On Navigability for Detecting Cartels in Peer-to-peer Networks

Martin Smit[a, 1]

[a]Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Rd, Oxford OX2 6GG

This manuscript was compiled on October 25, 2022

Navigability of a complex network is the key concept underpinning the efficient communication between its nodes. In peer-to-peer networks, where players are incentivised to minimise their *stretch* to other nodes, the optimal network structure is not supported by best response dynamics, which tends to break more links than it forms. This insight motivates our formulation of a method of identifying cartel-like behaviour in peer-to-peer networks, using best-responses as a null model which also helps explain *why* clusters can be suspicious. Our method is supported by empirical criminological studies into cartels which imply that communication and reciprocity, and therefore navigability, are requirements for their long-term success. We show that the presence of cartels can have implications for the effectiveness of peer-to-peer networks for non-cartel members and significantly change the structure of the network due to cartel members becoming highly central. On synthetic data, our method is able to correctly identify 23% of all cartel members, whereas random guessing only manages 9%.

network navigability | collusion | cartel | games of network formation

A number of experiments beginning in the late 60s observed that many networks found in real life exhibit small-world characteristics (1, 2), where the shortest paths between any two nodes are very short compared to the size of the network. Also observed was that these networks were highly navigable, meaning the shortest paths are very easy to find without global information about the structure of the network. This is to be expected given that many real networks are responsible for information routing such as (biological) neural networks, and protein interactions.

Games of network formation describe the process under which rational agents construct a network. In these games the utility or cost of each player depends on the structure of the resulting network, usually in terms of links with and distances to certain nodes. When players try to optimise the network for their own local navigability, the navigability of the network as a whole suffers (3). We use this to motivate a novel method of detecting potentially collusive structures in networks by analysing the degradation in navigability derived from allowing nodes to optimise their neighbourhood structure. In *navigation games* the strategy maximising the utility within a group of nodes leaves the nodes highly connected compared to the same group in competition. We show that high levels of collusion can be pinpointed by analysing the nodes or clusters where the most change in navigability occurs.

While the effects of cartels in network games has been examined in Ref. (4), to the best of our knowledge no method for detecting collusion based on best-response dynamics in navigation games has been proposed. Additionally, our method has some grounding in socioeconomic theory (5) and jurisprudence (6) where it is theorised that cartels maintain stability through personal relationships involving, "communication, reciprocity, and reputation"(5, p. 323). The swift dissemination of information is a requirement for effective communication and verifying the reciprocity of actions, as otherwise unfaithful actors may try to "cut and run" before other members of the cartel realise.

A major disadvantage of our method comes from its reliance on best-response dynamics and Nash equilibria, the computation of which is a major sub-problem of the open **P** vs **NP** problem (7). As such, a significant part of our paper is devoted to understanding the hyperbolic spaces that larger networks "reside" in: because we cannot apply our method to large graphs, we want to be able to generate smaller, but representative networks that allow us to test our method in as real a setting as possible. Moreover, in order to apply our method to a real network, one must first map the network to hyperbolic space, so we also detail this procedure.

Structure of the paper. We first discuss some existing work on detecting criminal activity, cartels, and collusion in network settings and give a brief overview of our method.

Next, we present the tools necessary to perform our method including latent spaces, generating hyperbolic random graphs, and embedding networks into hyperbolic spaces. We give peer-to-peer computing as an example of when navigability is the fundamental goal of a network. We put forward the game whose best response dynamics we apply in our method and, continuing with the peer-to-peer example, discuss how seemingly benign collusion could lead to security issues by potentially creating a single point of failure. We test our method on synthetic data constructed from hyperbolic geometric graphs made to be representative of potential peer-to-peer networks. Finally, we reflect on the hurdles of applying this method in

Significance Statement

Peer-to-peer computing is a way of distributing computational resources to users when needed without a central authority. However, without such oversight the system is left open to exploitation by cartels taking advantage of the lack of regulation. In this paper we motivate a new method of detecting cartel-like behaviour based on allowing nodes to rewire their connections. Cartels tend to be more connected than necessary due to wanting to help each other, and so will lose connections in this procedure. While the method can help identify collusion, it is severely limited in practice by its computational requirements which scale poorly as the network grows larger.

The authors declare no competing interest.

[1]Correspondence e-mail: jacobus.smit@stx.ox.ac.uk

practice, which we hope can provide some avenues of future exploration.

The code for scale-free model generation on a hyperbolic metric space, all plots in this paper, and the collusion detection method is available on GitHub.

Previous Work

Network and complexity science has been gaining popularity in the social sciences as tools to explain the complex systems which structure and are structured by rational agents. Despite this, limited work has been done on the use of network science to detect cartels, and much of the most similar material specifically focuses on the public procurement market.

An example of such work is Ref. (8), who, similarly to our paper, base their methods on the connective patterns required to sustain co-operation between cartel members. Building on their previous work (9), they analyse the cohesion and exclusivity of nodes in the network of firms' bidding behaviour and show that these meso-scale features correlate strongly with traditional signals of collusion.

While previous literature on governmental corruption is focused on the scale of the country, Ref. (10) finds interesting results about the groups of individuals involved in corruption in Brazil. They find that the number of people involved in any given corruption case is exponentially distributed, the mean being eight people. They also show that the average shortest path is approximately a third larger than a random network, reinforcing the "theory of secret societies"(11). Finally, they attempt to predict missing links in the corruption structure, but find that the vast majority of predicted links are false positives. We note that the collusion prediction in this network is based on prior knowledge of historical malpractice and may not be useful at identifying corruption in a general network

Our Contribution. In contrast to the previous methods, we present a more theoretical i.e. less empirical, due to computational considerations, general method for detecting deviations from expected selfish behaviour in network interaction settings.

We model agents as players in a game of network formation, where a player's utility is based on how easily they can send and receive messages, routed with greedy routing, to any other player in the network. Doing so allows us to apply the game's best response dynamics to an arbitrary network, which allows players to make and break directed connections, and measure the change in the rows of the adjacency matrix using some vector norm. As a row corresponds to a player, we can identify players who are behaving in non-strategic ways.

While this technique can be applied to any game of network formation, we believe the concept of navigability, which is present in the utility function, best explains the conditions required for collusion to occur. The reciprocity and communication aspects cited as key requirements of sustained co-operation in the absence of law in Ref. (5) are provided by a high degree of navigability amongst cartel members.

Navigability of Networks

Imagine yourself dropped in an unfamiliar city and trying to navigate using only a compass. The navigability of the city would be reflected in how many dead ends you encounter and how quickly you can find the correct path to your destination simply by following roads that point in its direction. These two human concepts are very similar to the mathematical concepts used to quantify the navigability of complex networks.

Latent metric structure. Streets create a topological structure in cities; navigation in a city is done on the connections created by streets. In contrast, the compass reveals the geometric information about how far any two nodes are from each other*, regardless of whether there is a street directly connecting the two.

For certain networks, the geometry underlying the topological structure may contain useful information which can be used to predict hidden (12) or future links (13). However, the networks for which this can be applied is not simply limited to those with an obvious geographic interpretation. Network embeddings attempt to explain the topological structure of arbitrary networks as the realisation of some underlying hidden metric space in which its nodes are located (14). Using their co-ordinates in this hidden space, nodes located closely are more likely to be connected than nodes that are far apart.

A *greedy* path is found by choosing the next node in the path to be the one that brings you the closest to your destination. Of course, this may not be the shortest possible path, and one may even fail to reach the destination if one encounters a node whose edges all end further away than the current node. These two inefficiencies are the ones we encountered in our city navigation metaphor. We can distil their essence into the definitions of *stretch* between two nodes: how much further the shortest path is than the geometric distance, and the *success ratio* of greedy paths reaching their destinations without encountering dead ends. When we describe the stretch or success ratio of a network, we refer to the empirical average of all nodes' stretches or success ratios respectively.

Clearly, a fully connected network would have a perfect score in both categories (for an accurate embedding): the shortest path would be the direct route, and greedy routing would never fail as you would immediately arrive at your destination. On the other hand, a network with a disconnected component could not have a perfect success rate, and the stretch determining would be infinite as the geometric distance would be finite, but the shortest path between nodes is generally defined to be infinite if there does not exist a path connecting these nodes.

Hyperbolic latent spaces. Some characteristics of real networks such as sparsity, high clustering, and small-worldness are found in networks generated by the Watts–Strogatz model (15) However the degree distribution it produces is not scale-free, which is where $P(k) \sim k^{-\gamma}$, where $P(k)$ is the probability of a node having degree k, and $\gamma \in (2, 3)$.

To generate such networks, we turn to hyperbolic geometry: we uniformly sample points in a circle in hyperbolic space, then connect two nodes based on their hyperbolic distance. This procedure produces networks that have the "correct" degree distribution when compared to real complex networks (16). We omit a proof but provide the intuition of this fact and pseudocode for generation in Alg. 1 adapted from Ref. (17). We note that more efficient algorithms exist (18, 19), but given that this is not a bottleneck for us, we chose not to focus on this.

*The author notes that they were born in 1999 and so have never actually used a compass.

The key property of hyperbolic space that explains why they produce scale-free degree distributions is that space expands exponentially in all directions. Due to the exponential expansion, circles in hyperbolic space depend exponentially on the radius:

$$\text{Area}(r) = 4\pi \sinh^2(\frac{r}{2}) = \Omega(e^r),$$

as opposed to polynomially in Euclidean space, where

$$\text{Area}(r) = \pi r^2 = \Omega(r^2).$$

The implication of this is that if one distributes points uniformly on a circle of radius R in hyperbolic space, from a Euclidean perspective almost all of the points will be near the perimeter of the circle, as illustrated in Fig. 1.

Similarly, if we connect vertices i and j based on their hyperbolic distance d_{ij} with probability

$$P(d_{ij}) = (1 + e^{\frac{1}{2T}(d_{ij}-R)})^{-1}, \qquad [1]$$

given a temperature parameter $T > 0$ influencing how likely a node near the boundary is to be connected, then the average node degree decreases exponentially the further the node is from the origin. This is because even though there are exponentially more nodes, the hyperbolic distances between these nodes are also exponentially larger due to their distance from the origin.

The combination of nodes becoming exponentially more dense and having exponentially fewer connections leads to a power-law degree distribution.

For a more thorough look into the properties of hyperbolic geometry in relation to complex networks, we refer the reader to Refs. (16, 20).

Algorithm 1 Generate a Hyperbolic Random Graph (HRG)

Require: $N \in \mathbb{N}, \gamma \in (2,3), C \in \mathbb{R}, T \in \mathbb{R}$
Ensure: $\mathcal{G}_{\alpha,C}(V,E)$ is a HRG
  $R \leftarrow 2\log N + C$
  $\alpha \leftarrow (\gamma - 1)/2$
  for $i \in \{1,\ldots,N\}$ **do**
    $\theta_i \sim \text{Uniform}([0,2\pi])$
    $r_i \sim P$ where $p(r) = \alpha \frac{\sinh(\alpha r)}{\cosh(\alpha R)-1}$
    $v_i \leftarrow (r_i, \theta_i)$
  for $i \in \{1,\ldots,N\}$ **do**
    for $j \in \{1,\ldots,i\}$ **do**
      $d \leftarrow d_{\mathbb{H}(v_i,v_j)}$
      $P_d \leftarrow (1 + e^{\frac{1}{2T}(d-R)})^{-1}$
      Append E with edge $\vec{v_i v_j}$ with probability P_d
      Append E with edge $\vec{v_j v_i}$ with probability P_d

Techniques for embedding networks in hyperbolic space. Network embeddings are important to our method for two reasons. The first is because geometric distances are required to calculate the stretch between two nodes and stretch features in the utility function of the game we present. Secondly, as previously mentioned the available data of P2P networks (21) is far too large to be used directly. For this reason we must generate smaller, representative datasets that have similar degree distributions and other properties.

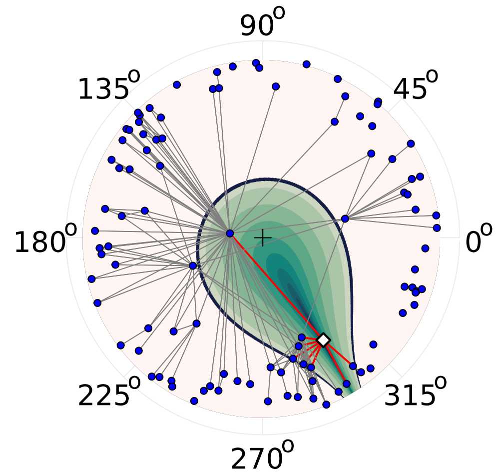

Fig. 1. The ball in hyperbolic space, centred at the white diamond node, is teardrop shaped from a Euclidean perspective. The further away from the origin, the harder it is to move across the angular dimension which causes its furthest point from the origin to appear spiky. The network is a realisation of Alg. 1 with $T = 0$ which shows that even when nodes are very close in the angular dimension, they may not be connected due to the shape of their hyperbolic ball.

Network embedding methods typically fall into two categories: likelihood maximising, and unsupervised. In a likelihood maximising setting (such as the one detailed in the appendices of (22)), we give the algorithm a probabilistic generative model (such as Alg 1) and a network, and tell it to find the geometric positions in hyperbolic space that maximises the likelihood that the model produces the input network.

In comparison, an unsupervised approach such as Isomap, among the techniques discussed in Ref. (23), embeds a network as follows: first, calculate a nearest neighbourhood for every node, then construct a new network where connections are precisely the neighbourhoods for each node, compute the geodesic (i.e. shortest-path) distance matrix of this new network, and finally, perform dimensionality reduction on the geodesic matrix by eigendecomposition.

One criticism of the second category of methods raised by the authors of Ref. (24) is that the approach does not explain the generation of the networks. For this reason, despite their relatively poor computational complexity, we make use of maximum likelihood methods. In particular, we use the NetHypGeom package.

Application of navigability: peer-to-peer (P2P) networks. By pooling together resources and sharing them when needed, P2P networks can be incredibly efficient. The key idea is that peers contribute some of their computing power (also simply called *compute*) to the network and in return they can draw from the collective compute in the network when they have large tasks to perform. There is no central server responsible for processing job requests and allocating resources to workstations, in its place every computer is both a server and a workstation.

The requirement to contribute in order to draw from the resources allows the network to provide for its members no matter how many there are, however the decentralised nature of tasks means that messages are constantly being passed between peers and so good communication infrastructure is key to the functionality of the network. The difference between the structure of centralised and decentralised networks is that in a centralised network it is important to have rapid information transfer to and from the main server, as all information must pass through it, whereas in a decentralised network each peer must be easily accessible by every other peer. This requirement leads naturally to the idea that peers want to maximise their *centrality* in some respect.

Especially relevant to our paper is that large P2P networks are scale-free (25), which implies they can be accurately represented by their embeddings into hyperbolic hidden metric spaces, and so we can calculate the stretch between points and therefore navigability of the network. Additionally, we can generate synthetic P2P networks on which to test our method and be reasonably confident that our results will generalise to larger networks, which helps to address one of the method's drawbacks.

It is natural to consider greedy packet routing as the protocol for P2P networks as it is both highly effective in real networks (22), and computationally inexpensive compared to computing shortest paths. In this setting, consider a member of a P2P network who would like to make effective use of the resources. Ideally, they would store the addresses of every other member, so they could directly send and receive information to every peer, but doing so would be impractical in a larger network. Hence, each peer chooses to store the addresses of only a subset of the entire network. We can model this decision process using game theory.

Network Formation to Maximise Local Navigability

Consider players as nodes on graph $\mathcal{G}(V, E)$, whose directed edges are defined as the result of a game. A strategy for player i is a vector of edges constructed by and whose source is i, and the cost to i that they[†] aims to selfishly minimise is

$$\alpha\, n_i + \sum_{j \neq i} \mathrm{Stretch}_{(\mathcal{G},\mathcal{E})}(i,j), \qquad [2]$$

where n_i is the number of edges constructed by i, α is the cost of building an edge, and

$$\mathrm{Stretch}_{(\mathcal{G},\mathcal{E})}(i,j) = \frac{\mathrm{ShortestPath}_{\mathcal{G}}(i,j)}{\mathrm{GeometricDistance}_{\mathcal{E}}(i,j)}. \qquad [3]$$

Note that the stretch may also be defined with the distance of the greedy path between i and j as the denominator, but defining it this way ensures that the stretch is always well defined given some embedding $\mathcal{E}$ of the network. We omit the subscripts from the stretch for clarity when the embedding or graph are not integral to the point being made or otherwise obvious from context.

A set of strategies constitutes a Nash equilibrium if no peer can change its outgoing edges and reduce its cost given that the other strategies, and thereby edges, remain the same. This game was defined and studied in Ref. (3), where they first show that its *Price of Anarchy* (*PoA*, defined as the ratio between the network with the lowest total cost and the Nash equilibrium with the highest total cost) is of order $\Theta(\min(\alpha, n))$, where $n = |V|$ is the number of nodes. They go on to show that there exist metric spaces and therefore embeddings for which there does not exist a Nash equilibrium, meaning that the best-response dynamics are stuck in an infinite loop. Additionally, they show that given an embedding, determining the existence of a Nash equilibrium is **NP**-hard.

From this we have two takeaways: firstly, networks formed by selfish agents will be less navigable than those by agents working together, and secondly, selfish agents may produce unstable cycles. We base our method for detecting cartel-like behaviour on the first observation, while taking the second into account when discussing its potential drawbacks.

The Dangers of Cartels. We assume for simplicity that players that form coalitions attempt to maximise the sum of utilities of players in the coalitions, disregarding their individual utility. This is equivalent to cession of control of strategy to the coalition, which can be modelled as a benevolent dictator to the members of the coalition.

In P2P networks, participation incentive schemes act to induce more-than-minimal resource sharing with the network to reduce free-riding among participants (26). If these coalitions exist, then the assumption that many schemes are based on, independent selfish agents, is broken, which can lead to exploits that try to gain the system for money, or tokens that allow users to make more use of the network.

In other cases, an over-reliance on the coalition for the navigability of the entire network could arise, leading to a single point of failure which could cause a huge deterioration in the quality of service if attacked.

Finally, the formation of links that would not occur as a result of strategic interactions of self-interested players, or the lack thereof, could indicate that the agents have ulterior motives not captured by the utility functions.

Our method

For a given P2P network, i.e. one that is directed, scale-free, and built around navigabiltiy, we embed the network into some hyperbolic space before applying best response dynamics and analysing the changes this makes to the network.

This analysis can be many things, and looking into how exactly the network changes with the influence of cartels and without is an interesting problem in and of itself. For this paper, we analyse the proportion of total difference in out-degree between each player's strategy (their row in the forward adjacency matrix) with the hypothesis that cartel members want to have a higher out-degree, which we find to be true.

Example. Let the adjacency matrices in a strategic and cartel graph respectively be:

$$S = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix},$$

where the cartel members are 1 and 4. Then using the method from before, the *suspiciousness* of each player is their number of new out-edges, divided by the total number i.e.

$$\mathrm{suspiciousness} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

[†] In more traditional Game Theory, convention has odd-numbered players as female and others as male. In this setting, it is unproductive to label and keep track of so many players, so we refer to all as a singular gender-neutral pronoun "they".

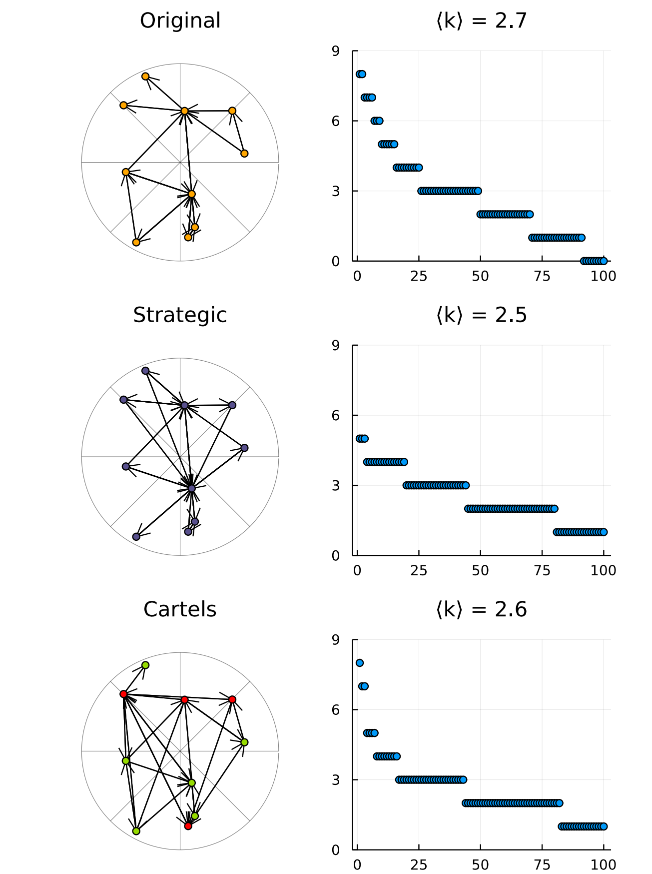

Fig. 2. Compare an instance of a generated network, the network where players are selfish, and the network where the nodes in red form a cartel. We also plot the out-degree distribution of **all** of the networks tested and note that the strategic networks have a noticeably more egalitarian degree distribution, whereas the cartel networks have one similar to the original networks.

Then, knowing that there are 2 members of the cartel, we select the players with the two highest suspiciousness: 1, and 4.

Results

Due to the computational complexity of solving for Nash equilibria, we restrict our findings to very small networks of 10 nodes. We generate a network and its embedding, decide on which nodes will form cartels, and apply best response dynamics. This leaves us with a simulated P2P network under the influence of cartels to which we can apply our method.

For simplicity we refer to an unaltered generated network as the *original* network, a network where best responses have taken place as a *strategic* network, and a network with a cartel where best responses have taken place as a *cartel* network.

Structural changes in the network. We can see in Fig 2 that the strategic network has very few nodes with out-degree one and most nodes have degree two, linking to both hubs. In the cartel network, cartel members (in red) try to create direct links with other cartel members, which turns them into pseudo-hubs and so other players also try to link with them. Overall, the cartels network looks more like the original network than the strategic one.

Connections	Cartel	Non-cartel
Cartel	0.41	0.52
Non-cartel	-0.15	-0.65

Greedy path length	Cartel	Non-cartel
Cartel	-0.049	-0.081
Non-cartel	-0.0072	0.0013

Table 1. The change in some attribute going from a strategic network to a cartel network. Connections in these tables are row → column.

We see in Table 1 that cartel members tended to help each other by building more connections than they would have otherwise, both amongst each other and to other nodes. This was accompanied by a drop in the number of nodes built by other players. We see that on average, the cartel to non-cartel connections increased the most, and this is counteracted by the decrease in non-cartel to non-cartel connections.

The greedy path length from any two cartel members, defined as the average number of steps taken by a greedy forwarding algorithm beginning and ending at a cartel member, is slightly higher for non cartel members. In a P2P setting, this would mean that the interactions, e.g. file sharing or messaging between non-cartel members would take slightly longer to complete. We do see that the overall path length seems to be decreased with the presence of the cartel, however we note that the cartel comprises of 4 players out of a total 10, and so the gains they see from their collusion is a significant portion of the total improvement.

Accuracy. Keeping a single cartel made up of the first 4 players, we generate a cartel network and apply best-responses to it. From here we can see which nodes have gained the most out-degree.

Applying this technique 150 times, we correctly identify 0.93 cartel members, or 23% of the cartel. A null model manages 0.35 members correctly identified, or 8.8%.

The accuracy of this technique is strongly influenced by the way in which the networks are analysed. For example, if we instead use the *Hamming distance* between the strategies (rows of the forward adjacency matrix) then our accuracy decreases to 15%.

Discussion and Future Work

The biggest challenge with developing and testing this method further is the computational challenge that computing best responses imposes. In our implementation, we do not even check whether the outcome of applying best response dynamics is indeed a Nash equilibrium. For N vertices, an exhaustive search of all possible combinations of strategies of every player is 2^{N^N}. Also, calculation of best responses is done by brute force; it may be possible to improve the complexity of the method by approximating best response behaviour by noting that the majority of nodes have a certain number of connections. However, even if best responses could be calculated quickly there is no guarantee that they will converge to the equilibrium even if one exists.

Another approach that could be taken is to somehow treat the Nash equilibrium as a null model. With this, statistical testing could be done to find the p-value that a network was generated by selfish agents i.e. no cartels given agents make mistakes with some probability p. Then, if the hypothesis were rejected, it would imply that there is some deviation in utility functions for some players.

Another issue with this method is that it relies on having an accurate embedding in order to properly calculate stretch. If the embedding is not very good, then the utility calculations will be wrong, so it is possible that false positives could arise as well as a reduction in accuracy.

One thing we did not test in this paper is the ability for our method to identify the number of members of a cartel. We always gave it the information that there was one cartel of four members and simply asked for which four were the most likely. We believe that our method would not perform well as it currently is, but this functionality could be implemented by having a suspiciousness threshold.

Finally, while we use one specific game to compute best responses, the choice may not be optimal. Moreover, while we believe that the concept of navigability is important to identifying cartels, the method could potentially be used in a wide variety of situations to identify deviation from selfish behaviour. This would indicate that the model being tested would require improvement in accuracy.

1. J Travers, S Milgram, An Experimental Study of the Small World Problem. *Sociometry* **32**, 425 (1969).
2. YC Yang, Terminal Road Bridges for San Francisco International Airport Airport,". *J. Proc.* **69**, 614–618 (1972).
3. T Moscibroda, S Schmid, R Wattenhofer, On the topologies formed by selfish peers in *Proceedings of the Twenty-Fifth Annual ACM Symposium on Principles of Distributed Computing*, PODC '06. (Association for Computing Machinery, New York, NY, USA), pp. 133–142 (2006).
4. A Hayrapetyan, É Tardos, T Wexler, The effect of collusion in congestion games in *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '06. (Association for Computing Machinery, New York, NY, USA), pp. 89–98 (2006).
5. JD Jaspers, Managing Cartels: How Cartel Participants Create Stability in the Absence of law. *Eur. J. on Crim. Policy Res.* **23**, 319–335 (2017).
6. A Stephan, Cartel Laws Undermined: Corruption, Social Norms, and Collectivist Business Cultures. *J. Law Soc.* **37**, 345–367 (2010).
7. C Daskalakis, P Goldberg, C Papadimitriou, The Complexity of Computing a Nash Equilibrium. *SIAM J. Comput.* **39**, 195–259 (2009).
8. J Wachs, J Kertész, A network approach to cartel detection in public auction markets. *Sci. Reports* **9**, 10818 (2019).
9. M Fazekas, J Skuhrovec, J Wachs, Corruption, Government Turnover, and Public Contracting Market Structure Insights Using Network Analysis and Objective Corruption Proxies. *SSRN Electron. J.* (2017).
10. HV Ribeiro, LGA Alves, AF Martins, EK Lenzi, M Perc, The dynamical structure of political corruption networks. *J. Complex Networks* **6**, 989–1003 (2018).
11. WE Baker, RR Faulkner, The Social Organization of Conspiracy: Illegal Networks in the Heavy Electrical Equipment Industry. *Am. Sociol. Rev.* **58**, 837–860 (1993).
12. M Kitsak, I Voitalov, D Krioukov, Link prediction with hyperbolic geometry. *Phys. Rev. Res.* **2**, 043113 (2020).
13. W Yu, W Cheng, CC Aggarwal, H Chen, W Wang, Link Prediction with Spatial and Temporal Consistency in Dynamic Networks. *Int. Jt. Conf. on Artif. Intell. Organ.* pp. 3343–3349 (2017).
14. F Papadopoulos, D Krioukov, M Boguna, A Vahdat, Greedy Forwarding in Dynamic Scale-Free Networks Embedded in Hyperbolic Metric Spaces. *2010 Proc. IEEE INFOCOM* pp. 1–9 (2010).
15. DJ Watts, SH Strogatz, Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
16. D Krioukov, F Papadopoulos, M Kitsak, A Vahdat, M Boguna, Hyperbolic Geometry of Complex Networks. *arXiv:1006.5169 [cond-mat, physics:physics]* (2010).
17. L Gugelmann, K Panagiotou, U Peter, Random Hyperbolic Graphs: Degree Sequence and Clustering - (Extended Abstract) in *ICALP.* (2012).
18. R Aldecoa, C Orsini, D Krioukov, Hyperbolic Graph Generator. *Comput. Phys. Commun.* **196**, 492–496 (2015).
19. T Bläsius, et al., Efficiently Generating Geometric Inhomogeneous and Hyperbolic Random Graphs. *arXiv:1905.06706 [cs]* (2019).
20. D Krioukov, F Papadopoulos, M Boguñá, A Vahdat, Greedy forwarding in scale-free networks embedded in hyperbolic metric spaces. *Sigmetrics Perform. Eval. Rev. - SIGMETRICS* **37**, 15–17 (2009).
21. M Ripeanu, I Foster, A Iamnitchi, SNAP: Network datasets: Gnutella peer-to-peer network, August 31 2002 (https://snap.stanford.edu/data/p2p-Gnutella31.html) (2002).
22. M Boguna, F Papadopoulos, D Krioukov, Sustaining the Internet with Hyperbolic Mapping. *Nat. Commun.* **1**, 62 (2010).
23. A Muscoloni, JM Thomas, S Ciucci, G Bianconi, CV Cannistraci, Machine learning meets complex networks via coalescent embedding in the hyperbolic space. *Nat. Commun.* **8**, 1615 (2017).
24. MÁ Serrano, M Boguñá, *The Shortest Path to Network Geometry: A Practical Guide to Basic Models and Applications.* (Cambridge University Press), First edition, (2022).
25. M Ripeanu, I Foster, A Iamnitchi, Mapping the Gnutella Network: Properties of Large-Scale Peer-to-Peer Systems and Implications for System Design. *arXiv:cs/0209028* (2002).
26. P Golle, K Leyton-Brown, I Mironov, Incentives for sharing in peer-to-peer networks in *Proceedings of the 3rd ACM Conference on Electronic Commerce*, EC '01. (Association for Computing Machinery, New York, NY, USA), pp. 264–267 (2001).