1$ for all $i, j \in V$. As a contrast, the general form is called *nonuniform* games.

A remark is now in order concerning the directions of edges in the B^{3}C (or ℓ-B^{3}C) games. By our definition, each player can build outgoing edges to other players. We may also define games in which each player can build incoming edges from other players. Building incoming edges matches the scenario of the Twitter follower network, in which users select a set of other users to follow, meaning to receive tweets from. We claim that these two formulations of the games are equivalent, by the following observation. Let G be a graph and G' be another graph obtained by reversing the directions of all edges of G. Let w be a weight function on pairs of vertices in G, and let w' be the weight function on pairs of vertices in G' such that $w'(u, v) = w(v, u)$ for all pair of vertices u and v. Then we observe that for any vertex i its betweenness (with path length constraint or not) in G with weight function w is the same as in G' with weight function w'. This observation implies that any game in which players build incoming edges is equivalent to another game in which players build outgoing edges. Henceforth, we focus on one type of games in which players build outgoing edges, and all of our results apply directly to games in which players build incoming edges.

3. Complexity of determining Nash equilibria in nonuniform games

Given the rich parameters, a nonuniform B^{3}C game may have complex behavior. In particular, it may not have any maximal (or strict) Nash equilibrium, and determining whether a game has a maximal (or strict) Nash equilibrium is NP-hard.

For simplicity, the main part of this section addresses the B^{3}C game without path length constraint. We address the ℓ-B^{3}C game after each main result for the B^{3}C game.

3.1. Nonexistence of maximal Nash equilibria

In this section, we show that maximal Nash equilibria may not exist in some version of B^{3}C games where edge costs are not uniform. By Lemma 2, it implies that strict Nash equilibria do not exist either in the same game.

We now construct a family of graphs, which we refer to as the gadget, and show that B^{3}C games based on the gadget do not have any maximal Nash equilibrium. The gadget is shown in Fig. 1. There are $5 + 3t + r$ nodes in the gadget, where $t \in \mathbb{N}$ and $r = 1, 2, 3$. The values of t and r allow us to construct a graph of any size great than 5. There are r nodes, denoted as A, A', A'' in the figure, which establish edges to B and C. Both B and C can establish at most one edge to a node in $\{D, E, F\}$ respectively. Each node in $\{D, E, F\}$ connects to a cluster of size t each (not shown in the figure). The only requirement for these three clusters is that they are identical to each other and are all strongly connected so D, E, F can reach all nodes in their corresponding clusters. Nodes in the three clusters do not establish edges to the other clusters or to $A, A', A'', B, C, D, E, F$.

We classify nodes and edges as follows. Nodes B and C are *flexible* nodes since they can choose to connect one node in $\{D, E, F\}$. Nodes D, E, F are *triangle* nodes, nodes in the clusters are *cluster* nodes, and nodes A, A', A'', are *additional* nodes. Edges (i, j) with $i \in \{B, C\}$ and $j \in \{D, E, F\}$ are *flexible* edges. Other edges shown in the figure plus the edges in the clusters are *fixed* edges. The remaining pairs with no edge connected (e.g. (A, D), (A, E), etc.) are referred to as *forbidden* edges.

We use the parameters (n, b, c, w) of a B^{3}C game to realize the gadget. In particular, (a) $n = 5 + 3t + r$; (b) $b(i) = 1$ for all $i \in V$; (c) $c(i, j) = 0$ if (i, j) is a fixed edge, $c(i, j) = 1$ if (i, j) is a flexible edge, $c(i, j) = M > 1$ if (i, j) is a forbidden edge; and (d) $w(i, j) = 1$ for all $i, j \in V$. Note that in the game only the edge costs are nonuniform.

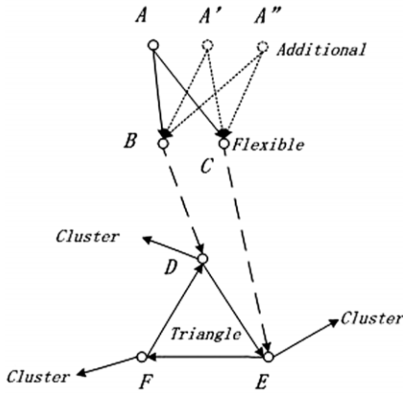

Fig. 1. Main structure of the gadget that has no maximal Nash equilibrium.

With the above construction, we can show the following theorem.

Theorem 3. *The B^3C game based on the gadget of Fig. 1 does not have any maximal Nash equilibrium. This implies that for any $n \geq 6$, there is an instance of B^3C game with n players that does not have any maximal Nash equilibrium.*

Proof. Note that in a maximal graph all fixed edges are included, and nodes B and C each selects one edge to connect to one node in $\{D, E, F\}$. Consider one maximal graph G in which B connects to D and C connects to E (as in Fig. 1). Node B is on all shortest paths from nodes in $\{A, A', A''\}$ to D and the cluster D points to, but it is not on any shortest paths from nodes in $\{A, A', A''\}$ to E and F and the two clusters they point to (these shortest paths all pass through C). Thus $btw_B(G) = r(t+1)$. In this case, B can change its strategy to connect to F instead of D, so that it will be on all shortest paths from those additional nodes to F and D and their clusters, and thus its betweenness is increased to $2r(t+1)$. Therefore, maximal graph G is not stable.

The second case to consider is that both B and C connect to the same node, say E. In this case, they split equally among all shortest paths from the additional nodes to the triangle nodes and the clusters nodes, giving each of them a betweenness $3r(t+1)/2$. In this case, each of them could improve their betweenness to $2r(t+1)$ by connecting to F instead of E. Hence, this maximal graph is not stable either.

All other maximal graphs are rotationally equivalent to one of the above two graphs. Therefore, we know that none of the maximal graphs is stable, and the theorem holds. □

For the ℓ-B^3C game with $\ell \geq 3$, the proof is similar to the B^3C game.

Lemma 4. *The ℓ-B^3C game with $\ell \geq 3$ based on the gadget of Fig. 1 does not have any maximal Nash equilibrium. This implies that for any $n \geq 6$, there is an instance of ℓ-B^3C game with n players that does not have any maximal Nash equilibrium.*

Proof. Consider the cluster connected to node D, we define t_k to be the number of nodes in the cluster with length at most k far away from D. Since three clusters are identical, t_k is the same in all clusters. Obviously, we have $t_k \geq t_{k-1}$.

Consider one maximal graph G in which B connects to D and C connects to E (as in Fig. 1). The betweenness of node B is that $btw_B(G) = r(t_{\ell-2}+1)$. In this case, B can change its strategy to connect to F instead of D, so that its betweenness is increased to $r(t_{\ell-2}+1) + r(t_{\ell-3}+1)$. Therefore, maximal graph G is not stable.

The second case to consider is that both B and C connect to the same node, say E. In this case, they split equally among all shortest paths from the additional nodes to the triangle nodes and the clusters nodes, giving each of them a betweenness $(r(t_{\ell-2}+1) + r(t_{\ell-3}+1) + r(t_{\ell-4}+1))/2$ for $\ell \geq 4$ or $(r(t_{\ell-2}+1) + r(t_{\ell-3}+1))/2$ for $\ell = 3$. In this case, each of them could improve their betweenness to $r(t_{\ell-2}+1) + r(t_{\ell-3}+1)$ by connecting to F instead of E. Hence, this maximal graph is not stable either.

All other maximal graphs are rotationally equivalent to one of the above two graphs. Therefore, we know that none of the maximal graphs is stable, and the theorem holds. □

However, the gadget in Fig. 1 does not work for the case of $\ell = 2$. We now construct a separate gadget for $\ell = 2$ in Fig. 2. The outgoing edges for nodes A, B, C, D and the two edges from X and Y point to each other are fixed as shown in the gadget. Node X can establish at most one edge to a node in $\{A, D\}$, while node Y can establish at most one edge to a node in $\{B, C\}$.

We classify nodes and edges as follows. Nodes X and Y are *flexible* nodes since they can choose to connect one node in $\{A, D\}$ and $\{B, C\}$ respectively. Nodes A, B, C, D are *rectangle* nodes. Edges (X, A), (X, D), (Y, B), (Y, C) are *flexible* edges (in the figure dotted arrows and dashed arrows represent conflicting choices of flexible edges, e.g. (X, A) and (X, D) cannot be selected at the same time). Other edges shown in the figure are *fixed* edges. The remaining pairs with no edge connected (e.g. (X, B), (X, C), etc.) are referred to as *forbidden* edges.

We use the parameters (n, b, c, w) of a 2-B^3C game to realize the gadget. In particular, (a) $n = 6$; (b) $b(i) = 1$ for all $i \in V$; (c) $c(i, j) = 0$ if (i, j) is a fixed edge, $c(i, j) = 1$ if (i, j) is a flexible edge, $c(i, j) = M > 1$ if (i, j) is a forbidden edge; and (d) $w(i, j) = 1$ for all $i, j \in V$.

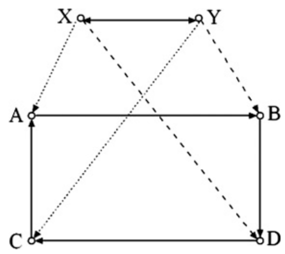

Fig. 2. Main structure of the gadget that has no maximal Nash equilibrium for 2-B^{3}C games, while dotted arrows and dashed arrows represent conflicting choices of flexible edges from a node.

With the above construction, we can show the following theorem.

Lemma 5. *The 2-B^{3}C game based on the gadget in Fig. 2 does not have any maximal Nash equilibrium. This implies that for any $n \geq 6$, there is an instance of ℓ-B^{3}C game with n players that does not have any maximal Nash equilibrium, and in the game only the edge costs are nonuniform.*

Proof. Note that in a maximal graph all fixed edges are included, and nodes X and Y each selects one edge to connect to one node in $\{A, D\}$ and $\{B, C\}$ respectively. We now show that this maximal graph is not stable, by discussing the following cases separately.

(1) Node X connects to A and node Y connects to B. In this case, the only path that can contribute betweenness to node Y is $X \to Y \to B$. But there is another shortest path $X \to A \to B$. So we have $btw_Y(G, 2) = 1/2$. However, if Y changes its strategy to connect to node C, it can gain betweenness 1 from the unique shortest path $X \to Y \to C$. So Y is not at its best response position.
(2) Node X connects to D and node Y connects to B. Here the only path that can contribute betweenness to node X is $Y \to X \to D$. But there is another shortest path $Y \to B \to D$ from Y to D. Thus $btw_X(G, 2) = 1/2$. Now if X changes its strategy to connect to node A, it can gain betweenness 1 from the unique shortest path $Y \to X \to A$. So X is not at its best response position.
(3) Node X connects to A and node Y connects to C. This case is equivalent to case (2), thus is not stable.
(4) Node X connects to D and node Y connects to C. This case is equivalent to case (1), which is also not stable.

In summary, each of X and Y uses the strategy such that its outgoing neighbor points to the outgoing neighbor of the other node, making an endless dynamic in the game.

Therefore, we know that none of the maximal graphs is stable, so the gadget of Fig. 2 does not have any maximal Nash equilibrium.

For $n > 6$, we can use 6 nodes of them to build the above gadget and make all other nodes' outgoing edges forbidden edges. It is easy to see that there is still no maximal Nash equilibrium in this graph, thus the theorem holds. □

Therefore, the following theorem is obtained by Lemmas 4 and 5.

Theorem 6. *For any $\ell \geq 2$ and $n \geq 6$, there is an instance of ℓ-B^{3}C game with n players that does not have any maximal Nash equilibrium.*

3.2. Hardness of determining the existence of maximal Nash equilibria

In this section we use the gadget given in Fig. 1 as a building block to show that determining the existence of maximal Nash equilibria given a nonuniform B^{3}C game is NP-hard. In fact, we use strict Nash equilibria to obtain a stronger result.

We define a problem TwoExtreme as follows. The input of the problem is (n, b, c, w) as the parameter of a B^{3}C game. The output of the problem is Yes or No, such that (a) if the game has a strict Nash equilibrium, the output is Yes; (b) if the game has no maximal Nash equilibrium, the output is No; and (c) for other cases, the output could be either Yes or No. It is easy to see that both deciding the existence of maximal Nash equilibria and deciding the existence of strict Nash equilibria is a stronger problem than TwoExtreme, because their outputs are valid outputs for the TwoExtreme problem by Lemma 2. The following theorem shows that even the weaker problem TwoExtreme is NP-hard.

Theorem 7. *The problem of* TwoExtreme *is NP-hard.*

Proof Outline. We reduce the 3-SAT problem to TwoExtreme. In particular, we provide a polynomial-time transformation from any 3-SAT instance to a B^{3}C game instance as the input of TwoExtreme. Given a 3-SAT instance, we construct the B^{3}C instance such that if a maximal graph G of the game is stable, G must be an 'assignment graph' (Lemma 8), which is a class of graphs in which the edge choices of some particular nodes 'encode' a truth assignment of variables of the 3-SAT instance.

We then show the following two properties:

(a) (Lemma 11) any non-satisfiable 3-SAT instance is transformed into a game that has no maximal Nash equilibrium. We prove this by showing that if the 3-SAT instance is non-satisfiable, then some part of the assignment graph G can be

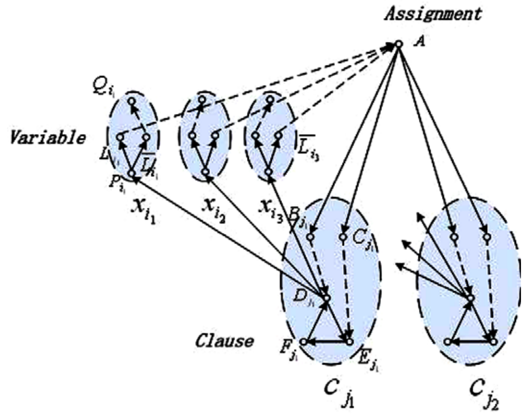

Fig. 3. The structure of the instance of a B^3C game corresponding to an instance of a 3-SAT problem.

reduced to the gadget in Fig. 1 (Lemmas 9 and 10). Thus by an argument similar to the one in the proof of Theorem 3, we can show that it has no maximal Nash equilibrium.

(b) (Lemma 14) any satisfiable 3-SAT instance is transformed into a game that must have a strict Nash equilibrium. Suppose the 3-SAT instance has a satisfying assignment f. We first show that there exists a maximal assignment graph G of the game corresponding to f (Lemma 12). Then based on G we will construct a new graph G' and show that G' is indeed a strict Nash equilibrium(Lemma 13).

The above two properties ensure that we can use the Yes/No answer of the TwoExtreme as the answer to the 3-SAT instance.

Proof. The detail of the transformation is as follows. Each 3-SAT instance has t variables $\{x_1, x_2, \ldots, x_t\}$ and m clauses $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_m\}$. Each variable x has two literals x and $\bar{x}$. Each clause has three literals from three different variables. We use the following construction to obtain an instance of a B^3C game with parameters (n, b, c, w) from the 3-SAT instance, which is illustrated by Fig. 3.

Each clause $\mathcal{C}_j$ is mapped to the core of gadget of Fig. 1, which is the substructure of the gadget excluding the additional nodes and the cluster nodes. We use B_j and C_j to represent the flexible nodes in the gadget and D_j, E_j and F_j to represent the triangle nodes in the gadget, all corresponding to the clause $\mathcal{C}_j$. This leads to $5m$ nodes in the graph. There is a special node A called the *assignment node*, with fixed edges pointing to all flexible nodes B_j and C_j in all gadgets corresponding to all clauses.

Each variable x_i is mapped to a structure with four nodes P_i, Q_i, L_i, and $\bar{L}_i$. Node P_i has two fixed edges pointing to L_i and $\bar{L}_i$. Node L_i and $\bar{L}_i$, called *literal nodes*, each may have one flexible edge pointing to either Q_i or the assignment node A. For each clause $\mathcal{C}_j$ with three variables x_{i_1}, x_{i_2} and x_{i_3}, we add one fixed edge from D_j to each of P_{i_1}, P_{i_2} and P_{i_3} respectively.

In order to realize the above structure, we set the parameters (n, b, c, w) of the B^3C game as follows. First, $n = 1+4t+5m$ and $b(i) = 1$ for all $i \in V$. Next, same as in Fig. 1, each fixed edge has cost 0, each flexible edge has cost 1 (so that the corresponding starting node can choose at most one flexible edge), and each forbidden edge has cost $M > 1$. Finally, the weight function has to be carefully set as follows to make the reduction work. For all $j \in \{1, \ldots, m\}$, $w(A, D_j) = w(A, E_j) = w(A, F_j) = w(D_j, F_j) = w(E_j, D_j) = w(F_j, E_j) = 1$; for all $i \in \{1, \ldots, t\}$, $w(P_i, A) = w(P_i, Q_i) = a$ for some constant $a > 2m$; for all $i \in \{1, \ldots, t\}$ and all $j \in \{1, \ldots, m\}$, (a) if clause $\mathcal{C}_j$ contains variable x_i, then $w(P_i, B_j) = w(P_i, C_j) = w(F_j, L_i) = w(F_j, \bar{L}_i) = 1$; and (b) if literal x_i (or $\bar{x}_i$) is in clause $\mathcal{C}_j$, then $w(L_i, D_j) = d$ (or $w(\bar{L}_i, D_j) = d$), for some constant $d > 1$. For all other pairs (u, v) not included above, $w(u, v) = 0$.

We consider maximal graphs of the game in which all fixed edges are present and exactly one flexible edge from each node in $\{L_i, \bar{L}_i \mid i = 1, 2, \ldots, t\} \cup \{B_j, C_j \mid j = 1, 2, \ldots, m\}$ is present. We say that a maximal graph G of the game is an *assignment graph* if for all $i \in \{1, \ldots, t\}$, there is exactly one edge from $\{L_i, \bar{L}_i\}$ to A in G. The following is a sequence of Lemmas that leads to the proof of the theorem.

Lemma 8. *If a maximal graph G of the game is stable, G must be an assignment graph.*

Proof. Suppose, for a contradiction, that G is not an assignment graph. Then for some $i \in \{1, \ldots, \ell\}$, both L_i and $\bar{L}_i$ connect to Q_i or to A. Suppose they both connect to Q_i. The only shortest paths that pass through L_i and $\bar{L}_i$ and have nonzero weights are $\langle P_i, L_i, Q_i\rangle$ and $\langle P_i, \bar{L}_i, Q_i\rangle$. Since $w(P_i, Q_i) = a$, we have $btw_{L_i}(G) = btw_{\bar{L}_i}(G) = a/2$. In this case, L_i can change its strategy to connect to A instead of Q_i to obtain G'. In G', L_i is on the only shortest path from P_i to A, and thus $btw_{L_i}(G') = a > btw_{L_i}(G)$. Therefore, G is not stable, contradicting to the assumption of the lemma.

Now suppose that both L_i and $\bar{L}_i$ connect to A. They split the shortest paths from P_i to A, which contributes $a/2$ to the betweenness of L_i and $\bar{L}_i$ each. Among other possible shortest paths that pass through L_i or $\bar{L}_i$, the only nonzero weight ones are from P_i to B_j and C_j for all $j \in \{1, \ldots, m\}$. Since L_i and $\bar{L}_i$ equally split these shortest paths, we have

$btw_{L_i}(G) \le a/2 + \sum_{j=1}^{m}(w(P_i, B_j) + w(P_i, C_j))/2 = a/2 + m$. In this case, L_i can change its strategy to connect to Q_i instead of A to obtain G'. In G', L_i is on the only shortest path from P_i to Q_i, so $btw_{L_i}(G') = a > a/2 + m$ since $a > 2m$. Therefore, G is not stable, again contradicting to the assumption of the lemma. Hence, G must be an assignment graph. □

Lemma 9. *If the 3-SAT instance does not have a satisfying assignment, then for any maximal assignment graph G, there always exists a $j \in \{1, \dots, m\}$ such that for all $i \in \{1, \dots, t\}$ and all literals $v \in \{L_i, \bar{L}_i\}$, edge (v, A) being in G implies $w(v, D_j) = 0$.*

Proof. Suppose that the 3-SAT instance does not have a satisfying assignment and G is a maximal assignment graph. The edges pointing to A in G correspond to a truth assignment to variables in the 3-SAT instance: If edge (L_i, A) is in G, assign variable x_i to true; if edge $(\bar{L}_i, A)$ is in G, assign variable x_i to false. Since the 3-SAT instance is not satisfiable, for the above assignment, there exists a clause $\mathcal{C}_j$ that is evaluated to false. For any variable x_i not in $\mathcal{C}_j$ we have $w(L_i, D_j) = w(\bar{L}_i, D_j) = 0$ by our definition of the weight function. So we only consider a variable x_i appearing in $\mathcal{C}_j$. If edge (L_i, A) is in G, we assign x_i to true, and since $\mathcal{C}_j$ is evaluated to false, we know that literal $\bar{x}_i$ is in $\mathcal{C}_j$. Then by our definition, $w(\bar{L}_i, D_j) = b$ but $w(L_i, D_j) = 0$. The case when $(\bar{L}_i, A)$ is in G has a symmetric argument. Therefore, the lemma holds. □

Lemma 10. *For a maximal assignment graph G, if there exists a $j \in \{1, \dots, m\}$ such that for all $i \in \{1, \dots, t\}$ and all literals $v \in \{L_i, \bar{L}_i\}$, edge (v, A) being in G implies $w(v, D_j) = 0$, then G is not a Nash equilibrium.*

Proof. Consider such a graph G with $j \in \{1, \dots, m\}$ satisfying the condition given in the lemma. Consider the shortest paths that pass through B_j and C_j. Since all literal nodes that connect to A have zero weights to D_j (and thus also to E_j and F_j), the only shortest paths passing through B_j and C_j that have nonzero weights are paths from A to D_j, E_j and F_j. This essentially reduces the gadget corresponding to $\mathcal{C}_j$ to the gadget in Fig. 1 with one additional node A and no cluster nodes. By an argument similar to the one in the proof of Theorem 3, no matter how B_j and C_j currently connect to nodes in $\{D_j, E_j, F_j\}$, one of them will always want to change its strategy to connect to one node in $\{D_j, E_j, F_j\}$ that is next to what the other current connects to (according to the direction of the triangle) to increase its utility. Therefore, G is not a Nash equilibrium. □

Lemma 11. *If the 3-SAT instance does not have a satisfying assignment, then the constructed B^3C game instance has no maximal Nash equilibrium.*

Proof. Suppose, for a contradiction, that the B^3C game instance has a maximal Nash equilibrium G. By Lemma 8 G must be an assignment graph. Since the 3-SAT instance does not have a satisfying assignment, by Lemmata 9 and 10 G is not stable, a contradiction. □

Lemma 12. *If the 3-SAT instance has a satisfying assignment, then there exists a maximal assignment graph G of the game in which for all $j \in \{1, \dots, m\}$, there exists $i \in \{1, \dots, t\}$ and literal $v \in \{L_i, \bar{L}_i\}$ such that the edge (v, A) is in G and $w(v, D_j) = d$.*

Proof. Suppose that the 3-SAT instance has a satisfying assignment f. construct a maximal assignment graph G such that for all $i \in \{1, \dots, \ell\}$, if variable x_i is assigned to true in the assignment f, then L_i connects to A; otherwise, $\bar{L}_i$ connects to A. For all $j \in \{1, \dots, m\}$, since clause $\mathcal{C}_j$ is evaluated to true under assignment f, there exists variable x_i whose corresponding literal in $\mathcal{C}_j$ is evaluated to true. If literal x_i is in $\mathcal{C}_j$, x_i is assigned to true. By the above construction of G, (L_i, A) is in G, and by the definition of the weight function, $w(L_i, D_j) = b$. The same argument applies to the case when literal $\bar{x}_i$ is in $\mathcal{C}_j$. Therefore, the lemma holds. □

Lemma 13. *Given a maximal assignment graph G in which for all $j \in \{1, \dots, m\}$, there exists $i \in \{1, \dots, t\}$ and literal $v \in \{L_i, \bar{L}_i\}$ such that the edge (v, A) is in G and $w(v, D_j) = d$, we construct a graph G' such that G' is the same as G except that for all $j \in \{1, \dots, m\}$, both B_j and C_j are connected to D_j in G'. The maximal graph G' must be a strict Nash equilibrium.*

Proof. We prove that in G' any strategy change strictly decreases the changers betweenness, and thus G' must be a nontransient Nash equilibrium.

We go through all nodes and check all possible strategy changes in the following list.

- For each node Q_i, $i \in \{1, \dots, \ell\}$, it has only the empty strategy so there is no strategy change for Q_i.
- For each node P_i, $i \in \{1, \dots, \ell\}$, the only change of the strategy is to remove one or both of the edges (P_i, L_i) and $(P_i, \bar{L}_i)$. Suppose variable x_i appears in clause $\mathcal{C}_j$. Then we know that D_j connects to P_j (since G' is maximal). By the definition of the weight function $w(F_j, L_i) = w(F_j, \bar{L}_i) = 1$. Thus paths from F_j to L_i and $\bar{L}_i$ through P_i contribute positive values to the betweenness of P_i. If P_i were to remove edge (P_i, L_i) or $(P_i, \bar{L}_i)$ or both, P_i's betweenness would strictly decrease.
- For each node L_i, $i \in \{1, \dots, \ell\}$, its strategy change is either removing its flexible edge or changing its flexible edge. If it removes its flexible edge, it loses the shortest path from P_i to Q_i or A, and since $w(P_i, Q_i) = w(P_i, A) = a$, its betweenness strictly decreases. If it changes its flexible edge, then both L_i and $\bar{L}_i$ connects to Q_i or A. By the same argument as in the proof of Lemma 8, its betweenness strictly decreases.
- For each node $\bar{L}_i$, $i \in \{1, \dots, \ell\}$, the argument is the same as the argument for L_i.
- For node A, it can remove any of edges (A, B_j) or (A, C_j), for $j \in \{1, \dots, m\}$. Suppose it removes edge (A, B_j) in G'. Let x_i be a variable in $\mathcal{C}_j$. Since G' is an assignment graph, P_i has a shortest path connecting to B_j through L_i or $\bar{L}_i$ and A. Since $w(P_i, B_j) = 1$, this shortest path contributes 1 to the betweenness of A in G'. If A removes edge (A, B_j) in G', there will be no path from P_i to B_j and A's betweenness will decrease by 1. Therefore, any strategy change of A strictly decrease its betweenness.

- For each node $B_j, j \in \{1, \ldots, m\}$, it can either remove its flexible edge or change its flexible edge. By the assumption of the Lemma, there exists $i \in \{1, \ldots, \ell\}$ and literal node $v \in \{L_i, \bar{L}_i\}$ such that the edge (v, A) is in G and $w(v, D_j) = b$. Suppose that there are t such literal nodes v. By the definition of w, we know that $t \leq 3$. Since B_j at least splits the shortest paths from v and A to D_j, $btw_{B_j}(G') = (tb + 3)/2 \geq (b + 3)/2$. If B_j removes its flexible edge (B_j, D_j), it will not connect to any node and its betweenness will decrease to zero. If B_j changes its flexible edge to (B_j, E_j) to obtain a graph G'', it loses the share on the shortest paths from v and A to D_j but gain the full share on the shortest paths from A to E_j and F_j. Then $btw_{B_j}(G'') = 2 < (b + 3)/2 \leq btw_{B_j}(G')$ since $b > 1$. So B_j's betweenness strictly decreases. If B_j changes its flexible edge to (B_j, F_j), it loses the share on the shortest paths from v and A to D_j and E_j and only gains the full share on the shortest paths from A to F_j, so it is worse than the above case. Therefore, all strategy changes on B_j strictly decreases B_j's betweenness.
- For each node $C_j, j \in \{1, \ldots, m\}$, the argument is the same as the argument for B_j.
- For each node $D_j, j \in \{1, \ldots, m\}$, it can change its strategy by removing its fixed edge to E_j and/or removing some of its fixed edges to some P_i's. If it removes its edge to E_j, it loses the shortest path from F_j to E_j with weight 1, so its betweenness strictly decreases. If it removes any edge to some node P_i, it loses shortest paths from F_j to L_i and $\bar{L}_i$ with weight 1, so its betweenness strictly decreases. Therefore, D_i cannot change its strategy.
- For each node $E_j, j \in \{1, \ldots, m\}$, it can change its strategy by removing its fixed edge to F_j. This however will cause E_j losing the shortest path from D_j to F_j with weight 1, so its betweenness strictly decreases.
- For each node $F_j, j \in \{1, \ldots, m\}$, it can change its strategy by removing its fixed edge to D_j. This however will cause F_j losing the shortest path from E_j to D_j with weight 1, so its betweenness strictly decreases.

By the above argument exhausting all possible cases, we show that graph G' is indeed a nontransient Nash equilibrium. □

Lemma 14. *If the 3-SAT instance has a satisfying assignment, then the constructed B^3C game instance has a strict Nash equilibrium.*

Proof. This is immediate from Lemmata 12 and 13. □

The entire proof for Theorem 7 is now complete with Lemmata 11 and 14. □

The immediate consequence of the above theorem is:

Corollary 15. *Both deciding the existence of maximal Nash equilibria and deciding the existence of strict Nash equilibria of a B^3C game are NP-hard.*[3]

We now address the NP-hardness for the ℓ-B^3C game. By a close inspection of the proof above, we see that all critical paths that matter are of length at most 3. Therefore, we know that for $\ell \geq 3$, both deciding the existence of maximal Nash equilibria and deciding the existence of strict Nash equilibria for an ℓ-B^3C game are also NP-hard.

For the case of $\ell = 2$, we rely on the gadget we developed for $\ell = 2$ that we mentioned in Section 3.1 to show that the decision problem is still NP-hard

Theorem 16. *For $\ell = 2$, both deciding the existence of maximal Nash equilibria and deciding the existence of strict Nash equilibria in an ℓ-B^3C game are NP-hard.*

Proof. We reduce the problem from the 3-SAT problem. Each 3-SAT instance has k variables $\{x_1, x_2, \ldots, x_k\}$ and m clauses $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_m\}$. Each variable x has two literals x and $\bar{x}$. Each clause has three literals from three different variables. We use the following construction to obtain an instance of a 2-B^3C game with parameters (n, b, c, w) from the 3-SAT instance, which is illustrated by Fig. 4.

The overall idea of the reduction is as follows. First, each clause $\mathcal{C}_j$ is mapped to the gadget similar to the gadget in Fig. 1 while each literal x_i and $\bar{x}_i$ are mapped to the gadget containing nodes $L_i, \bar{L}_i, P_i, Q_i$. We call nodes L_i's and $\bar{L}_i$'s *literal nodes*. Nodes L_i and $\bar{L}_i$ can either point to node Q_i or all of the nodes X_j. We make sure that those literal nodes pointing to nodes X_j's correspond to an assignment. Next, if the 3-SAT instance has a satisfying assignment, we show that for each clause $\mathcal{C}_j$, there exist shortest paths from some literal nodes to A_j with significant weights. We show that these paths make the gadget for clause $\mathcal{C}_j$ stable. Thus all gadgets are stable and the configuration is a maximal Nash equilibrium. We further argue that it is a strict Nash equilibrium by examining all other alternatives of all nodes and showing that they strictly decrease nodes' betweenness. Finally, if the 3-SAT instance has no satisfying assignment, there must exist at least one clause $\mathcal{C}_j$ such that there is no path from the literal nodes to A_j with nonzero weights. When this is the case, the gadget corresponding to $\mathcal{C}_j$ will not be stable and thus the game has no Nash equilibrium.

All of the solid arrows in the graph are called *fixed edges*. They are $\{(P_i, L_i), (P_i, \bar{L}_i), (X_j, Y_j), (Y_j, X_j), (A_j, B_j), (B_j, D_j), (D_j, C_j), (C_j, A_j), (X_j, P_i), (D_j, Y_j) \mid \forall 1 \leq i \leq k, 1 \leq j \leq m\}$. All of the dashed arrows and dotted arrows represent conflicting

[3] In fact, the decision problem for any intermediate concept between maximal Nash equilibrium and strict Nash equilibrium is also NP-hard. For example, deciding the existence of nontransient Nash equilibria [10] is also NP-hard because any strict Nash equilibrium is a nontransient Nash equilibrium while the existence of a nontransient Nash equilibrium implies the existence of a maximal Nash equilibrium in B^3C games.

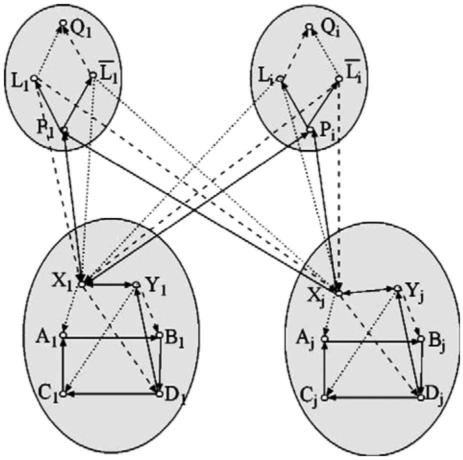

Fig. 4. The structure of the instance of a 2-B^{3}C game corresponding to an instance of a 3-SAT problem. Solid arrows represent fixed edges, while dotted arrows and dashed arrows represent conflicting choices of flexible edges from a node.

choices of *flexible edges* starting from one node (e.g. edge (L_1, Q_1) cannot be selected together with any edge (L_1, X_j)). They are $\{(L_i, Q_i), (\bar{L}_i, Q_i), (L_i, X_j), (\bar{L}_i, X_j), (X_j, A_j), (X_j, D_j), (Y_j, B_j), (Y_j, C_j) \mid \forall 1 \le i \le k, 1 \le j \le m\}$.

We set the parameters (n, b, c, w) of the ℓ-B^{3}C game as follows. First, $n = 4k + 6m$. The budgets of all nodes are 0 except $b(L_i) = b(\bar{L}_i) = m$ and $b(X_j) = b(Y_j) = 1$. The costs of all fixed edges are 0. The costs of all flexible edges are 1 except $c(L_i, Q_i) = c(\bar{L}_i, Q_i) = m$. The costs of all other edges (which is forbidden edges) are larger than m. Finally, the weight function has to be carefully set as follows to make the reduction work. For all $1 \le i \le k$, $1 \le j \le m$, $w(X_j, L_i) = w(X_j, \bar{L}_i) = w(Y_j, P_i) = w(L_i, Y_j) = w(\bar{L}_i, Y_j) = 1$,; for all $1 \le i \le k$, $1 \le j \le m$, $w(P_i, Q_i) = ma$, $w(P_i, X_j) = w(P_i, Y_j) = a$ for some constant a; for all $1 \le j \le m$, $w(X_j, B_j) = w(X_j, C_j) = w(Y_j, A_j) = w(Y_j, D_j) = w(C_j, B_j) = w(B_j, C_j) = w(A_j, D_j) = w(D_j, A_j) = w(B_j, Y_j) = w(D_j, X_j) = 1$; for all $i \in \{1, \ldots, k\}$ and all $j \in \{1, \ldots, m\}$, if literal x_i (or $\bar{x}_i$) is in clause $\mathcal{C}_j$, then $w(L_i, A_j) = b$ (or $w(\bar{L}_i, A_j) = b$), for some constant $b > 1$. For all other pairs (u, v) not included above, $w(u, v) = 0$.

We consider maximal graphs of the game in which all nodes exhaust their budget. Then, for all nodes L_i and $\bar{L}_i$, they point to Q_i or the nodes X_j for all $1 \le j \le m$ in G. We call the second case *pointing to the clause nodes*. We say that a maximal graph G of the game is an *assignment graph* if for all $1 \le i \le k$, there is exactly one node from $\{L_i, \bar{L}_i\}$ pointing to Q_i in G. Thus, the other node points to the clause nodes.

Lemma 17. *If a maximal graph G of the game is stable, G must be an assignment graph.*

Proof. Suppose, for a contradiction, that G is not an assignment graph. Then for some $i \in \{1, \ldots, k\}$, both L_i and $\bar{L}_i$ connect to Q_i or to X_j. Suppose they both connect to Q_i. The only shortest paths that pass through L_i and $\bar{L}_i$ and have nonzero weights are $\langle P_i, L_i, Q_i\rangle$ and $\langle P_i, \bar{L}_i, Q_i\rangle$. Since $w(P_i, Q_i) = ma$, we have $btw_{L_i}(G) = btw_{\bar{L}_i}(G) = ma/2$. In this case, L_i can change its strategy to connect to the clause nodes instead of Q_i to obtain G'. In G', L_i is on the only shortest path from P_i to X_j, and thus $btw_{L_i}(G') = m \times a > btw_{L_i}(G)$. Therefore, G is not stable, contradicting to the assumption of the lemma.

Now suppose that both L_i and $\bar{L}_i$ connect to the clause nodes. They split the shortest paths from P_i to X_j, which contributes $ma/2$ to the betweenness of L_i and $\bar{L}_i$ each. By the same reason, L_i can change its strategy to connect to Q_i instead of X_j to obtain betweenness value ma. Therefore, G is not stable, again contradicting to the assumption of the lemma. Hence, G must be an assignment graph. □

Lemma 18. *If the 3-SAT instance does not have a satisfying assignment, then for any maximal assignment graph G, there always exists a $j \in \{1, \ldots, m\}$ such that for all $i \in \{1, \ldots, k\}$ and all literals $v \in \{L_i, \bar{L}_i\}$, edge (v, X_j) being in G implies $w(v, A_j) = 0$.*

Proof. Suppose that the 3-SAT instance does not have a satisfying assignment and G is a maximal assignment graph. The edges pointing to the clause nodes in G correspond to a truth assignment to variables in the 3-SAT instance: If the node L_i points to the clause nodes in G, assign variable x_i to be true; otherwise, assign variable x_i to be false. Since the 3-SAT instance is not satisfiable, for the above assignment, there exists a clause $\mathcal{C}_j$ that is evaluated to false. For any variable x_i not in $\mathcal{C}_j$ we have $w(L_i, A_j) = w(\bar{L}_i, A_j) = 0$ by our definition of the weight function. So we only consider a variable x_i appearing in $\mathcal{C}_j$. If the node L_i points to the clause nodes in G, we assign x_i to true, and since $\mathcal{C}_j$ is evaluated to false, we know that literal $\bar{x}_i$ is in $\mathcal{C}_j$. Then by our definition, $w(\bar{L}_i, A_j) = b$ but $w(L_i, A_j) = 0$. The case when $\bar{L}_i$ points to the clause nodes in G has a symmetric argument. Therefore, the lemma holds. □

Lemma 19. *For a maximal assignment graph G, if there exists a $j \in \{1, \ldots, m\}$ such that for all $i \in \{1, \ldots, k\}$ and all literals $v \in \{L_i, \bar{L}_i\}$, node v pointing to the clause nodes in G implies $w(v, A_j) = 0$, then G is not a Nash equilibrium.*

Proof. Consider such a graph G with $j \in \{1, \ldots, m\}$ satisfying the condition given in the lemma. Consider the shortest paths that pass through X_j and Y_j. Since all literal nodes that connect to the clause nodes have zero weights to A_j, the only shortest paths passing through X_j and Y_j that have nonzero weights are paths from X_j to B_j, C_j, from Y_j to A_j, D_j, from L_i, $\bar{L}_i$ to Y_j and from D_j to X_j. The betweenness of pairs from L_i, $\bar{L}_i$ to Y_j and from D_j to X_j are only affected by whether X_j points to Y_j and vice verse. Since these two edges are cost 0, they are always connected in a stable graph. For other pairs, it essentially reduces the gadget corresponding to $\mathcal{C}_j$ to the gadget in Fig. 1. The only difference is that here we have an additional edge (D_j, Y_j) compare to Fig. 1. But the additional edge does not have any infection to the betweenness value of node X_j and node Y_j. It only helps to make the graph a strict Nash equilibrium when needed. We will explain this later in Lemma 22. Therefore, by an argument similar to the one in the proof of Theorem 3, no matter how X_j and Y_j currently connect to nodes in $\{A_j, B_j, C_j, D_j\}$, one of them will always want to change its strategy to increase its utility. Therefore, G is not a Nash equilibrium. □

Lemma 20. *If the 3-SAT instance does not have a satisfying assignment, then the constructed 2-B^{3}C game instance does not have maximal Nash equilibrium.*

Proof. Suppose, for a contradiction, that the 2-B^{3}C game instance has a maximal Nash equilibrium. Then there exists a maximal graph G that is stable. By Lemma 17, G must be an assignment graph. Since the 3-SAT instance does not have a satisfying assignment, by Lemmata 18 and 19, G is not stable, a contradiction. □

Lemma 21. *If the 3-SAT instance has a satisfying assignment, then there exists a maximal assignment graph G of the game in which for all $j \in \{1, \ldots, m\}$, there exists $i \in \{1, \ldots, k\}$ and literal $v \in \{L_i, \bar{L}_i\}$ such that the node v points to the clause nodes in G and $w(v, A_j) = b$.*

Proof. Suppose that the 3-SAT instance has a satisfying assignment f. construct a maximal assignment graph G such that for all $i \in \{1, \ldots, k\}$, if variable x_i is assigned to true in the assignment f, then L_i connects to the clause nodes; otherwise, $\bar{L}_i$ connects to the clause nodes. For all $j \in \{1, \ldots, m\}$, since clause $\mathcal{C}_j$ is evaluated to true under assignment f, there exists variable x_i whose corresponding literal in $\mathcal{C}_j$ is evaluated to true. If literal x_i is in $\mathcal{C}_j$, x_i is assigned to true. By the above construction of G, L_i points to the clause nodes in G, and by the definition of the weight function, $w(L_i, A_j) = b$. The same argument applies to the case when literal $\bar{x}_i$ is in $\mathcal{C}_j$. Therefore, the lemma holds. □

Lemma 22. *Given a maximal assignment graph G in which for all $j \in \{1, \ldots, m\}$, there exists $i \in \{1, \ldots, k\}$ and literal $v \in \{L_i, \bar{L}_i\}$ such that the node v points to the clause nodes in G and $w(v, A_j) = b$, we construct a graph G' such that G' is the same as G except that for all $j \in \{1, \ldots, m\}$, X_j connects to A_j and Y_j are connected to C_j in G'. The maximal graph G' must be a strict Nash equilibrium.*

Proof. We prove that in G' any strategy change strictly decreases the changers betweenness, and thus G' must be a strict Nash equilibrium.

We go through all nodes and check all possible strategy changes in the following list.

- For each node Q_i, $i \in \{1, \ldots, k\}$, it has only the empty strategy so there is no strategy change for Q_i.
- For nodes other than $L_i, \bar{L}_i, X_j, Y_j$ $(1 \le i \le k, 1 \le j \le m)$, they only have fixed edge to choose, so we only need to prove that for each fixed edge, there exists a pair with nonzero weight such that if the node removes this fixed edge, the betweenness value will decrease. We call this pair *pushes* such fixed edge.

  For node P_i, pair (X_j, L_i) pushes edge (P_i, L_i) while pair $(X_j, \bar{L}_i)$ pushes edge $(P_i, \bar{L}_i)$.

  For node A_j, pair (C_j, B_j) pushes edge (A_j, B_j). For node B_j, pair (A_j, D_j) pushes edge (B_j, D_j). For node C_j, pair (D_j, A_j) pushes edge (C_j, A_j). For node D_j, pair (B_j, C_j) pushes edge (D_j, C_j) while pair (B_j, Y_j) pushes edge (D_j, Y_j).
- For each node L_i, $i \in \{1, \ldots, k\}$, its strategy change is either removing its flexible edge or changing its flexible edge. If it removes its flexible edge, it loses the shortest path from P_i to Q_i or X_j, and since $w(P_i, Q_i) = a$ and $w(P_i, X_j) = a/m$, its betweenness strictly decreases. If it changes its flexible edge, then both L_i and $\bar{L}_i$ connects to Q_i or X_j. By the same argument as in the proof of Lemma 17, its betweenness strictly decreases. For each node $\bar{L}_i$, $i \in \{1, \ldots, k\}$, the argument is the same as the argument for L_i.
- For each node X_j, $j \in \{1, \ldots, m\}$, it can remove its fixed edge or remove its flexible edge or change its flexible edge. For the fixed edge, pair (Y_j, P_i) pushes edge (X_j, P_i) and pair (L_i, Y_j) or $(\bar{L}_i, Y_j))$ pushes edge (X_j, Y_j). Then, we only consider the betweenness value caused by the flexible edge. By the assumption of the Lemma, there exists $i \in \{1, \ldots, k\}$ and literal node $v \in \{L_i, \bar{L}_i\}$ such that the node v points to the clause nodes G and $w(v, A_j) = b$. Suppose that there are t such literal nodes v. By the definition of w, we know that $t \le 3$. Since X_j splits the shortest paths from v to A_j and Y_j to A_j $btw_{X_j}(G', 2) = tb + 1/2 \ge b + 1/2$. If X_j removes its flexible edge (X_j, A_j), it will not connect to any node and its betweenness will decrease to zero. If X_j changes its flexible edge to (X_j, D_j) to obtain a graph G'', it does not connect nodes v and A_j but gain the full share on the shortest paths from Y_j to D_j. Then $btw_{X_j}(G'', 2) = 1 < b + 1/2 \le btw_{X_j}(G', 2)$ since $b > 1$. So X_j's betweenness strictly decreases. Therefore, all strategy changes on X_j strictly decreases X_j's betweenness.
- For each node Y_j, $j \in \{1, \ldots, m\}$, it can remove its fixed edge or remove its flexible edge or change its flexible edge. For the fixed edge, pair (D_j, X_j) pushes edge (Y_j, X_j). For the flexible edge, by the same argument in Theorem 3, all strategy changes on Y_j strictly decreases Y_j's betweenness.

By the above argument exhausting all possible cases, we show that graph G' is indeed a strict Nash equilibrium. □

Lemma 23. *If the 3-SAT instance has a satisfying assignment, then the constructed 2-B^{3}C game instance has a strict Nash equilibrium.*

Proof. This is immediate from Lemmata 13 and 21. □

The entire proof of Theorem 16 is now complete with Lemmata 20 and 23. □

Note that in the above proof, edge costs and weights are nonuniform while node budgets are uniform. And the following theorem summarizes the hardness for the ℓ-B^{3}C game.

Theorem 24. *For any $\ell \geq 2$, both deciding the existence of maximal Nash equilibria and deciding the existence of strict Nash equilibria in an ℓ-B^{3}C game are NP-hard.*

4. Complexity of computing best responses

The *best response* of a node in a configuration of the uniform game is the strategy of the node that gives the node the best utility (i.e. best betweenness). In this section, we show the complexity of computing best responses first for uniform games and then extend it for nonuniform games.

In a uniform game with parameters (n, k), one can exhaustively search all $\binom{n-1}{k}$ strategies and find the one with the largest betweenness. Computing the betweenness of nodes given a fixed graph can be done by all-pair shortest paths algorithms in polynomial time (e.g. [11]). Therefore, the entire brute-force computation takes polynomial time if k is a constant. However, if k is not a constant, the result depends on ℓ, the parameter bounding the shortest path length in the ℓ-B^{3}C game.

For $\ell = 2$, we show that there exists a polynomial-time algorithm to compute a best response in a uniform ℓ-B^{3}C game. To reach this result, we first need the following lemma.

Lemma 25. *Let $G = (V, E)$ be a directed graph. For a node v in G, let $G_{v,S}$ be the graph where v has outgoing edges to nodes in $S \subseteq V \setminus \{v\}$ and all other nodes have the same outgoing edges as in G. Then we have for all $S \subseteq V \setminus \{v\}$, $btw_v(G_{v,S}, 2) = \sum_{u\in S} btw_v(G_{v,\{u\}}, 2)$.*

Proof. Let $S = \{v_1, \ldots, v_k\}$. Consider any shortest path of length 2 from a node u to a node u' that passes through node v. The path must be $u \to v \to u'$, which means $u' = v_i$ for some i. So $btw_v(G_{v,S}, 2)$ can be written as

$$btw_v(G_{v,S}, 2) = \sum_{v_i\in S}\left(\sum_{u\neq v\neq v_i, m(u,v_i,2)>0} \frac{m_v(u, v_i, 2)}{m(u, v_i, 2)}\right).$$

Suppose now we change S to a single vertex set $\{v_i\}$ for some i. The value of $m_v(u, v_i)$ and $m(u, v_i)$ will not change because none of these paths goes through any other edges that start from v. On the other hand, $m_v(u, v_j)$ will become 0 if $j \neq i$. So we have

$$btw_v(G_{v,\{v_i\}}, 2) = \sum_{u\neq v\neq v_i, m(u,v_i,2)>0} \frac{m_v(u, v_i, 2)}{m(u, v_i, 2)}.$$

Compare the formulas of $btw_v(G_{v,S}, 2)$ and $btw_v(G_{v,\{v_i\}}, 2)$. We know that

$$btw_v(G_{v,S}, 2) = \sum_{1\leq i\leq k} btw_v(G_{v,\{v_i\}}, 2)$$

Therefore, the lemma holds. □

The lemma shows that for 2-B^{3}C game, the betweenness of a node can be computed by a simple sum of the its betweenness when adding each of its outgoing edges alone into the graph.

Theorem 26. *Computing a best response in a uniform ℓ-B^{3}C game when $\ell = 2$ can be done in $O(n^3)$ time.*

Proof. Consider a graph G with n nodes and k outgoing edges for each node. For any node v in G, let $btw_v(u, 2)$ be the betweenness value of node v if v chooses $\{u\}$ as its strategy. We can compute $btw_v(u, 2)$ using the following method: for each node w where $(w, v) \in G$, if $(w, u) \in G$, then node v will not get any betweenness value from the path from w to u. If $(w, u) \notin G$, let $m(w, u, 2)$ be the number of length-two paths (which are the shortest paths) from w to u. Notice that $m(w, u, 2)$ can be computed in $O(n)$ time by enumerating the intermediate node of the path. Node v gains $\frac{1}{m(w,u,2)}$ betweenness value from these paths. Adding such values for all node w where $(w, v) \in G$ together, we can get $btw_v(u, 2)$ in $O(n^2)$ time.

Then we can compute $btw_v(u, 2)$ for all nodes $u \neq v$ in $O(n^3)$ time, and by Lemma 25, the top k nodes with the largest $btw_v(u, 2)$ values will form the best response for node v. The sorting and selecting only cost $O(n \log n)$ time. Thus the whole algorithm can be done in time $O(n^3)$. □

For $\ell \geq 3$, we show that the task of computing a best response in a uniform ℓ-B^{3}C game is NP-hard. This also implies that the task is NP-hard in the B^{3}C game without path length constraint. To show the result, we define its decision problem version below.

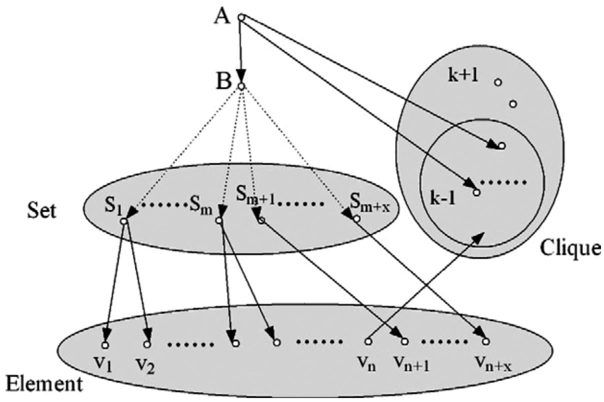

Fig. 5. Structure corresponding to a set cover instance.

For $\ell \in \mathbb{N}$, we define a decision problem ℓ-BestResponse as follows. The input of the problem includes (a) a directed graph $G = (V, E)$ with n nodes and each node has k outgoing edges; (b) a natural number k, (c) one node v in G, and (d) a natural number b. The output is Yes or No. Let S_v be a strategy of v (i.e., $S_v \subseteq V \setminus \{v\}$ and $|S_v| = k$). Let G_{v,S_v} be the graph where v uses strategy S_v and all other nodes have same outgoing edges as in G. The output of the problem is Yes if and only if there exists a strategy S_v of node v such that $btw_v(G_{v,S_v}, \ell) \geq b$.

Theorem 27. *For all $\ell \geq 3$, problem ℓ-BestResponse is NP-hard.*

Proof. We reduce this problem from the set cover problem. Given an instance of the set cover problem $\langle U, S, t\rangle$, in which U is a universe and S is a family of subsets of U with $|U| = n$, $|S| = m$, and t is a natural number. The problem is to determine whether there are at most t subsets in S whose union is the universe. We construct an instance of the betweenness problem as follows (see Fig. 5).

- Let r be the maximum size of subset in T, i.e., $r = \max\{|s| \mid s \in S\}$.
- Let $t' = \min(t, m), x = \max(r - t', 0), k = t' + x$.
- We use $k + 1$ nodes to form a clique so that each node has out degree k. These nodes are used to absorb links from other nodes that would otherwise do not have k outgoing edges.
- We set node B to be the one we need to compute the best response for.
- We set node A to connect to B and another $k - 1$ nodes in the clique;
- We set n element nodes $v_1, \ldots, v_n$ to correspond to n elements in U, and they connect to arbitrary k nodes in the clique;
- We add x new elements to U to form a new universe U'. Then set x new elements nodes $v_{n+1}, v_{n+2}, \ldots, v_{n+x}$ to correspond to them. Connect these new nodes to arbitrary k nodes in the clique;
- We add x new subsets to S to form a new family of subsets S', where the ith subset contains only one element v_{n+i} ($1 \leq i \leq x$). Then set $m + x$ subset nodes $s_1, \ldots, s_{m+x}$ to correspond to $m + x$ subsets in S' (here $s_{m+1}, s_{m+2}, \ldots, s_{m+x}$ are set to correspond to the new added subsets). For a slight abuse of notation, we use v_i to denote both the node in the graph and the element in U', and s_j to denote both the node in the graph and the subset in S'. We connect s_j to all node v_i if $v_i \in s_j$, because $|s_j| \leq r \leq k$, we can always make such connections. For subsets have less than k nodes, we connect them arbitrarily to nodes in the clique to increase their out-degree to k.

The decision problem in the game is to determine whether node B can choose a set of edges of size at most k that make its betweenness at least $n + x + k$.

Lemma 28. *If there is a cover of size at most t whose union is the universe U, then node B can choose a set of edges of size at most k that makes its betweenness to be at least $n + x + k$.*

Proof. Suppose that the cover which satisfies the requirement is C. Without loss of generality, we can assume that $|C| = t'$. Let node B connects to the subset nodes s_i for all $s_i \in C$ and all the new added subset nodes s_{m+j}, where $1 \leq j \leq x$. In this case, B stands on the shortest paths from A to the k subset nodes, and thus gains betweenness k from these shortest paths. Since $\cup_{s_i \in C} s_i = U$ and $\cup_{1 \leq i \leq x} s_{m+i} = U' \backslash U$, according to the construction of the structure, B can reach all $n+x$ element nodes $\{v_1, \ldots, v_{n+x}\}$ and B stands on all the paths from A to the elements nodes. Hence they contribute $n + x$ to the betweenness of B. So betweenness of B is at least $n + x + k$. This concludes the proof. □

Lemma 29. *If node B can find a set of edges of size at most k that makes its betweenness to be at least $n + x + k$, then there is a cover of size at most t whose union is the universe U.*

Proof. We first prove that B can achieve the best betweenness by connecting to k subset nodes s_j's.

Node B's betweenness comes from the shortest paths from A to other nodes. If B connects to a node L not in the clique, it will not gain any betweenness from the paths from A to B to L and then to any clique node, because A can reach $k - 1$ clique nodes directly and the remaining two clique nodes in one more step, but the paths through B and L have length at least 3.

Since $k \le m + x$, if B connects to any nodes other than the subset nodes, it will lose a connection to some subset node s_j. We argue case by case below that it will not give a better betweenness than B connecting to some subset node instead.

- Node B connects to some element node v_i. It can gain at most 1 by the shortest path from A to v_i via B, since v_i only connects to clique nodes, and we have already argued that the path from A to B to v_i and then to clique nodes are not shortest paths. In this case, B can instead connect to an available subset node s_j not yet connected, by which it gains betweenness of at least 1, no worse than the connection to v_i.
- Node B connects to some clique node L. If A has a direct connection to L, B will not gain any betweenness by this connection. If A does not have direction connection to L, $\langle A, B, L\rangle$ is a shortest path of length 2, but there are $k - 1$ other shortest paths from A to L. Thus B gains betweenness of at most $1/k$. In this case, B is better off connecting to an available subset node s_j.
- Node B connects to node A. This does not contribute any betweenness to B, so B is better off connecting to an available subset node s_j.

Therefore, node B can achieve the best betweenness by connecting to k subset nodes. Let these k subset nodes form a set C'. In this case, the betweenness of B is $k + |\cup_{s_i \in C'} s_i|$, because only through B node A can reach all k subset nodes in C' plus nodes in $\cup_{s_i \in C'} s_i$, but for the clique nodes A has shorter paths to reach them not through B. Since B can achieve a betweenness of at least $n + x + k$, we know that $|\cup_{s_i \in C'} s_i| \ge n + x$, which means that C' must cover all the element nodes. Also notice that s_{m+i} is the only subset that contains element v_{m+i}. So C' must all the new added subset nodes s_{m+i} $(1 \le i \le x)$. Then let $C = C' \backslash \{s_{m+1}, \ldots, s_{m+x}\}$. We know C must have $k - x = t'$ elements and can cover $\{v_1, \ldots, v_m\}$. Thus C is a solution to the set cover instance. □

The proof of Theorem 27 is now complete with Lemmata 28 and 29. □

Remark 1. Although it is well known that set cover problem cannot be approximated in polynomial time to within a factor of $c \cdot \ln n$, where c is a constant. We cannot directly apply this inapproximability result to the ℓ-BestResponse problem using a similar proof. Whether problem ℓ-BestResponse can be approximated efficiently is still left open.

Theorem 27 can be directly applied to both B^3C games without path length constraint and the nonuniform ℓ-B^3C games for $\ell \ge 3$. For nonuniform ℓ-B^3C games with $\ell = 2$, however, our polynomial-time algorithm does not work any more. In the non-uniform version, we define a decision problem 2-NBestResponse as follows. The input of the problem contains (a) a 2-B^3C game with parameter (n, b, c, w); (b) a configuration s of the game; (c) one node v in graph G; and (d) a natural number A. The output is Yes or No. Let S_v be a strategy of v and G_{v,S_v} be the graph that node v uses strategy S_v and all other nodes use the same strategies in configuration s. The output of the problem is Yes if and only if there exists a strategy S_v of node v such that $btw_v(G_{v,S_v}, 2) \ge A$. In Lemma 30 we prove the problem is reducible from the knapsack problem.

Lemma 30. *Problem* 2-NBestResponse *is NP-hard.*

Proof. We reduce this problem from the knapsack problem. Given an instance of the knapsack problem $\langle U, \widetilde{w}, value\rangle$, in which set U contains m items. Each item $U_i = (w_i, value_i)$ has its weight w_i and its value $value_i$. The problem is to determine whether we can pick items from the set U such that the total weight does not exceed w but the total value is at least $value$.

We construct an instance of the betweenness problem as follows. There are $m + 2$ nodes $u, v, v_1, \ldots, v_m$ in the graph. The edge (u, v) is the fixed edge, and the edge (v, v_i) for $i = 1, \ldots, m$ are flexible edges. Other edges in the graph are forbidden edges. We use the parameters (n, b, c, w) of 2-B^3C game as follows. In particular, (a) $n = m + 2$; (b) $b(v) = \widetilde{w}$, $b(u) = b(v_i) = 0 (i = 1, \ldots, m)$; (c) $c(v, v_i) = w_i (i = 1, \ldots, m)$, $c(u, v) = 0$ and $c(i, j) = M > \widetilde{w}$ for all other edges; and (d)$w(u, v_i) = value_i (i = 1, \ldots, m)$. The knapsack instance has a solution exceeding $value$ if and only if the 2-B^3C instance has a configuration such that the betweenness of node v exceeds $value$. □

Therefore, for nonuniform ℓ-B^3C games, computing a best response is NP-hard even for $\ell = 2$. Combined these results, we summarize as follows.

Theorem 31. *It is NP-hard to compute the best response in either a nonuniform* 2-*B^3C game, or an ℓ-B^3C game with $\ell \ge 3$ (uniform or not), or a B^3C game without path length constraint (uniform or not).*

5. Nash equilibria in uniform games

In this section we focus on uniform ℓ-B^3C games. we first define a family of graph structures called *shift graphs* and show that they are able to produce Nash equilibria for B^3C games. We then study some properties of Nash equilibria in uniform games.

5.1. Construction of Nash equilibria via shift graphs

We first define *shift graphs* and *non-rotational shift graphs*. Then we show that for any ℓ, k and any $\ell' \ge \ell$, the *non-rotational shift graphs* with $n = (\ell' + k)!/k!$ nodes are all Nash equilibria in the uniform ℓ-B^3C game with parameter n and k. Moreover, we use shift graphs to construct *strict* Nash equilibria for both ℓ-B^3C games and B^3C games without path length constraint, for certain combinations of n and k where $k = \Theta(\sqrt{n})$.

Definition 4. A *shift graph* $G = (V, E)$ with parameters $m, t \in \mathbb{N}_+$ and $t \ge m$, denoted as $SG(m, t)$, is defined as follows. Each vertex of G is labeled by an m-dimensional vector such that each dimension has t symbols and no two dimensions have the same symbol appeared in the label. That is, $V = \{(x_1, x_2, \ldots, x_m) \mid x_i \in [t]$ for all $i \in [m]$, and $x_i \neq x_j$ for all $i, j \in [m], i \neq j\}$. A vertex u has a directed edge pointing to a vertex v if we can obtain v's label by shifting u's label to the left by one digit and appending the last digit on the right. That is, $E = \{(u, v) \mid u, v \in V, u[2 : m] = v[1 : (m-1)]\}$, where $u[i : j]$ denote the sub-vector $(x_i, x_{i+1}, \ldots, x_j)$ with $u = (x_1, x_2, \ldots, x_m)$.

In the shift graph $SG(m, t)$, we know that the number of vertices is $n = t \cdot (t-1) \cdots (t-m+1) = t!/(t-m)!$, and each vertex has out-degree $t - m + 1$. Notice that the definition requires that m dimensions have all different symbols. If they are allowed to be the same, then the graphs are the well-known De Bruijn graphs, whereas if we require only that the two adjacent dimensions have different symbols, the graphs are Kautz graphs, which are iterative line graphs of complete graphs.

Definition 5. A *non-rotational shift graph* with parameter $m, t \in \mathbb{N}+$ and $t \geq m + 1$, denoted as $SG_{nr}(m, t)$, is a shift graph with the further constraint that if (u, v) is an edge, then v's label is not a rotation of u's label to the left by one digit. That is, $E = \{(u, v) \mid u, v \in V, u[2 : m] = v[1 : (m-1)]$ and $u[1] \neq v[m]\}$, where $u[i]$ denotes the i-th element of u.

Graph $SG_{nr}(m, t)$ also has $t!/(t-m)!$ vertices but the out-degree of every vertex is $t - m$. A simple non-rotational shift graph $SG_{nr}(2, 4)$ is given in Fig. 6 as an example. Non-rotational shift graphs have the following basic properties.

Proposition 1. *Non-rotational shift graph $SG_{nr}(m, t)$ satisfies the following properties:*

(1) *It is Eulerian, i.e., every vertex has the same in-degree $t - m$.*

(2) *It is vertex-transitive.*

(3) *When $t \geq m + 2$, it is strongly connected, with diameter at most $2m(m + 1)$.*

(4) *For $m \geq 2$, it is the line graph of $SG_{nr}(m - 1, t)$ with all edges on the smallest circles of the line graph removed; for $m = 1$, it is simply t-clique (completely connected t-vertex directed graph with no self-loop).*

Proof. (1) and (2) are straightforward by definition.

(3): We first prove the following claim.

Claim 1. *For any node $v = (x_1, \ldots, x_m)$, there exist a length $m + 1$ path from v to node $u = (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_m)$ for every $1 \leq i \leq m$ and $y \neq x_j$, $1 \leq j \leq m, j \neq i$.*

Proof. Notice that $t \geq m + 2$, so there must exist a symbol t such that $t \neq x_i (1 \leq i \leq m)$ and $t \neq y$. Then we can construct the following path:

$$
\begin{aligned}
v &= (x_1, x_2, \ldots, x_m) \\
&\to (x_2, x_3, \ldots, x_m, t) \\
&\to (x_3, \ldots, x_m, t, x_1) \\
&\to (x_4, \ldots, x_m, t, x_1, x_2) \\
&\quad \cdots \\
&\to (x_{i+1}, \ldots, x_m, t, x_1, \ldots, x_{i-1}) \\
&\to (x_{i+2}, \ldots, x_m, t, x_1, \ldots, x_{i-1}, y) \\
&\to (x_{i+3}, \ldots, x_m, t, x_1, \ldots, x_{i-1}, y, x_{i+1}) \\
&\quad \cdots \\
&\to (x_1, x_2, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_m) = u
\end{aligned}
$$

It is easy to check that each step here is a valid edge in G and the total length is $m + 1$. □

Having this claim, now we can use it as a subroutine. Consider two nodes $v = (x_1, x_2, \ldots, x_m)$ and $u = (y_1, y_2, \ldots, y_m)$ in G. In order to find a path from v to u, we first reach a node v_1 that satisfy $v_1[1] = y_1$ from node v using the following way: if y_1 exists in node v's label, namely $x_j = y_1$ for some j, then we first go from node v to node $w_1 = (x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_m)$ using $m + 1$ steps, here t is a symbol that does not appear in node v's label. If y_1 does not appear in v's label, we can just let $w_1 = v$. Then from node w_1, we can reach node $v_1 = (y_1, x_2, \ldots, x_{j-1}, t, x_{j+1})$ in $m + 1$ steps. Thus total length from v to w_1 is no more than $2(m + 1)$.

Using the similar way, we can find a path from v_i to some node v_{i+1} with length no more than $2(m+1)$, where v_j satisfies $v_j[t] = y_t$ for all $1 \leq t \leq j$. Thus finally we will reach $v_m = u$, and the total path length is no more than $2m(m + 1)$.

(4): According to the definition of line graph, each edge (u, v) in $SG_{nr}(m - 1, t)$ will become a new vertex t. Suppose the labels for u, v in $SG_{nr}(m - 1, t)$ are $u = (x_1, x_2, \ldots, x_{m-1})$ and $v = (x_2, \ldots, x_{m-1}, y)$ where $y \neq x_1$. Then we can label the new vertex $t = (x_1, \ldots, x_{m-1}, y)$, which is a valid label in $SG_{nr}(m - 1, t)$. And it is easy to check that every edge (s, t) in this line graph satisfies $s[2 : m] = t[1 : m - 1]$. Thus the line graph is just the shift graph $SG(m, t)$. Since the smallest circles in $SG(m, t)$ have length m and every edge (s, t) in such circles has form $s[2 : m] = t[1 : m - 1], s[1] = t[m]$. Thus after removing these edges, we get exactly the non-rotational shift graph $SG_{nr}(m, t)$. □

Moreover, non-rotational shift graphs have one important property that leads to their being Nash equilibria of ℓ-B^{3}C games, as we now explain.

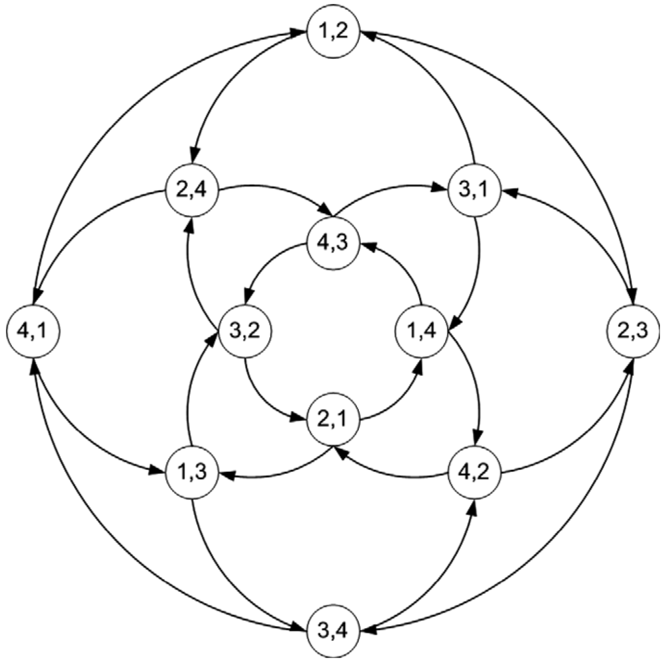

Fig. 6. Non-rotational shift graph $SG_{nr}(2, 4)$.

We say that a vertex v in a graph G is *ℓ-path-unique* if any path that passes through v (neither starting nor ending at v) with length no more than ℓ is the unique shortest path from its starting vertex to its ending vertex. A graph is *k-out-regular* if every vertex in the graph has out-degree k. A k-out-regular graph is an *ℓ-path-unique graph* (or *ℓ-PUG* for short) if every vertex in the graph is ℓ-path-unique.

Lemma 32. *Non-rotational shift graph $SG_{nr}(\ell, k+\ell)$ is an ℓ-PUG.*

Proof. Suppose for a contradiction that there exist two nodes s and t, such that there are two paths from s to t which both have length no more than ℓ, which are denoted as below:

$$s = a_1 \to a_2 \to \cdots \to a_{\ell_1} = t \quad \text{and}$$
$$s = b_1 \to b_2 \to \cdots \to b_{\ell_2} = t,$$

where (a_i, a_{i+1}) and (b_i, b_{i+1}) are all edges in this graph, $1 < \ell_1, \ell_2 \le \ell + 1$. Let i be the smallest index such that $a_i \ne b_i (1 < i \le \ell)$. Since $a_{i-1} = b_{i-1}$, we have $a_i[1 : \ell-1] = a_{i-1}[2 : \ell] = b_{i-1}[2 : \ell] = b_i[1 : \ell-1]$. So it must be that $a_i[\ell] \ne b_i[\ell]$. We also know that $a_i[\ell] = a_{i+1}[\ell-1] = \cdots = a_{\ell_1}[\ell-\ell_1+i] = t[\ell-\ell_1+i]$. Similarly we have $b_i[\ell] = t[\ell-\ell_2+i]$. So $t[\ell-\ell_1+i] \ne t[\ell-\ell_2+i]$, which means $\ell_1 \ne \ell_2$. So one of them must be less than $\ell+1$. Suppose $\ell_1 < \ell+1$, we have

$$s[\ell_1] = a_1[\ell_1] = a_2[\ell_1 - 1] = \cdots = a_{\ell_1}[1] = t[1]$$

If $\ell_2 < \ell+1$, use the same way we can get $s[\ell_2] = t[1] = s[\ell_1]$. But this cannot be true since $\ell_1 \ne \ell_2$ and the symbols must be all different in one label. So $\ell_2 = \ell + 1$. Then we have $t[1] = b_{\ell+1}[1] = b_2[\ell]$. But according to the definition, $b_2[1 : \ell-1] = b_1[2 : \ell] = s[2 : \ell]$ and $t[1] = b_2[\ell] \ne b_1[1] = s[1]$ (the no rotation requirement in $SG_{nr}()$ graphs). This implies that $t[1]$ cannot be same with any symbol in s's label. So $t[1] \ne s[\ell_1]$, which is a contradiction. Therefore, the lemma holds. □

The following lemma shows the importance of ℓ-PUG to uniform ℓ-B3C games.

Lemma 33. *If a directed graph G has n nodes and is k-out-regular and ℓ-path-unique, then G is a maximal Nash equilibrium for the uniform ℓ-B^{3}C game with parameter n and k.*

Proof. For any node v in G, we want to show that v is at its best response in the current configuration.

Suppose $total(v)$ is the total number of paths with length no more than ℓ that pass through node v (neither starting nor ending at v) in the current configuration. Note that here we consider all paths, including paths may visit some node multiple times. We first show that $total(v)$ is invariant with respect to node v's strategy and it is an upper bound of v's betweenness if v can only change its own strategy.

Let $start_x(v)$ be the number of paths with length x that start from node v. Since every node has out-degree k, we know $start_x(v) = start_{x-1}(v) * k = \cdots = start_0(v) * k^x = k^x$, which only depends on x and k and is invariant to the choice of v's k outgoing edges.

Let $end_x(v)$ be the number of paths with length x that end at v. Notice that every path with length no more than l that ends at v will not contain v's outgoing edges. Otherwise there will be a path from v to itself with length no more than ℓ, which is not a shortest path (the shortest path is just the node v itself). So $end_x(v)$ is independent of node v's strategy.

Now consider the number of paths with length x that pass through node v (neither starting or ending at v), denoted as $pass_x(v)$. We know $pass_x(v) = \sum_{1 \le i \le x-1} end_i(v) * start_{x-i}(v) = \sum_{1 \le i \le x-1} end_i(v) * k^{x-i}$. Thus $total(v) = \sum_{2 \le x \le \ell} pass_x(v)$ is also independent of v's strategy. At the same time, notice that these are the only paths that can contribute to v's betweenness. Thus for any strategy s_v of node v, we have $btw_v(s_v) \le total(v)$.

On the other hand, in the current configuration G every path with length no more than ℓ that passes v is a unique shortest path, thus will contribute one to v's betweenness. So we get $btw_v(G) = total(v)$, which means that node v is at its best response. Therefore the lemma holds. □

With the above result, we immediately have

Theorem 34. *For any $\ell \ge 2$, $\ell' \ge \ell$, $k \in \mathbb{N}_+$, graph $SG_{nr}(\ell', k + \ell')$ is a maximal Nash equilibrium of the uniform ℓ-B^{3}C game with parameters $n = (k + \ell')!/k!$ and k.*

Proof. This is immediate from Lemmata 32 and 33, and from the fact that any ℓ'-PUG is an ℓ-PUG for $\ell' \ge \ell$. □

The above construction of maximal Nash equilibria is based on path-unique graphs. Next we show that shift graphs also lead to another family of Nash equilibria not based on path uniqueness. In fact, we show that they are strict Nash equilibria for uniform ℓ-B^{3}C games for every $\ell \ge 2$ as well as B^{3}C games without path length constraint.

Definition 6. Given a graph $G = (V, E)$, a *vertex-duplicated graph* $G' = (V', E')$ of G with parameter $d \in \mathbb{N}_+$, denoted as $D(G, d)$, is a new graph such that each vertex of G is duplicated to d copies, and each duplicate inherits all edges incident to the original vertex. That is, $V' = \{(v, i) \mid v \in V, i \in [d]\}$, and $E' = \{((u, i), (v, j)) \mid u, v \in V, (u, v) \in E, i, j \in [d]\}$.

Theorem 35. *For any $t \ge 2$, $d \ge 2$, graph $D(SG(2, t), d)$ is a strict Nash equilibrium of the uniform ℓ-B^{3}C game with parameters $n = dt(t-1)$ and $k = d(t-1)$. It is also a strict Nash equilibrium of the uniform B^{3}C game without the path length constraint.*

Proof. Let G be the graph $D(SG(2, t), d)$. The nodes in G can be represented as (i, j, δ) where $1 \le i \ne j \le t$ and $1 \le \delta \le d$. The strategy of each node $v = (i, j, \delta)$ in configuration graph G is $s^*_v = \{(j, i', \delta') \mid 1 \le i' \ne j \le t, 1 \le \delta' \le d\}$.

Claim 1. *For any node v in G, $G \backslash \{v\}$ has diameter* 2.

Proof: notice that $d \ge 2$, and thus for any two nodes $u = (i, j, \delta)$ and $u' = (i', j', \delta')$ in $G \backslash \{v\}$, there are at least two length-2 paths from u to u' in G: one goes through $(j, i', 1)$ and the other goes through $(j, i', 2)$. Thus, after removing one node v, u and u' are still connected with at least one length-2 path. Claim 1 holds.

With Claim 1, it is immediate that for any possible strategy s_v of v and the graph G' that differs from G only in v's outgoing edges, all shortest paths that can contribute to the betweenness $btw_v(G')$ are of length 2. Therefore, $btw_v(G') = btw_v(G', \ell)$ for all $\ell \ge 2$. Hence in the following we only show that G is a strict Nash equilibrium for the uniform B^{3}C game without the path length constraint, and the result immediately applies to the uniform ℓ-B^{3}C games for all $\ell \ge 2$.

Given a vertex v, we fix the strategies for all of the vertices other than v and consider the betweenness value of v under different choice of v's strategy. By Lemma 1, we only need to consider maximal strategies of v when computing its best response. Let $s_v = \{v_1, v_2, \ldots, v_k\}$ be a maximal strategy of v. Let $btw_v(s_v)$ be the betweenness value of vertex v if v chooses s_v as its strategy, and $btw_v(u)$ be the betweenness value of v if v changes its strategy to $s_v = \{u\}$ (a non-maximal strategy). By Lemma 25 and the fact that $btw_v(G') = btw_v(G', 2)$ for all G' that differs from G only in v's outgoing edges, we have

$$btw_v(s_v) = \sum_{i=1}^{k} btw_v(v_i)$$

Thus for any vertex $v = (i, j, \delta)$, we only need to compare $btw_v(u)$ for all of the other vertices u and prove that the largest $d(t-1)$ values are exactly from the vertices in s^*_v.

By symmetry, we only need to consider vertex $v = (2, 1, 1)$. There are $(t-1)d$ vertices $(i', 2, \delta')$ with $1 \le i' \le t, i' \ne 2, 1 \le \delta' \le d$) connecting to vertex v. We divide the outgoing edges of v into seven cases based on their end points $u = (i, j, \delta)$ to compute the corresponding betweenness value $btw_v(u)$. We assume $1 \le \delta \le d$ and $i \ne j$.

$u = (2, 1, \delta)$: $btw_v(u) = 0$, since there is already an edge from $(i', 2, \delta')$ to $(2, 1, \delta)$;
$u = (i, 1, \delta), i \ge 3$: $btw_v(u) = \frac{(t-1)d}{d+1}$;
$u = (1, 2, \delta)$: $btw_v(u) = \frac{(t-1)d-1}{d}$;
$u = (i, 2, \delta), i \ge 3$: $btw_v(u) = \frac{(t-1)d-1}{d+1}$;
$u = (1, j, \delta), j \ge 3$: $btw_v(u) = \frac{(t-1)d}{d}$;
$u = (2, j, \delta), j \ge 3$: $btw_v(u) = 0$, since there is already an edge from $(i', 2, \delta')$ to $(2, j, \delta)$;
$u = (i, j, \delta), i, j \ge 3$: $btw_v(u) = \frac{(t-1)d}{d+1}$.

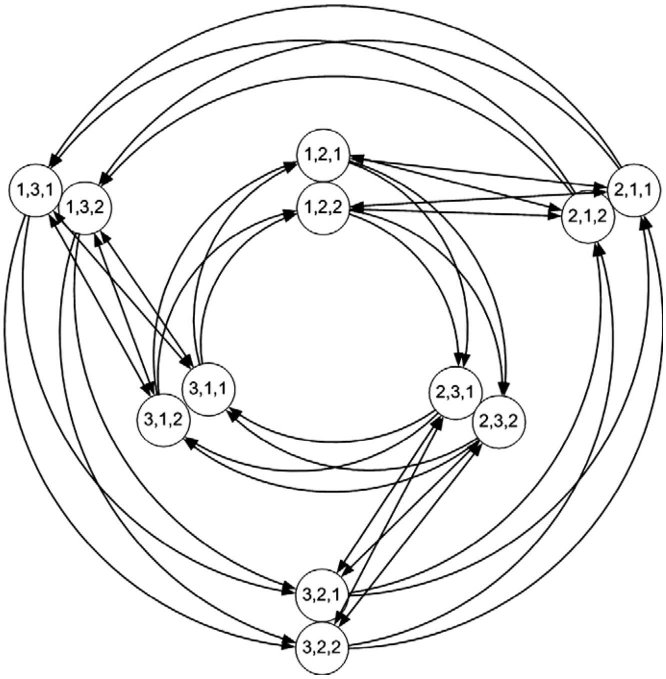

Fig. 7. Vertex duplicated shift graph $D(SG(2, 3), 2)$.

When $t \geq 3$, we have $\frac{(t-1)d}{d} > \frac{(t-1)d-1}{d} > \frac{(t-1)d}{d+1} > \frac{(t-1)d-1}{d+1} > 0$. Thus the top $k = d(t-1)$ vertices with the best $btw_v(u)$ values are $(1, j, \delta)$ with $2 \leq j \leq t$ and $1 \leq \delta \leq d$, which is exactly s_v^*. Moreover, the sum of $btw_v(u)$'s of these vertices are strictly larger than the sum of any other subsets of k vertices. Therefore, s_v^* is a strict best response and the graph is a strict Nash equilibrium.

When $t = 2$, only two cases $u = (2, 1, \delta)$ and $u = (1, 2, \delta)$ are left, and the $k = d$ best choices are $u = (1, 2, \delta)$ with $1 \leq \delta \leq d$, again exactly s_v^*. Any other subset of k nodes give strictly lower betweenness. Therefore, the graph is a strict Nash equilibrium too when $t = 2$. □

In the simple case of $t = 2$, graph $D(SG(2, 2), d)$ is the complete bipartite graph with d nodes on each side. For larger t, $D(SG(2, t), d)$ is a t-partite graph with more complicated structure. Fig. 7 shows an example of graph $D(SG(2, 3), 2)$. When $d = 2$, we have $n = 2t(t-1)$ and $k = 2(t-1)$. Thus, we have found a family of strict Nash equilibria with $k = \Theta(\sqrt{n})$.

An important remark is that when $d \geq 2$, each node is split into at least two nodes inheriting all incoming and outgoing edges, and thus graphs $D(SG(2, t), d)$ for all $t \geq 2$ and $d \geq 2$ are not ℓ-PUGs for any $\ell \geq 2$. Therefore, the construction by splitting nodes in shift graphs $SG(2, t)$ are a new family of construction not based on path-unique graphs.

5.2. Properties of Nash equilibria

From Lemma 33, we learn that ℓ-PUGs are good sources for maximal Nash equilibria for uniform ℓ-B^{3}C games. Thus we start by looking into the properties of ℓ-PUGs to obtain more ways of constructing Nash equilibria. The following lemma provides a few ways to construct new ℓ-PUGs given one or more existing ℓ-PUGs.

Lemma 36. *Suppose that G is a k-out-regular ℓ-PUG. The following statements are all true:*

(1) *If G' is a k'-out-regular subgraph of G for some $k' \leq k$, then G' is an ℓ-PUG.*

(2) *Let v be a node of G and $\{v_1, v_2, \ldots, v_k\}$ be v's k outgoing neighbors. We add a new node u to G to obtain a new graph G'. All edges in G remains in G', and u has k edges connecting to $v_1, v_2, \ldots, v_k$. Then G' is also an ℓ-PUG.*

(3) *If G' is another k-out-regular ℓ-PUG and G' does not shared any node with G, then the new graph G'' simply by putting G together with G' is also an ℓ-PUG.*

The proof of the lemma is straightforward by definition and is omitted. Lemma 36 has several important implications. First, by repeatedly applying Lemma 36(2) on an existing ℓ-PUG, we can obtain an ℓ-PUG with an arbitrary size. Combining it with Theorem 34, it immediately implies the following theorem.

Theorem 37. *For any $\ell \geq 2$, $k \in \mathbb{N}_+$, and $n \geq (k+\ell)!/k!$, there is a maximal Nash equilibrium in the uniform ℓ-B^{3}C game with parameters n and k.*

Next, Lemma 36 implies that there exist rich structures among the Nash equilibria of uniform ℓ-B^{3}C games. In particular, Lemma 36(3) implies that Nash equilibria may be disconnected, while Lemma 36(2) implies that Nash equilibria may be weakly connected but not strongly connected. Furthermore, by repeatedly adding new nodes based on Lemma 36(2) such that all new nodes connected to the same set of $\{v_1, v_2, \ldots, v_k\}$ nodes, we may have very unbalanced Nash equilibria in which some nodes have zero in-degree while other nodes have in-degree close to n. This also implies that Nash equilibria may have some nodes with zero betweenness while other nodes have very large betweenness, that is, we have very unfair Nash equilibria. Note that Nash equilibria based on shift graphs given in Theorems 34 and 35 are all fair in that all nodes have the same betweenness.

Finally, we investigate non-PUG maximal Nash equilibria in the uniform 2-B^{3}C game with parameters (n, k), which by Theorem 26 is the most interesting case since its best response computation is polynomial. We want to see that when we fix k, whether we can find non-PUG maximal Nash equilibria for arbitrarily large n. Let $maxInd(G)$ denotes the maximum in-degree in graph G. The following result provides the condition under which all maximal Nash equilibria are PUGs.

Theorem 38. *Let G be a k-out-regular graph with n nodes. If $maxInd(G) \leq \frac{n-k}{k^2+k+1}$, then G is a maximal Nash equilibrium for the uniform 2-B^{3}C game with parameter n and k if and only if G is a 2-PUG.*

Proof. Lemma 33 already shows the part of sufficient condition. Thus we only need to prove that if G is not a 2-PUG, some node will have better response in G.

Suppose node v is a node in G that is not 2-path unique. Let $S = \{u | (u, v) \in G\}$. We know that $|S| \leq maxInd$. Then let S' be the set of nodes that can be reached from any node in S in no more than 2 steps. Since every node has out-degree k, we know that $|S'| \leq |S| + |S| \times k + |S| \times k \times k = |S| \times (1 + k + k^2) \leq maxInd \times (1 + k + k^2)$. Also notice that $n \geq maxInd \times (1+k+k^2)+k$, so there exist at least k nodes that are not in S'. If we let node v connect to these k nodes, then every length 2 path that passes through v in the form $x \to v \to y$ will be the unique shortest path from x to y, because y is not reachable from x within 2 steps in any other ways. So v is 2-path unique now, and this will give it a better response. □

The above theorem implies that non-PUG equilibria is only possible if $maxInd(G) = \Theta(n)$ when k is a constant, which means that non-PUG equilibria must have very unbalanced in-degrees when n is large. In the following, we show as an example how to construct such non-PUG equilibria for the case of $k = 2$.

First, we introduce a general scheme of adding nodes, similar to the one in Lemma 36 (2), such that if the original graph is non-PUG Nash equilibria with certain properties, then the new graph is still a non-PUG Nash equilibria with the same properties.

We say that an edge (v, u) in G is *shortcut* by a node w if (w, v) and (w, u) is in G. Then we have the following lemma.

Lemma 39. *Suppose that G is a k-out-regular graph in which only one node v is not 2-path unique, and every edge (v, u) in G is shortcut by at most one node. Let (w, v) be an edge in G, and w has k outgoing neighbors $v_1, v_2, \ldots, v_k$ including v. We add a new node x to G to obtain a new graph G' such that x connects to $v_1, v_2, \ldots, v_k$ and all edges in G remains in G'. Then in G' only node v is not 2-path unique. If G is a maximal Nash equilibria for the uniform 2-B^{3}C game, then G' is also a maximal Nash equilibria.*

Proof. First it is easy to see that v is also not 2-path unique in G' because every path in G is still a path in G'.

For every length 2 path that passes through node v, suppose it is $x \to v \to y$. If it is not the unique shortest path from node x to node y, we must have $(x, y) \in G$, i.e. (v, y) is shortcut by node x. Because otherwise there must exist another node u such that $x \to u \to y$ is also a shortest path. But then u is also not 2-path unique, and that will contradict the fact that v is the only node in G that is not 2-path unique.

Now suppose that G is a maximal Nash equilibria. Let a be the in-degree of node v. Then for each $(v, u) \in G$, since it is shortcut by at most one node, which means there is at most node w such that $(w, v) \in G$ and $w \to v \to u$ is not the unique shortest path from w to u. Therefore we have $a - 1 \leq btw_v(G_{\{u\}}, 2) \leq a$. Since node v is at its best response in G, along with Lemma 25 we know that $btw_v(G_{\{t\}}, 2) \leq a - 1$ for every node t where $(v, t) \notin G$

Now consider G' with new node x in it. First is obvious to see that every node in G' except node v is still at its best response. And we have $btw_v(G'_{\{u\}}, 2) = btw_v(G_{\{u\}}, 2) + 1$ for every node $u \neq x, u \neq v$. Because there is exactly one more unique shortest path $x \to v \to u$ that contribute betweenness value to edge (v, u). Thus we have $a \leq btw_v(G'_{\{u\}}, 2) \leq a + 1$ when $(v, u) \in G$ and $btw_v(G'_{\{u\}}, 2) \leq a$ when $(v, u) \notin G$. Also notice that $btw_v(G'_{\{x\}}, 2) \leq a$ because path $x \to v \to a$ is not a shortest path. So we know node v is also at its best response in graph G', thus G' is a maximal Nash equilibria too. □

Fig. 8 shows a Nash equilibria for the uniform 2-B^{3}C game with $n = 10$, $k = 2$, which we found by our experiments. In this graph, only one node is not 2-path unique and every edge out of this node is shortcut by at most one node. This means that, at least for $k = 2$, we apply the scheme of Lemma 39 to Fig. 8 to generate arbitrarily large graphs that are still non-PUG Nash equilibria.

Theorem 38 can also be used to eliminate some families of graphs with balanced in-degrees as maximal Nash equilibria. We now show that a family of symmetric graphs called Abelian Cayley graphs cannot be Nash equilibria of uniform 2-B^{3}C games. An *Abelian Cayley graph* $G = (V, E)$ is a graph generated by the additive group $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ and a generating set $A \subseteq \mathbb{Z}_n$ of size k, such that $V = \mathbb{Z}_n$ and $E = \{(x, y) \mid x, y \in \mathbb{Z}_n, \exists z \in A, y = x + z \mod n\}$. We denote such a graph by $\langle \mathbb{Z}_n, A \rangle$.

It is easy to see that Abelian Cayley graphs are not 2-PUGs when $k \geq 2$. Let $z_1, z_2 \in A$, and $y = x + z_1 + z_2 \mod n$ for some $x \in \mathbb{Z}_n$. Then from node x to node y, there are at least two length-two paths, one passing through $w_1 = x + z_1 \mod n$ and the other passing through $w_2 = x + z_2 \mod n$. Therefore none of the nodes in an Abelian Cayley graph is

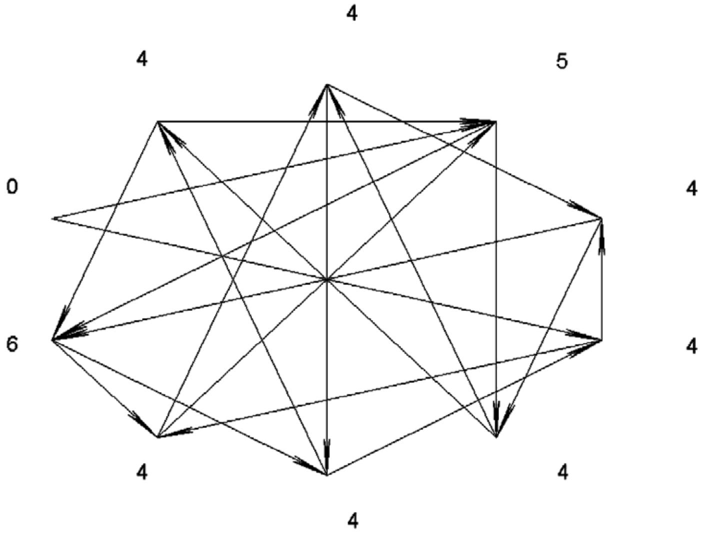

Fig. 8. A Nash equilibrium for the uniform 2-B^{3}C game with $n = 10$, $k = 2$ that is not a 2-PUG. The number next to a vertex is its betweenness value. The vertex with betweenness 5 is not 2-path-unique.

2-path unique. Moreover, it is clear that every node in the Abelian Cayley graph has in-degree k. Therefore, by Theorem 38 we have the following result.

Corollary 40. *For any $n \geq k^3 + k^2 + 2k$, any Abelian Cayley graph $\langle \mathbb{Z}_n, A\rangle$ with $|A| = k$ is not a maximal Nash equilibria for the uniform 2-B^{3}C game with parameters n and k.*

6. Conclusion and future work

In this paper, we present results on bounded budget betweenness centrality (B^{3}C) game, a type of network formation games in which nodes in the network try to strategically select other nodes to connect subject to the budget constraint in order to maximize their betweenness centrality in the network. We focus on ℓ-B^{3}C game, where shortest paths contributing to betweenness have path length constraint of at most ℓ, which matches realistic scenarios and generalizes the work of [15]. We present both hardness results for the nonuniform version of the game and constructive existence results for the uniform version of the game. We also study the complexity of computing best response in the game.

There are still a number of issues and open problems to be address in future research on betweenness-centrality (or other network measures) based network formation games. One issue is that the Nash equilibrium structures found so far is rather limited and unnatural. This is often a limitation in other studied network formation games as well. For example, Nash equilibria studied in [15] have to be dense graphs and the structures found are layered complete graphs, and the Nash equilibria found in the game of [16] are tree-like structures. One important open question is whether more natural phenomena of networks, such as scale-free and small-world phenomena, can naturally emerge from such network formation games. The second issue is related to solution concepts. In our betweenness game, computing best response is NP-hard except for the case of uniform game with $\ell = 2$. This means that individuals cannot in general find their best responses. Thus a future direction is to study equilibria that based on feasible computations of individuals. Moreover, individual's computation is better be local since it is unlikely that individuals know the entire network structure. This is also an issued shared by other studies of network formation games and should be improved upon. Another issue is to model interaction and information diffusion in a more realistic way. For example, pure undirected links and pure directed links are two extremes of real-world interactions. Can we model unsymmetric and fractional interaction between individuals? Also, in this paper as well as a number of other papers information diffusion between a pair of nodes are only through their shortest paths. A future direction would be to study other diffusion models such as random walks and their corresponding network formation games.

Acknowledgements

We would like to thank Heiko Röglin for his helpful discussions on this research topic. This work was supported in part by the National Natural Science Foundation of China Grant 60553001, and the National Basic Research Program of China Grants 2007CB807900, 2007CB807901.

References

[1] Susanne Albers, Stefan Eilts, Eyal Even-Dar, Yishay Mansour, Liam Roditty, On nash equilibria for a network creation game, in: Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms, 2006, pp. 89–98.
[2] Ronald S. Burt, Structural Holes: The Social Structure of Competition, Harvard University Press, 1992.

[3] Ronald S. Burt, Secondhand brokerage: evidence on the importance of local structure for managers, bankers, and analysts, The Academy of Management Journal 50 (2007) 119–148.
[4] Vincent Buskens, Arnout van de Rijt, Dynamics of networks if everyone strives for structural holes, American Journal of Sociology 114 (2) (2008) 371–407.
[5] Bram Cohen, Incentives build robustness in BitTorrent, 1st Workshop on Economics of Peer-to-Peer Systems, 2003.
[6] Jacomo Corbo, David C. Parkes, The price of selfish behavior in bilateral network formation, in: Proceedings of the 24th ACM Symposium on Principles of Distributed Computing, 2005, pp. 99–107.
[7] DARPA, DARPA network challenge project report, Tech. Report, Defense Advanced Research Projects Agency, February 2010.
[8] N.G. de Bruijn, A combinatorial problem, Koninklijke Nederlandse Akademie v. Wetenschappen 49 (1946) 758–764.
[9] Robin I.M. Dunbar, Neocortex size as a constraint on group size in primates, Journal of Human Evolution 22 (6) (1992) 469–493.
[10] Alex Fabrikant, Ankur Luthra, Elitza Maneva, Christos H. Papadimitriou, Scott Shenker, On a network creation game, in: Proceedings of the 22nd ACM Symposium on Principles of Distributed Computing, 2003, pp. 347–351.
[11] Robert W. Floyd, Algorithm 97 (SHORTEST PATH), Communications of the ACM 5 (6) (1962) 345.
[12] L. Freeman, Centrality in social networks: conceptual clarification, Social Networks 1 (1979) 215–239.
[13] Sanjeev Goyal, Fernando Vega-Redondo, Structural holes in social networks, Journal of Economic Theory 137 (1) (2007) 460–492.
[14] Matthew Jackson, Asher Wolinsky, A strategic model of social and economic networks, Journal of Economic Theory 71 (1) (1996) 44–74.
[15] Jon Kleinberg, Siddharth Suri, Éva Tardos, Tom Wexler, Strategic newtwork formation with structural holes, in: Proceedings of the 9th ACM Conference on Electronic Commerce, 2008.
[16] Nikolaos Laoutaris, Laura Poplawski, Rajmohan Rajaraman, Ravi Sundaram, Shang-Hua Teng, Bounded budget connection (bbc) games or how to make friends and influence people, on a budget, in: Proceedings of the 27th ACM Symposium on Principles of Distributed Computing, 2008.
[17] Thomas Moscibroda, Stefan Schmid, Roger Wattenhofer, On the topologies formed by selfish peers, in: Proceedings of the 25th ACM Symposium on Principles of Distributed Computing, 2006, pp. 133–142.
[18] M.E.J. Newman, J. Park, Why social networks are different from other types of networks, Physical Review E 68 (2003) 036122.
[19] Christos H. Papadimitriou, Algorithms, games, and the internet, in: Proceedings of the 33rd ACM Symposium on Theory of Computing, 2001, Invited talk, pp. 749–753.
[20] Galen Pickard, Iyad Rahwan, Wei Pan, Manuel Cebrián, Riley Crane, Anmol Madan, Alex Pentland, Time critical social mobilization: the DARPA network challenge winning strategy, Tech. Report arXiv:1008.3172, Computing Research Repository, August 2010.
[21] John Scott, Social Network Analysis, Sage, 1991.
[22] Daniel Stutzbach, Subhabrata Sen, Characterizing unstructured overlay topologies in modern p2p file-sharing systems, ACM/USENIX Internet Measurement Conference, 2005.