A Delivery Network Creation Game

Georgios Piliouras

Department of Computer Science
Cornell University
qpil@cs.cornell.edu

Ymir Vigfusson
Department of Computer Science
Cornell University
ymir@cs.cornell.edu

ABSTRACT

We study a non-cooperative network creation game where players, represented by nodes, can build edges to other players for a cost of α , and strive to maintain short paths to other players while minimizing cost. Players incur a penalty of β for each unreachable node in addition to the charges for constructing edges. Specifically, each player i optimizes

$$cost(i) = \alpha e_i + \beta n_i + \sum_j d(i, j)$$

where e_i is the number of edges built by i, n_i is the number of unreachable nodes from i by paths in the network, and α and β are parameters.

The model generalizes previous work such that it provides an abstraction for describing the synthesis of various economic networks. For instance, in a network for the transportation of goods between facilities, the α cost parameter can intuitively be viewed as the price of establishing a route between facilities, and β is the value (or incentive) to have access to goods at a remote site.

We observe sharp changes in optima as the α and β parameters vary. Furthermore, we bound the price of anarchy of the game for all values of α , β and n.

We identify surprising properties in the structure of Nash equilibria. We show that not only do there exist zero-incentive strict Nash equilibria of arbitrarily large size but they also exhibit properties such as constant diameter and resilience to any single-edge deletion. Lastly, we identify the first superconstant lower bound on the price of anarchy in this line of research and prove that it is persistent even if we incorporate in our model coalitions of size up to $o(\sqrt{n})$.

1. INTRODUCTION

The theory of network formation is fundamental in computer science, economics, operations research, and the social sciences. One of the goals is to predict the behavior of the

network systems by using only structural properties of the formed graphs and rules defined on the vertices.

The methods employed in network design have generally assumed central coordination during the synthesis of the network. This approach breaks down for explaining distributed and uncoordinated behavior in several large-scale networks, such as linkage of the backbones of the Internet [11] and goods exchange networks [13].

Fabrikant, Luthra, Maneva, Papadimitriou and Shenker [11] propose a game-theoretic model for network creation that addresses the uncoordinated and selfish behavior of the participating agents. Their model involves a game with nplayers who build edges to form a connected, undirected graph. Each player is denoted by a vertex and is allowed to buy link to the other players in the game at a cost of $\alpha > 0$ per link. The players' goal is to minimize the number of edges they build while maintaining short paths to all other members of the network. Specifically, each player i attempts to optimize an associated cost function $\alpha e_i + \sum_i d(i,j)$ where e_i is the number of edges she builds and d(i,j) is the shortest distance between players i and j. Players can use any edge in the network when calculating distances, but they can only remove the edges they built themselves. The social cost of a network is the sum of the costs of all its players.

Among the key concepts used in game-theoretic analysis is the *Nash equilibrium* [19]. It is a combination of strategies in which no player can improve his situation single-handedly by changing his strategy. In the aforementioned model, a Nash equilibrium is a stable solution of the network creation game in which no player can lower her cost by building new edges or removing any of those she already had constructed, or both. The resulting network is called an *equilibrium graph*.

In [16], Koutsoupias and Papadimitriou studied the impact on performance due to selfish behavior of players. They coined the term "price of anarchy" which is defined to be the ratio of the cost of the worst-case Nash equilibrium to the cost of the global optimum solution. Essentially, the price of anarchy measures the degradation of performance due to the lack of coordination between players. The concept has been used in literature on routing [22, 23], congestion games [8, 9, 12] and network design [2, 3, 4, 10, 15].

1.1 Previous work

One of the earliest papers that employs game theory to study network formation games is that of Myerson [18]. His model consists of a cooperative game in which players can create links to communicate or cooperate such that players in the same component can act together as a coalition. How-

^{*}Supported in part by NSF ITR grant CCR-0325453.

ever, his model is indifferent of the structure of the coalition as long as players are connected.

Jackson and Wolinsky [14] address this concern, again using cooperative game-theory. They present a model in which players pay for having direct links to other players, and are rewarded for having good connectivity in the network. The authors show that the only efficient networks obtainable are the complete graph, the star and the empty graph, and that the star may fail to be pairwise stable even if it is efficient.

Among the earlier work in non-cooperative network formation games is the work of Bala and Goyal [5]. Their model is similar to the one that is the subject of this paper, in that it has unilateral and simultaneous link creation with direct link penalty, but they instead focus on information flow in the network.

Anshelevich et al. [2, 3] analyze a network connection game of a different flavor where each player desires to connect a pair of terminals in a flow graph. Players are allowed to contribute to the cost of any edge on paths between their terminals. They analyze cost-allocation schemes and how they can affect the quality of the Nash equilibria.

Johari, Mannor and Tsitsiklis [15] investigate a network formation game where players also have a trade-off between connectivity cost and traffic flow, and nodes pay for their own links. However, it differs from our model in that it focuses on bargaining, negotiation and contracts between players for building edges, routing of flow in the network and employs a different cost scheme.

More general models of network formation that involve transfer payments between players have been considered by Anshelevich et al. [4], and Bloch and Jackson [6].

Carbo and Parkes [10] study an extension of the model of Fabrikant *et al.* where edge costs are shared between players and the creation of an edge requires consent from both parties. They show that the worst-case price of anarchy in their model is worse than in the unilateral one, but provide experimental results suggesting that the bilateral model performs better on average.

Another extension on the Fabrikant et al. model was done by Moscibroda et. al. [17] who study the effects on the topology of a P2P network imposed by the fact peers selfishly select the peers to connect to. They model the peers of a P2P network as points in a metric space and they prove that resulting topologies may be much worse than if peers collaborated and also that the overall system may not never stabilize.

The original model that we extend in this paper is that of Fabrikant et al. [11]. They derive the upper bound of $O(\sqrt{\alpha})$ for $2 \le \alpha < n^2$, and O(1) for other values of α . Their paper shows that when $\alpha < 2$, the social optimum is a complete graph, and when $\alpha \ge 2$ it is a star. They also give the lower bound of $3 - \varepsilon$ on the price of anarchy for any $\varepsilon > 0$. Additionally, the authors conjecture that there is a constant A such that for $\alpha > A$ all non-transient equilibrium graphs are trees. A transient Nash equilibrium is one where there exists a sequence of strategy changes that retains the social cost in every step but leads to a non-equilibrium position.

More recently, Albers, Eilts, Even-Dar, Mansour, and Roditty [1] improve on these bounds. They derive a constant upper bound on the price of anarchy for $\alpha \in O(\sqrt{n})$ and $\alpha \geq 12n\lceil\log n\rceil$, and give the improved bound of $O(1+(\min\{\frac{\alpha^2}{n},\frac{n^2}{\alpha}\})^{1/3})$ for the intermediate values of α . For $\alpha \geq 12n\lceil\log n\rceil$ they show that every equilibrium graph is a

tree. Finally, they disprove the aforementioned conjecture by constructing Nash equilibria for arbitrarily many players and for $1 < \alpha \le \sqrt{n/2}$ that contain cycles. The construction has the properties that the shortest paths between players are unique, and have a diameter of 2.

1.2 The Model

In this paper we extend the aforementioned network creation model by introducing a parameter $\beta \geq 0$ to the cost function that controls the penalty for players being in different connected components. Intuitively, a high value of β denotes a high incentive to be connected to a large part of the network. Before giving a motivating example, let us thoroughly define the details of the model that we analyze in this paper.

Consider the following game with $n \geq 2$ players $V = \{1, 2, \ldots, n\}$. Each player i has a strategy $E_i \subseteq V \setminus \{i\}$ which denotes the players to whom i will build edges. If the combination of all strategies $S = (E_1, E_2, \ldots, E_n)$ is known, we can create the corresponding undirected graph

$$G(S)=(V,E)$$
 where $E=\bigcup_{i=1}^n\bigcup_{j\in E_i}\{i,j\}$. Suppose that $G(S)$ has the connected components C_1,C_2,\ldots,C_k and that

G(S) has the connected components C_1,C_2,\ldots,C_k and that $i\in C_t$ for some t. Let $e_i=|E_i|$ for all i. Player i incurs a cost of

$$cost^{\beta}(i) = \alpha e_i + \beta n_i + \sum_{j \in C_t} d(i, j)$$

where $n_i = n - |C_t|$ and d(i, j) is the shortest distance between players i and j in G(S). Here $\alpha \geq 0$ and $\beta \geq 0$ are parameters. The social cost of the strategy is the aggregate cost of all players, in other words $\cot^{\beta}(G(S)) = \sum_{i=1}^{n} \cot^{\beta}(i)$. A configuration for the game is a tuple (α, β, n) .

Notice that the model of Fabrikant *et al.* is the special case $\beta = \infty$ in our model. Thus $\cot^{\infty}(G(S))$ is equal to the social cost incurred for G(S) in the aforementioned model. Also, if the graph is connected, $\cot^{\beta}(G(S)) = \cot^{\infty}(G(S))$ since the β -term is zero.

A Nash equilibrium is a combination of strategies $S = (E_1, E_2, \ldots, E_n)$ such that for each player i and any other profile of strategies S', identical to S except in i's component, we have $\cot^{\beta}(G(S)) \leq \cot^{\beta}(G(S'))$. The network corresponding to a Nash equilibrium is called an equilibrium graph. A combination of strategies S is optimal if $\cot^{\beta}(G(S)) \leq \cot^{\beta}(G(S'))$ for all other profiles of strategies S'. The corresponding network is called the optimum graph, or simply the optimum.

To motivate the extension, we remark that when $\beta=\infty$ the social cost is finite if and only if the network is connected. As mentioned in [11], this is unrealistic because it forces players to take part in expensive networks. By introducing the β parameter, we also allow for exploration of the dynamics of network formation. The fundamental difference is that disconnected graphs are now valid outcomes of the game since players can opt to not take part in the formation of the network by building no edges. We can now start from an empty graph and watch the network evolve over time, whereas when $\beta=\infty$ none of the intermediate stages corresponding to networks with multiple components would be valid. This uncovers the magic behind the creation of the connected networks in the model of Fabrikant et~al.

Our model also provides an abstraction for describing the synthesis of various economic networks. For instance, in a

network for the transportation of goods between facilities, the α cost parameter can intuitively be viewed as the price of establishing a route between facilities, and β is the value or incentive to have access to the goods of a remote site. By increasing the dimensionality of the economic model (use of two parameters) we gain the ability to describe the evolution of more realistic economic procedures, as for examples partly regulated economies where one of the parameters (i.e. the cost of creating connections) can be fine tuned be some external entity (i.e. government) in order to influence the overall behavior of the system. Of course, this extra power comes at the cost of more complicated proofs and finer partitioning of the α - β plane in terms of what the exact behavior of the resulting system looks like.

1.3 Our results

Our main contribution is to derive upper bounds on the price of anarchy for any configuration (α, β, n) of the game.

We determine the optimum networks for any configuration (α, β, n) . We show that the unique optimum is a star for $\beta >$ $\frac{1}{2}(\alpha-2)+2$ if $\alpha\geq 2$, and a complete graph for $\beta>\frac{1}{2}(\alpha+2)$ if $\alpha \leq 2$. For other values of β , the empty graph is optimal.

If $\beta > \alpha + 1$, we show that the equilibrium graph is connected and give bounds on the price of anarchy in this region. It is constant $\alpha \in O(\sqrt{n})$ or $\alpha \geq 12n\lceil \log n \rceil$, and $O(1 + (\min\{\frac{\alpha^2}{n}, \frac{n^2}{\alpha}\})^{1/3})$ for other values of α . Now consider $\beta \leq \alpha + 1$. The following regions correspond

to areas B, C and D on figure (2).

- If $\beta > (\alpha + 2)/2$ the price of anarchy is constant.
- Assuming $\alpha \geq 2$, the price of anarchy when $\beta \geq (\alpha -$ 2)/n + 2 is

$$O\left((\min\{\frac{\alpha^2}{n},\frac{n^2}{\alpha}\})^{1/3}+\frac{\beta n}{\alpha+2n-2}\right).$$

Furthermore, when $\alpha \geq 12n\lceil \log n \rceil$ we give a tight bound of $\Theta\left(\frac{\beta n}{\alpha+2n-2}\right)$ on the price of anarchy.

• For the remaining region, the price of anarchy is 1 when $\alpha \geq 12n\lceil \log n \rceil$, it is $\Theta\left(\frac{1}{\beta}\right)$ when $\alpha < \sqrt{n/2}$ and for other values of α it is

$$O\left(\left(1+(\min\{\frac{\alpha^2}{n},\frac{n^2}{\alpha}\})^{1/3}\right)\frac{\alpha+2n-2}{\beta n}\right).$$

We conclude by exploring the implications of allowing coalitions in the model, and show that even if we allow collusions of size $o(\sqrt{n})$, there are instances in the network formation model where the price of anarchy remains unbounded.

2. **OPTIMA**

In order to determine the price of anarchy in our model, we will need to bound the social cost of optimum graphs and equilibrium graphs for all (α, β, n) configurations. Generally speaking problems have optima that are hard to find or characterize. However, as was shown in [11], the optimum in the case where $\beta = \infty$ is extremely easy to compute. In fact, for $\alpha < 2$ it is a complete graph on *n*-vertices, and for $\alpha \geq 2$ it is a star. With this in mind, we will commence our analysis by precisely determining the optima for our model.

The main obstacle is that we can no longer rely on the optimum graph to be connected like is implicitly assumed in the old model. However, the following lemma shows that each component in an optimum in must be optimal when viewed in isolation for $\beta = \infty$.

Lemma 2.1. Suppose that C is a connected component with k nodes in an optimum solution C^* for some (α, β, n) . Then C in isolation is an optimum for (α, ∞, k) .

PROOF. Suppose C was suboptimal. Substitute the Cpart of C^* with an optimum of the same size (namely k) for the same α . This is necessarily an improvement because the β -term of the cost function is unaffected. This contradicts the optimality of C^* .

We are now ready to exactly characterize the optimum graphs of our model. The result is summarized on figure 2.

Theorem 2.2. For $\alpha \geq 2$ the optimum network is a star if $\beta > \frac{1}{n}(\alpha - 2) + 2$ and the empty graph otherwise, and for $\alpha \leq 2$ the optimum is a complete graph if $\beta > \frac{1}{2}(\alpha + 2)$ and the empty graph otherwise. Furthermore, for $\beta > \alpha + 1$, all optima C^* are connected.

PROOF. Suppose there are k components in the optimum

Assume $\alpha \geq 2$. By the previous lemma the optimum C^* is a collection of stars. Let x_i denote the number of nodes in the ith star. The cost of star i is $\alpha(x_i - 1) + 2(x_i - 1)^2 +$ $\beta x_i(n-x_i)$. The total cost of C^* is thus

$$cost^{\beta}(C^{*}) = \sum_{i=1}^{k} (\alpha(x_{i}-1) + 2(x_{i}-1)^{2} + \beta x_{i}(n-x_{i}))$$

$$= \sum_{i=1}^{k} (\beta n - 4 + \alpha)x_{i} + \sum_{i=1}^{k} (2 - \alpha) + \sum_{i=1}^{k} (2 - \beta)x_{i}^{2}$$

$$= (\beta n - 4 + \alpha)n + (2 - \alpha)k + \sum_{i=1}^{k} (2 - \beta)x_{i}^{2}.$$

Let us determine the characteristics of the collection of stars that minimizes the total cost.

(a) For $\beta > 2$, which includes the case $\beta > \alpha + 1$, we minimize $(\beta n - 4 + \alpha)n - (\alpha - 2)k - \sum_{i=1}^k (\beta - 2)x_i^2$. Equivalently, since $(\beta n - 4 + \alpha)n$ is constant, we maximize $(\alpha - 2)k + \sum_{i=1}^k (\beta - 2)x_i^2$. For any k, the optimum network that maximize this state $(\alpha - 2)k + \sum_{i=1}^k (\beta - 2)x_i^2$. work that maximizes this quantity consists of a big star of n - k + 1 nodes, and k - 1 isolated vertices. We now determine the number of components k that maximizes

$$(\beta - 2) ((n - k + 1)^2 + k - 1) + (\alpha - 2)k.$$

Since $\beta - 2 > 0$, this is a convex function for continuous $k \in [1, n]$ and therefore the optimal k is either 1 or n. If $\beta > \frac{1}{n}(\alpha - 2) + 2$, such as when $\beta > \alpha + 1$, then the optimal k is 1, in other words the optimum network is a single star. When $\beta < \frac{1}{n}(\alpha - 2) + 2$ the optimal k is n so the optimum network has n isolated vertices. Finally, when $\beta = \frac{1}{n}(\alpha - 2) + 2$ then both the star and the empty graph are optimal.

(b) Suppose $\beta = 2$. Then the last term of the social cost function $cost^{\beta}(C^*) = (\beta n - 4 + \alpha)n + (2 - \alpha)k + \sum_{i=1}^{k} (2 - \alpha)k + \sum_{i=1}^{k}$ β) x_i^2 is zero. Since $\alpha \geq 2$ we see that $(\beta n - 4 + \alpha)n + \alpha$ $(2-\alpha)k$ is minimized for k=n so the optimum network is the empty graph.

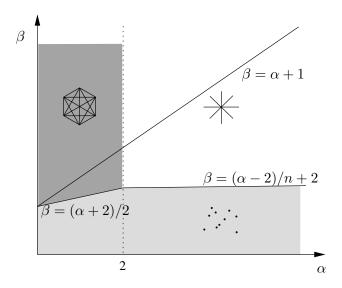


Figure 1: Optima in our model

(c) For $\beta < 2$ we minimize $(\beta n - 4 + \alpha)n - (\alpha - 2)k - \sum_{i=1}^k (\beta - 2)x_i^2$. Since $\sum_{i=1}^k x_i = n$, the third term is minimized when $x_i = 1$ for all i. This implies that k = n which in turn minimizes the second term. Since the first term is constant, the overall minimum is indeed achieved at k = n. Hence the optimum network is the empty graph.

Now assume $\alpha \leq 2$. By the previous lemma the optimum C^* is a collection of cliques. Let x_i denote the number of nodes in the i^{th} clique. The cost of clique i equals $(\alpha + 2)\binom{x_i}{2} + \beta x_i(n-x_i)$, so the total cost of C^* is thus

$$cost^{\beta}(C^*) = \sum_{i=1}^{k} (\alpha + 2) {x_i \choose 2} + \beta x_i (n - x_i)$$

$$= \sum_{i=1}^{k} \left(\beta n - \frac{\alpha + 2}{2}\right) x_i + x_i^2 \left(\frac{\alpha + 2}{2} - \beta\right)$$

$$= \left(\beta n - \frac{\alpha + 2}{2}\right) n + \sum_{i=1}^{k} x_i^2 \left(\frac{\alpha + 2}{2} - \beta\right)$$

Let us determine the characteristics of the clique collection that minimizes the total cost.

- (a) Suppose $\beta > \frac{a+2}{2}$, i.e. $\frac{a+2}{2} \beta < 0$. Then the function is minimized when $\sum_{i=1}^k x_i^2$ is maximized, which happens when k=1 and $x_1=n$. Hence the optimum is an n-clique.
- (b) When $\beta < \frac{a+2}{2}$, the function is maximized when $\sum_{i=1}^k x_i^2$ is minimized. We can derive by a similar argument that the empty graph is optimal.
- (c) When $\beta = \frac{a+2}{2}$, the cost function is constant so any collection of cliques is optimal.

Note that for $\alpha=2$, analysis of both cases hold. Hence in this case, for $\beta>2$, the optimum is either a complete graph or a star. \square

3. PRICE OF ANARCHY

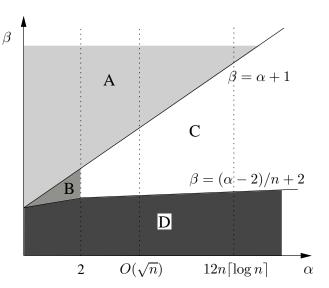


Figure 2: Price of anarchy

In the previous section we determined the optimum solutions for all configurations of (α, β, n) . This is a key step in determining the worst-case Nash equilibrium in the game. We will show that the equilibrium graphs are connected if $\beta > \alpha + 1$. This result prompts the analysis of the price of anarchy to be split up by cases.

Definition 1. We will be referring to the following regions, which can also been seen graphically on figure 2.

- Region A is the area defined by $\beta > \alpha + 1$.
- Region B is defined by $(\alpha + 2)/2 \le \beta \le \alpha + 1$ and $0 \le \alpha < 2$.
- Region C is defined by $(\alpha 2)/n + 2 \le \beta \le \alpha + 1$ and $2 \le \alpha$.
- Finally, region D is the remaining area, that is $0 \le \beta < (\alpha + 2)/2$ when $0 \le \alpha < 2$ and $0 \le \beta < (\alpha 2)/n + 2$ when $\alpha \ge 2$.

3.1 Region A

Lemma 3.1. When $\beta > \alpha + 1$, all equilibrium graphs are connected.

PROOF. Suppose that some equilibrium graph consists of two or more connected components. Let i be the node corresponding to some player. The β -term of the cost function for this node equals βn_i where n_i is the total number of nodes in the other components. If i creates an edge to each of the n_i nodes, we will argue that the cost decreases and thus reach a contradiction. The β -term of i's cost function becomes 0 whereas the α -term increases by αn_i for creating the edges. Additionally, the distance term increases by n_i . In summary, i's cost decreases by $n_i(\beta - (\alpha + 1)) > 0$. \square

Theorem 3.2. For $\beta > \alpha + 1$, a network is at equilibrium in our model if and only if it is an equilibrium graph in the old model for the same α and n.

PROOF. We know from (3.1) that equilibrium graphs in this region are connected and hence viable outcomes for a (α, ∞, n) configuration. When $\beta < \infty$, the strategies of a node in an equilibrium graph are more than those available when $\beta = \infty$. No node can thus decrease its cost, so this network is also at Nash equilibrium for $\beta = \infty$.

Suppose that there is a Nash equilibrium in the an (α, ∞, n) configuration that is not a Nash equilibrium in for (α, β, n) for any $\beta > \alpha + 1$. This implies that some player i in our model can improve her cost by disconnecting parts of the network. Assume that after doing this, there are n_i nodes are in other connected components. We will prove that i can improve her situation in both models, which contradicts the assumption of Nash equilibrium in the $\beta = \infty$. The lack of connectivity to the n_i vertices costs the player βn_i , whereas by building edges to each of them and end up with a connected graph, she would be charged $(\alpha + 1)n_i$. Since $\beta > \alpha + 1$ by assumption, the latter option is preferable and the contradiction is reached. \square

The following corollary is immediate.

COROLLARY 3.3. The price of anarchy is equal for (α, ∞, n) and (α, β, n) for any $\beta > \alpha + 1$. \square

3.2 Region B

LEMMA 3.4. Let $\beta \leq \alpha + 1$. A connected component C of size k in an equilibrium graph for (α, β, n) is an equilibrium graph for (α, ∞, k) .

PROOF. Suppose that some equilibrium graph N contains a connected component C of size k, such that C is not an equilibrium for the (α, ∞, k) . Then there is at least one player i in C who can deviate from his strategy to another that does not alter the connectivity of the graph in any way (and hence the β -term of his cost), but improves his overall cost. Hence, N is not an equilibrium graph for (α, β, n) either and we're done. \square

We will now give a rather technical lemma that holds for regions B and C.

Definition 2. Given a configuration (α, β, n) of our game we define S to be the set of all component sizes that occur in equilibria in the game. More specifically, $k \in S$ iff there exists a component of size k in some equilibrium graph of game.

LEMMA 3.5. Let (α, β, n) be a configuration such that the parameters are within regions B or C. The price of anarchy ρ is $O\left(\rho' + \frac{\beta n(n-1)}{\cos t^{\beta}(OPT)}\right)$ where ρ' is the maximum price of anarchy for (α, ∞, n') where $n' \in S$.

PROOF. Let OPT denote the optimum network on n nodes and N be an equilibrium graph. N is the union of components N_1, \ldots, N_k , each of which is at Nash equilibrium for (α, ∞, n_i) where n_i is the size of N_i as was seen in (3.4).

When $\alpha \geq 2$ the optimum is a star, so by ignoring β -terms

(considering $cost^0(\cdot)$) we obtain

$$\sum_{i=1}^{k} \cot^{0}(OPT_{i}) = \sum_{i=1}^{k} 2(n_{i} - 1)^{2} + \alpha(n_{i} - 1)$$

$$= 2\sum_{i=1}^{k} n_{i}^{2} + (\alpha - 4)\sum_{i=1}^{k} n_{i} + \sum_{i=1}^{k} (2 - \alpha)$$

$$\leq 2\left(\sum_{i=1}^{k} n_{i}\right)^{2} + (\alpha - 4)n + k(2 - \alpha)$$

$$\leq 2n^{2} + (\alpha - 4)n + 2 - \alpha = \cot^{\beta}(OPT).$$

When $\alpha < 2$ the optimum is a clique, so $\cos^0(OPT_i) = (\alpha + 2)\binom{n_i}{2}$ and thus

$$\sum_{i=1}^{k} \operatorname{cost}^{0}(\operatorname{OPT}_{i}) = \frac{\alpha+2}{2} \left(\sum_{i=1}^{k} n_{i}^{2} - \sum_{i=1}^{k} n_{i} \right)$$

$$\leq \frac{\alpha+2}{2} (n^{2} - n) = \operatorname{cost}^{\beta}(\operatorname{OPT}).$$

Since $\operatorname{cost}^{\beta}(N) \leq \beta n(n-1) + \sum_{i=1}^{k} \operatorname{cost}^{0}(N_{i})$ we now deduce that the price of anarchy ρ in the region is bounded by

$$\rho = \max_{N} \frac{\cot^{\beta}(N)}{\cot^{\beta}(OPT)}$$

$$\leq \max_{N} \frac{\sum_{i=1}^{k} \cot^{0}(N_{i}) + \beta n(n-1)}{\cot^{\beta}(OPT)}$$

$$\leq \max_{N} \frac{\sum_{i=1}^{k} \cot^{0}(N_{i})}{\sum_{i=1}^{k} \cot^{0}(OPT_{i})} + \frac{\beta n(n-1)}{\cot^{0}(OPT)}$$

We conclude that

$$\rho \le \max_{N} \frac{\cos^{0}(N_{i})}{\cot^{0}(OPT_{i})} + \frac{\beta n(n-1)}{\cot^{\beta}(OPT)} = \rho' + \frac{\beta n(n-1)}{\cot^{\beta}(OPT)}.$$

Theorem 3.6. The price of anarchy in region B is O(1).

PROOF. We have already proven that in region B the optimum is the complete graph. Since $\alpha \leq 2$ we know from earlier work [11] that the price of anarchy in this region is at most constant, namely $\frac{4}{3}$. Using the previous lemma we can now deduce that

$$\rho \le \rho' + \frac{\beta n(n-1)}{\cot^{\beta}(OPT)} \le \frac{4}{3} + \frac{\beta n(n-1)}{(\alpha+2)n(n-1)/2} = \frac{4}{3} + \frac{2\beta}{\alpha+2}.$$

However, we have that $\beta \leq \alpha + 1$ and $\alpha \leq 2$ in this region, so we find that

$$\rho \le \frac{4}{3} + \frac{2\alpha + 2}{\alpha + 2} \le \frac{4}{3} + \frac{3}{2} = \frac{17}{6}$$

Moving on to region C, we apply different results to the region.

LEMMA 3.7. The price of anarchy ρ in the C region is bounded from below by $\beta n/(2n + \alpha - 2)$.

PROOF. The empty graph is trivially an equilibrium graph in the region since $\beta \leq \alpha + 1$. The cost of an empty graph is $\beta n(n-1)$, whereas the cost of the optimum, a single star,

is $\alpha(n-1) + 2(n-1)^2$. Therefore the price of anarchy is bounded from below by

$$\max_{N} \frac{\cot^{\beta}(N)}{\cot^{\beta}(OPT)} \geq \frac{\beta n(n-1)}{\alpha(n-1) + 2(n-1)^{2}} = \frac{\beta n}{2n + \alpha - 2}.$$

An immediate corollary of (3.5) is an upper bound on the price of anarchy in this region. Indeed, merely applying the cost of the optimum graph (the star) to the lemma we obtain the following bound.

COROLLARY 3.8. For any configuration (α, β, n) in region C, the price of anarchy ρ in region C is $O(\rho' + \frac{\beta n}{\alpha + 2n - 2})$ where ρ' is the maximum price of anarchy for (α, ∞, n') where $n' \in S$. \square

Combining this corollary with the previous lemma gives a tight characterization of the price of anarchy in the area where $\alpha > 12n\lceil\log n\rceil$ in the C region.

Theorem 3.9. The price of anarchy in the part of region C where $\alpha > 12n\lceil \log n \rceil$ is $\Theta\left(\frac{\beta n}{\alpha + 2n - 2}\right)$.

PROOF. Lemma (3.7) provides the desired lower bound for the price of anarchy. Hence, in order to complete the proof we need a corresponding upper bound of $O(\frac{\beta n}{\alpha+2n-2})$. However, according to lemma (3.4), any Nash equilibria in this area will consist of connected components which in isolation would be Nash equilibria for $\beta = \infty$. Since $\alpha > 12n\lceil\log n\rceil$, and the size n_i of any such connected component is less than n, we immediately get that $\alpha > 12n_i\lceil\log n_i\rceil$. The price of anarchy for any such component is constant, as was seen in the introduction, and thus ρ' as defined in (3.5) will be constant. This fact, along with the corollary above, implies the upper bound and the proof is complete. \square

3.3 Region D

Before we delve into the specifics of the price of anarchy for the remaining region where the empty graph is optimal, we point out that the β -term in the cost function does not alter the price of anarchy significantly. In fact, it contributes at most 1 to it because the β -term is at most equal to the cost of the optimum graph.

Upper Bounds

LEMMA 3.10. Given (α, β, n) in region D, the price of anarchy is

$$O\left(\rho_{\infty}\frac{cost^{\beta}(OPT_{\infty})}{\beta n(n-1)}\right),$$

where ρ' is the price of anarchy for the same (α, ∞, n) α and n, and $\beta = \infty$.

PROOF. As mentioned above, the β -term of the cost function contributes at most 1 to the price of anarchy in this region. Hence, we can disregard it altogether in this analysis since it will not cause differences in the asymptotical characteristics of the system. Instead we focus on identifying the equilibria that maximize the other two terms of the cost, the α -term and the sum of distances. Naturally, both are maximized when the equilibrium graph has a single component, so it suffices to explore connected equilibria to bound the price of anarchy. We will denote by OPT the optimum

graph for the given instance of the game and by OPT_{∞} the optimum graph for the instance of the game with (α,∞,n) . As usual, N denotes any connected equilibrium graph for the given instance of the model. By definition, we can bound the the price of anarchy ρ as follows.

$$\rho \leq \max_{N} \frac{\cot^{\beta}(N)}{\cot^{\beta}(OPT)} + 1 = \max_{N} \frac{\cot^{\beta}(N)}{\beta n(n-1)} + 1$$

$$= \max_{N} \frac{\cot^{\beta}(N)}{\cot^{\beta}(OPT_{\infty})} \cdot \frac{\cot^{\beta}(OPT_{\infty})}{\beta n(n-1)} + 1$$

$$= \rho' \frac{\cot^{\beta}(OPT_{\infty})}{\beta n(n-1)} + 1.$$

We have shown that for $\alpha \leq 2$, the optimum when $\beta = \infty$ is the complete graph, and therefore $\mathrm{cost}^{\beta}(\mathrm{OPT}_{\infty}) = \frac{\alpha+2}{2}n(n-1)$, whereas the price of anarchy ρ' is at most 4/3. Combining these bounds with the previous lemma gives an upper bound on the price of anarchy.

COROLLARY 3.11. The price of anarchy in the part of region D where $\alpha \leq 2$ is $O(1/\beta)$. \square

Similarly, we have shown that for $\alpha>2$, the optimum when $\beta=\infty$ is the star, and hence $\cot^\beta(\mathrm{OPT}_\infty)=\alpha(n-1)+2(n-1)^2$ in this case. As far as the price of anarchy ρ' is concerned, it is constant for $2<\alpha<\sqrt{n}$, whereas for $\sqrt{n}<\alpha<12n\lceil\log n\rceil$ the best known upper bound for it is $O(1+(\min\{\frac{\alpha^2}{n},\frac{n^2}{\alpha}\})^{1/3})$. Hence, by using (3.10) we obtain the following bounds.

COROLLARY 3.12. The price of anarchy in the part of region D, where $2 < \alpha < \sqrt{n}$ is $O(1/\beta)$. In the same region for $\sqrt{n} \le \alpha < 12n\lceil \log n \rceil$ the price of anarchy is $O((1+(\min\{\frac{\alpha^2}{n},\frac{n^2}{\alpha}\})^{1/3})\frac{\alpha+2n-2}{\beta n})$. \square

The following lemma gives some insight into the structural characteristics of the equilibrium graphs in this area.

Lemma 3.13. The connectivity of equilibrium graphs in the D region is resilient to any move by a single player. Namely, if a node removes any number of the edges it bought, the number of components will remain the same.

PROOF. Suppose that some player could by delete k of the edges he bought and split his component into two, E and E'. We can assume he stays in E', and that none of the k edges are redundant, i.e. E and E' remain connected if k-1 or fewer edges are removed. Note that these edges all stretch from E' into E.

Now, after deleting his edges, the player will incur a cost of $\beta|E|$ due to the lack of connectivity to the E component. However, before the move his he was paying at least $k\alpha+k+2(|E|-k)$ for his connections to E. By the definition of the D region, we have that $\beta<\frac{k}{|E|}(\alpha-1)+2,$ so $\beta|E|< k\alpha+k+2(|E|-k).$ This means that the player can profitably deviate from his strategy, which in turn implies that the original graph was not an equilibrium. \Box

One immediate implication of the above lemma is that equilibrium graphs in this area are resilient to the removal of any single edge. It is obvious that some graphs, for instance non-trivial trees (i.e. trees with at least 2 nodes), do not have this property. Hence we deduce the following.

Corollary 3.14. An equilibrium graph contains no non-trivial trees. \square

Theorem 3.15. In region D where $\alpha \geq 12n\lceil \log n \rceil$ the price of anarchy is 1.

PROOF. According to lemma (3.4), any Nash equilibrium graph in this area consists of connected components which, if viewed in isolation, would be equilibria for $\beta = \infty$. Since $\alpha > 12n\lceil\log n\rceil$, and the size of any such connected component n_i is less than n, we immediately get that $\alpha > 12n_i\lceil\log n_i\rceil$ and thus any component in the equilibrium graph has to be a tree. However, by corollary (3.14) above we conclude that the only equilibrium graph is actually the empty graph, which happens to be optimal. The result follows. \square

Lower Bounds

It will be useful in the following analysis to have handy notation for determining the owner of an edge. We will thus give edges direction such that a directed edge (v,u) indicates that player v built an edge to u.

LEMMA 3.16. In the D region for $\alpha < \sqrt{n/2}$, the price of anarchy is $\Omega(\frac{1}{\beta})$.

The proof is given in the appendix.

Combining propositions (3.11), (3.12) and (3.16) yields the following theorem.

THEOREM 3.17. The price of anarchy is
$$\Theta\left(\frac{1}{\beta}\right)$$
 in region D for $\alpha < \sqrt{n/2}$. \square

The graph described in the proof of (3.16) has several interesting properties. It is a strict Nash equilibrium graph in our model, even when there is no incentive ($\beta = 0$) which follows easily from the proof. It has diameter of 2, a unique shortest path between pairs of vertices [1] and is resilient to the removal of edges owned by a single player.

4. FURTHER EXTENSIONS

The network formation game that was defined by Fabrikant et al. and which we have extended was unable to handle networks that consist of multiple components. As mentioned earlier, this drawback deprives it of the opportunity to express interesting dynamical procedures and phenomena which are inherent in the process of network formation. In this section, we will attempt a shallow exploration of these deep waters.

4.1 Coalitions

Moving back to our model, we have the luxury of being able to consider every possible configuration as a starting configuration. The starting position of the empty graph has special significance though, since it actually simulates the birth of a network. Given α and n we have that for any $\beta > \alpha + 1$, the network creation process starts off and the Nash equilibria that it reaches are not only connected but also coincide with the equilibria of the old model where essentially $\beta = \infty$. Although this is interesting in its own right as it suggests an incentive threshold (that we express by β), above which there are no new properties of the final stable network, the region below the $\beta = \alpha + 1$ is really interesting.

For $\beta \leq \alpha + 1$, the starting configuration of the empty graph is not only a valid strategy profile for our players but also a Nash equilibrium. Basically, the game gets stuck in the initial, empty network due to the selfish behavior of the players. Although this kind of behavior is not only desired but optimal in region D and its effects are insignificant in region B since it only increases the total social cost by a constant, it severely deteriorates the social welfare of the network in the C region.

As we have already seen, the cost of the empty graph in this region can become arbitrarily worse than the cost of the optimal graph, the star. Hence we can have a number of short-sighted selfish players who basically force themselves into paying arbitrarily higher cost than the optimal. In real-life scenarios, this would be the ideal setting to start talking to your neighbors and this is what we will examine next by introducing coalitions to the game.

We will model coalitions in our game based on a collusion framework that is closely related to the one used by Hayrapetyan et al. in [12]. In essence, the members of a coalition cooperate so as to selfishly maximize their collective welfare. In other words, the coalitions may be viewed as super-players that control all of the members of their respective coalitions and choose strategies for them. Their cost function is the aggregate cost of all members of its coalition. The purpose of the super-player is naturally to maximize his utility, i.e. minimize his cost. Lastly, the coalitions are static and defined as a part of the description of the game. In our specific model of network creation, we start from an empty graph and all the costs and incentives are uniform. It is evident that all players within a coalition are isomorphic, so all we need to define are the respective sizes of the coalitions. We will partition the players into coalitions of equal sizes for our purposes, although allowing differently sized coalitions would produce similar results.

Returning to the specific problem in region C where individual selfish players are trapped in a setting where they can be be forced to pay arbitrarily higher costs than they optimally would. A rather natural question that arises asks whether we can overcome this deadlock situation by allowing coalitions which are not large. The answer to this question is negative.

THEOREM 4.1. Even if $S \in o(\sqrt{n})$ players collude, there are instances in the network formation model where the price of anarchy remains unbounded in region C.

PROOF. We will prove that even if we allow coalitions of size $o(\sqrt{n})$, then we can still get equilibria of this new game, where the price of anarchy is unbounded. Specifically, we can set the parameters of the model α , β and n such that its price of anarchy can be bounded from below by any k where k is a positive number k > 1.

Let's examine the outcome of the game where the only edges that are bought are between nodes in the same coalition and every coalition forms a star. We will prove that we can set the parameters of the model α, β and n such that this outcome constitutes an equilibrium graph then its cost could become arbitrarily worse than the cost of the optimum in the C region.

We start by setting α big enough so that it is too expensive for any super-player to buy more than one edge towards another component. Once an edge is bought towards another component, any other edge towards the same component has

an effect only on the distance-term of the cost to that superplayer. In order to make sure that α would be too expensive for any super-player to buy more than one edges towards another component all we need to do is set α to be greater than improvement that can be caused to the distance-term of his cost by the creation of a single edge. A trivial lower bound of α that would achieve this would be n^3 . Now, we want to be the case that the super-player has no incentive to actually create any edges towards other components. However, the cost that a disconnected components force upon him is $\sum_{\text{all his nodes}} \beta S = \beta S^2$. Now, we want this cost to less than the smallest possible cost that the super-player would pay if we actually created that single edge to him. Naturally, the super-player would build that edge towards the center of the other star-component. In this case the cost to this player would be at least

$$\alpha + \sum_{\substack{\text{members } i \\ j \text{ in other components}}} \operatorname{dist}(i,j) \geq \alpha + 3S^2 - 2S$$

which is the cost if the super-player's component was a star and bought an edge connecting the two centers. We want the cost of being disconnected to any component to be less (or equal) to the minimal cost of being connected to that. By the analysis above however it is sufficient to have $\beta S^2 = \alpha + 3S^2 - 2S$ or in other words to set β equal to

$$\beta = \frac{\alpha}{S^2} + 3 - \frac{2}{S}.$$

However, since $S=o(\sqrt{n})$, for any positive number $k\geq 1$ we can find a big enough n such that $\beta\geq 2k(\frac{\alpha-2}{n}+2)$. Regardless of what is the exact relation of S to n, since $S=o(\sqrt{n})$, for big enough n the slope of the line $\beta=2k(\frac{\alpha-2}{n}+2)$ for any given k will become less than the slope of the line $\beta=\frac{\alpha}{S^2}+3-2\frac{1}{S}$, which in turn is less than the slope of the line $\beta=\alpha+1$. If this price of n is less than 2S we set it equal to 2S. Given, the fact that we have only a lower bound on α we can finally set α big enough, such that the α,β,n parameters of our model are such that $\alpha>2$, $\beta\leq\alpha+1$ and $\beta\geq\frac{\alpha-2}{n}+2$). In other words, we have created an instance of a network creation game in the C region of our new model (with collusions) such that the graph that we defined earlier and in which every coalition formed a star and no other edges were formed is a Nash equilibrium.

All we need to prove in order to finish the proof of this theorem is that the cost of this graph is at least k times the cost of the optimum graph which in this region is the star. However, the social cost of the graph we have described is $\beta n(n-S) + \frac{n}{S}(\alpha(S-1)+2(S-1)^2)$ and the cost of the star is $\alpha(n-1)+2(n-1)^2$. Hence, we conclude that

$$\frac{\operatorname{cost}^{\beta}(\operatorname{Nash graph})}{\operatorname{cost}^{\beta}(\operatorname{OPT})} = \frac{\beta n(n-S) + \frac{n}{S}(\alpha(S-1) + 2(S-1)^{2})}{\alpha(n-1) + 2(n-1)^{2}}$$

$$\geq \frac{(\beta/2)n(n-1)}{\alpha(n-1) + 2(n-1)^{2}}$$

$$= \frac{\beta n/2}{\alpha + 2n - 2} \geq k$$

and the proof is complete. \Box

5. ACKNOWLEDGMENTS

We are most grateful to Éva Tardos for her invaluable guidance. We would also like to thank Ara Hayrapetyan, Tom Wexler and Raghuram Ramanujan for helpful discussions.

6. REFERENCES

- S. Albers, S. Eilts, E. Even-Dar, Y. Mansour and L. Roditty. On Nash Equilibria for a Network Creation Game. ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 89–98, 2006.
- [2] E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler and T. Roughgarden. The Price of Stability for Network Design with Fair Cost Allocation. Symposium on Foundations of Computer Science (FOCS), pages 295–304, 2004.
- [3] E. Anshelevich, A. Dasgupta, E. Tardos and T. Wexler. Near-optimal network design with selfish agents. ACM Symposium on Theory of Computing (STOC), pages 511–520, 2003.
- [4] E. Anshelevich, B. Shepherd and G. Wilfong. Local Peering and Service Contracts in Strategic Network Formation. Symposium on Foundations of Computer Science (FOCS), 2006.
- [5] V. Bala and S. Goyal. A Non-Cooperative Model of Network Formation. In *Econometrica 68*, pages 1181–1229, 2000.
- [6] F. Bloch and M. O. Jackson. The Formation of Networks with Transfers among Players. To appear in Journal of Economic Theory, 2006.
- [7] A. Blokhuis and A. E. Brouwer. Geodetic graphs of diameter two. *Geometricae Dedicata*, 25, pages 526–533, 1988.
- [8] G. Christodoulou and E. Koutsoupias. On the Price of Anarchy and Stability of Correlated Equilibria of Linear Congestion Games. European Symposium on Algorithms (ESA), pages 59–70, 2005.
- [9] G. Christodoulou and E. Koutsoupias. The Price of Anarchy of Finite Congestion Games. ACM Symposium on Theory of Computing (STOC), pages 67–73, 2005.
- [10] J. Corbo and D. Parkes. The Price of Selfish Behavior in Bilateral Network Formation. Symposium on Principles of Distributed Computing (PODC), pages 99–107, 2005.
- [11] A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou and S. Shenker. On a Network Creation Game. Symposium on Principles of Distributed Computing (PODC), pages 347–351, 2003.
- [12] A. Hayrapetyan, E. Tardos and T. Wexler. The Effect of Collusion in Congestion Games. ACM Symposium on Theory of Computing (STOC), pages 89–98, 2006.
- [13] M. O. Jackson. A Survey of Models of Network Formation: Stability and Efficiency. In Group Formation in Economics: Networks, Clubs and Coalitions. G. Demange and M. Wooders, eds. Cambridge, 2004.
- [14] M. O. Jackson and A. Wolinsky. A Strategic Model of Social and Economic Networks. *Journal of Economic Theory*, 71, pages 44–74, 1996.
- [15] R. Johari, S. Mannor and J. N. Tsitsiklis. A Contract-Based Model for Directed Network Formation. To appear in *Games and Economic Behavior*, 56(2), 2006.

- [16] E. Koutsoupias and C. Papadimitriou. Worst-case Equilibria. Symposium on Theoretical Aspects of Computer Science (STACS), 1999.
- [17] T. Moscibroda, S. Schmid, and R. Wattenhofer. On the topologies formed by selfish peers. In Proc. 25th ACM Symp. on Principles of Distributed Computing, 2006.
- [18] R. Myerson. Graphs and Cooperation in Games. Mathematics of Operations Research, 2, pages 225–229, 1977.
- [19] J. F. Nash. Non-cooperative games. Annals of Mathematics 54, pages 289–295, 1951.
- [20] M. E. J. Newman. The structure and function of complex networks. SIAM Review, 45, pages 167–256, 2003.
- [21] C. Papadimitriou. Algorithms, Games and the Internet. ACM Symposium on Theory of Computing (STOC), pages 749–753, 2001.
- [22] T. Roughgarden. The Price of Anarchy is Independent of the Network Topology. ACM Symposium on Theory of Computing (STOC), pages 428–437, 2002.
- [23] T. Roughgarden and E. Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2), pages 236–259, 2002.

APPENDIX

Proof of Lemma 3.16. Suppose $\alpha \leq 1$.

We will prove that in this area a clique can be built from the players participating in it in such a manner that it is a Nash equilibrium for every $\alpha \leq 1$. Indeed, consider a directed n-cycle where each player has bought exactly one edge to his neighbor. Convert this construct into a complete graph by filling in the remaining edges at random. This yields an equilibrium graph. Truly, no player can disconnect himself from the graph or partition it in any way by dropping some of his edges. The only possible deviation is for him to destroy some of his edges, but by dropping any of his edges he faces the penalty of paying for the extra distance cost of 1 for every edge. However, we have that $\alpha \leq 1$, and therefore the specific clique with the orientation that we described above constitutes a Nash equilibrium. This in turn implies the desired bound.

$$\rho = \frac{\mathrm{cost}^{\beta}(\mathrm{worst\ Nash})}{\mathrm{cost}^{\beta}(\mathrm{OPT})} \ge \frac{\frac{\alpha+2}{2}n(n-1)}{\beta n(n-1)} = \frac{\alpha+2}{2\beta} \ge \frac{1}{\beta}.$$

Now suppose $1 < \alpha < \sqrt{n/2}$.

We will use the following construction of Albers et al. [1] to prove this claim. It yields a connected graph that we will show is a Nash equilibrium in our model. The distance between distinct nodes is at least 1, so the sum of distances over all nodes is at least n(n-1)/2. The optimum in this region is the empty graph, so we can bound the price of anarchy trivially by $\rho \geq \frac{n(n-1)/2}{\beta n(n-1)} = \frac{1}{2\beta}$. Let q be a prime power, and F denote the finite field of q

Let q be a prime power, and \hat{F} denote the finite field of q elements. Set $A = F \times F$ and let $\mathcal{L} = \bigcup_{a,b \in A} \{a+bi \mid i \in F, b \neq 0\}$ denote the set of $\binom{q^2}{2}/\binom{q}{2} = q(q+1)$ lines. (A,\mathcal{L}) defines an affine plane of order q. Two lines are parallel if they are disjoint or equal, and so parallelism defines an equivalence relation on the lines in \mathcal{L} . Each of the q+1 equivalence classes contains q lines that in turn contain q

points each.

The simple graph G=(V,E) that we construct contains a vertex for each point and every line, that is $V=A\cup\mathcal{L}$. There is an edge in E between every pair of parallel lines, and between a line and each vertex it contains. These are the only edges in the graph.

We will now determine who constructed what edge. Firstly, since all parallel lines are interconnected G has a q-clique for every equivalence class of lines. It is possible to orient the edges in each of the cliques so that the difference between the out-degree and in-degree of every line in the clique is 0 if q is odd and 1 if q is even. Secondly, to orient edges between points and lines, we start by letting L^0, \ldots, L^q denote representatives from each of the equivalence classes. We arbitrarily order the lines in the equivalence class of L^q as L_0^q, \ldots, L_{q-1}^q . Edges between a line L_i^q for some i and its points are bought by the points. When $0 \le i \le q - 1$, each line L in the equivalence class of L^i builds edges towards the point in which L intersects with L_i^q , and the point in which it intersects with $L_{i+1 \mod q}^q$. The remaining edges are purchased by the points. From the point's point of view, it is contained in L_i^q for exactly one j, and had to buy all of its edges except for two – the edges to the unique lines in the equivalence classes of L^j and $L^{j-1 \mod q}$ that pass through the point.

Note that for q=2 this construction yields the Petersen graph.

Albers et al. prove that this construction is an equilibrium graph for the old model for the aforementioned values of α and n. The only case where it would fail to be an equilibrium graph for our model is if some player could break the graph up into smaller components by removing some of his edges. Without going into detail as this can be easily verified by the avid reader, Albers et al. argue in their proof that no player can increase the number of components even if he removes all of his edges. The scenario above is thus impossible and the proof is complete. \square