A hierarchical network formation model

Atabati, Omid and Farzad, Babak

University of Calgary, Brock University

February 2015

A hierarchical network formation model[★]

Omid Atabati [a,1]   Babak Farzad [b,2]

[a] *University of Calgary, Calgary, Canada*
[b] *Brock University, St Catharines, Canada*

Abstract

We present a network formation model based on a particularly interesting class of networks in social settings, where individuals' positions are determined according to a topic-based or hierarchical taxonomy. In this game-theoretic model, players are located in the leaves of a complete b-ary tree as the seed network with the objective of minimizing their collective distances to others in the network. In the grid-based model of Even-Dar and Kearns [3], they demonstrate the existence of small diameter networks with the threshold of $\alpha = 2$ where the cost of a new link depends on the distance between the two endpoints to the power of α. We show the appearance of small diameter equilibrium networks with the threshold of $\alpha = 1/4$ in the hierarchical or tree-based networks. Moreover, the general set of equilibrium networks in our model are guaranteed to exist and they are pairwise Nash stable with transfers [2].

Keywords: Network formation, Hierarchical networks, Linking game
JEL Classification: D85, C79

---

1 Introduction

The role of network structures in determining the outcome of many social and economic settings has captured increasing attention in recent years. In

---

particular, different researches have demonstrated this significant role in the wide range of problems [4]. As a result, it is crucial to study the formation process of social and economic networks and to know their characteristics, as there is a frequent appearance of certain properties such as small diameter and large clustering coefficient in naturally occurring networks [7]. A growing body of works in strategic or game-theoretic network formation[3] considers the underlying network structure as a set of self-interested individuals and their connections. In a previous effort [3], Even-Dar and Kearns proposed a model that starts from a $\sqrt{n} \times \sqrt{n}$ grid network. Players' objective is to minimize their collective distances to all other players. Forming a link is costly at the fixed price of the initial grid distance between the endpoint players of that link to the power of α. Also, equilibrium analysis is relaxed from Nash equilibrium to link stability. Link stability implies that the network is stable under single link unilateral deviations.[4] They showed the appearance of small diameter link stable networks within the threshold of $\alpha = 2$. Further, Atabati and Farzad [1] expand this result by showing the existence of a small diameter in the general set of equilibrium networks under the new assumption of *dynamic link-pricing* that better represents the dynamic nature of network formation.[5]

Hierarchical networks are known as credible social structures based on observations in social sciences [8]. In particular, previous studies [5,8] have analyzed the hierarchical networks by applying stochastic models. However, less is known from the strategic perspective. We aim to fill this gap by presenting our game-theoretic hierarchical network formation model. The rest of this paper is organized as follows. In Section 2, we explain our model. We then provide our analysis on the diameter of equilibrium networks in Section 3 and the conclusion in Section 4.

2   The model

Let $N = \{1, ..., n\}$ be the set of n players being located at the leaves or the lowest level of a complete b-ary tree, denoted by G_0 as the starting or seed network. Typically in these networks, the distance of each two players is defined to be the height of their lowest common ancestor in the tree.[6] Note

---

[3] See [4] for a comprehensive survey in this topic.
[4] This is as opposed to Nash, which allows for arbitrary link unilateral deviations.
[5] Dynamic link-pricing updates the applied distances of each pair of players in the related link-prices from the current network rather than sticking with the initial grid distances.
[6] Considering a total of n players in the network, the upper bound for the distance between two players is $\log_b n$.

that in the beginning, there is no additional link between the players in the network and distances are defined by the structure of the tree.

Players seek to minimize their collective distances to others in the network. To follow this goal, player i is able to establish links by announcing a vector of transfer payments $\mathbf{t}^i \in \mathbb{R}^{n(n-1)/2}$. The entry t^i_{jk} is i's offer (to pay) or demand (to gain) on the link jk. Direct (indirect) transfer payment is implied when $i \in jk$ ($i \notin jk$). Typically, individuals can make offers (positive transfers) for direct and indirect payments, but only can make demands (negative transfers) for direct transfers. In the equilibrium, there is no excess of payments allowed. Thus, a link exists in the equilibrium network, if the summation of all transfers on that link equals zero. Also, $G(\mathbf{t})$ is the resulted network associated with the strategy vector $\mathbf{t}$. Further, network $G + ij$ is obtained by adding a new link ij and $G - ij$ is resulted by removing an existing link ij.

As there will be added links between the leaves of the starting tree network during the game, we need to define the generalized notion of distance between any two players u and v as follows:
$d_G(u, v) = min\{\big(\delta(u, x_1) + \delta(x_1, x_2) + \cdots + \delta(x_{k-1}, x_k) + \delta(x_k, v)\big)$, the height of lowest common ancestor of u and $v\}$, where $x_1, \ldots, x_k$ are the intermediary players in the shortest path between u and v according to the definition of $\delta(x_i, x_{i+1})$.

$$\delta(x_i, x_{i+1}) = \begin{cases} \text{height of lowest common ancestor} & \text{if } x_i x_{i+1} \notin G(\mathbf{t}) \\ 1 & \text{otherwise} \end{cases}$$

Finally, the *utility function* of player i in the network $G(\mathbf{t})$ is defined as follows:

$$u_i(G(\mathbf{t})) = -\sum_{j \in N} d_{G(\mathbf{t})}(i, j) - \sum_{j \in \mathcal{N}_i} c_{ij} - \sum_{j \in \mathcal{S}_i} p_{ij} - \sum_{jk \in G(\mathbf{t})} t^i_{jk}, \qquad (1)$$

where $\mathcal{N}_i$ denotes the set of i's neighbors, and $\mathcal{S}_i$ is a subset of i's neighbors whose links to i are initiated by i. Also, $p_{ij} = b^{\alpha d_{G(\mathbf{t})}(i,j)}$ denotes the link-price for the link ij.[7] Link-price is charged to the player who initiates the link formation for each link. Moreover, c_{ij}, as it is introduced in [1], illustrates the fixed maintenance cost of link ij in which both i and j should consider this charge for maintaining their links at each of point of analysis during the game. In addition, the process of network formation is defined in terms of *improving paths*.[8] Pairwise stability with transfers is used as the equilibrium

---

[7] This price structure is inspired by the Kleinberg's stochastic model [5], where he provides a probability function for creating a link ij that grows asymptotically as $b^{-\alpha d_G(i,j)}$.

[8] It represents a sequence of changes from one network to another. The changes can emerge when individuals create or sever one link based on the improvement relative to the current

notion in our model. A network $G(\mathbf{t})$ is Pairwise Stable with transfers (PS^t), if (i) $ij \in G \implies u_i(G) \geq u_i(G - ij)$ as well as $u_j(G) \geq u_j(G - ij)$, and (ii) $ij \notin G \implies u_i(G) \geq u_i(G + ij)$ as well as $u_j(G) \geq u_j(G + ij)$.

3 Threshold of $\alpha = 1/4$

Proposition 3.1 introduces the domain of α that is the main interest in our model, as it is the necessary condition for the process of network formation to evolve. The proof of Proposition 3.1 is provided in the Appendix.

Proposition 3.1 *If $0 < \alpha < 1$, the seed network cannot be in equilibrium.*

Theorem 3.2 shows that the hierarchical model generates small diameter networks with the threshold of $\alpha = 1/4$.

Theorem 3.2 *If $\alpha \leq 1/4$, the diameter of any equilibrium network is at most two.*

Proof. Frist, we introduce the two sets S_u^1 and $T_{i,j}^d$. The set S_u^1 contains u and its neighbors in $G(\mathbf{t})$. The set $T_{i,j}^d$ contains the participant players in the benefit of establishing a link from i to j with distance $d = d_G(i,j)$ when this choice of linkage is not beneficial. [9] Therefore, $|T_{i,j}^d| \leq b^{\alpha d} + c_{ij} + t_{ij}^i$, where the left hand side is a lower bound for the benefit of this link creation.

Now, we construct an upper bound for the network size n. For each player such as $u' \in S_u^1$, we allocate $|T_{u',W_{u'}}|$ number of players such as w to the player u' such that $w \notin S_u^1$ and $d_G(u',w) \geq 2$. This allocation must satisfy Inequality 2.

$$|S_u^1|(1 + |T_{u',W_{u'}}|) \geq n \tag{2}$$

$$W_{u'} = \{w \mid w \notin S_u^1, d_G(u',w) \geq 2\}, T_{u',W_{u'}} = \sum_{w \in W_{u'}} T_{u',w}$$

Because of the distance structure in the hierarchical networks, we need to allocate a set $T_{u',w}$ to u' for several vertices like w such that $w \in W_{u'}$ in order to construct an upper bound for n.

Players such as w have different distances in the range of $[2, \Delta]$ from u'. In addition, we can point out that by creating a link from u' to w all players with distance of at most $x = \left\lfloor \frac{d_G(u',w)}{2} \right\rfloor - 1$ from w participate in the benefit of this

---

network.

[9] The participation of these players implies that creating a link from i to j decreases their distances to i.

linkage. This is because, $d_G(u', w) \leq d_G(u', w') + x$. Also, after creation the link $u'w$, by considering the closest possible player w to player u', it can be implied that $1 + x < d_G(u', w) - x$. According to the structure of the seed tree, $b^{\left\lfloor \frac{d_G(u',w)}{2} \right\rfloor - 1}$ is a lower bound for the number of players whose distances to u' are decreased by creation of link $u'w$. Consequently, it can be implied that if we repeat this enumeration "k" times for every such player u', we can build an upper bound for the term $|T_{u', W_{u'}}|$. This upper bound is $k|T^\Delta_{u', w}|$. Further, we can obtain an upper bound of $(n - b)/(b^{\left\lfloor \frac{\Delta}{2} \right\rfloor})$ for k. [10]

The next step is building an upper bound for $|S^1_u|$. Consider an equilibrium network G and two players $u, v \in N$ such that $d_G(u, v) = \Delta$, where Δ is the diameter of G. Therefore, without loss of generality, it is not beneficial for v to establish a link to u. This is noted in Inequality 3, where its left side represents a lower bound for the benefit of player v in establishing link uv by considering the reduced distances from players in the set S^1_u to player v. There are two types of players in the set S^1_u in terms of having distance Δ and $\Delta - 1$ from v and also these players can have new updated distance 1 or 2 from v after creating a link between u and v. As a result, there are three cases of $\Delta - 1, \Delta - 2$, and $\Delta - 3$ for players in S^1_u to participate in the benefit of linkage from u to v.

Note that we consider an upper bound of c for both maintenance costs and direct transfer payments [11] in the following inequality.

$$|S^1_u|\Big(\Delta - (f_1 + 2f_2 + 3f_3) \Big) \leq b^{\alpha\Delta} + 2c \qquad (3)$$

where f_i is the fraction of players with one the above-mentioned possibilities in the reduced distances. Also, $f_1 + f_2 + f_3 = 1$, and $0 < f_i < 1$. As a result, we can rewrite the Inequality 2 as follows:

$$\left(\frac{b^{\alpha\Delta} + 2c}{\Delta - (f_1 + 2f_2 + 3f_3)} \right)\left(1 + (\frac{n - b}{b^{\left\lfloor \frac{\Delta}{2} \right\rfloor}})(\frac{b^{\alpha\Delta} + 2c}{B}) \right) \geq n \qquad (4)$$

Assuming the maximum possible diameter (Δ), the magnitude of left hand side of Inequality 4 is $n^{2\alpha + 1/2}$. Hence, given a sufficiently large size n and if $\alpha \leq 1/4$, any distance bigger than $f_1 + 2f_2 + 3f_3$ clearly contradicts Inequality 4.

---

[10] For each player $u' \in S^1_u$, there are $n - b$ players outside of this set that must be allocated in the set $T^\Delta_{u', w}$. As, there are at least b players in the set S^1_u, an upper bound for $|T^\Delta_{u', w}|$ can be achieved.

[11] Direct transfer covers the loss of utility for a player from the maintenance cost of a link.

As a result, the largest value the diameter can take is 2. [12]   □

4 Conclusion

In this hierarchical network formation model, we show the appearance of small diameter pairwise stable networks with the threshold of $\alpha = 1/4$. We can also go further and show that the set of pairwise stable networks in our model coincides with the broader set of pairwise Nash stable networks with transfers [2]. In addition, the convergence of our model to equilibrium is guaranteed, since the existence of cycles is ruled out. [13]

References

[1] Atabati, O., and B. Farzad, *A strategic model for network formation*, Computational Social Networks **2** (1) (2015), 1–14.

[2] Bloch, F., and M. O. Jackson *Definitions of equilibrium in network formation games*, Int J Game Theory **34** (3) (2006), 305–318.

[3] Even-Dar, E., and M. Kearns, *A small world threshold for economic network formation*, Advances in Neural Information Processing Systems **19** (2007), 385–392.

[4] Jackson, M. O., Social and Economic Networks, Princeton U. Press, Princeton, N.J., 2008.

[5] Kleinberg, J., *Small-world phenomena and the dynamics of information*, Advanced Neural Information Processing Systems **14** 2001, 431–438.

[6] Milgram, S., *The small world problem*, Psychology Today **1** (1967), 61–67.

[7] Watts, D., Six Degrees: The Science of a Connected Age, W. W. Norton, Cambridge, Mass. 2003.

[8] Watts, D., P. Dodds, and M. Newman, *Identity and search in social networks*, Science **296** (2002), 1302–1305.

---

[12] Intuitively, complete hierarchical networks would not be stable for a large range of b. Hence, diameter of two can be met strictly given the large size of network.

[13] This can be implied from our previous work [1], as its arguments do not depend on the structure of the seed network. Thus, $0 \leq \alpha \leq 1/4$ is the necessary and sufficient condition for the emergence of small diameter pairwise Nash stable networks.

A   Appendix

Proof of Proposition 3.1: Let u and v be two players in the tree G such that $d_G(u,v) = d$. By forming the link uv, player u can benefit of reduced distances not only from those players such as w in the neighborhood of v (type i), but also player u can obtain a reduction in her distances to some neighbors of w which formed a link to w during the game (type ii). We can accurately compute the benefit that comes from the players in type i based on the structure of the seed tree as follows: [14]

$$d - 1 + (b-1)(\sum_{i=2}^{d-1} b^{i-2}(d-i)) = \frac{b^{d-1} - 1}{b - 1} \tag{A.1}$$

The decision for creating a link at the beginning of the game only depends on the gained benefit from the players in type i. As a result, we can determine the criteria for the parameter α that allows the game to progress. This criteria is achieved by satisfying Inequality A.2 that indicates the condition for a link creation at the beginning of the game.

$$\frac{b^{d-1} - 1}{b - 1} \geq b^{\alpha d} + c_{uv} + t_{uv}^u \implies 1 \geq \frac{b^{\alpha d}(b-1)}{b^{d-1} - 1} + \frac{(c_{uv} + t_{uv}^u)(b-1)}{b^{d-1} - 1} \tag{A.2}$$

Furthermore, at the beginning of the game, it is necessary for establishing any link such as uv that $t_{uv}^u \geq 0$ as well as $t_{uv}^v \geq 0$. Now, we should find the choice of linkage that provides the largest possible increase in the the utility for the initiator. In the case of $\alpha \leq 1$, with respect to Inequality A.2, the largest utility can be achieved for the most distant choices, $d = \log_b n$. By using this distance in Inequality A.2, we reach to the Inequality A.3. It can be seen that for the case of $\alpha < 1$, Inequality A.3 and consequently Inequality A.2 can be satisfied for a sufficiently large number of values of n.

$$1 \geq \frac{b^{\alpha \log_b n}(b-1)}{b^{\log_b n - 1} - 1} + \frac{(c_{uv} + t_{uv}^u)(b-1)}{b^{\log_b n - 1} - 1} \implies$$
$$0 \geq (b^2 - b)n^\alpha - n + (c_{uv} + t_{uv}^u)(b^2 - b) + b \tag{A.3}$$

It is also worth noting that regarding the Inequality A.3, there is no possible link formation in case of $\alpha \geq 1$, as the cost of linking become too large. Hence, the interesting and meaningful domain for α that defines the link-prices in our model is between zero and one. □

---

[14] Equation A.1 can be realized, since in the structure of the complete b-ary tree, there are $b^{k-1}(b-1)$ players in which their distance to u can be reduced by $d - k$ for $k = 1, 2, ..., (d-1)$.