HAL
open science

Designing virus-resistant, high-performance networks: a game-formation approach

Stojan Trajanovski, Fernando A Kuipers, Yezekael Hayel, Eitan Altman, Piet Van Mieghem

▶ **To cite this version:**

Stojan Trajanovski, Fernando A Kuipers, Yezekael Hayel, Eitan Altman, Piet Van Mieghem. Designing virus-resistant, high-performance networks: a game-formation approach. IEEE Transactions on Control of Network Systems, 2017. hal-01575073

HAL Id: hal-01575073

https://inria.hal.science/hal-01575073v1

Submitted on 17 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Designing virus-resistant, high-performance networks: a game-formation approach

Stojan Trajanovski, *Member, IEEE,* Fernando A. Kuipers, *Senior Member, IEEE,* Yezekael Hayel, *Member, IEEE,* Eitan Altman, *Fellow, IEEE,* and Piet Van Mieghem *Member, IEEE*

***Abstract*—Designing an optimal network topology while balancing multiple, possibly conflicting objectives like cost, performance, and resiliency to viruses is a challenging endeavor, let alone in the case of decentralized network formation. We therefore propose a game-formation technique where each player aims to minimize its cost in installing links, the probability of being infected by a virus and the sum of hopcounts on its shortest paths to all other nodes.**

In this article, we (1) determine the Nash Equilibria and the Price of Anarchy for our novel network formation game, (2) demonstrate that the Price of Anarchy (PoA) is usually low, which suggests that (near-)optimal topologies can be formed in a decentralized way, and (3) give suggestions for practitioners for those cases where the PoA is high and some centralized control/incentives are advisable.

***Keywords*—Game theory; Virus spread; Network performance; Network design; Networks of Autonomous Agents.**

I. Introduction

Designing communication and computer networks are complex processes in which careful trade-offs have to be made with respect to performance, resiliency/security and cost investments. For instance, if a host in a computer network wants to route traffic to multiple other hosts, it could directly connect to those other hosts, in this way increasing its expenses in installing and maintaining these connections and at the same time also becoming more susceptible to viruses from those other hosts. In return, it would obtain a better and faster performance with minimum delays, compared to when it would have used intermediate hosts as relays. Although in this example, both installation costs and risk to viruses are increasing, they are linearly independent and they do not necessarily optimize together. Indeed, reducing the number of direct connections would reduce the cost and the host would be less vulnerable to viruses. However, even when being connected to a few high-degree nodes with direct connections, the host would still be seriously imposed to a virus.

In practice, hosts often are autonomous, act independently and do not coordinate as in P2P networks [2], peering relations between Autonomous Systems [3], overlay networks [4], wireless [5], [6], [7] and mobile [8] networks, resource sharing in VoIP networks [9], social networks [10], [11] or the Internet [12]. Their aim is to optimize their own utility functions, which are not necessarily in accordance to the global optimum. To study global network formation under autonomous actors, the network formation game (NFG) framework [13] has been proposed. However, resilience and notably virus protection have not been taken into account in that NFG context, even though their importance is undisputed. In this paper, we therefore take the NFG framework one step further by including performance and virus protection as key ingredients. Virus propagation will be modeled by the Susceptible-Infected-Susceptible (SIS) model [14] and we will evaluate the effect of uncoordinated autonomous hosts *versus* the optimal network topology via standard game-theoretic concepts, such as Nash Equilibria and the Prices of Anarchy and Stability.

Our network formation game is called the *Virus Spread-Performance-Cost* (VSPC) game. Each node (i.e., autonomous player) attempts to minimize both the cost and infection probability, while still being able to route traffic to all the other nodes in a small number of hops. When the hopcount performance metric is irrelevant, the game is driven by the cost and virus objectives; a scenario we studied in [1]. That particular scenario resulted in sparse graphs, which may not always represent real-world networks, but it helped to understand the process of virus spread better. In this paper, we generalize those results by also including the hopcount performance metric. The probability of the node being infected and the hopcounts to the other nodes change in a different direction, for example adding a link reduces the former, but increases the latter. Therefore, there is a tradeoff in the number of added links and how these new links are best added. Moreover, the two metrics are linearly independent and closed-form expressions do not exist, which makes the problem complex. Finally, the inclusion of the hopcount allows us to better capture realistic networks. In particular, our main contributions are the following:

- We provide a complete characterization of the various relevant parameter settings and their impact on the formation of the topologies.
- We show that depending on the input, the Nash Equi-

S. Trajanovski is now with Philips Research, Eindhoven and Delft University of Technology, The Netherlands. The research was done while S. T. was with Delft University of Technology (stojan.trajanovski@philips.com).

F. A. Kuipers and P. Van Mieghem are with Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, P.O. Box 5031, 2600 GA Delft, The Netherlands ({F.A.Kuipers, P.F.A.VanMieghem}@tudelft.nl).

Y. Hayel is with University of Avignon, Avignon, France (yezekael.hayel@univ-avignon.fr).

E. Altman is with University of Cote d'Azur, INRIA, BP95, 06902 Sophia Antipolis, France. He is also member of LINCS, 23 Ave. d'Italie, 75013 Paris, France and an associate member of LIA, University of Avignon, France (eitan.altman@inria.fr).

This article builds on our work [1] published in Proc. of IEEE CDC 2015, Osaka, Japan.

libria may vary from tree graphs, via graphs of different diameters, to complete graphs.

- We demonstrate, both via theory and simulations, that the Price of Anarchy (PoA) is small in most of the cases, which implies that (near-)optimal topologies can be formed in a decentralized non-cooperative manner. We will also identify for which scenarios the PoA may be high. In those cases a central point of control would be desirable to limit/steer the players' decisions.

This paper is organized as follows: The SIS-virus spread model and the network-formation game model are introduced in Section II. The *Virus Spread-Performance-Cost* (VSPC) game formation is analyzed in Section III. Related work on game formation and protection against viruses is discussed in Section IV. The conclusion and directions for future work are provided in Section V.

II. Models and problem statements

A. Virus-spread model

The spread of viruses in communication and computer networks can be described, using virus-spread epidemic models [14], [15], [16]. We consider the Susceptible-Infected-Susceptible (SIS) NIMFA model [14], [17],

$$\frac{dv_i(t)}{dt} = \beta(1 - v_i(t))\sum_{j=1}^{N} a_{ij}v_j(t) - \delta v_i(t) \tag{1}$$

where N is the number of network nodes and $v_i(t)$ is the probability of node i being infected at time t, for all $i \in \{1, 2, \ldots, N\}$. If a link is present between nodes i and j, then $a_{ij} = 1$, otherwise $a_{ij} = 0$. In (1), a host with a virus can infect its direct healthy neighbors with rate β, while an infected host can be cured at rate δ, after which the node becomes healthy, but susceptible again to the virus. The probability $v_i(t)$ depends on the probabilities $v_j(t)$ of the neighbors j of node i and there is no trivial closed form expression for $v_i(t)$. The model incorporates the network topology and is thus more realistic than the related population dynamic models. The model relies on the network topology, which makes it more realistic than the related population dynamic models. The goodness of the model has been evaluated in [18]. The probability of a node being infected in the metastable regime, denoted by $v_{i\infty}$, where $\frac{dv_i(t)}{dt} = 0$ and $v_{i\infty} \neq 0$, follows from (1) as [14],

$$v_{i\infty} = 1 - \frac{1}{1 + \tau\sum_{j=1}^{N} a_{ij}v_{j\infty}} \tag{2}$$

where $\tau = \frac{\beta}{\delta}$ is called the *effective infection* rate. The epidemic threshold τ_c is defined as a value of τ, such that $v_{i\infty} > 0$ if $\tau > \tau_c$, and otherwise $v_{i\infty} = 0$ for all $i \in \{1, 2, \ldots, N\}$. The value of $v_{i\infty}$ depends of the values of all $v_{j\infty}$ for all the neighbors j of i, so the network topology and the interconnectivity have impacts on $v_{i\infty}$s.

B. Game-formation model

In our network formation game, each player i (a node in the network) aims to minimize its own *cost function* J_i, and the *social cost* J is defined as $J = \sum_{i=1}^{N} J_i$. Specifically, the *optimal social cost* is the smallest social cost over all possible connected topologies. We look for the existence, uniqueness, and characterization of (*pure*) *Nash Equilibria*[1]. The *Price of Anarchy* (*PoA*) and the *Price of Stability* (*PoS*) are defined as the ratio of social cost in the worst-case Nash Equilibrium (the one with highest social cost) and the optimal social cost, and the ratio of the social cost in the best-case Nash Equilibrium (the one with lowest social cost) and the optimal social cost, respectively:

$$\text{PoA} = \frac{J(\text{worst NE})}{\min J},\ \text{PoS} = \frac{J(\text{best NE})}{\min J}. \tag{3}$$

PoA is an efficiency measure, illustrating how bad selfish playing is, in comparison to the global optimum. PoS, on the other hand, reflects the best possible performance without coordination in comparison to the global optimum. The network about to be designed, is empty and every node in the network is a player. We assume the cost of building one (communication) link between two nodes is fixed. Every player i can install a link from itself to another node j. Installing a link between i and j means that both i and j can utilize it, but only one pays for the cost, like often assumed in NFG models [12], [19], [4]. Several examples fit this scenario: (i) a friend request is initiated by one node in a social network, but both read the posts from one another; (ii) a new road connecting two cities is built by one city in a road network, but both utilize it; and (iii) in a hand-shake protocol in a computer network one node initiates a connection used by two nodes.

We consider a ***Virus Spread-Performance-Cost*** **(VSPC)** network formation game, where player i aims to reduce its cost and the probability $v_{i\infty}$ of being infected, but concurrently also wants to improve its performance by shortening the hopcounts $h(i, j)$ of the shortest paths to all the other nodes j. The cost function of player i that unites these objectives is given by:

$$J_i = \alpha \cdot k_i + \gamma\sum_{j=1}^{N} h(i, j) + v_{i\infty}. \tag{4}$$

Function J_i involves the cost k_i of installing all the links from node i, weighted by a coefficient α. The hopcounts $h(i, j)$ are weighted by γ. Opposing goals meet in this game: the more links are installed, the shorter the paths, but the higher the probability of being infected and the higher the cost.

The social cost J for the whole network is a weighted sum over all nodes

$$J = \sum_{i=1}^{N} J_i = \alpha L + \gamma\sum_{i=1}^{N}\sum_{j=1}^{N} h(i, j) + \sum_{i=1}^{N} v_{i\infty}, \tag{5}$$

where L denotes the number of links.

III. Virus spread-performance-cost (VSPC) game

A. Optimal social cost, Nash Equilibria and the PoA for $\gamma \to 0$

In order to understand the effect of the virus protection, we start by setting γ to an infinitely small number (approaching

[1] A Nash Equilibrium is the state of the players' network strategies, where none of the players can reduce its cost by unilaterally changing its strategy.

zero[2]). As a result, the hopcounts are of no influence anymore, while network connectivity is still guaranteed (the hopcount between two disconnected nodes is assumed to be infinity). Lemma 1 limits the possible Nash Equilibria.

Lemma 1. *The probability $v_{i\infty}(G)$ of node i being infected in the metastable state in network G does not exceed the probability $v_{i\infty}(G+l)$ of node i being infected in the metastable state in network $G+l$ obtained by adding a link l to G.*

Proof: The newly added link $l=(a,b)$ is between nodes a and b. We make use of the canonical infinite form [14],

$$v_{i\infty} = 1 - \cfrac{1}{1+\tau d_i - \tau \sum_{j=1}^{N} \cfrac{a_{ij}}{1+\tau d_j - \tau \sum_{k=1}^{N} \cfrac{a_{jk}}{1+\tau d_k - \ddots}}}. \tag{6}$$

After the addition of link $l=(a,b)$, the expression (6) for $v_{i\infty}(G+l)$ has all the terms the same as in $v_{i\infty}(G)$, except the following differences: $d_a \to d_a+1$; $d_b \to d_b+1$ and the presence of the adjacency entry $a_{ab}=0 \to a_{ab}=1$ in the canonical representation. The last statement implies that its contribution is a part that is the same as in $v_{i\infty}(G)$ until it "reaches" nodes a and b, where the change (at a certain depth of the canonical form) is:

$$\tau(d_a+1) - \frac{\tau}{\tau(d_b+1)-\dots} + U = \tau d_a + U + \tau\left(1 - \frac{1}{\tau(d_b+1)-\dots}\right), \tag{7}$$

where d_a and d_b are the degrees of a and b in G, while U is the remaining part in the canonical form. In (7), the term $\tau(1-\frac{1}{\tau(d_b+1)-\dots})$ is positive and U increases with d_a, d_b and a_{ab}. U increases with d_a, d_b and a_{ab} as itself is an infinite canonical form with any of these three variables being in the numerator or in the denominator with a negative sign in front, in the same way as in (7). Therefore, the whole term in (7) increases, which implies that $v_{i\infty}(G+l) > v_{i\infty}(G)$ also increases for each node i. ■

We start by looking into the possible Nash Equilibria.

Theorem 1. *If a Nash Equilibrium is reached, then the constructed graph is a* tree.

Proof: If G is connected and each node can reach every other node, then changing the strategy of node i from the current one to *investing in an extra link*, will increase both its cost (by 1, scaled by α) and $v_{i\infty}$ (by Lemma 1). Hence, unilaterally *investing in an extra link* is not beneficial for a node.

We now assume that G is not a tree. Then, there is at least one cycle in this graph. If a node i in that cycle changes its strategy from *investing* in a link in that cycle to *not investing*, the cost is decreased by 1 (weighted by α) and all the other nodes in the graph are still reachable from i. Moreover, by *not investing* in that link, node i decreases its probability $v_{i\infty}$ of being infected in the metastable state, according to Lemma 1. Hence, by unilaterally changing its strategy, node i decreases its cost utility J_i, which is in contradiction with a Nash Equilibrium. ■

Observation 1. *A Nash Equilibrium is achieved for both the* star *graph and the* path *graph, but not all trees are Nash Equilibria.*

Proof: Let us consider a *star graph*, where all the links are installed by the root node as shown in Fig. 1a. (A link is installed and paid for by the node marked with p.) The root node cannot unilaterally decrease its cost, because cutting at least one of its installed links would disconnect it, while installing a link from a leaf node i would increase both k_i and $v_{i\infty}$ (Lemma 1). Hence, the *star graph* is a Nash Equilibrium.

Let us now assume that a *path graph* (Fig. 1b) is constructed, such that $(N-1)$ nodes invest in exactly one link and one of the leaves does not invest in installing a link. Similarly as for a *star graph*, none of the nodes can unilaterally decrease their cost by just installing extra links or cutting some of them. A "re-wiring"[3] from one of the nodes by re-directing its installed links to another node may be in order. In such a case, if node i "re-wires" its installed link to another node, then J_i would not decrease. 1) If it is installed to one of the leaves, such that the graph is connected, we end up with an isomorphic graph, where the position of i is the same as in the initial graph, so J_i stays the same. 2) If i "re-wires" to one of the other nodes j (w.l.o.g., $i<j$) as visualized in Fig. 1e, i would have the same degree, but its "new neighbor" would have a degree 3 instead of 2. The degree of j increases by 1 to 3 and the degree of $(i+1)$ decreases by 1 to 1 (node $(i+1)$ will become terminal and "far" from i), while all the other degrees remain the same. Moreover, i would be equally close to any of the nodes "behind" $\{1,\dots,i-1\}$, closer to the nodes "at the end" $\{j+1,\dots,N\}$ and equally close to the nodes in the set $\{i+1,\dots,j-1\}$, but just in a reverse order. Based on the canonical infinite form (6), $v_{i\infty}$ would increase[4]. Therefore, the *path graph* is also a Nash Equilibrium.

There are also other *trees* that are Nash Equilibria (e.g., T' given in Fig. 1c). Moreover, there are values of τ such that worst- and best-case Nash Equilibria are achieved for trees different from star $K_{1,N-1}$ and path P_N graphs. For $\tau \in [1.475, 1.589]$, tree T' is the best-case Nash Equilibrium and has optimal social cost.

However, not all the trees are Nash Equilibria. For example, the tree given in Fig. 1d. Here, whomever pays for the "central" link between a and b, can reduce its cost utility by "re-wiring" to c or d. ■

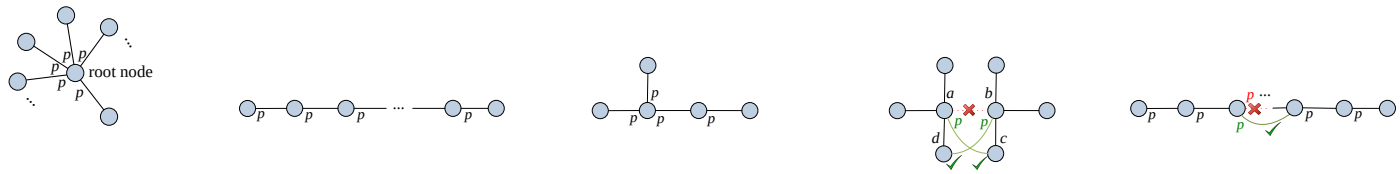
We proceed by characterizing the worst- and best-case Nash Equilibria.

Theorem 2. *For sufficiently high effective infection rate τ, the optimal social cost and the* best-case *Nash Equilibrium are achieved by the* star graph $K_{1,N-1}$, *while the* worst-case

[2]The case of $\gamma=0$ is either trivial or debatable. By neglecting the hopcounts, the optimal topology would be the (non-realistic) empty graph with no links (cost) and no epidemic to be propagated. Moreover, infinite hopcounts will be multiplied by $\gamma=0$ which is undefined.

[3]"Re-wiring" is a process of removing a link to node k initiated by node i and establishing a new link to another node j. The degree of node i does not change, while the degrees of k and j are decreased and increased, respectively.

[4]$v_{i\infty}$ in (6) would have bigger values by having nodes with "bigger degrees" as as close as possible (i.e. in fewer hops) to the node.



(a) Star $K_{1,N-1}$. (b) Path P_N. (c) Tree T'. (d) Tree T''. (e) Re-wiring increases $v_{i\infty}$.

Fig. 1: A link is installed by the end-node marked with p. Trees in (a), (b), and (c) are Nash Equilibria. (d) Tree T'' cannot be a Nash Equilibrium.

Nash Equilibrium is achieved for the path graph P_N,

$$J(K_{1,N-1}) \leq J \leq J(P_N).$$

Proof: According to Theorem 1, in a Nash Equilibrium the graph is a tree, hence it has $N-1$ links. In a general case, from a tree in which there are two nodes i and k, connected to one another, for which $d_i \geq 3$ and $d_k = 1$ (i.e. k is a leaf), by breaking the connection between i and k and connecting k to another leaf j instead, we have: the degree of k is 1 (remains the same); the degree of node i becomes $d_i - 1 \geq 2$ (decreased by one); and the degree of j is 2 (increased by one). The process can be repeated until there exists a node of degree at least 3 in the tree. At the end, we end up with a tree with no degree bigger than 2 and this is a path P_N. The social cost J is increased in each step [1, Lemma 2]. In this way, the process converges to a path P_N.

In a very similar (but reverse) process, starting from any tree G, we can decrease J at each step, ending up with a star $K_{1,N-1}$ with a maximum $J(G)$ in the final step. ■

However, what would be the optimal social cost, and the worst- and best-case Nash Equilibria highly depends on the effective infection rate τ.

Theorem 3. *For low values of the effective infection rate τ, above but sufficiently close to the epidemic threshold τ_c, the optimal social cost and the* best-case *Nash Equilibrium are achieved by the* path graph P_N, *while the* worst-case *Nash Equilibrium is achieved by the* star graph $K_{1,N-1}$,

$$J(P_N) \leq J \leq J(K_{1,N-1}).$$

Proof: We consider a spectral approach [20] and denote $y(\tau) = \sum_{i=1}^{N} v_{i\infty}(\tau)$ the infection probability of all nodes in the metastable state. The probabilities of a node in the graph being infected are non-zero and $y(\tau) > 0$ if $\tau > \tau_c = \frac{1}{\lambda_1}$, where λ_1 is the largest eigenvalue of the adjacency matrix in the graph [16]. For $\tau < \frac{1}{\lambda_1}$, $y(\tau) = 0$.

Lovász and Pelikán [21] ordered all the trees with N nodes by the largest eigenvalues of the adjacency matrices. It turns out that, the path P_N and star $K_{1,N-1}$ are the trees with the minimum $\lambda_1(P_N)$ and maximum $\lambda_1(K_{1,N-1})$ largest eigenvalues, respectively.

For values $\tau = \frac{1}{\lambda_1(K_{1,N-1})} + \varepsilon = \frac{1}{\sqrt{N-1}} + \varepsilon$, it holds that $y_{K_{1,N-1}}(\tau) > y_T(\tau) = 0$, where T is any tree different from $K_{1,N-1}$, therefore $J(K_{1,N-1})$ is the largest.

For values $\tau = \frac{1}{\lambda_1(P_N)} - \varepsilon = \frac{1}{2\cos(\frac{\pi}{N+1})} - \varepsilon$, we have $y_T(\tau) > y_{P_N}(\tau) = 0$, where T is any tree different from P_N, hence $J(P_N)$ is the smallest. ■

Theorems 2 and 3 show opposite behavior depending on whether the value τ is in the high or low regime, although both revolve around the path and star graphs. For τ in the intermediate regime, different trees may give the best-/worst-case Nash Equilibrium.

Corollary 1. *For both high and low effective infection rate τ, PoS $= 1$ and PoA $= \max\left\{\frac{J(P_N)}{J(K_{1,N-1})}, \frac{J(K_{1,N-1})}{J(P_N)}\right\}$.*

Proof: Based on Theorems 2 and 3, for high (low) τ, tree $K_{1,N-1}$ (P_N) is both optimal in social cost and the best-case Nash Equilibrium, while P_N ($K_{1,N-1}$) is the worst-case Nash Equilibrium. Based on the definitions for PoS and PoA in (3), PoS $= \frac{J(K_{1,N-1})}{J(K_{1,N-1})} = 1 (= \frac{J(P_N)}{J(P_N)})$; and PoA$= \frac{J(P_N)}{J(K_{1,N-1})}$ for large enough τ and PoA$= \frac{J(K_{1,N-1})}{J(P_N)}$ for τ close to the epidemic threshold τ_c. ■

Corollary 2. *For sufficiently high effective infection rate τ, in the virus spread-cost game formation,*

$$\textit{PoA} < 1 + \frac{1}{2\big(\tau(\alpha+1)-1\big)}, \textit{ where } \tau(\alpha+1) > 1.$$

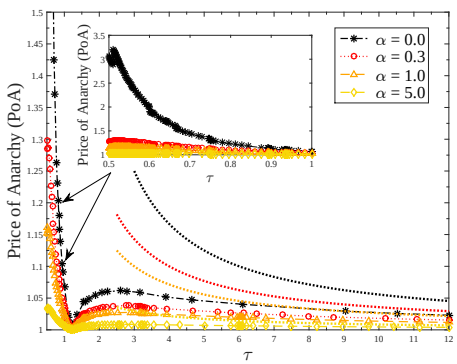
Proof: The proof is provided in [1]. ■

The exact value of the PoA is given in Fig. 2 by making use of Corollary 1. It is highest (~ 3.3) for small τ, above the epidemic threshold and it further sharply decreases reaching 1 for a unique Nash Equilibrium. For higher τ, the PoA increases towards its maximum around 1.1 and then it slowly decreases approaching 1.

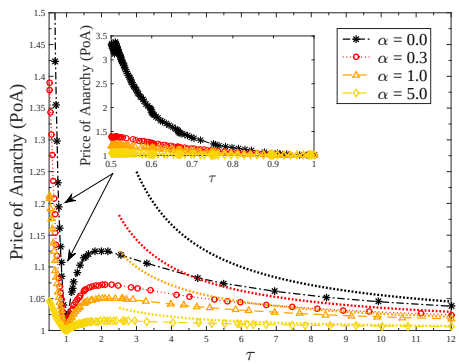
We have observed that the equilibria tree topology in which a virus thrives is not always a star (i.e., the tree with the smallest diameter), but that it may differ with the virus infection rate. For most of the τ values (except maybe small τ), a small value for the Price of Anarchy (PoA) means that a topology close to optimal can be obtained in a decentralized manner, even when the individual players play selfishly.

B. Optimal social cost, Nash Equilibria and the PoA for $\gamma > 0$

We start by analyzing the social cost (5). Node i is one hop away from its d_i neighboring nodes, while it is at least 2 hops away from the other $N-1-d_i$ nodes, hence



(a) $N = 10$.



(b) $N = 1000$.

Fig. 2: The Price of Anarchy (PoA). The dotted lines represent the bound from Corollary 2.

$\sum_{j=1}^{N} h(i,j) \geq d_i + 2(N-1-d_i)$. Using this, for large enough τ when $\sum_{i=1}^{N} v_{i\infty}$ can be approximated[5] by using truncation of Maclaurin seria [17, Lemma 1], the social cost in (5) is lower bounded as

$$J \geq N + 2\gamma N(N-1) + (\alpha - 2\gamma)L - \frac{1}{\tau}\sum_{i=1}^{N}\frac{1}{d_i}. \tag{8}$$

The following bound is due to Cioabă [22, Theorem 9],

$$\sum_{i=1}^{N}\frac{1}{d_i} \leq \frac{N^2}{2L} + (\frac{1}{d_{\min}} - \frac{1}{d_{\max}})(N-1-\frac{2L}{N}),$$

where the equality holds for regular graphs and the star graph. Based on this, $d_{\min} \geq 1$, and $d_{\max} \leq N-1$, we obtain

$$\begin{aligned}\sum_{i=1}^{N}\frac{1}{d_i} &\leq \frac{N^2}{2L} + (1-\frac{1}{N-1})(N-1-\frac{2L}{N}) \\ &= \frac{N^2}{2L} + \frac{N-2}{N(N-1)}(N(N-1)-2L). \end{aligned}\tag{9}$$

Equality in (9) is achieved only for the star $K_{1,N-1}$, where $d_{\max} = N-1$ and $d_{\min} = 1$, or for the complete graph K_N (where $2L = N(N-1)$). (The equality for other regular graphs is ruled out because of the inequality in (9).) Using (9) into (8) yields

$$J \geq N + 2\gamma N(N-1) - \frac{N-2}{\tau} + (\alpha - 2\gamma + \frac{2(N-2)}{\tau N(N-1)})L - \frac{N^2}{2\tau L} \tag{10}$$

Let us consider two regimes:

1) If $\alpha - 2\gamma + \frac{2(N-2)}{\tau N(N-1)} \geq 0$, then the bound in (10) is an increasing function in L, hence the optimal social cost is achieved for $L = N-1$. The bound in (10) is tight for such L, because the bounds in (9) and (8) become equalities for $K_{1,N-1}$ and any graph with a diameter at most two, respectively. Hence, $J \geq J(K_{1,N-1})$ and equality is achieved only for the star graph $K_{1,N-1}$.
2) If $\alpha - 2\gamma + \frac{2(N-2)}{\tau N(N-1)} < 0$, then the bound in (10) increases for $L < \frac{N}{\sqrt{2\tau(2\gamma-\alpha)-\frac{4(N-2)}{N(N-1)})}}$ and decreases for $L > \frac{N}{\sqrt{2\tau(2\gamma-\alpha)-\frac{4(N-2)}{N(N-1)})}}$. Hence, the optimal social cost is achieved in one of two boundary cases: $L = N-1$ and $L = \binom{N}{2}$. For $L = N-1$, similarly as in 1), we obtain that the only possibility is the star graph $K_{1,N-1}$, while for $L = \binom{N}{2}$ it is the complete graph K_N. Finally, $J \geq \min\{J(K_{1,N-1}), J(K_N)\}$.
It remains to compare $J(K_{1,N-1})$ and $J(K_N)$: $J(K_{1,N-1}) = N + \alpha(N-1) + 2\gamma(N-1)^2 - \frac{(N-1)^2+1}{\tau(N-1)}$ and $J(K_N) = N + \alpha\frac{N(N-1)}{2} + \gamma N(N-1) - \frac{N}{\tau(N-1)}$. Hence,

$$J(K_N) - J(K_{1,N-1}) = (N-1)(N-2)(\frac{\alpha}{2} - \gamma + \frac{1}{\tau(N-1)}).$$

If $\alpha \leq 2\gamma - \frac{2}{\tau(N-1)}$, then $J(K_N) \leq J(K_{1,N-1})$ and the optimal social cost is achieved for the complete graph K_N. If $\alpha \geq 2\gamma - \frac{2}{\tau(N-1)}$, then $J(K_N) \geq J(K_{1,N-1})$ and the optimal social cost is achieved for the star graph $K_{1,N-1}$. The last also covers case 1), because $2\gamma - \frac{2}{\tau(N-1)} < 2\gamma - \frac{2(N-2)}{\tau N(N-1)}$.

Now, for the optimal social cost, Theorem 4 follows.

Theorem 4. *For sufficiently high τ, the optimal social cost is achieved for the star $K_{1,N-1}$ if $\alpha \geq 2\gamma - \frac{2}{\tau(N-1)}$, and for the complete graph K_N, otherwise.*

We proceed with characterization of the Nash Equilibria and the Price of Anarchy for sufficiently high τ. In the VSPC game, Nash Equilibria topologies can be complex, while the star and the complete graph can appear as extreme cases:

- The complete graph K_N is a Nash Equilibrium, if and only if $\alpha \leq \gamma - \frac{1}{\tau(N-1)(N-2)}$. Since new links cannot be added, changing the strategy for a node i means deleting k of its links ($1 \leq k \leq N-2$). The corresponding change would increase the cost J_i of i by $k(\gamma-\alpha) - \frac{1}{\tau(N-1)(N-1-k)} \geq \frac{(k-1)(N-2-k)}{\tau(N-1)(N-2)(N-1-k)} \geq 0$. Hence, node i has no interest to deviate from its current strategy. On the other hand, if $\alpha > \gamma - \frac{1}{\tau(N-1)(N-2)}$ and node i changes its strategy by cutting one link, the change

[5] In fact, the sum can be lower bounded [14, p. 10] by $\sum_{i=1}^{N} v_{i\infty} > N - \sum_{i=1}^{N}\frac{1}{1+(\tau-1)d_i}$, which is meaningful for $\tau > 1$.

in J_i is equal to $\gamma - \alpha - \frac{1}{\tau(N-1)(N-2)} < 0$, which will reduce its cost.

- The star graph $K_{1,N-1}$ is a Nash Equilibrium, if and only if $\alpha \geq \gamma - \frac{1}{\tau(N-1)}$. The root node cannot delete a link, because this would make its cost infinity. If i is a leaf, for some $k \geq 0$, changing its strategy means: (i) adding k links, then the hopcounts to these nodes are reduced from 2 to 1, hence the contribution from the hopcounts is changed by $-k\gamma$; or (ii) deleting the link installed by him (if any) and installing $(k+1)$ links, where $k+1 < N-2$. In (ii), the hopcount to the root node is increased from 1 to 2, the hopcount to $(k+1)$ links is decreased from 2 to 1, and the hopcounts to the other $\big((N-2)-(k+1)\big)$ nodes are increased from 2 to 3. The change in the sum of hopcounts is: $-(k+1)\gamma + 1\cdot\gamma + \big((N-2)-(k+1)\big)\gamma = -k\gamma + \big((N-2)-(k+1)\big)\gamma \geq -k\gamma$, hence the change of the hopcount is again at least $-k\gamma$. Thus, the change in J_i is at least $k\alpha - k\gamma - \frac{1}{\tau(k+1)} + \frac{1}{\tau} = k(\alpha - \gamma + \frac{1}{\tau(k+1)}) \geq k(\alpha - \gamma + \frac{1}{\tau(N-1)}) > 0$. On the other hand, if $\alpha < \gamma - \frac{1}{\tau(N-1)}$, the change in J_i by adding $(N-2)$ links from one leaf to all the other leaves in $K_{1,N-1}$, is $(N-2)(\alpha-\gamma) - \frac{1}{\tau(N-1)} + \frac{1}{\tau} = (N-2)(\alpha-\gamma+\frac{1}{\tau(N-1)}) < 0$, i.e. it is not a Nash Equilibrium.

As becomes clear, the Nash Equilibria and the Price of Anarchy may vary on different intervals and a case analysis, as provided in the following, is required. We will consider the case $\alpha < 2\gamma - \frac{1}{\tau}$ and the case $\alpha > 2\gamma - \frac{1}{\tau}$.

Case $\alpha < 2\gamma - \frac{1}{\tau}$.

Now, $\gamma > \frac{1}{2\tau}$. A Nash Equilibrium is achieved only for graphs with a diameter at most 2 - an argument used in the later points (b) and (c). The proof is by contradiction. Let us assume node i is at least 3 hops away from another node. Clearly, $d_i \leq (N-2)$ and if i installs a link from i to j, the difference in J_i is at least $\alpha - 2\gamma + \frac{1}{\tau d_i(d_i+1)} \leq -\frac{1}{\tau} + \frac{1}{\tau d_i(d_i+1)} < 0$. Hence, i reduces its cost and the graph is not a Nash Equilibrium. We consider three sub-intervals (a), (b) and (c):

(a) If $\alpha < \gamma - \frac{1}{2\tau}$, adding a link from i will change J_i by at least $\alpha - \gamma + \frac{1}{\tau d_i(d_i+1)} < -\frac{1}{2\tau} + \frac{1}{\tau d_i(d_i+1)} \leq 0$. Therefore, the complete graph K_N is the only Nash Equilibrium. Because, $\alpha < \gamma - \frac{1}{2\tau} \leq 2\gamma - \frac{2}{\tau(N-1)}$ for $N \geq 3$ and, according to Theorem 4, it also has optimal social cost. Finally, PoA $=$ PoS $= 1$.

(b) If $\gamma - \frac{1}{2\tau} \leq \alpha \leq \gamma - \frac{1}{\tau(N-1)}$ and we assume, by contradiction, that there is a Nash Equilibrium different from K_N, we have the following:

- If there is a link in the graph, installed by node i such that its deletion increases the sum of hopcounts from i by only 1, then J_i is increased by: $\gamma - \alpha - \frac{1}{\tau d_i(d_i-1)} > 0$. On the other hand, adding a link would change J_i to: $\alpha - \gamma + \frac{1}{\tau d_i(d_i+1)} > 0$. The last two inequalities imply, $0 < \alpha - \gamma + \frac{1}{\tau d_i(d_i+1)} < -\frac{1}{\tau d_i(d_i-1)} + \frac{1}{\tau d_i(d_i+1)} = -\frac{2}{\tau(d_i-1)d_i(d_i+1)} < 0$, which is a contradiction. Hence, there is no other Nash Equilibrium different from K_N and PoA $=$ PoS $= 1$.
- If deleting any of the links installed by i would increase the sum of hopcounts by at least 2; by link deletion, the difference in J_i is at least $2\gamma - \alpha - \frac{1}{\tau d_i(d_i-1)}$ and we have $2\gamma - \alpha - \frac{1}{\tau d_i(d_i-1)} \geq \gamma + \frac{1}{\tau(N-1)} - \frac{1}{\tau d_i(d_i-1)} \geq \frac{1}{2\tau} - \frac{1}{\tau d_i(d_i-1)} + \frac{1}{\tau(N-1)} \geq \frac{1}{\tau(N-1)} > 0$. We proceed by considering the properties of the possible Nash Equilibria in particular sub-intervals: $-\frac{1}{\tau(k-1)k} \leq \alpha - \gamma < -\frac{1}{\tau k(k+1)}$ for $k \in \{2, 3, \ldots, \lfloor\sqrt{N - \frac{3}{4}} - \frac{1}{2}\rfloor\}$. By link addition, the difference in J_i is $\alpha - \gamma + \frac{1}{\tau d_i(d_i+1)}$ and a necessary condition for a Nash Equilibrium is $d_i < k$. On the other hand, $k \leq \sqrt{N - \frac{3}{4}} - \frac{1}{2} \leq \sqrt{N-1}$, hence $d_i < \sqrt{N-1}$. Therefore, we have less than $\sqrt{N-1}$ nodes that are on a distance 1 from a node i. Each of these nodes is directly connected by less than $\sqrt{N-1} - 1$ nodes different from i. Hence, there less than $\sqrt{N-1} + \sqrt{N-1}(\sqrt{N-1} - 1) = N - 1$ nodes that are at most 2 hops from i, hence at least one node that is more than 2 hops away from i, a contradiction to the general claim (before (a))! Hence, K_N is the only Nash Equilibrium and PoA $=$ PoS $= 1$.

(c) If $\gamma - \frac{1}{\tau(N-1)} \leq \alpha < 2\gamma - \frac{1}{\tau}$, then $K_{1,N-1}$ is a Nash Equilibrium. Graphs that are of diameter at most 2 are also candidates for a Nash Equilibrium.

Because the diameter of the graph is not bigger than 2, (8) becomes an equality $J = N + 2\gamma N(N-1) + (\alpha - 2\gamma)L - \frac{1}{\tau}\sum_{i=1}^{N}\frac{1}{d_i}$ for sufficiently large τ. Applying the condition of (c) leads to

$$\begin{aligned} J(\text{worst NE}) &< N + 2\gamma N(N-1) - \frac{1}{\tau}L - \frac{1}{\tau}\sum_{i=1}^{N}\frac{1}{d_i} \\ &= N + 2\gamma N(N-1) - \frac{1}{\tau}\sum_{i=1}^{N}(\frac{d_i}{2} + \frac{1}{d_i}) \\ &\leq N + 2\gamma N(N-1) - \frac{3N}{2\tau} = N(1 + 2\gamma(N-1) - \frac{3}{2\tau}) \end{aligned} \tag{11}$$

due to the fact that $\frac{d_i}{2} + \frac{1}{d_i} \geq \frac{3}{2}$ (equivalent to $(d_i-1)(d_i-2) \geq 0$). Equality holds (only) in the last line of (11) if $d_i = 1$ or $d_i = 2$ for all i (e.g., the ring C_N or the path P_N graphs), otherwise a strict inequality in the second part also holds. Finally, knowing that the optimal social cost is attained by the complete graph K_N and the condition inequality condition in (c) for $J(K_N)$: PoA $= \frac{J(\text{worst NE})}{J(K_N)} < \frac{1+2\gamma(N-1)-\frac{3}{2\tau}}{1+\frac{3\gamma(N-1)}{2}-\frac{1}{2\tau}-\frac{1}{\tau(N-1)}} = \frac{\frac{4}{3}(1+\frac{3\gamma(N-1)}{2}-\frac{1}{2\tau}-\frac{1}{\tau(N-1)})-(\frac{1}{3}+\frac{5}{6\tau}-\frac{4}{3\tau(N-1)})}{1+\frac{3\gamma(N-1)}{2}-\frac{1}{2\tau}-\frac{1}{\tau(N-1)}}$ $\leq \frac{\frac{4}{3}(1+\frac{3\gamma(N-1)}{2}-\frac{1}{2\tau}-\frac{1}{\tau(N-1)})}{1+\frac{3\gamma(N-1)}{2}-\frac{1}{2\tau}-\frac{1}{\tau(N-1)}} = \frac{4}{3}$ for each $N \geq 3$; because $\frac{1}{3} + \frac{5}{6\tau} - \frac{4}{3\tau(N-1)} > 0$. This bound is approached, for instance, when α and γ are large and bigger than τ: K_N is the social optimum and $K_{1,N}$ is the worst-case Nash Equilibrium and the bounds in (11) and the inequality for PoA are closely approached. If $\alpha < \gamma - \frac{1}{\tau(N-1)(N-2)}$, K_N is a Nash Equilibrium and PoS $= 1$, otherwise PoS > 1.

Case $\alpha > 2\gamma - \frac{1}{\tau}$.

We first consider the links, whose deletion leaves the graph connected. For any node i, we focus on the links installed by i. Let $l = (i, j)$ be one such link and the number of all nodes q that use l as a link for the shortest paths from j to q is z. According to Schoone *et al.* [23, Theorem 2.1., case $k = 1$], all the distances from i to the other nodes are increased by at most $2d$, where d is the diameter in the original graph. In a Nash Equilibrium, $2dz\gamma - \alpha - \frac{1}{\tau d_i(d_i-1)} > 0$ for any possible value of $d_i \geq 2$, i.e. we obtain $2dz\gamma - \alpha - \frac{1}{2\tau} > 0$. Hence, $z > \frac{\alpha+\frac{1}{2\tau}}{2d\gamma}$ and then the number of such links to node j is not bigger than $\frac{2d\gamma N}{\alpha+\frac{1}{2\tau}}$. Taking into account all possible nodes, the number of links whose deletion does not disconnect the graph is not bigger than $\frac{2d\gamma N^2}{\alpha+\frac{1}{2\tau}}$. On the other hand there are at most $(N-1)$ links such that a removal of any of those links disconnects the graph. Indeed, a connected graph has a spanning tree T and a link removed from T disconnects the graph, while a removal of a link that is not in T leaves the graph connected. Therefore,

$$L \leq N - 1 + \frac{2d\gamma N^2}{\alpha + \frac{1}{2\tau}} \tag{12}$$

If two nodes i and j are a hops apart from each other, adding a link from i to j would reduce the hopcounts from i to all the nodes in the "second half" along the previous path to j by at least half of their lengths, by $a-1, a-3, \ldots, 1$ for a even or by $a-1, a-3, \ldots, 2$ for a odd. Hence, the total reduction in the sum of shortest paths from i is $\sum_{i=1}^{\frac{a}{2}}(2i-1) = \frac{a^2}{4}$ or $\sum_{i=1}^{\frac{a-1}{2}}(2i) = \frac{a^2-1}{4}$ for a even or odd, respectively. Assuming a Nash Equilibrium and i is a starting node on the diameter, considering the change in cost J_i, the following inequality would hold for any d: $\alpha - \frac{d^2-1}{4}\gamma + \frac{1}{\tau d_i(d_i+1)} > 0$. Using $d_i \geq 1$, and the absolute maximum for d being $N-1$, we arrive at

$$d < \min\{\sqrt{1 + \frac{4}{\gamma}(\alpha + \frac{1}{2\tau})}, N\} \tag{13}$$

Each node i has at least one neighbor and all the others are no more than d hops apart, hence: $\sum_{j=1}^{N} h(i,j) \leq 1 + (N-1)d$. Applying the arithmetic-harmonic mean inequality leads to $\frac{1}{\tau}\sum_{i=1}^{N}\frac{1}{d_i} \geq \frac{1}{\tau}\frac{N^2}{\sum_{i=1}^{N} d_i} = \frac{N^2}{2L\tau}$. We proceed by upper bounding J in (8),

$$J \leq N + \alpha L + \gamma N(1 + (N-1)d) - \frac{N^2}{2L\tau} \tag{14}$$

Applying (12) modifies (14), into

$$\begin{aligned} J \leq & N + \alpha(N-1) + \gamma N + N\big((N-1) + \frac{2\alpha N}{\alpha + \frac{1}{2\tau}}\big)\gamma d \\ & - \frac{N^2}{2\tau(N - 1 + \frac{2d\gamma N^2}{\alpha+\frac{1}{2\tau}})} \end{aligned} \tag{15}$$

We distinguish two sub-cases, (a) and (b):

(a) If $\alpha \geq 2\gamma - \frac{2}{\tau(N-1)}$, then the optimal social cost (and a Nash Equilibrium) is achieved for the star graph $K_{1,N-1}$, hence PoS=1. Now, using (15) for PoA,

$$\text{PoA} \leq \frac{N + \alpha L + \gamma N(1 + (N-1)d) - \frac{N^2}{2L\tau}}{N + \alpha(N-1) + 2\gamma(N-1)^2 - \frac{(N-1)^2+1}{\tau(N-1)}} \tag{16}$$

and applying (12),

$$\text{PoA} \leq \frac{N + \alpha(N-1) + N\gamma + N\gamma(N - 1 + \frac{2\alpha N}{\alpha+\frac{1}{2\tau}})d - \frac{N^2}{2\tau(N-1+\frac{2d\gamma N^2}{\alpha+\frac{1}{2\tau}})}}{N + \alpha(N-1) + 2\gamma(N-1)^2 - \frac{(N-1)^2+1}{\tau(N-1)}} \tag{17}$$

- $\gamma \cdot d$ is not infinitesimally small. For sufficiently large[6] N, after some algebraic transformation, division by N^2 in both the numerator and denominator of (17) and applying (12),

$$\text{PoA} \leq O\big((\frac{1}{2} + \frac{\alpha}{\alpha + \frac{1}{2\tau}})\sqrt{1 + \frac{4}{\gamma}(\alpha + \frac{1}{2\tau})}\big) \tag{18}$$

- $\gamma \cdot d$ is infinitesimally small. According to (12), $L = O(N)$. Now, (16) yields

$$\text{PoA} \leq O\Big(\frac{N + \alpha L - \frac{N^2}{2L\tau}}{N + \alpha(N-1) - \frac{(N-1)^2+1}{\tau(N-1)}}\Big) \tag{19}$$

(b) If $2\gamma - \frac{2}{\tau(N-1)} \geq \alpha \geq 2\gamma - \frac{1}{\tau}$, then the optimal social cost is achieved for the complete graph K_N, and using (15), $\text{PoA} \leq \frac{N+\alpha L+\gamma N(1+(N-1)d)-\frac{N^2}{2L\tau}}{N+\alpha\frac{N(N-1)}{2}+\gamma N(N-1)-\frac{N}{\tau(N-1)}}$. Now,

- $\gamma \cdot d$ is not infinitesimally small. Using the bound for L in (12), for sufficiently high enough[6] N, (17) is transformed into $\text{PoA} \leq O\Big(2\frac{\sqrt{\gamma^2+4\gamma(\alpha+\frac{1}{2\tau})}}{\alpha+2\gamma}(1+\frac{2\alpha}{\alpha+\frac{1}{2\tau}})\Big)$ and we have a constant value for PoA.
- $\gamma \cdot d$ is infinitesimally small. Then α is small and PoA has a value close to 1.

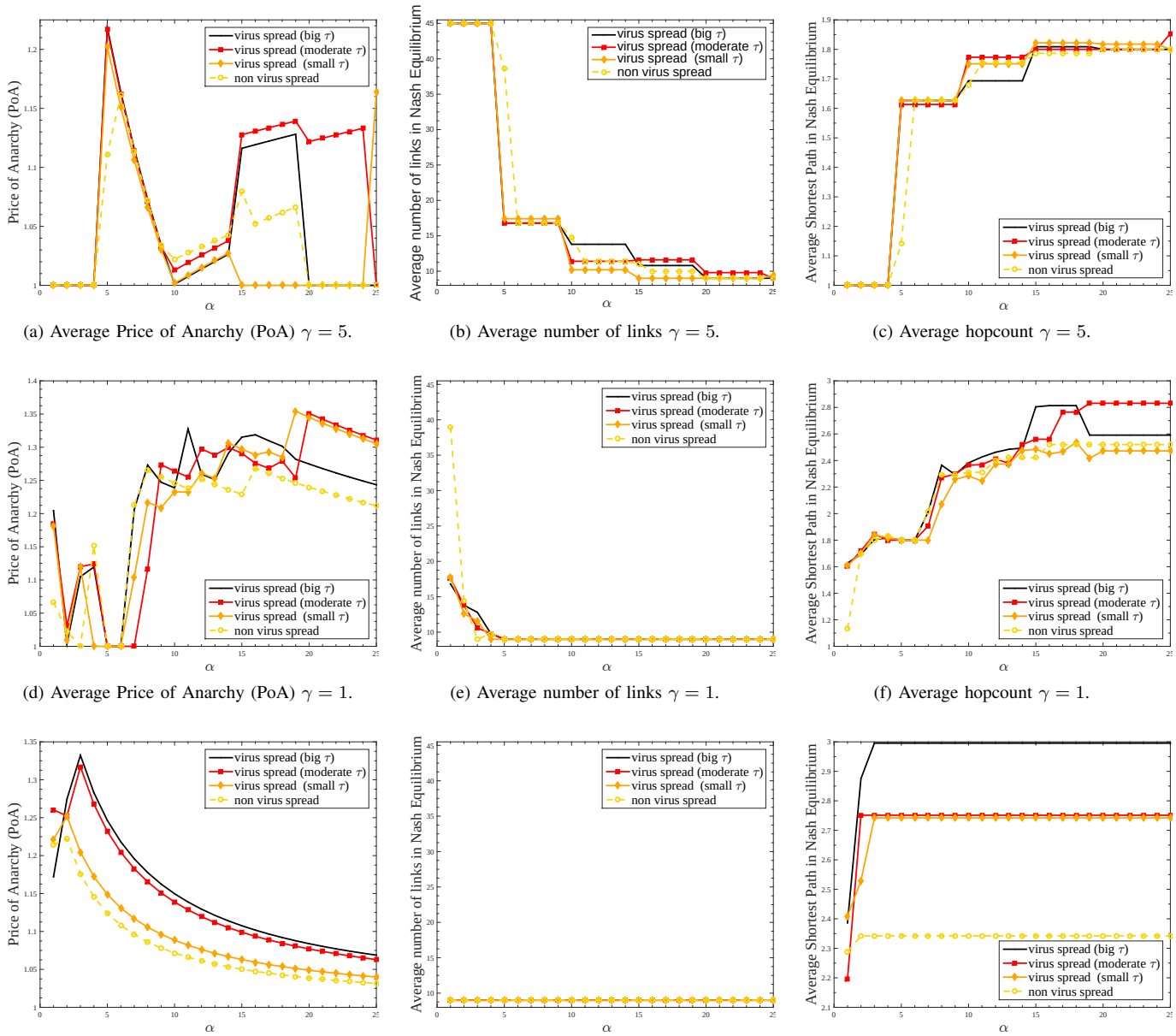
Based on these results, we present Theorem 5.

Theorem 5. *For sufficiently high τ in the* VSPC *game, the PoA depends on the parameters α, γ and τ,*

1) *if $\alpha \geq 2\gamma - \frac{2}{\tau(N-1)}$, then PoS=1 and if*
   - *γd is not small, then PoA $\leq O\big((\frac{1}{2} + \frac{\alpha}{\alpha+\frac{1}{2\tau}})\sqrt{1 + \frac{4}{\gamma}(\alpha + \frac{1}{2\tau})}\big)$.*
   - *$\gamma \cdot d$ is small, then PoA is given by Corollary 1.*
2) *if $2\gamma - \frac{2}{\tau(N-1)} \geq \alpha \geq 2\gamma - \frac{1}{\tau}$ and*
   - *γd is not small, then PoA $\leq O\Big(2\frac{\sqrt{\gamma^2+4\gamma(\alpha+\frac{1}{2\tau})}}{\alpha+2\gamma}(1+\frac{2\alpha}{\alpha+\frac{1}{2\tau}})\Big)$.*
   - *$\gamma \cdot d$ (and γ) is small, then α is also small and we have a constant value for PoA close to 1.*
3) *if $2\gamma - \frac{1}{\tau} > \alpha \geq \gamma - \frac{1}{\tau(N-1)}$, then Nash Equilibria graphs have diameters at most two and PoS $\leq$ PoA $< \frac{4}{3}$.*
4) *if $\gamma - \frac{1}{\tau(N-1)} > \alpha$, then K_N is the only Nash Equilibrium and PoA = PoS = 1.*

Theorem 5 and Corollaries 1 and 2 are compatible for small γ.

[6]Significantly larger compared to the coefficients α, γ and τ.



(a) Average Price of Anarchy (PoA) $\gamma = 5$. (b) Average number of links $\gamma = 5$. (c) Average hopcount $\gamma = 5$.

(d) Average Price of Anarchy (PoA) $\gamma = 1$. (e) Average number of links $\gamma = 1$. (f) Average hopcount $\gamma = 1$.

(g) Average Price of Anarchy (PoA) $\gamma = 0.1$. (h) Average number of links $\gamma = 0.1$. (i) Average hopcount $\gamma = 0.1$.

Fig. 3: Simulation results of the heuristic algorithm for the obtained networks in a Nash Equilibrium. The three regimes big, moderate and small τ are represented with values 5.2, 1.4 and 1, respectively. The number of nodes is $N = 10$.

C. Computational aspects and simulations

Since, (i) for small τ below the epidemic threshold τ_c, $v_{i\infty} = 0$ and (ii) for $\tau = \infty$, $v_{i\infty} = 1$ for all $i \in \{1, 2, \ldots, N\}$, the problem of finding a best response in the VSPC game includes the best-response problem described in [12], which is NP-hard. We therefore use a best-response heuristic algorithm, as in [4]. The steps of the algorithm are the following:

1) We start with an initial random graph $G = G_1^1$.
2) Time t is slotted and the first time slot is $t = 1$.
3) Each node takes only two actions at each time slot t. We fix the order of actions from node 1 to node N. The possible actions for each node are: dropping a link (D); adding a link (A); or doing nothing (N).
4) We denote by G_t^i the graph at time t before the action of node i.
5) Starting from node 1, each node i first computes the maximum reduction of its cost J_i induced by dropping a link (D) from graph G_t^i, or takes action (N) if no

reduction could be realized. Taking the obtained graph, node i computes the maximum reduction of its cost J_i induced by adding a link (A), or takes action (N) if no reduction could be realized.

6) After the decision of node i at time t, the graph becomes G_t^{i+1}. After the decision of node N, the algorithm moves to time $t+1$ (i.e., to graph G_{t+1}^1).
7) An equilibrium is reached at time t when all the nodes take the action (N) or the algorithm stops after a certain number of iterations $t_{\max}$ is reached.

In Fig. 3, results are given for the Price of Anarchy (PoA), the average number of links and the average hopcount as a function of installation cost α for different effective infection rates τ (namely big, moderate and small τ with values 5.2, 1.4 and 1, respectively that well represents the 3 regimes) and different weights (costs) for the hopcounts γ in a graph with $N = 10$ nodes. We also display some typical outcomes of the algorithm for different values of α and τ and three different values of γ in Fig. 4, Fig. 5 and Fig. 6 (see the visualizations at the end of the paper, after the bibliography). For all the metrics shown in Fig. 3, there is an interesting behavior for the curve with "no virus," in the sense that it follows the same shape as the curves where the virus is present, but is often shifted/delayed from them. This is due to the "enhancing" effect from the virus spread on the installation cost contribution.

For small values of α, due to the resulting cheap installation cost, and for non-negligible performance values γ, the Nash Equilibrium is a very dense graph, often the complete graph K_N, for all τ (Fig. 4). This reflects case 4) in Theorem 5, although the interval for α, where K_N is the only Nash Equilibrium would shrink (and may vanish) for small γ (Fig. 5 and Fig. 6). In the latter case, the corresponding PoAs in Fig. 3g and Fig. 2 have comparable shapes and the obtained topologies are trees (see Fig. 3h), although not necessarily star graphs (see Fig. 6).

Because of the higher installation cost (higher α) in intervals 3) and 2) from Theorem 5, the Nash Equilibria topologies are sparse: (i) in particular non-tree networks for $\gamma = 5$ (constant average number of links and sum of hopcounts) and mostly trees for γ equal to 1 or 0.1. Consequently, the PoA linearly decreases with α on this interval. For the interval 1) in Theorem 5 (high installation cost), the PoA increases with α (and τ), reaching a local maximum and then have a different behavior for larger α. Namely, it decreases towards 1 for small γ, while it is unpredictable for higher γ (the left column, subfigures (a), (d) and (g), in Fig. 3). But most importantly the effect of the epidemics part is noticeable and higher τ introduces inefficiency, which is reflected by a high PoA. It is also important to note that for comparable α and γ, the algorithm displays somewhat fluctuating behavior in terms of PoA (middle row of Fig. 3) due to the heuristic nature of the algorithm.

Being able to detect the intervals with high PoA, as we have done in this section, means that for those intervals some coordination/incentives of and for the players is needed. Since the PoS is generally low, in the best case with Nash Equilibrium a small amount of coordination likely suffices. For the intervals where the PoA is low, the selfish behavior of the players might still lead to (near)-optimal topologies without any coordination.

IV. Related work

Virus spread in networks has been thoroughly explored during the last decades [14], [15], [16], [17], [24]. These works involve studies ranging from virus-spread propagation, the computation of the number of infected hosts [14] to the epidemic threshold [15] in various epidemic models on networks. There is a large body of literature on game formation, that mostly minimizes a cost utility based on hopcount and the cost for installing links [12], [3], [19], [4], [25], [26], [27]. Fabrikant *et al.* [12] have studied the case, where a node's utility is a weighted sum of the installed links and the sum of hopcounts from each node in an undirected graph. The follow up work by Albers *et al.* [19] resolved some open questions from [12]. Chun *et al.* [4] have conducted extensive simulations on the same type of game formation. A game formation problem involving hopcounts and costs, applied to P2P networks has been considered by Moscibroda *et al.* [25]. Meirom *et al.* [28], [29] have provided dynamic and data analyses (apart from their static analysis) in an NFG setting with heterogenous players and robustness objectives. Nahir *et al.* [27] have considered similar NFG problems in directed graphs. A coalition and bilateral agreements between players in NFG and game-theory, in general, have been considered in [30], [8], [31], [3]. In order to evaluate "the goodness" of the equilibria, the prices of anarchy and stability [26], [32] have been used.

In this work, we have considered virus protection aspects together with cost and the length of shortest paths. In this sense, our work extends (with virus spread) and generalizes the related work [12], [19], [27]. However, to the best of our knowledge, network formation games concerning virus spread and protection both with or without the performance aspects have not been considered in the NFG framework, although *security games* [33], [34], [35], [36], [37], [38] have been used in modeling the virus spread suppression and network immunization. Performance aspects, represented by the hopcounts are linearly independent from the resilience to virus spread– the two metrics do not possess closed-form expressions– making the NFG problem challenging, apart from the novelty.

V. Conclusion

We have considered a novel network formation game (NFG), called the *virus spread-performance-cost* (VSPC) game, for communication networks in which the aspects link installation costs, virus infection probability, and performance in terms of the number of hops needed to reach other nodes in the network, all need to be balanced. We have characterized the Nash Equilibria and the Price of Anarchy (PoA) for various cases. In most of the cases, the PoA is not high, often close to 1, which implies that the decisions of non-cooperative players would lead, in a decentralized way, to an optimal topology.

When the aspect of the shortest hopcounts is not important, we have found that only *trees* (but not all) could be Nash Equilibria. In that case, surprisingly, a *path graph* is the worst- and the *star graph* is the best-case Nash Equilibrium for big virus infection rate τ, while it is the opposite for small τ. For intermediate values of τ, other trees are optimal. The PoA is the highest for values of τ just above the epidemic threshold. However, the PoA is generally small and close to 1, does not depend on the number of players, and is inversely proportional to τ and the installation cost α.

When the hopcounts do matter, the Nash Equilibria might be formed by more complex topologies. The PoA highly depends on τ, the installation links and hopcount costs α and γ, respectively, as shown by both theory and simulation. Although the PoA is small for most of the cases, for some intervals of those parameters, the PoA could be high. Hence, a central control and regulatory mechanism should be in place in such cases. Being able to detect those intervals, as we have done, helps in the design of optimal, efficient, virus-free and cheap overlay, P2P or wireless networks by limiting the non-cooperative freedom of the hosts' decisions.

There are several possibilities for follow-up work, such as a study on mixed Equilibria, player coalitions, inhomogeneous costs, or time-varying networks.

References

[1] S. Trajanovski, F. A. Kuipers, Y. Hayel, E. Altman, and P. Van Mieghem, "Designing virus-resistant networks: a game-formation approach," in *Proc. of IEEE CDC (54th IEEE Conference on Decision and Control)*. Osaka, Japan: IEEE, December 2015, pp. 294–299.

[2] G. Neglia, G. Lo Presti, H. Zang, and D. Towsley, "A network formation game approach to study bittorrent tit-for-tat," in *NetgCoop*, 2007.

[3] J. Corbo and D. Parkes, "The price of selfish behavior in bilateral network formation," in *PODC*. ACM, 2005, pp. 99–107.

[4] B.-G. Chun, R. Fonseca, I. Stoica, and J. Kubiatowicz, "Characterizing selfishly constructed overlay routing networks," in *IEEE INFOCOM*, vol. 2, March 2004, pp. 1329–1339.

[5] S. Shakkottai, E. Altman, and A. Kumar, "Multihoming of users to access points in wlans: A population game perspective," *IEEE Jour. on Sel. Areas in Communications*, vol. 25, no. 6, pp. 1207–1215, 2007.

[6] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Game theoretic anti-jamming dynamic frequency hopping and rate adaptation in wireless systems," in *WiOpt*, 2014, pp. 247–254.

[7] A. MacKenzie and L. DaSilva, *Game Theory for Wireless Engineers*, ser. Synth. lect. on communications. Morgan & Claypool Publ., 2006.

[8] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, "Enabling crowd-sourced mobile internet access," in *INFOCOM*, 2014, pp. 451–459.

[9] E. Watanabe, D. Menasché, E. de Souza e Silva, and R. Leão, "Modeling resource sharing dynamics of voip users over a wlan using a game-theoretic approach," in *INFOCOM*, 2008, pp. 1588–1596.

[10] B. Jiang, N. Hegde, L. Massoulie, and D. Towsley, "How to optimally allocate your budget of attention in social networks," in *INFOCOM*. IEEE, April 2013, pp. 2373–2381.

[11] F. Ming, F. Wong, and P. Marbach, "Who are your friends? - a simple mechanism that achieves perfect network formation," in *INFOCOM*. IEEE, 2011, pp. 566–570.

[12] A. Fabrikant, A. Luthra, E. Maneva, C. H. Papadimitriou, and S. Shenker, "On a network creation game," in *PODC*. ACM, 2003, pp. 347–351.

[13] R. Aumann and R. Myerson, *Endogenous Formation of Links Between Players and of Coalitions: an Application of the Shapley Value*. Stanford University, 1988.

[14] P. Van Mieghem, J. Omić, and R. Kooij, "Virus spread in networks," *IEEE/ACM Trans. on Netw.*, vol. 17, no. 1, pp. 1 –14, feb. 2009.

[15] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic thresholds in real networks," *ACM Trans. Inf. Syst. Secur.*, vol. 10, no. 4, pp. 1:1–1:26, Jan. 2008.

[16] A. Ganesh, L. Massoulie, and D. Towsley, "The effect of network topology on the spread of epidemics," in *IEEE INFOCOM*, vol. 2, 2005, pp. 1455–1466.

[17] P. Van Mieghem, "The N-intertwined SIS epidemic network model," *Computing*, vol. 93, no. 2-4, pp. 147–169, 2011.

[18] P. Van Mieghem and R. van de Bovenkamp, "Accuracy criterion for the mean-field approximation in susceptible-infected-susceptible epidemics on networks," *Phys. Rev. E*, vol. 91, p. 032812, Mar 2015.

[19] S. Albers, S. Eilts, E. Even-Dar, Y. Mansour, and L. Roditty, "On nash equilibria for a network creation game," in *SODA*, 2006, pp. 89–98.

[20] P. Van Mieghem, *Graph Spectra of Complex Networks*. Cambridge University Press, UK, 2011.

[21] L. Lovász and J. Pelikán, "On the eigenvalues of trees," *Periodica Mathematica Hungarica*, vol. 3, no. 1-2, pp. 175–182, 1973.

[22] S. M. Cioabă, "Sums of powers of the degrees of a graph," *Discrete Mathematics*, vol. 306, no. 16, pp. 1959 – 1964, 2006.

[23] A. A. Schoone, H. L. Bodlaender, and J. van Leeuwen, "Diameter increase caused by edge deletion," *Journal of Graph Theory*, vol. 11, no. 3, pp. 409–427, 1987.

[24] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Rev. Mod. Phys.*, vol. 87, pp. 925–979, Aug 2015.

[25] T. Moscibroda, S. Schmid, and R. Wattenhofer, "On the topologies formed by selfish peers," in *PODC*. ACM, 2006, pp. 133–142.

[26] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge University Press, 2007.

[27] A. Nahir, A. Orda, and A. Freund, "Topology design and control: A game-theoretic perspective," in *IEEE INFOCOM*, april 2009, pp. 1620 –1628.

[28] E. A. Meirom, S. Mannor, and A. Orda, "Network formation games with heterogeneous players and the internet structure," in *ACM EC (Conference on Economics and Computation)*. New York, NY, USA: ACM, 2014, pp. 735–752.

[29] ——, "Formation games of reliable networks," in *IEEE INFOCOM*, April 2015, pp. 1760–1768.

[30] K. Avrachenkov, J. Elias, F. Martignon, G. Neglia, and L. Petrosyan, "A nash bargaining solution for cooperative network formation games," in *IFIP Networking*, ser. LNCS. Springer, 2011, vol. 6640, pp. 307–318.

[31] S. Yerramalli, R. Jain, and U. Mitra, "Coalitional games for transmitter cooperation in mimo multiple access channels," *IEEE Trans. on Signal Processing*, vol. 62, no. 4, pp. 757–771, Feb 2014.

[32] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," *Computer Science Review*, vol. 3, no. 2, pp. 65 – 69, 2009.

[33] J. Omić, A. Orda, and P. Van Mieghem, "Protecting against network infections: A game theoretic perspective," in *IEEE INFOCOM*, 2009, pp. 1485–1493.

[34] J. Aspnes, K. Chang, and A. Yampolskiy, "Inoculation strategies for victims of viruses and the sum-of-squares partition problem," *Jour. of Computer and System Sciences*, vol. 72, no. 6, pp. 1077 – 1093, 2006.

[35] D. Acemoglu, A. Malekian, and A. Ozdaglar, "Network security and contagion," MIT/LIDS intern report, Tech. Rep., 2013.

[36] M. Lelarge and J. Bolot, "Network externalities and the deployment of security features and protocols in the internet," in *SIGMETRICS*. ACM, 2008, pp. 37–48.

[37] S. Trajanovski, Y. Hayel, E. Altman, H. Wang, and P. Van Mieghem, "Decentralized Protection Strategies against SIS Epidemics in Networks," *IEEE Trans. Control of Network Systems*, vol. 2, pp. 406–419, 2015.

[38] Y. Hayel, S. Trajanovski, E. Altman, H. Wang, and P. Van Mieghem, "Complete game-theoretic characterization of sis epidemics protection strategies," in *Proc. of IEEE CDC (53rd IEEE Conference on Decision and Control)*, Los Angeles, CA, USA, December 2014, pp. 1179–1184.

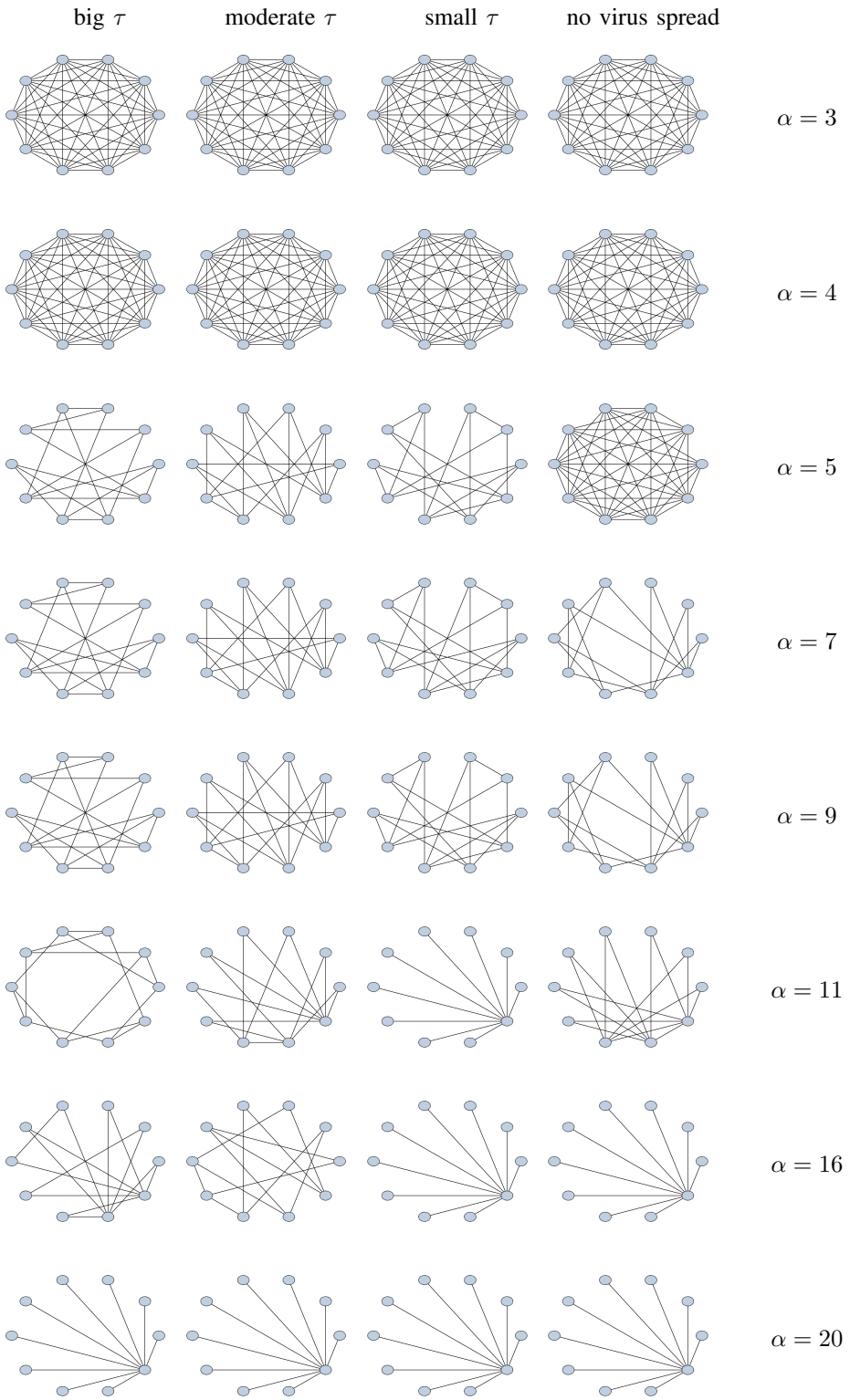


Fig. 4: Obtained topologies by the algorithm for $\gamma = 5$.

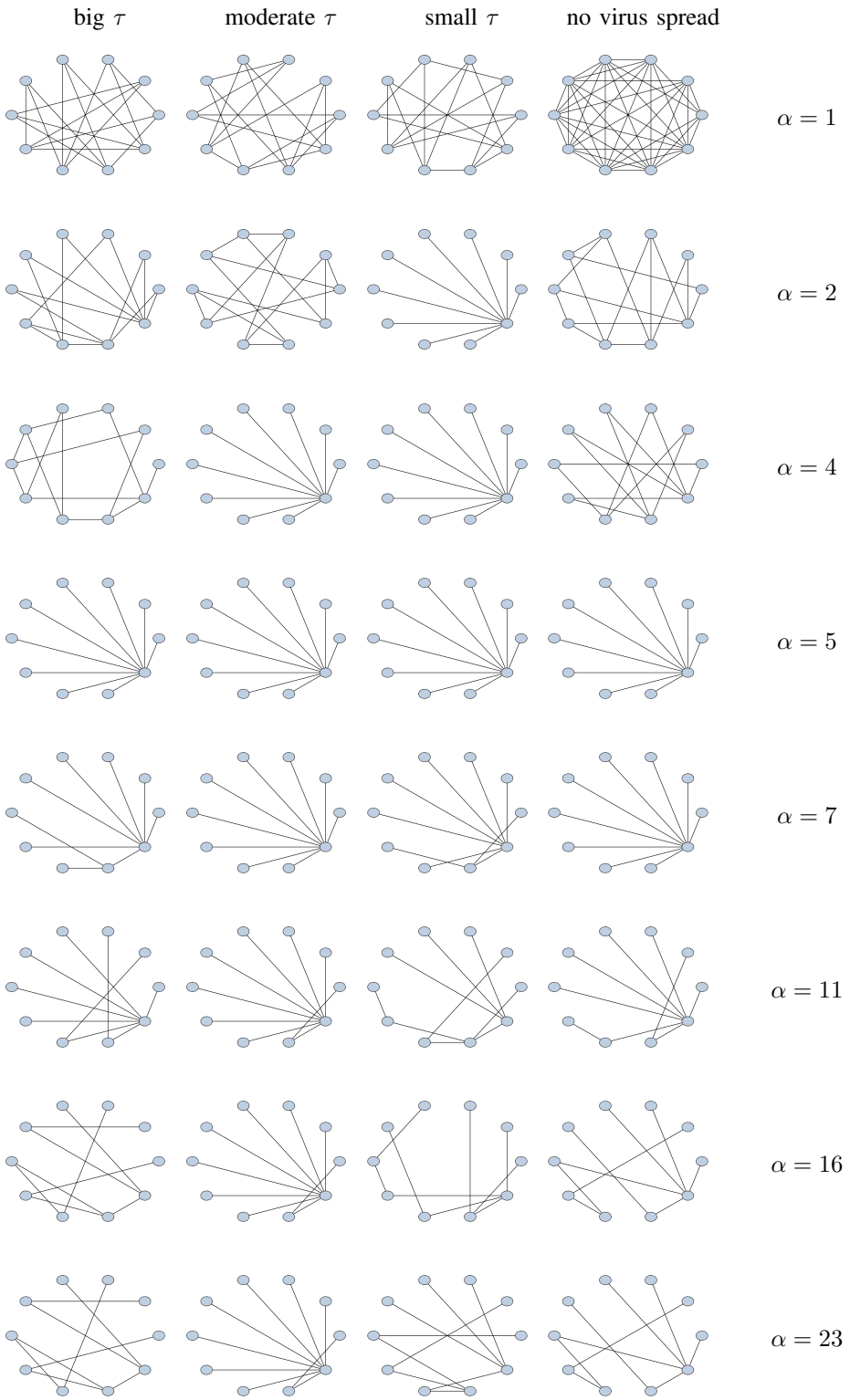


Fig. 5: Obtained topologies by the algorithm for $\gamma = 1$.

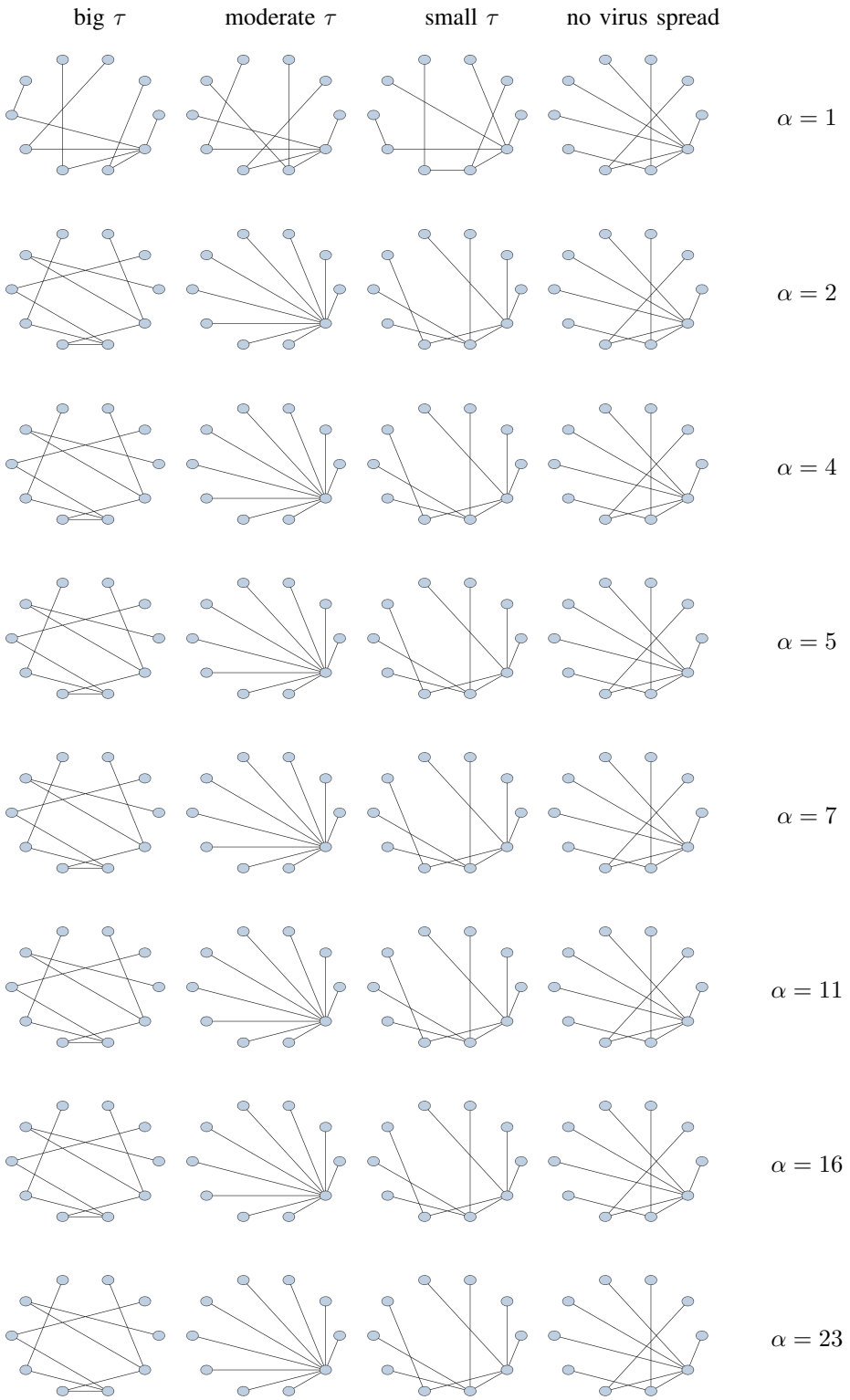


Fig. 6: Obtained topologies by the algorithm for $\gamma = 0.1$.