Selfish Behavior and Compact Representation in Routing and Information Networks

Aikaterini Papakonstantinopoulou*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
katia@di.uoa.gr

Abstract. In this short note we summarize our results on the study of selfish user behavior in modern networks and the compact representation of such networks in systems' level. In the first part of the dissertation we use principles of Game Theory to build frameworks for the theoretical study of problems arising in routing networks and the worldwide web due to the selfish nature of their users, while in the second one we move to the important technical issue of representing such networks so that they can fit in the computer's main memory but also boost the efficiency of critical applications that run over these networks. The results include the introduction and study of the class of congestion games with time-dependent strategies, the analysis of game-theoretic aspects of link placement in the worldwide web and the design and evaluation of a network compression algorithm that outperforms the state-of-the-art method by achieving a better compression ratio and retrieval time of the network's elements.

1 Introduction

During the last fifteen years, the explosive growth of the internet and its use in everyday life has given rise to various types of networks created by human activity, without any central design or control. The most well known example is the worldwide web (WWW), which became an object of study when it started being used by a huge number of users and thus attracted a lot of information-oriented and economic activity. More recently a number of other networks available through internet have been developed, that also offer abstractions of the social context of the users, referred to as social networks. The development of the internet has also motivated the study of routing networks, as there are usually various alternates for the communication between two network entities.

The popularity of such networks has clearly affected the incentives of their users. For instance, web page authors are interested in establishing hyperlinks that increase not only the quality but also the *popularity* of their pages. However, attracting users is not desirable in all types of networks. For example, in routing networks it is usually in users' interest to *avoid* crowded communication

* Dissertation Advisor: Elias Koutsoupias, Professor

channels. Apparently, since we need to be able to understand the mechanisms behind the creation of such networks, network analysis has become an intriguing object of study.

Apart from the user point of view consequences of the networks' popularity, there are also consequences from the system point of view. Perhaps the most basic such issue is that the network representations cannot fit in a computer's main memory, leading to two drawbacks: a restriction on the network size that can be handled, or poor performance of the critical applications that run over such network infrastructures, implied by the usage of secondary memory. Neither of them is acceptable, so we need compressed representations that allow fast access to the network's elements.

The main questions that motivated our research are the following:

– What is the optimal *timing* for a user to enter a *routing network*?
– What is the optimal *linking* strategy in the *worldwide web*?
– How can we efficiently *represent* the WWW, social or routing networks?

This dissertation has been conducted along the two directions sketched above. The first one is the study of *selfish behavior* in various networks of special interest in our days, in particular routing networks and information networks. In both types the users are trying to maximize their own payoff from the network; the different incentives, however, make the individual study of each network type necessary. In the case of routing networks, which we study theoretically using congestion games, we introduce and study the class of congestion games with time-dependent strategies. We model the routing networks using congestion games, because it is considered the most appropriate approach for the theoretical study of congestion networks. In a more practical level, we model and study the worldwide web, which is the largest information network in our days, from a game-theoretic point of view that captures naturally the users' incentives. The second direction has to do with an important technical issue, the representation of these network structures and more specifically with their *compact-yet-efficient* representation, that can boost the performance of critical applications that run on such networks. We study the common properties of the networks of interest and exploit them to design an efficient lossless compression method.

The contribution of the thesis [35] is threefold, and we outline it in the following sections. We model the selfish behavior of the users of such systems using the Game Theory framework. Regarding the class of congestion games with time-dependent strategies, we describe its structural properties and compute its Nash equilibria and prices of anarchy and stability. A brief illustration of these results can be found in section 2.1. Then, in section 2.2, we give an idea of our analysis of the game-theoretic aspects of the worldwide web, considering a link placement game among the web authors and computing their exact and approximate best responses. Both these objects are frameworks for studying classes of problems. Finally we focus on the representation of networks created by human activity, with routing networks, social networks and the worldwide web being the most popular examples. We propose a network representation method that outperforms the state-of-the-art method in terms of compressed size and access time to the elements of the network. Section 3 outlines our approach, its analysis and experimental evaluation.

Most of these results have appeared in [24,25,26]. What follows is a brief presentation of the research topics and results of the thesis, avoiding technical details.

2 Selfish Behavior in Routing and Information Networks

We first consider routing (or traffic) networks and study the effect of the participation timing in such systems (Section 2.1). We then consider information networks, focus on WWW and highlight its game-theoretic aspects (Section 2.2).

2.1 Congestion games with time-dependent strategies

In the last dozen years, the concepts of the price of anarchy (PoA) and stability (PoS) have been successfully applied to many classes of games, most notably to *congestion* games and its relatives [23,40,30]. In congestion games, the players compete for a set of resources, such as facilities or links; the cost of each player depends on the number of players using the same resources; the assumption is that each resource can be shared among the players, but with a cost. Another interesting class of games are the *contention* games [16] in which the players again compete for resources, but the resources cannot be shared. If more than one players attempt to share a resource at the same time, the resource becomes unavailable and the players have to try again later. There are however interesting games that lie *between* the two extreme cases of the congestion and contention games. For example, the game that users play for dealing with congestion on a network seems to lie in between—the TCP congestion control policy is a strategy of this game. Timing is part of the strategy of the players (as in contention games) and the latency of a path depends on how many players use its edges (as in congestion games).

In our work, we attempted to abstract away the essential features of these games, to model them, and to study their properties, their Nash equilibria, and their price of anarchy and stability. The games that we consider are essentially congestion games with the addition of time dimension. The difference with congestion games is that players now don't simply select which path to use, but they also decide *when* to enter the system, and hence the delay they suffer depends on the number of users that use *the same path edges at the same time* as they do.

Related work. Contention resolution in communication networks is a problem that has attracted the interest of diverse communities of Computer Science, due to the fact that contention is inherent in many critical network applications. One of them is the design of multiple access protocols for communication networks, such as Slotted Aloha. In such protocols, the increase of users of the network incurs a large number of collisions and subsequently poor utilization of the system's resources.

During the last four decades many more refined multiple access protocols have been proposed to increase the efficiency of Aloha, the vast majority of which assume that the agents follow the protocol, even if they might prefer not doing so. Recently, slotted Aloha has been studied from a game-theoretic point of view, trying to capture the selfish nature of its users [2,3,27,16,15].

Routing in networks by selfish agents is another area that has been extensively studied based on the notion of the price of anarchy (PoA) [23] and the price of stability (PoS) [4]. The PoA and the PoS compare the social cost of the worst-case and best-case equilibrium to the social optimum. Selfish routing is naturally modeled as a congestion game. The class of congestion or potential games [38,29] consists of the games where the cost of each player depends on the resources he uses and the number of players using each resource.

All these models share with our work the interest in game-theoretic issues of timing in routing, but they differ in an essential ingredient: in our games, timing is the most important part of the players strategy, while in the previous work, time delays exist because of the interaction of the players; in particular, *in all these models the strategy of the players is to select only a path*, while in our games the strategy is essentially the timing.

Our contribution. We introduce and study the class of congestion games with time-dependent strategies. Consider a link or facility e of a congestion game with latency function ℓ_e. In the congestion game the latency that a player experiences on the link is $\ell_e(k)$, where k is the number of players that use the link. In our model however, since the players may also decide when to start, we have to redefine the latency.

We propose the following latency models for the links:

The boat model, in which only the group of players that start together affect the latency of the group. Imagine that one boat departs from the source of the link at every time step; all players that decide to start at time t enter the boat which takes them to their destination; the speed of the boat depends only on the number of players in the boat and it is independent of the players on the other boats.

The conveyor belt model, in which the latency of a player depends on the number of other players using the link at the same time, regardless if they started earlier or later. Specifically, the link resembles a conveyor belt from the source to the destination; the speed of the belt at every time depends on the number of people on it.

We consider only symmetric strategies based on the assumption that these games are played by many players with no coordination among them. Moreover we assume non-adaptive strategies, in which the players decide on their strategy before the game starts. *A pure strategy of a player consists of a path and a starting time.*

We first study the structural properties of the boat and conveyor belt games. We prove that the boat games are congestion games; in contrast, we give examples which show that conveyor belt models are not in general congestion games with the exception of the case of two players. In fact, even simple games with 3 players may not even possess pure Nash equilibria.

We characterize the symmetric Nash equilibria of the boat model game for parallel links of affine latency functions, i.e., $\ell_e(k) = a_e\, k + b_e$, and any number of players. We show that there is a unique symmetric mixed Nash equilibrium for these games. At the Nash equilibrium the probability that a player starts at time t drops linearly on t.

We also compute the optimal symmetric solution. Interestingly, in both the boat and conveyor belt model, the optimal symmetric strategy has exactly the same form with the Nash equilibria but it is less aggressive. That is, in the optimal symmetric strategy the probabilities drop also linearly in time but they are spread out to more strategies. In particular, the optimal strategy is a Nash equilibrium of a game with higher latency functions (by almost a factor of 2). A similar bicriteria relation between the Nash equilibria and the optimal solution has been observed in simple congestion games before [40].

From the characterization of the Nash equilibria and the optimal strategy, we get that the price of anarchy and stability is very low $3\sqrt{2}/4 \approx 1.06$. This is the price of anarchy (and stability) when we fix the latencies and let the number of players tend to infinity; when the latency function is tailored to the number of players n, the price of anarchy can be as high as $8n/(7n + 1)$.

We also study the class of conveyor belt games. These are more complicated games and here we consider only two players and arbitrary latency functions; for two players the class of affine and the class of arbitrary latency functions are identical. We again characterize the Nash equilibria, the optimal solution, and we compute the PoA and the PoS. Specifically, we show that there exists a unique symmetric Nash equilibrium for the conveyor belt model in which the players assign non-zero probabilities to multiples of $\ell_e(1)$ and these probabilities drop linearly. The explanation of the nature of these equilibria is the following: a player attempts with some probability to start at time $t = 0$; the probability has to balance the risk of the other player starting also at time $t = 0$ and the delay incurred by waiting. The interesting property of the Nash equilibrium is that the player waits enough time steps in order to avoid interference with the other player, given that he had started at time $t = 0$. After exactly $\ell_e(1)$ steps, with some probability, the player attempts again and the process is repeated.

The price of anarchy and stability is (for large latencies) again approximately $3\sqrt{2}/4 \approx 1.06$. This is the price of anarchy we computed for the boat model, but the relation is not as straightforward as it may appear: in the boat model we take the limit as the number of players tends to infinity, while in the conveyor model, we take the limit as the latencies tend to infinity. In fact, the latter is the same limit as keeping the latencies steady and letting the time step to tend to 0 (thus approximating a continuous-time protocol).

To our knowledge, these games differ significantly from the classes of congestion games that have been studied before. Also, the techniques developed for bounding the PoA and the PoS for congestion games do not seem to be applicable in our setting. In particular, the smoothness analysis arguments [13,39,14] do not seem to apply because we consider symmetric equilibria. In fact, the focus and difficulty of our analysis is to characterize the Nash equilibria and not to bound the PoA (or PoS).

The results of this work are detailed in chapter 3 of the thesis [35] and have been published in [24]. In the context of our ongoing research on this problem, we have extended our results to more general network models.

2.2 Game-theoretic aspects of the Worldwide Web

The worldwide web has been the focus of an enormous amount of research in the last fifteen years and several models have been proposed for it. These models aim at our understanding of the properties and evolution of the web, and assist us in designing more efficient web algorithms and applications, for example search engines. Recently, the exploitation of web's link structure by the search engines as well as the emergence of advertising links have given new incentives to link placement: strategic web page owners now explicitly attempt to boost their reputation and monetary revenue by careful selection of links, and Search Engine Optimization (SEO) has grown into a billion-dollar industry. Therefore Game Theory seems to provide the appropriate framework for studying the evolution of the web. Moreover, the impact of advertising links on the link structure of the web, and consequently on the relative importance of web pages, is unknown.

In our work we introduce a game-theoretic model for the worldwide web that captures the selfish nature of web page authors. In our model the page authors decide which advertising links to buy in order to maximize their revenue, which depends on the traffic their page attracts. We use Google PageRank as a measure of traffic. Based on this model, we study the game-theoretic aspects of the worldwide web.

Related work. The first attempts to model the web graph coincide temporally with the development of successful web search algorithms which were based, partially, on the link structure of the web, with PageRank [34] and HITS [21] being the most well-known examples.

Since then many models have been proposed for the web graph, which try to predict its structural properties. Classifying the models according to the deemed linking incentive, we get the following classes: *random graph models*, in which new nodes link to existing ones with high degree or PageRank; *economic models*, in which the nodes endorse existing ones that are regarded as good web search results; and *game theoretic models*, in which nodes explicitly try to maximize their own PageRank and/or revenue.

The decisions on the link structure of each web page are made locally, with each page owner trying to maximize the value and importance of her own page.

Therefore game theory appears as the proper framework for modeling the link establishment process, and we can think of the web as the equilibrium of some network creation game among its users. During the last eight years, several aspects of linking in the web have been studied from a game theoretic point of view. The inner structure of a web site with strategic owner is studied in [18]. All other game theoretic models for the web concern the establishment of external links; either *reference only* [17,11] or *reference and advertising* [20,22].

Our contribution. We introduce and study a model for the web graph, in which selfish page owners aim at maximizing their PageRank and revenue by purchasing the appropriate incoming links to their page. We assume two different approaches for link pricing: fixed-prices and prices-per-click.

Our model is a game played among the web page owners, in which each one aims at purchasing the set of new incoming links to her page that maximizes the revenue from the page, i.e., the page's PageRank minus the prices of these links. We call this game the *web game*. We show that the web game is not a potential game for three or more players. The problem of finding the best response for each player given the strategies of the rest players is formulated as follows: Given a directed graph $G = (V, E)$, a node $u \in V$ and the price p_i of the links emanating from node i for all $i \in V$, compute the set of new backlinks of u that maximize $\pi_u - \sum_{j:j \to u} p_j$, where π is the PageRank vector. We show that the best response computation is NP-hard, since the LINK BUILDING problem, which is known to be NP-hard (in fact it has no FPTAS unless $P = NP$) [31], reduces to it. It follows that verifying a Nash equilibrium is NP-hard as well, since we have to verify that each player plays her best response given the strategies of the rest players, so we are interested in approximate best responses. In [33,32] the authors propose a constant factor approximation algorithm for the link building problem. We employ it to compute an *approximate best response* for the prices-per-click model in our web game.

The results of this work are presented in detail in chapter 4 of the thesis [35].

3 Compact Representation of Routing and Information Networks

Real-world systems and phenomena that involve interactions among various entities are being modeled using *graphs* for decades now. The recent explosive growth of large-scale systems that are traditionally modeled as graphs, the worldwide web and social networks being typical examples, has intensified the need for compact-yet-efficient representations of graphs. In particular, we need compressed graph representations that allow *mining* without decompressing the whole graph. In this way, algorithms and applications with tasks that correspond to graph mining problems, can take advantage of such representations to boost their performance, as they can run in main memory over much larger graphs using their compressed representations instead of the plain ones. For example, serving adjacency queries or maintaining and querying low-cost snapshots for archival purposes are common operations in such critical applications, and can benefit from the use of in-memory representations of graphs.

The graphs we are interested in representing share some common features. First, they represent huge networks extending to millions of nodes, but the degrees (in/out-degrees) of the latter are power law distributed [12,9], rendering the graphs to be rather *sparse*. Moreover, the graphs exhibit the *locality of reference* property: nodes tend to have successors that are 'close' to them in a sense that depends on the context and the nature of the network. For instance, web pages often contain links to pages of the same web site or domain, and people in social networks are often friends with individuals from the same neighborhood. Furthermore, these graphs exhibit the *copy* property (or *similarity* property), which denotes that nodes occurring close to each other tend to have many common successors.

These properties induce various types of redundancy in the graphs' representations, and are taken into account when designing compression methods. The state-of-the-art approach to the compact representation of graphs is the method of Boldi and Vigna [5], further improved using a reordering of the graph [6] before compressing it. The reorderings can favor any compression algorithm that takes the aforementioned properties into account.

The web and social graphs may share the above properties, but feature a substantial difference in the way they are represented: while it is easy to order the nodes of a web graph in a meaningful way which favors its compression (i.e., lexicographically by URL), there is no such obvious ordering for general networks, including social ones. This makes social and general networks less amenable to compression than web graphs and their compression is a challenging issue.

Related work. The need for compact representations of graphs emerged with the explosion of the size of the worldwide web, so the first such attempts focused on compressing web graphs. In the last dozen years graph compression has turned into a very active research area and many algorithms have been proposed, some of them designed for more general graphs like the social network ones. Most algorithms in this direction try to offer a good space/time trade-off.

The graph compression algorithms that have been proposed so far can be classified in the following three main categories: (*i*) algorithms for compressing web graphs, (*ii*) algorithms for compressing (also) more general graphs (mostly social network graphs), and (*iii*) algorithms that include or employ reordering of the graph in order to favor higher degree of compression. It is also very often the case that specific web graph compression algorithms were later enriched with new techniques in order to be able to compress social graphs as well.

In [37] the authors take into account the locality of reference and the copy properties for the case of the web and initiate research on web graph compression by maintaining compressed forms of the graph's adjacency lists. The highest compression ratios are achieved by the method of Boldi and Vigna [5], combined with a reordering using label propagation [6]. The WebGraph compression method introduced in [5] is indeed the most successful member of a family of approaches [36,9,1,42] for compressing *web* graphs based on the statistical properties described in the introduction. In another line of work, Brisaboa et al. [8] propose a compact representation of the adjacency matrix that represents the graph. The approach we propose is to some extent similar to [8], in the sense that we represent parts of the adjacency matrix of a given graph. The main difference with our approach is that we represent only some *dense* parts of the graph, those that are close to the main diagonal. In [12] Chierichetti et al. view the problem of graph compression from a theoretical point of view and study the extent to which a large social network can be compressed. Their proposed method, however, is a compression scheme rather than a compressed data structure, as noted in [6], i.e., it aims solely at minimizing the size of the compressed graph (bits/edge) instead of providing fast access to each edge.

The locality of reference property of a graph reflects on its adjacency matrix in the following way: using a proper ordering of the nodes' labels, i.e., an ordering in which labels of densely connected nodes are close to each other, many edges fall close to the main diagonal of the adjacency matrix. Such orderings are preferred in practice, but finding the ordering that minimizes the distance of the edges from the main diagonal is NP-hard [41]. In [7] Boldi et al. test some known orderings of the nodes and propose some new ones, and study their effect on the compression of web and social graphs.

Some methods that claim to yield lower bits/edge ratios [12,10,28] do not address the issue of retrieving the edges fast. In [19] the authors introduce SlashBurn, an ordering method that offers the best bits/edge ratio according to the information theoretic lower bound, among other competing methods.

Our contribution. We concentrate on the compression of web, social network and other similar graphs by exploiting the locality of reference property. After observing that all real graphs of the above types we tested, as well as most graphs created by human activity, demonstrate the locality property, i.e., they can be represented by adjacency matrices with high concentration of edges around the main diagonal of the matrix, we exploited this fact to improve their compression.

Since the highest compression ratios are achieved by the state-of-the-art algorithm of [5] (denoted as BV) after applying the *Layered Label Propagation* (LLP) algorithm [6] on the input graphs, we decided to build on it, achieving the following:

– we improved BV by exploiting the locality property in a different way than in [5] and, thus, went beyond the state-of-the-art in graph compression
– we evaluated experimentally our method using real datasets and clearly verified that it achieves a better compression ratio than BV, while allowing faster retrieval of the elements of the graph than BV.

We assume that we apply our algorithm on a reordered version of the input graph, using for example the reordering algorithm of [6].

In our proposed method [25] we isolate the dense part of the graph's adjacency matrix, which lies around its main diagonal (referred to as the diagonal stripe), and use a bit vector to represent it, while we resort to BV to compress the remaining edges. Every possible pair of nodes lying in the diagonal stripe is mapped through a simple function to the bit vector. Thus, the existence of an edge there can be verified in constant time. A large percentage of these pairs represent edges absent from the graph; however, including them in our representation allows us to be aware of the position of every pair and not resort to using an index as in [8], which would not only introduce a similar space overhead, but would dramatically increase the retrieval time as well. By using BV to compress the rest, sparse part, of the graph, we manage to provide a full graph compression framework and perform comparisons over the whole graph, not only the diagonal stripe. The computational complexity of this approach is approximately equal to the complexity of BV alone, as mapping the diagonal stripe to a bit vector is linear in the number of diagonal edges. Furthermore, this mapping can only decrease the query time on the compressed graph's elements when compared with the query time of BV alone. We tested our approach on a dataset of six real social network graphs. Our method outperformed BV for the whole dataset, showing that the effect of our observation is powerful on social network graphs.

In order to make our method efficient for a wider class of graphs, including web graphs and routing networks, we refined it further [26]. Using data compression techniques that exploit the redundancy of the diagonal stripe, allows us to reduce the size of the stripe significantly. Shannon's source coding theorem states that it is impossible to compress with an average number of bits per symbol less than the entropy of the source. We present a proposition that imposes an upper bound to that limit, and provides us with an estimation of the space requirements of our method for the dense part of the graph. Comparing this estimation for various widths of the diagonal stripe of a graph, to the compression ratio of the state-of-the-art method, we may assess the overall room for improvement and the optimal width of the stripe. However, the estimation on the latter is far from accurate due to the delicate balance between easing the task of compressing the rest of the graph by including as many edges as possible in the diagonal stripe, on one hand, and minimizing its ratio, on the other. We still wish to retain the ability to access the elements of the stripe in constant time after compressing them. Therefore, we encode them using a form of lossy, but fixed-length encoding to preserve the direct access of the edges. The edges that are excluded during this step are added to the ones existing outside the diagonal stripe. These edges will be then compressed using BV, thus, our overall method is lossless. The dataset that we used to apply and test our compression technique, comprises nine well-studied *web, road network, citation, and social network* graphs. Our method still outperforms BV, with impressive results for road networks (>20% decrease in compressed size), a large impact on social network graphs (approximately 16.6% decrease) and improvements even on web graphs. There are significant improvements in access times as well.

The results of this work are presented in chapter 5 of the thesis [35] and have been published in [25,26]. The developed software is open-source, publicly available at: `https://bitbucket.org/network-analysis/bvplus`

References

1. Micah Adler and Michael Mitzenmacher. Towards Compressing Web Graphs. In *DCC*, 2001.
2. Eitan Altman, Rachid El Azouzi, and Tania Jiménez. Slotted aloha as a game with partial information. *Comput. Netw.*, 45:701–713, August 2004.
3. Eitan Altman, Dhiman Barman, Rachid El Azouzi, and Tania Jiménez. A game theoretic approach for delay minimization in slotted ALOHA. In *IEEE International Conference on Communications*, 2004.
4. Elliot Anshelevich, Anirban Dasgupta, Jon M. Kleinberg, Éva Tardos, Tom Wexler, and Tim Roughgarden. The price of stability for network design with fair cost allocation. In *FOCS*, 2004.
5. P. Boldi and S. Vigna. The WebGraph Framework I: Compression Techniques. In *WWW*, 2004.
6. Paolo Boldi, Marco Rosa, Massimo Santini, and Sebastiano Vigna. Layered Label Propagation: A MultiResolution Coordinate-Free Ordering for Compressing Social Networks. In *WWW*, 2011.
7. Paolo Boldi, Massimo Santini, and Sebastiano Vigna. Permuting Web and Social Graphs. *Internet Mathematics*, 6(3):257–283, 2009.
8. Nieves R. Brisaboa, Susana Ladra, and Gonzalo Navarro. k2-Trees for Compact Web Graph Representation. In *SPIRE*, 2009.
9. Andrei Z. Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet L. Wiener. Graph structure in the Web. *Computer Networks*, 33(1-6):309–320, 2000.
10. Gregory Buehrer and Kumar Chellapilla. A scalable pattern mining approach to web graph compression with communities. In *WSDM*, 2008.
11. Wei Chen, Shang-Hua Teng, Yajun Wang, and Yuan Zhou. On the alpha-sensitivity of nash equilibria in pagerank-based network reputation games. In *FAW*, 2009.
12. Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, Michael Mitzenmacher, Alessandro Panconesi, and Prabhakar Raghavan. On compressing social networks. In *KDD*, 2009.
13. G. Christodoulou and E. Koutsoupias. The price of anarchy of finite congestion games. In *STOC*, 2005.
14. G. Christodoulou, E. Koutsoupias, and P. Spirakis. On the performance of approximate equilibria in congestion games. In *ESA*, 2009.
15. George Christodoulou, Katrina Ligett, and Evangelia Pyrga. Contention resolution under selfishness. In *ICALP*, 2010.
16. Amos Fiat, Yishay Mansour, and Uri Nadav. Efficient contention resolution protocols for selfish agents. In *SODA*, 2007.
17. John Hopcroft and Daniel Sheldon. Network reputation games, 2008.
18. Nicole Immorlica, Kamal Jain, and Mohammad Mahdian. Game-theoretic aspects of designing hyperlink structures. In *WINE*, 2006.
19. U. Kang and Christos Faloutsos. Beyond "caveman communities": Hubs and spokes for graph compression and mining. In *ICDM*, 2011.

20. Zsolt Katona and Miklos Sarvary. Network formation and the structure of the commercial world wide web. *Marketing Science*, 27(5):764–778, 09-10 2008.
21. Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *J. ACM*, 46(5):604–632, September 1999.
22. Scott Duke Kominers. Sticky content and the structure of the commercial web. Technical report, Harvard University, 2009.
23. Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *STACS*, 1999.
24. Elias Koutsoupias and Katia Papakonstantinopoulou. Contention issues in congestion games. In *ICALP*, 2012.
25. Panagiotis Liakos, Katia Papakonstantinopoulou, and Michael Sioutis. On the Effect of Locality in Compressing Social Networks. In *ECIR*, 2014.
26. Panagiotis Liakos, Katia Papakonstantinopoulou, and Michael Sioutis. Pushing the envelope in graph compression. In *CIKM*, 2014.
27. Allen B. MacKenzie and Stephen B. Wicker. Stability of multipacket slotted aloha with selfish users and perfect information. In *IEEE INFOCOM*, 2003.
28. Hossein Maserrat and Jian Pei. Neighbor query friendly compression of social networks. In *KDD*, 2010.
29. D. Monderer and L.S. Shapley. Potential games. *Games and economic behavior*, 14:124–143, 1996.
30. N. Nisan, T. Roughgarden, É. Tardos, and V.V. Vazirani. *Algorithmic game theory*. Cambridge Univ Pr, 2007.
31. Martin Olsen. Maximizing pagerank with new backlinks. In *CIAC*, 2010.
32. Martin Olsen and Anastasios Viglas. On the approximability of the link building problem. *Theoretical Computer Science*, 518(0):96 – 116, 2014.
33. Martin Olsen, Anastasios Viglas, and Ilia Zvedeniouk. A constant-factor approximation algorithm for the link building problem. In *COCOA*, 2010.
34. Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999.
35. Katia Papakonstantinopoulou. *Selfish Behavior and Compact Representation in Routing and Information Networks*. PhD thesis, National and Kapodistrian University of Athens, Department of Informatics and Telecommunications, Jan. 2015.
36. Sriram Raghavan and Hector Garcia-Molina. Representing Web Graphs. In *ICDE*, 2003.
37. Keith H. Randall, Raymie Stata, Janet L. Wiener, and Rajiv G. Wickremesinghe. The Link Database: Fast Access to Graphs of the Web. In *DCC*, 2002.
38. R. W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.
39. Tim Roughgarden. Intrinsic robustness of the price of anarchy. In *STOC*, 2009.
40. Tim Roughgarden and Éva Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, March 2002.
41. Jirí Síma and Satu Elisa Schaeffer. On the NP-Completeness of Some Graph Cluster Measures. In *SOFSEM*, 2006.
42. Torsten Suel and Jun Yuan. Compressing the Graph Structure of the Web. In *DCC*, 2001.