CHALMERS

Emergence of Small World Networks

Network formation through link selection by selfish nodes

Master of Science Thesis

MARTIN BÖRJESSON

Chalmers University of Technology
University of Gothenburg
Department of Computer Science and Engineering
Göteborg, Sweden, October 2012

Emergence of Small World Networks
Network formation through link selection by selfish nodes

Martin Börjesson

Examiner: Devdatt Dubhashi

Chalmers University of Technology
University of Gothenburg
Department of Computer Science and Engineering
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Department of Computer Science and Engineering
Göteborg, Sweden October 2012

Abstract

A game theoretic approach is taken to the Small World Phenomenon. A standard model for the algorithmic approach is generalised in a game theoretic framework, and the corresponding normal form game is defined and analysed. The thesis shows that when nodes in the network select long range contacts based on incentives that promote participation in short message chains the game admits a potential function. This implies the existence and optimality of pure Nash equilibria as well as convergence of some known game dynamics. The structure of Nash equilibria is further characterised by showing, theoretically and experimentally, the existence of a symmetric equilibrium. Drawing on results in the literature allows the Price of Anarchy to be lower bounded by $\Omega(\log n)$. Thus showing that while there are globally optimal equilibrium points the selfish actions of nodes have a large negative impact on the performance of the network. As an example of this, the convergence behaviour of regret minimizing agents in a fully dynamic version of the game is investigated experimentally.

Contents

1 Introduction

1.1 Background

In the 1960's the social psychologist Stanley Milgram performed a now famous experiment. Randomly selected individuals from a few geographically dispersed cities in the United States were asked to deliver a letter to a target person in Boston, but with the restriction that they were only allowed to forward the letter to people they knew on a first name basis. The recipients, in turn, could only forward the letter to people they knew, and so on.

Through analysing the resulting message chains Milgram reached the conclusion that citizens of the United States of America were on average connected by a path of as few as 6 people. The experiment, or at least the phenomenon it describes, has become a part of public conciousness, perhaps mostly in the form of the phrase “6 degrees of separation”.

Formally, a network is said to be a Small world network if the length of the path connecting two randomly selected nodes is expected to be polylogarithmic in the size of the network (number of nodes, n). Other properties that are typically associated with small world networks are a high degree of clustering (networks tend to have subsets of nodes that are tightly interconnected) and a prevalence of ”hubs” (nodes with a large amount of incoming connections).

Naturally, the problem has seen further study from a sociology perspective, even today where the advent of online Social Networks such as Facebook provides ample opportunity to investigate social structures. Sociology is however not the only interesting perspective on this phenomenon, as networks with this characteristic property are found in such diverse areas as biology, technology, and infrastructure. Consequently, much effort have been put into the study of Small World Networks and their mathematical properties.

A good way to understand something is to learn how to create it, and there have been attempts in the literature to create mathematical models for Small World Networks. In their paper “Collective dynamics of small-world networks” [17] Watts and Strogatz showed that a certain class of random graphs are small world networks by proposing a rewiring model that creates graphs with the required properties.

Although there are short paths between nodes in such graphs, a distributed node with only local information tend to not be able to find them. That is a severe drawback since in many situations where you find small world networks there is no central planner; graphs such as social networks or the link structure on the web are prominent examples where the structure has arisen through the intelligent actions of distributed agents. If the long range connection is subject to intelligent rational, it is safe to assume that the choice matters (a person does not select friends randomly, nor do websites link to unrelated content).

A node contemplating the decision of which other node in the network to select as a long range contact presumably does so with the goal of optimizing some utility. The empirically observed existence of short paths indicate that the distributed agents derive utility from being closely connected to other nodes. An example would be the facilitation of information propagation in a social network (i.e the spreading of a rumour or an idea). If short paths can not be found using local information only, what utility can be derived from them is called into question, but in particular it suggests that nodes are employing different utility functions. That is troubling since it does not describe the behaviour of nodes in many of the known instances of Small world networks.

For a self organizing network to have the small world property it seemingly must make sense to embed the nodes in some metric space, so one can define a distance function that has something to say about the utility gained from establishing a particular long range contact. For example, in the original small world experiment participants appeared to use a composite function, where the factors deciding whom the current message holder would send the message to was a function of geographical distance as well as social structure.

1.2 Navigable Small World Networks

The Small World Phenomenon rose to prominence in computer science when Jon Kleinberg took an algorithmic approach to the problem, seeking to address some of the drawbacks of the Wattz-Strogatz model [10]. In particular, Kleinberg addressed the notion of navigability; the experimentally observed property of small world networks that distributed agents are able to find short paths between nodes with only local information.

Kleinberg defined a model (elaboration in section 2) where nodes are located in a d-dimensional lattice and connected to their immediate neighbours. The graph was then augmented with long range contacts sampled from a governing distribution. The question is then what type of distribution generates navigable small worlds? In the Wattz-Strogatz model the augmentation is basically uniform, which accounts for some of it's drawbacks as only highly decreasing distributions are expected to have good performance. Kleinberg focused instead on power law distributions. More precisely, when the probability that a node in a d-dimensional lattice has a long range contact of length x is proportional to $1/x^\alpha$ the networks generated by such distributions were navigable small worlds exactly when $\alpha = d$. Such networks exhibit a type of "scale free" behaviour, where if you partition the lattice into exponentially larger sets with increasing distance from some node, it is equally likely to have a long range contact in any of the sets (any "scale of resolution"). When nodes select long range contacts according to the distribution with $\alpha = d$ the expected length of a greedy path connecting two random nodes is upper bounded by $O(\log^2 n)$.

1.3 Purpose

Kleinberg's model in some respect describe the behaviour of nodes in a Small World Network, since short paths are found greedily using only local information. However, it does not explain how such networks can arise, nor if symmetric behaviour is reasonable in a competitive model. Given the prevalence of networks with the small world property in diverse fields and under a wide range of different circumstances this is a highly interesting question.

Following the premise described in the previous section, where a small world network is a network where distributed agents gain utility from participating in short paths, a game theoretic generalisation of Kleinberg's model is considered. In this game each node maintains it's own distribution over link lengths (strategy) rather than sampling from one, single governing distribution. Study of this game might yield insights into Small World Networks and answer questions not addressable in Kleinberg's model, as it allows for individual behaviour and therefore study of how and why such networks emerge.

In the words of Kamal Jain: "If your laptop cannot find it, neither can the market". The market forces in this case are the distributed nodes in the network, working with only local knowledge, striving to maximize some utility derived from their participation in short message chains (the incentive on the nodes are to select long range contacts as efficiently as possible so as to minimize the routing time of the messages that it will be asked to route). How are they able to produce such efficient networks?

Some questions that would almost completely characterize how Small World Networks can emerge, if one were able to answer them, are listed below (readers unfamiliar with game theory might want to refer to sections 3 and 3.3):

- Where are the equilibrium points located?

  Are there Pure Nash Equilibria?

  Are there Mixed Nash Equilibria (i.e not pure)?

  Does the set of Correlated Equilibria extend the set of Nash Equilibria in an interesting way?
- How hard are they to find (compute)?
- How hard are they to approximate?
- What type of learning dynamics are able to exploit the structure of the game and approximate some equilibria?
- What is the relative performance difference between equilibrium points?
- What is the price of selfishness?

Naturally, a complete characterization lies beyond the scope of the project and focus is on a normal form version of the game, where the geometry of the game can be analysed, and a fully dynamic version in an experimental setting with nodes employing no-regret learning.

1.4 Delimitations

- The only aspect that is of interest here is routing performance, and no consideration is given to whether the networks exhibit other characteristic properties of Small World Networks (such as the prevalence of hubs).

- A variety of solution concepts have been defined in the literature, only those described in section 3.3 are given any consideration. (regret minimization outcomes are also considered in an experimental setting)
- Behaviour not captured by the model defined in section 3 is not investigated.

1.5 Notes on methodology

Other than the body of literature pertaining to Small World Networks and the augmented ring, the project intersects areas of Decision Learning, Networks, and Algorithmic Game Theory. The scope of the relevant literature is too large to summarize accurately but brief statements of relevant results are dispersed in the text.

In the early stages of the project heuristic experiments where carried out with the purpose of gaining intuition about the problem. Some relevant insights and conjectures were made, but as these where generally subsumed by later work, they are not expressly mentioned.

1.6 Overview

The over arching structure of the thesis is in five parts. The relevant concepts are first developed in section 2 where Kleinberg's model is described in more detail and a small part of the relevant research is summarized. Section 3 proceeds in the preliminary work by defining the game theoretic generalisation of the model as well as addressing questions of how to evaluate the game.

After the preliminaries the initial focus is on the structure of Nash equilibria, examined in section 4, since the results detailed there have implications for any type of question one might ask about Small World Networks.

Next,the symmetric strategy profiles are optimized over to find strategy profiles the existence of which is implied by the results in section 4

The optimized distributions then serve as a baseline when evaluating the routing performance of regret minimizing agents in a fully dynamical version of the game (6).

Since the three main sections of the thesis have their own result discussions, the final section 7 only briefly summarizes the main result and describes their relation to the research in the field. A few of the tantalising problems and questions outside the scope of the thesis are then mentioned to give a glimpse of some of the future possibilities.

2 Preliminaries

2.1 Model definition

In his original paper Kleinberg used a 2-dimensional grid as the base graph, but much of the following research has been focused on the equivalent 1-dimensional augmented ring. The model is illustrated through an example in figure 1.

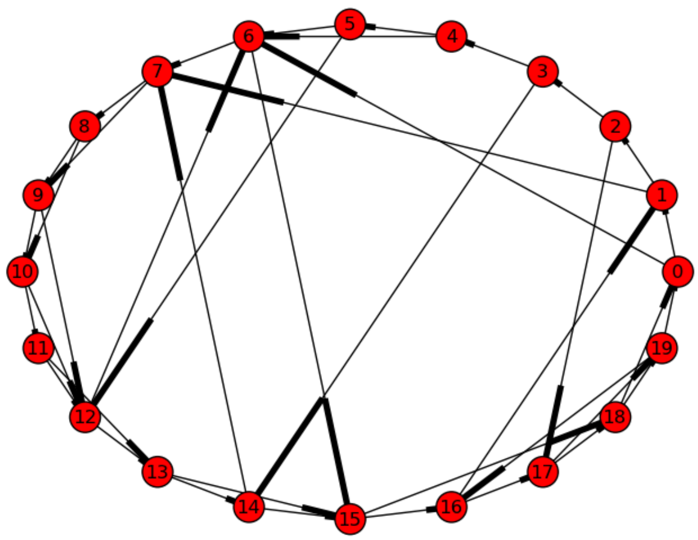

Figure 1: Example of the network model, augmented with a shortcut configuration where nodes selected their long range contacts using the 1-harmonic distribution

- The base graph is a directed circle with n sequentially numbered nodes, so that node v follows $((v-1) \mod n)$ and precedes $((v+1) \mod n)$.
- Every node has one additional outgoing directed edge (referred to as long range wire, long range contact, or shortcut).
- The probability that a node will have a long range wire of length a is given by a distribution over link lengths, σ. All nodes sample link lengths from the same distribution.

The purpose of this model is to find what characteristics the distribution σ has when it yields navigable small world networks. This is evaluated through routing messages in the network, as minimizing the expected greedy routing time (for suitable environmental conditions) is equivalent with minimizing the expected greedy pathlength between nodes.

The environmental conditions are that a message source is selected uniformly at random from the set of nodes and the target is selected uniformly at random from the remaining nodes. Thus we have a set of messages where a member $m \in M$ is a source-target tuple, $m = (s, t), s \neq t$, and every message is equally likely to occur. A message is routed in the network using a greedy protocol.

Tasked with routing a message, a node v with a long range wire of length a will route the message to it's next neighbour if $((t - v) \mod n) < a$, otherwise to $((v + a) \mod n)$ via it's long range wire. Consequently, the distance to the target decreases at every step and the message never overshoots the target.

Henceforth $GR(m)$ will be used to denote the expected length of the path when routing message m using greedy routing.

2.2 Brief review of related results in the literature

Kleinberg showed that in a d-dimensional model the expected routing time was upper bounded by $O(\log^2 n)$ for the d-harmonic distribution, where the probability that a node has a long range wire of length x is proportional to $\frac{1}{x^d}$. In his experimental analysis he focused on the 2-dimensional case where the base graph is a grid, but the vast majority of subsequent work has focused on the 1-dimensional model; the directed ring. Kleinberg's original upper bound was quickly and in two independent papers, proven for the directed ring as well as matching lower bounds [4] [13]. With tight bounds for the 1-harmonic distribution, the natural follow up question is if there exists other distributions with potentially greater performance. The only distributions expected to have good performance are the highly decreasing ones, and most of the research effort focused on the class of distributions generated by non-increasing functions. Flammini et al. (2005) showed a $\Omega(\log^2 n)$ lower bound on the expected diameter (length of the longest path) for such distributions [7]. Although indicative of things to come, this result is not sufficient grounds to claim that the harmonic distribution is asymptotically optimal as it makes no claim on the average length of paths.

Aspnes et al (2002), [2] showed the first general lower bound, $\Omega(\frac{\log^2 n}{\log \log n})$ for the expected routing time, i.e a bound that holds without any restrictions on the distribution from which link lengths are sampled. This was later tightened to $\Omega(\frac{\log^2 n}{a^{log * n}})$ by Giakkoupis et al (2007), where a is some constant [9].

In the paper "Tight bounds for blind search on the integers" Dietzfelbinger et al. (2008) showed a lower bound of $\Omega(\log^2 n)$ for a problem that is essentially equivalent with the one under investigation here, and that work was later expanded in "Tight lower bounds for greedy routing in uniform small world rings" to show the bound $\Omega(\frac{\log^2 n}{l})$ where l is the number of long range wires a node is allowed.

3 Game theoretic generalisation

In this section the definition of a game and associated notation is first given, after which it is shown how to interpret the generalised model as a game.

3.1 Definition of normal form game and notation

A finite n-player game in normal form can be defined in abbreviated form by the tuple $< N, A, u >$:

- $N = \{1, \dots, n\}$ is the set of of players, indexed by i.
- $A = \{A_1, A_2, \dots, A_n\}$ is a set of finite strategy spaces. A_i is the strategy space for player i and contain all player i's possible actions (pure strategies).

  A pure strategy profile is a vector $a = (a_1, a_2, \dots, a_n)$ where a_i is the action taken by player i ($a_i \in A_i$).

  A mixed strategy is a probability distribution over actions. If S_i denotes all distributions over A_i then S_i is player i's set of mixed strategies.

  The set $P = S_1 \times S_2 \times \dots \times S_n$ is called the set of strategy profiles. Note that the set of pure strategy profiles is a subset of P.
- For every player $i \in N$ there is an associated utility function $u_i : P \to R$. $u_i(p)$ is thus player i's utility under strategy profile $p \in P$.

3.2 Game interpretation

The set of players N are the nodes in the ring, and are indexed correspondingly. All n players have identical strategy spaces, $A_1 = A_2 = \dots = A_n = A$, with $A = \{1, \dots, n-1\}$. A pure strategy $a \in A$ corresponds to a long range wire of length a. A strategy is thus directly interpretable as a distribution over link lengths, and the set of strategy profiles captures all network behaviour that is not time dependent.

The game can be seen as a generalisation of the 1-dimensional Kleinberg model since the set of all symmetric strategy profiles exactly captures the networks generated when all nodes use a single, governing distribution.

3.2.1 The utility function

For the game to exhibit the desired macro behaviour the utility function must have the property that an agent optimizes it's utility when it minimizes the length of the paths in it's hitting set. Such a function ensures that the set of Nash equilibria contains all strategy profiles that are local optima for greedy routing in the network, and thus it sheds the most light on the dynamics of the formation of small world networks.

For a given strategy profile p and any message $m \in M$ one can compute the probability that some agent i will be asked to route m; the probability that the path of m 'hits' i. We denote the set of paths that hit a node as that agents hitting set; H_i.

From the perspective of a distributed agent, minimizing the expected routing time is equivalent with selecting a strategy that minimizes the expected routing time of the paths in H_i, since a players strategy choice has no immediate impact on it's hitting set. That is, if player i in strategy profile p switches strategy, then in the resulting strategy profile p' we have $H_i = H'_i$ due to the properties of greedy routing (a message never overshoots it's target). A consequence of this is that minimizing the expected routing time is equivalent with minimizing the weighted sum of all such paths.

Thus, the utility of player i is the sum of all pathlengths in it's hitting set H_i, where $r_i(h)$ denotes the probability that $h \in H_i$ hits i :

$$u_i = \sum_{h \in H_i} r_i(h) GR(h)$$

In abuse of notation the utility function defined above and the term hitting set will be used throughout the text without further mention of $r_i(h)$. Note also that the game is defined with a minimization objective, which will be reflected in some future definitions, but otherwise is of no import since any minimization game can be reformulated as a maximization game (by negating the utilities, for example).

3.2.2 Objective function

Although the game is defined in terms of distributed agents trying to further their own selfish good, the main point of interest is the macro behaviour. Will agents with the above defined micro behaviour generate good global performance (low expected routing time)?

To evaluate the global performance a function $\Phi : P \to R$ is defined. A local optima of Φ should have the property that no neighbouring strategy profile generates networks with lower expected routing time, and the global optimum should correspondingly generate networks that minimizes the routing time.

Since the cardinality of the set of message tuples remain unchanged throughout the lifetime of the game, and every message is given equal weight, a function with this property is the sum of the length of all paths:

$$\Phi = \sum_{m \in M} GR(m)$$

3.2.3 Computing expected routing time

Computing the expected routing time for a message in this model is achievable in polynomial time. To see this consider an arbitrary message (s, t). The routing for any node with distance d to the target can be expressed as a random variable X_d, with the desired property that $\mathbb{E}(X_d) = GR(t - d, t)$. Trivially $\mathbb{E}(X_1) = 1$. By the principle of deferred decision X_d can be reformulated as a random variable ranging over $(1, 1+\mathbb{E}(X_1), \ldots, 1+\mathbb{E}(X_{d-1}))$. The probabilities for the various choices are given by the strategy employed by the node at distance d, and X_d can thus be computed using straightforward dynamical programming in $O(n^2)$ time using the following relation (π denotes the strategy):

$$\mathbb{E}(X_d) = \sum_{1}^{d} \pi(i)\mathbb{E}(X_{d-i}) + \mathbb{E}(X_{d-1}) \sum_{d+1}^{n} \pi(i))$$

3.2.4 Symmetry in games

Within the scope of this project, strategy profile $p \in P$ is said to be a symmetric if all agents employ identical strategies. This differs from the general meaning of symmetry in games where it's usually defined to mean invariant under any permutation of the players. A game is said to be symmetric if every strategy profile is symmetric in the general sense. This means that in a Symmetric Game only the strategies being played matters, and not who plays them.

The Small World Game is obviously not symmetric since it matters greatly which agent plays which strategy (a permutation that is not a translation would redefine the hitting sets and thus the players utilities). Indeed, the importance of who plays what is so large that the strategy profiles where agents employ identical strategies is presumably isomorphic to the set of symmetric strategy profiles (in the general sense).

3.3 Solution concepts

Definition 1 (Pure Nash Equilibrium)**.** *A pure Nash Equilibrium is a pure strategy profile where no agent have incentive to unilaterally deviate and switch to a new action. Formally, a pure strategy profile* a^* *is a pure Nash equilibrium if:*

$$u_i(a_i^*, a^*_-i) \leq u_i(a_i, a^*_-i), \text{ for all } a_i \in A_i$$

Pure Nash equilibria is perhaps the most natural solution concept in non-cooperative games, but they do not always exist. A natural extension is to consider equilibria for mixed strategies, as defined above.

Definition 2 (Mixed Nash Equilibrium)**.** *A mixed Nash Equilibrium is a strategy profile where no agent has incentive to unilaterally deviate and switch to a new strategy. Formally, a strategy profile* p^* *is a pure Nash equilibrium if:*

$$u_i(s_i^*, p^*_-i) \leq u_i(s_i, p^*_-i), \text{ for all } s_i \in S_i$$

All Nash equilibria are product distributions since all players maintain independent distribution over their set of actions and thus they can be modelled as independent random variables. A commonly studied generalisation that relaxes this constraint are correlated equilibria. This solution concept considers all distributions over the set of action profiles, and a distribution is a correlated equilibrium if no agent has incentive to deviate from the strategy suggested by the distribution.

A common way do describe correlated equilibria in an intuitive manner is as a trusted intermediary that chooses an action profile according to some distribution and then tell each local agent what their component in the strategy is (what action they should take). A distribution over the sets of action profiles is a correlated equilibria if no agent can improve their expected utility by deviating from the recommended strategy. Intuitively, the power of correlated equilibrium is derived from it being a distribution over outcomes (action profiles/cells in the matrix representation of the game) while in a Nash equilibrium the probability that an agent takes a particular action is a function of the entire corresponding row or column (in a two player game).

Correlated equilibria are of interest partly because it's a natural concept in the sense that it captures some real world phenomenon such as a traffic light at an intersection, partly because correlated equilibria can sometimes outperform Nash equilibria (again, traffic light) and partly because they are easier to compute. Though, for the purpose of this thesis the interest stems mainly from the fact that agents employing no regret learning strategies are known to converge to some approximate correlated equilibrium.

Note that every Nash equilibria is also a correlated equilibria since, by definition, there is no incentive to deviate. See figure 2 for how the various solution concepts relate to each other.

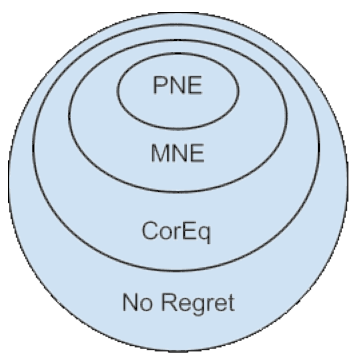

Figure 2: Comparison between regret minimizing agents, Kleinbergs 1-harmonic distribution, and optimized distributions (previous section)

3.4 Evaluating macro behaviour

3.4.1 Cost of selfishness

In his seminal 1951 paper John Nash showed that the set of equilibrium points is non-empty for all non-cooperative games [15]. However, there are no guarantees on the efficiency of equilibria, that is, that a Nash equilibrium constitutes an efficient solution from the perspective of macro behaviour. Many games also have a plurality of equilibria, with potentially very different global performance. So even in a game with distributed agents that exhibits convergence behaviour, the equilibrium point finally reached might yield global performance much worse than a global optimum, or indeed, it might get stuck on an equilibrium point that is significantly worse than other equilibria. To provide a framework to deal with these types of questions, Koutsoupias and Papadimitriou introduced the Price of Anarchy [12]. This is a performance metric that aims to capture the cost of selfishness through comparing the worst equilibria with the socially optimal solution (best global performance)

Definition 3 (Price of Anarchy)**.**

$$\textit{Price of Anarchy} = \frac{\textit{worst Nash equilibrium}}{\textit{socially optimal outcome}}$$

A related concept is that of Price of Stability, which quantifies the loss in performance due to the dynamics of the game situation. That is, in general there is no reason to expect the socially optimal solution to be a Nash equilibrium, so any distributed process that converges to a Nash equilibrium in such a case must necessarily imply suboptimal macro behaviour.

Definition 4 (Price of Stability)**.**

$$\textit{Price of Stability} = \frac{\textit{best Nash equilibrium}}{\textit{socially optimal outcome}}$$

Note that here these concepts are defined for a minimization objective.

4 Structure of Nash Equilibria

In this section the normal form version of the Small World Network game is analysed. First pure strategy profiles are considered, whereafter attention turns to mixed strategies. Lastly, the results claimed in the section are used characterize some aspects of the cost of selfishness.

4.1 Pure Nash Equilibria

Combining the definition of Φ with the previously defined utility function, an interesting property is derivable.

Property 1. *When a selfish agent unilaterally deviates and improves it's utility, the global performance of the network improves.*

Justification. Given a strategy profile p let p' denote the resulting strategy profile after agent i has changed it's strategy, and u_i' the resulting payoff for player i in profile p'.

The set of all messages M can be partitioned into two disjoint subsets; the set A of all paths that hit i and the set $B = M \setminus A$.

The sum of the paths in B are thus unaffected.

So we must have $\Phi(p) - \Phi(p') = u_i - u_i'$ □

Games with this type of property are called potential games, and were first defined in [14]. In this case we have an equivalence relation on the objective function difference and the deviating players difference, and it's therefore an exact potential game. Formally:

Definition 5 (Exact Potential Game). *A game is an exact potential game if there exists a function $\Phi : P \rightarrow R+$ such that $\Phi(p) - \Phi(p') = u_i(p) - u_i'(p)$*

Claim 1. *The Small World Game is a potential game with the potential function Φ.*

Proof. This follows straight from the definition coupled with property 1 defined above. □

Another interesting implication is that the Small World Game has at least one pure Nash equilibria; the socially optimal solution.

This follows from property 1: No agent can unilaterally increase it's score in a socially optimal configuration since that would improve the value of the objective function. Thus the socially optimal solution must be a Nash equilibrium.

Potential games are fairly well studied and have several interesting properties. Many dynamics are known to converge, most notably myopic best response sequences. To see this consider the corresponding state graph for the game. Every pure strategy profile is represented as a node, and the edge set is defined by adding a directed edge if the node at the endpoint is the resulting strategy profile after some agent has played a best response. The set of pure Nash equilibria are the nodes with no outgoing edges, and we know this set to be non-empty. There are no circles in this graph, since that would contradict property 1. Indeed, for all nodes in the graph the edge set defines directed paths that always end in a pure Nash equilibria.

Claim 2. *A pure Nash equilibrium for the small world game can be found in time polynomial in the number of players n.*

Justification. For any pure strategy profile the expected path length for a message is always an integer in the range $[1 \ldots n)$. The sum of all path lengths is thus upper bounded by a polynomial $p(n)$. Trivially we have $p(n) < n^3$.

By Property 1 we know that an agent that unilaterally improves it's score decreases the sum of the paths in it's hitting set. Here, the sum is an integer so a strategy improvement would imply a discrete decrease. So every time an agent switches to a better strategy the sum of all paths in the network decreases by at least 1. Thus $p(n)$ is an upper bound on the number of local strategy improvements required to reach a Nash equilibrium.

Any synchronization mechanism that ensures that no node must wait more than a polynomial amount of time between turns ensures that the network will reach a pure Nash equilibrium within polynomial time. □

Although a synchronization protocol as used above is an artificial construct and does not accurately represent the dynamics of repeated play of the small world game, the insight into the geometry of the game is still valid. Furthermore, an example of how similar behaviour might arise in a practical situation is given below.

Consider an idealised model of a network where nodes arrive over time, stay for a while, and then leave. For the purpose of this example consider a protocol where a newly arrived agent assumes the network is stationary and proceeds to try and learn it's optimal long range connections, which is fairly easy if the network is stationary, after that it does not change strategy again until it leaves the network (exploration is after all potentially costly as it entails playing suboptimal strategies). Furthermore lets say the stream of new nodes is equal in magnitude to the stream of nodes leaving the network, so the size of the network remains constant. Over time the network will presumably mutate, and old strategy choices grow inefficient, but if all nodes stay roughly the same time in the network then the nodes with the most incentive to change strategy are the nodes most likely to leave and be replaced by a new node. The macro behaviour becomes that of a best response sequence with a synchronization protocol where the agent with the most incentive to change is allowed to do so.

The network in question could be a peer-to-peer network that uses a hashing function to assign identities to nodes in the network in a virtual ring, and then the contacts for the nodes are defined over this virtual ring. This is a well defined concept called Virtual Ring Network, and Kleinberg's work has seen application in [18] where nodes select long range contacts based on the 1-harmonic distribution.

One can stretch the example even further, as the situation described is not wholly unlike a social network. People are born at a steady stream and enter an already formed social structure, live for about the same time, then die and get replaced.

4.2 Mixed Nash Equilibria

Most of the research following Kleinberg's work has focused on the class of distributions generated by some non-increasing function $f : N \rightarrow R+$. In the game theoretic setting this corresponds to the subset of strategy profiles where all agents employ the same strategy and that strategy assigns a non-zero probability to all pure strategies, and the probability of a link does not increase with distance.

In the previous section the existence and optimality of pure Nash Equilibria was established, as well as how agents can agree to some type of synchronization protocol to reach a pure Nash equilibria efficiently. However, a feature of pure Nash equilibria for the Small World Game is that the expected utility is not the same for all agents. Such a disparity in the utilities of the players is presumably strictly necessary for close to optimal performance, but at least in the above case it can be seen as an artefact of the synchronization protocol. In general, in a game such as this, with identical players and identical utility functions, we have no reason a priori to expect that one agent will outperform another.

If we denote the set of strategy profiles where all agents achieve the same score as the Egalitarian Line, then an interesting subset of this line consists of all strategy profiles where the only restriction is that all agents employ the same strategy, henceforth denoted as the Symmetric Line.

$$\text{Non-increasing} \subset \text{Symmetric Line} \subseteq \text{Egalitarian Line}$$

The set of symmetric strategy profiles is of particular interest since it generalizes the class of non-increasing distributions that has been the predominant focus in the literature, and since agents are identical we would expect the Symmetric Line to contain at least one Nash Equilibria.

This notion is formalized in section 4.2.2 and experimentally confirmed in 5. Initially some observations about the incentives in a symmetric strategy profile are presented. As a result, the existence of Nash equilibria in the set of non-increasing strategy profiles is precluded, and some insight is gained into what characteristics an optimal symmetric strategy profile would have.

4.2.1 Non-increasing distributions are suboptimal

If all agents in the network employ the same strategy the expected delivery time for an arbitrary message is only dependent on the distance between sender and receiver. Let the function $g : N \rightarrow R+$ denote the expected delivery time as a function of the distance (i.e $g(d) = GR(0, d)$).

From the perspective of an arbitrary, self interested agent the problem of maximizing it's score amounts to playing a best response to current network conditions. If one assumes that $g(d)$ is non-decreasing, which holds for the class of distributions generated by non-increasing functions, it is never in the best interest of the agent to include link lengths above $n/2$ in it's strategy.

Claim 3. *No generating function with range $R+$ can yield a Nash equilibrium*

Justification. This can be seen by a decomposition of the agents hitting set into subsets, grouping messages from the same sender together. The subset with the highest cardinality is the one where the agent itself is the sender, and it is also the subset with the largest influence on the economic incentive. Optimizing the strategy choice for this subset alone, we seek to find a link length m that minimizes

$$\sum_{1,m} g(i) + \sum_{1,n-m} g(i)$$

.

Clearly, there is no incentive to go beyond $n/2$ since g is non-decreasing. Combining this with the observation that all other subsets exert an influence that will decrease the maximum link lengths further, since a subset with cardinality x contains the paths with destinations in the range $1 \dots x$. □

The performance of any distribution generated by some non-increasing function f can be improved through iterated applications of two transforms, one that excludes link lengths above some m from the support and another that imposes a minimum link length k. Link length 1 is trivial to exclude since it is strictly dominated by all other strategy choices, but to show that the first few pure strategies, that correspond to very short link lengths, should not be included in the support seems slightly more involved.

Claim 4 (Unsubstantiated)**.** *There exists a link length k such that it is suboptimal to include any links shorter than k in the support of a symmetric strategy profile.*

Intuition. If the generating function is non-increasing the shortest link length is assigned a probability at least as high as any other strategy choice. Since only highly decreasing functions are expected to have good performance, this implies that every time a message is routed through a long range link it is most likely to be the shortest one. However, the only time we would actually wish to use a link of, say, length 2, is if the distance from the current message holder to the target is exactly 2. At distance $d > 2$ all other viable strategies are more preferable (under the assumption that $g(d)$ is non-decreasing), yet the situations were we have $d > 2$ far outnumber the situations where we have $d = 2$ (for reasonably large n). The accumulated regret from including link length 2 in the support should thus be much higher than the regret experienced from being distance 2 from the target when it is excluded from the support. □

From the above it is clear that any symmetric distribution that is also a Nash equilibrium would assign probability zero to all link lengths greater than m or shorter than k. Thus the class of distributions that has seen the most study in the literature does not accurately capture the behaviour of individual agents.

4.2.2 Symmetric Mixed Nash Equilibria

Since the agents and utility functions are identical it is reasonable to expect that somewhere on the Symmetric Line there exists a Nash equilibrium. To see this, consider the following adaptation of Nash's original proof of existence for Nash equilibria. First, stating Brouwer's fixed point theorem on which the proof relies:

Theorem 1 (Brouwer's fixed point theorem). *If B is a compact and convex subset of Euclidean space, then any continuous function $f : B \to B$ has a fix point.*

Note that a fix point is an element $b \in B$ such that $f(b) = b$.

Claim 5. *The Small World Network Game has a symmetric mixed Nash equilibria.*

Proof. The subset of symmetric strategy profiles, L, is compact since the set of strategies is an Euclidean space and the complement of L is unbounded. L is also convex, since for any two symmetric strategy profiles x and y, we have that $\alpha x + (1 - \alpha)y$ is necessarily a symmetric strategy profile and thus a member of L (for any $\alpha \in (0, 1)$).

Consequently, any continuous function that maps elements of L to itself will have a fix point.

Nash's original proof proceeded by defining a continuous function whose fix points are the set of Nash equilibria. The proof below follows a similar procedure, in that any fixpoint of the defined function is also a fixpoint of the function defined by Nash, and thus it is an equilibrium point. This is accomplished by defining the mapping $T : L \to L$ using the same methods Nash employed, and showing that when the input strategy profile is symmetric the modified strategy profile will also be symmetric. Thus, the transform has the desired property of mapping elements of L to itself.

First, a gain function is defined (for a minimization objective):

$$\varphi_i(p, a) = \max\{0, u_i(p) - u_i(a, p_{-i})\}$$

$\varphi_i(p, a)$ signifies the performance gain for player i switching to strategy a when all other players adhere to strategy profile p.

Using the gain function a transform that assigns higher probability to strategies with positive gain is defined:

$$s'_i = \frac{s_1 + \sum_a \varphi_i(p, a)\pi_i a}{1 + \sum_a \varphi_i(p, a)}$$

The mapping $T : L \to L'$ is thus defined as

$$T(p) = p' = (s'_1, s'_2, s'_3, \ldots, s'_n)$$

For a symmetric strategy profile p the transformed strategy profile p' is also symmetric since we have $s_1 = s_2 = s_3 = \ldots = s_n$ we must have $s'_1 = s'_2 = s'_3 = \ldots = s'_n$. More succinctly, $L = L'$ and T is a mapping from L to L.

Note that in a Nash equilibrium p^* we have $\varphi_i(p^*, a) = 0, \forall i, a$. In that case the modification above does nothing and a Nash equilibrium is therefore mapped to itself.

□

To elaborate on the above it shall be noted that for very small n the equilibrium can be pure. The best pure strategy profile that is also symmetric is one where all agents have long range wires of length $\sqrt{n}$. When n is sufficiently small so $\sqrt{n}$ is roughly proportional to $(\log^2 n)$ there is no gain in employing mixed strategies, but for any network of interesting size the equilibrium will be mixed.

An alternative way to prove claim 5 would be to lean on Nash's other famous theorem that says every finite game has a symmetric equilibrium (in the sense of being invariant under any permutation), and to show equivalence between the Symmetric Line and the set of strategy profiles invariant under permutation.

4.3 Equilibrium performance

Claim 6. *The best Nash equilibrium and socially optimal solution achieves* $\Theta(\log n)$ *routing time.*

Justification. Since we know the set of Nash equilibria to contain the socially optimal solutions we can borrow results from parts of the literature concerned with deterministic constructions facilitating greedy routing, and [1] [7] gives asymptotically optimal constructions for rings. Namely, the $O(\log n)$ upper bound for networks where nodes have one long range connection. □

Claim 7. *The Price of Stability is 1 for the small world game.*

Justification. Proof is immediate since the socially optimal strategy profile is a Nash equilibrium. □

Claim 8. *Expected greedy routing for the symmetric mixed Nash equilibrium is lower bounded by* $\Omega(\log n^2)$

Justification. Since the strategy profile is symmetric we can interpret it as a distribution over link lengths. In [6] a $\Omega(\log^2 n)$ lower bound for all distributions of this kind is shown. □

Claim 9. *The Price of Anarchy is lower bounded by* $\Omega(\log n)$

Proof. Combining claim 6 with 8, we get:

$$\text{Price of Anarchy} = \frac{\text{worst Nash equilibrium}}{\text{socially optimal outcome}} = \frac{\Omega(\log^2 n)}{\Theta(\log n)} = \Omega(\log n)$$

□

4.4 Extensions

Note that the argument for the existence of symmetric equilibria relies only on identical agents and identical utility functions. Thus one could relax the constraints on the message generation. As long as the source is picked uniformly at random an arbitrary distribution can be used to pick the target. Other utility functions could also be defined.

As described in the preliminaries Correlated Equilibria and No Regret learning equilibria are generalizations of Nash equilibria, and every Nash equilibria is thus a Correlated Equilibria etc. This implies that the lower bound on the Price of Anarchy shown above applies to the vast majority of commonly discussed game dynamics.

5 Optimizing over the Symmetric Line

In the previous section it was established that the small world game has a symmetric equilibrium, and that the class of the most commonly studied distributions don't contain any Nash equilibria. In this section experimental evidence of a symmetric strategy profile that outperforms Kleinberg's 1-harmonic distribution is given, and it also appears to be a Nash equilibrium.

5.1 Optimization problem

The local minima of the potential function Φ accurately captures networks that minimizes the expected routing time, and all such strategy profiles are Nash equilibria. To minimize the expected routing time one needs to find a distribution μ_n that minimizes the following expression:

$$h(\mu_{\mathbf{n}}) = \sum_{1,n} GR_{\mu_n}(0,i)$$

.

From the preliminaries described in section 3.2.3 we can compute h for any μ_n in $O(n^2)$ time, and it's clear that $h(\mu_n)$ imposes an ordering on the set of symmetric strategy profiles such that a profile on this line with a zero-gradient has a lower expected routing time than any neighbouring symmetric strategy profile. Such a local optima can be found fairly easily using standard optimization techniques.

With distributions μ_n modelled as real valued vectors the optimization problem becomes:

$$\begin{aligned}
\underset{\mu_{\mathbf{n}}}{\text{minimize}} \quad & h(\mu_{\mathbf{n}}) && (1)\\
\text{subject to} \quad & \mu_n(i) \geq 0,\ i \in 1,\ldots,n, && (2)\\
& \mu_n(i) \leq 1,\ i \in 1,\ldots,n, && (3)\\
& \sum_{i=1}^{n} \mu_n(i) = 1. && (4)
\end{aligned}$$

Presumably any suitable non-linear optimization method may be used as the point is merely to find a solution with a zero gradient, but for the results reported in this thesis Matlab's implementation of the Interior Point method was used. However, similar results were also achieved using sequential quadratic programming (SQP), and Active Set as well.

5.2 Resulting distribution

Optimizing for various network sizes all yield distributions that are decidedly similar. The resulting distributions all have an intermittent pattern, where only

$$I_n(d) \approx \begin{cases} 2/d & \text{if d is a (rounded) multiple of k} \\ 0 & \text{otherwise} \end{cases}$$

n	mean routing time	k	max
100	8.67	5.57	0.40*n
200	11.24	5.69	0.375*n
300	12.93	6.22	0.373*n
400	14.24	6.5	0.36*n
800	17.7	8.3	0.356*n

Figure 3: Parameter values for some optimized distributions

a few strategies are included in the support, and all of them are multiples of the minimum link length k (which seems to be roughly proportional to $log(n)$).

Figure 4 is the optimized distribution for a network with 100 agents. As can be seen it is intermittent, and the interval varies between 5 and 6 (minimum link length $k \approx 5.6 \approx \log n + 1$). The strategies include in the support have roughly the same relative relationship as in Kleinberg's 1-harmonic distribution, with the notable exception of the longest link. This pattern reoccurs for other network sizes.

If $I_n(d)$ denotes the probability that the optimized distribution for a network of size n has a link of length d, then the distributions take roughly the shape given in figure 5.2 (the table in the figure contains the parameter values for various n).

5.3 Distribution comparison

If the 1-harmonic distribution can be said to yield networks where a message is guaranteed to not have to make too many small jumps before a large one is found, then distributions of the above kind yields a slightly different behaviour. It will approach the target with medium large jumps until it is within link range, then the mean jump length will decrease, until the message is within some distance k from the target, whereupon it will proceed purely in the base graph (step size 1). But as seen, the rate at which the probability of establishing a link decreases with the distance is approximately the same as Kleinberg's (disregarding the intermittent behaviour).

In figure 5 is a plot of the mean routing time for some optimized distributions compared to the 1-harmonic distribution. The optimized distributions outperform Kleinberg's distribution measurably, but in accordance with the theoretical results it's only with a constant factor, as can be more clearly seen in figure 6.

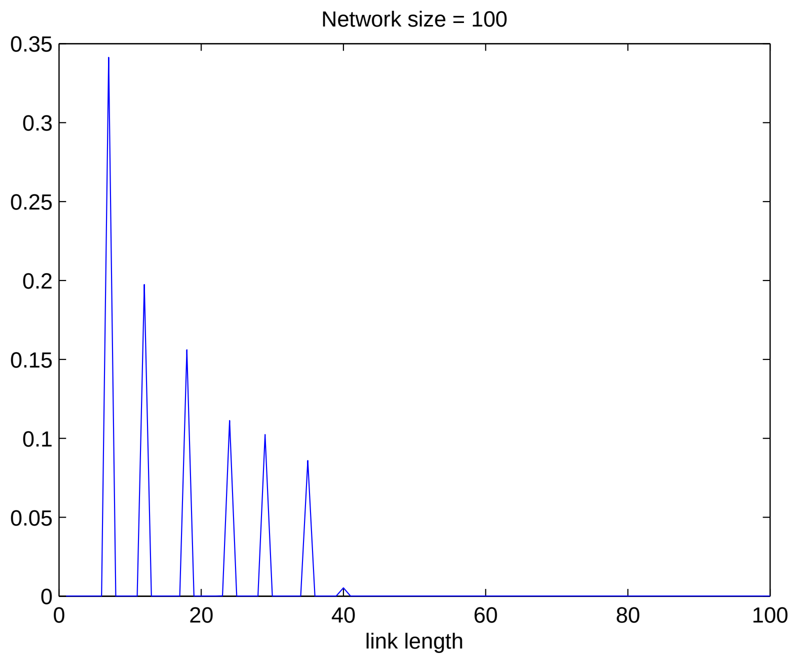

Figure 4: Plot of optimized distribution for n=100

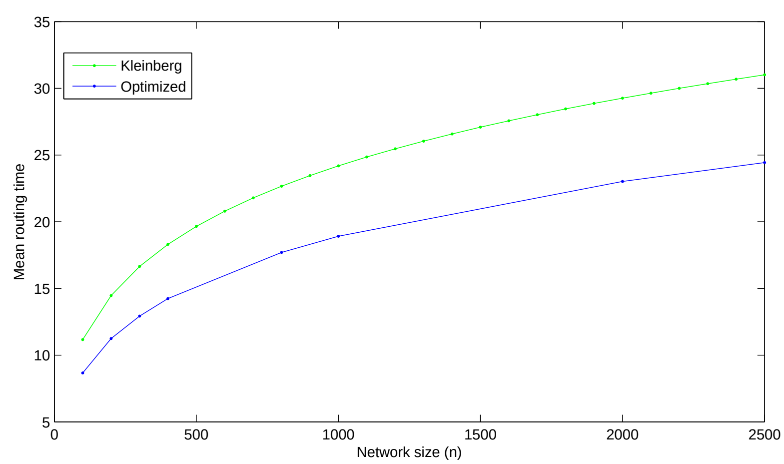

Figure 5: Kleinberg's 1-harmonic distribution compared to optimized distributions

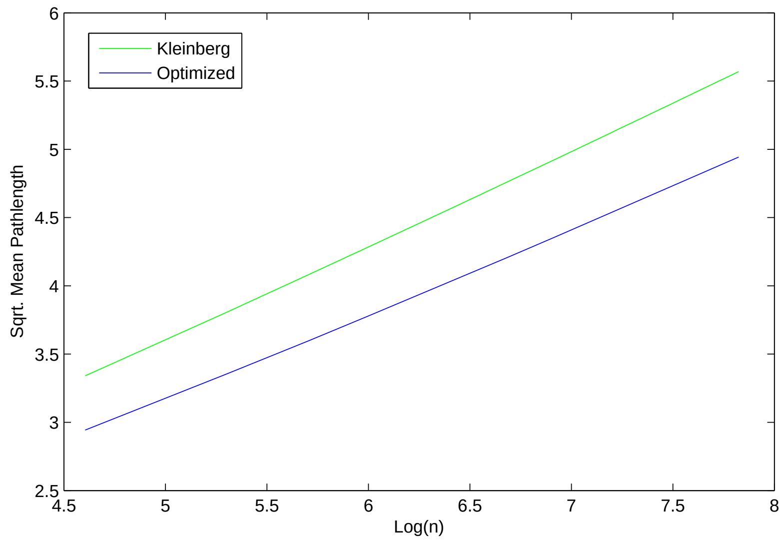

Figure 6: Same as figure 5

5.4 Nash equilibrium

Due to the game being of potential type we know that every Nash equilibrium is a local optima for the objective function with respect to individual agents deviating, but that does not mean a locally optimal symmetric strategy profile, with respect to other neighbouring symmetric strategy profiles, is necessarily a Nash equilibrium.

The natural question is of course then to ask if dropping the symmetry constraint means it's possible for an agent to improve it's score by switching strategy? Or, more succinctly, are the optimized distributions above examples of symmetric equilibria? As it turns out, yes.

Claim 10. *A distribution of this kind is a symmetric mixed Nash Equilibrium*

Experimental evidence. If the symmetric strategy profile described above constitutes a mixed Nash equilibrium, then every pure strategy in the support should yield the same return for a deviating agent, and no other pure strategy should yield a higher return. Figure 7 is a plot of the mean routing time for a network where one agent deviates. As can be seen, the routing time is minimized exactly when the strategies included in the optimized symmetric strategy profile are employed.

By Property 1 the impact on the routing time is effected purely by the sum of the deviating agents hitting set, and consequently we can conclude that there is no pure strategy deviation that would decrease the sum of the paths in the deviating agents hitting set. Thus, there is no incentive to deviate, and the symmetric strategy profile appears to be a Nash equilibrium.

Or, to phrase it in terms of the proof for the existence of symmetric equilibria; for every agent and every action the gain function is zero, so the distribution is a fixpoint and therefore a Nash equilibrium. □

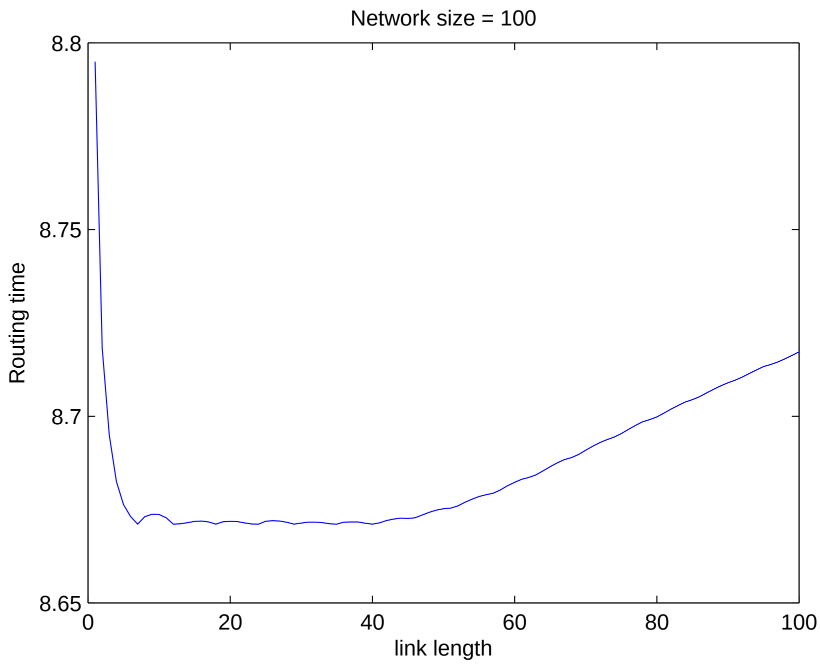

Figure 7: Plot of expected routing time when one agent deviates from the symmetric strategy profile and employ a pure strategy. Routing time is minimized exactly when the same pure strategies as those in the support of the symmetric strategy profile is employed.

6 Regret minimizing

Computing a Nash equilibrium is intractable in general, but there are a variety of algorithms achieving arbitrarily close approximations in some scenarios as well as real world examples of distributed agents achieving good approximations. Regret minimizing is a fairly natural concept in that an agent, based on the outcome of previous rounds of play, seek to employ a strategy that would minimize regret in hindsight. There are various notions of regret used in the literature, but the general concept is that a player has regret if there exist some other strategy that would have returned a higher payoff (i.e I should have played strategy x instead). Regret minimization is of particular interest since it can be shown that as time goes to infinity the regret minimizing strategies converges to an arbitrarily close approximation of the set of correlated equilibria.

The general applicability of regret minimization and the relatively weak assumptions it makes about the behaviour of selfish agents in games has led [5] to suggest a generalisation of the concept of Price of Anarchy: The Total Price of Anarchy, which takes all zero-regret outcomes into account.

6.1 Game model

In the previous section the geometry of the game was studied through the analysis of a normal form version of the game where the utility function was an accurate measure of the employed strategy. For the purpose of this experimental investigation a dynamical version of the small world game is considered in order to approximate real world conditions as closely as possible. This is also the version of the game with the highest computational difficulty. In this game a message tuple is generated by the environment which then gets routed through the network. The nodes that are hit by the message path have gained new information about the network and are thus allowed to update their strategy. The constraints on the information available to agents as well as the dynamical nature of the game means agents will be unable to directly infer the utility of a particular strategy, since that would require that the network remain static for a sufficiently long period of time for an agent to route all the messages in it's hitting set (and the cardinality of the hitting set is $O(n^2)$).

The utility function used for this experiment is $1/pathlength$. Intuitively the nodes that participate in the routing of a message split a unit reward equally among themselves. Thus as agents seek to maximize their payoffs the incentive is to take part in short message chains. Thus this utility function achieves a similar incentive as the previously discussed game, but with a maximization objective and codomain $[1/n, \ldots, 1]$ (with superset $[0, \ldots, 1]$).

6.2 Learning algorithm

The challenge, from a theoretical perspective, when attempting to learn an optimal strategy during repeated play of this game is that actions taken by node v at time t will influence the economic incentives of the other nodes. Thus the actions of node v influences the future development of the network (this is of course true for all nodes).

The problem of selecting among a set of strategies and observing the corresponding reward in repeated play naturally invokes online decision learning, and the approach taken here is that the nodes in the network view the problem as a generalized multi armed bandit.

The multi armed bandit framework in it's original formulation is concerned with independent stochastic processes, such as finding the most lucrative one armed bandit in a casino, but has subsequently been generalised in the context of online decision learning to extend to problems where the assumption of independent processes does not hold: the adversarial multi armed bandit. In an adversarial bandit problem the rewards are allowed to be non-stochastic, indeed it is named adversarial because it assumes that the rewards are selected by an opposing player with knowledge of all past actions.

The feedback model is opaque, in that the player only receives the reward corresponding to the action taken, and not the full reward vector as in traditional best expert algorithms. Since the only way to learn the expected reward from a particular strategy is to play it, and the reward can change over time, a multi armed bandit algorithm can not be satisfied with finding what appears to be a lucrative strategy but must always enforce some measure of exploration. Algorithms in the full feedback model such as Weighted Majority Algorithm and Hedge [8] are problems of exploitation, where multi armed bandit problems can be said to weigh exploration versus exploitation.

The algorithm used for this experiment is the Exponential-weight algorithm for Exploration and Exploitation (Exp3) from "The non-stochastic multi-armed bandit problem" by Auer et. al. [3]. The Exp3 algorithm maintains a distribution over strategies that is calculated the same way as the full feedback algorithm Hedge does (exponential weights), but to compensate for the opaque feedback the received reward at time t is divided by the probability of that action being taken. This ensures that the algorithm can quickly recover if a strategy that previously was assigned a low probability begins to yield larger rewards. Furthermore, to ensure that the algorithm spends sufficient time exploring the distribution is mixed with the uniform distribution before sampling a strategy. The input to Exp3, $\gamma \in (0, 1]$, is the exploration parameter which determines the mixing ratio (see step 1 in the loop body of the algorithm in figure 6.2).

6.3 Experiment parameters

For the purpose of this experiment the agents in the network employ a version of the Exp3 algorithm described in 6.2 modified for unknown time horizons (called Exp3.1 in [3]). It is essentially an application of the doubling technique in that it proceeds in a series of epochs, where the Exp3 algorithm is reinitialized with smaller exploration parameter γ when some strategy has achieved a sufficiently large sum of rewards.

The experiments are run until such a time as all nodes have exploration parameter $\gamma \leq 0.0625$ (preferably smaller, but even such small networks as in

Algorithm Exp3
Parameters Real $\gamma \in (0, 1]$
Initialization $w_i(1) = 1$ for $i = 1, \ldots, n$
For each $t = 1, 2, \ldots$

1. Set
$$p_i(t) = (1 - \gamma)\frac{w_i(t)}{\sum_1^d w_j(t)} + \frac{\gamma}{n} \text{ for } i = 1, \ldots, n$$

2. Draw action i_t randomly according to the probabilities $p_i(t), \ldots, p_n(t)$

3. Receive reward $x_{i_t} \in [0, 1]$

4. For $j = 1, \ldots, n$ set
$$\hat{x}_j(t) = \begin{cases} x_j(t)/p_j(t) & \text{if } j = i_t \\ 0 & \text{otherwise} \end{cases}$$
$$w_j(t + 1) = w_j(t)exp(\gamma \hat{x}_j(t)/n)$$

Figure 8: Pseudo-code for the regret minimizing algorithm Exp3 [3]

the rage 50-100 requires several million routed messages, making a higher degree of accuracy computationally intensive). After all nodes have sufficiently small exploration parameters the reported mean pathlengths are computed from the next 100 000 messages. This implies that when the routing performance is being evaluated most agents sample strategies from distributions where roughly 3 to 6 percent of the weight is exploration.

Furthermore the internal distributions (unmixed) are extracted from all agents and a distribution over link lengths is formed. If $p_i(d)$ denotes the probability that agent i has a long range wire of length d, then:

$$D_n(d) = \frac{\sum_1^n p_i(d)}{n}$$

Thus $D_n(d)$ denotes the probability that an anonymous node has a link of length d in the network populated by the regret minimizing agents from an experimental run as above. The mean routing time for the network where all agents sample link lengths from this distribution is calculated exactly as described in previous sections (3.2.3 etc.).

Due to the large time requirements for networks with many agents the experimental results reported are from single runs and not the average values of many runs. The time taken (the number of messages routed) in the experimental runs for which results are reported in the next section can be seen in figure 9, as well as fitted curve as reference for approximating time complexity.

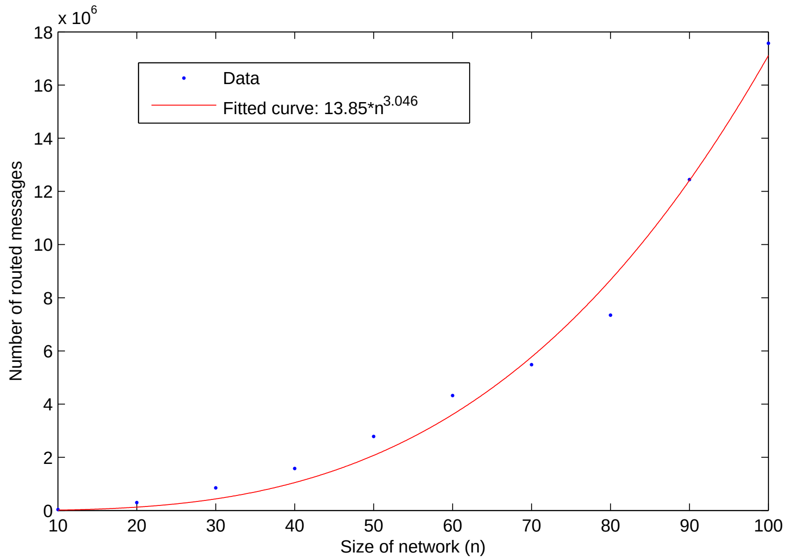

Figure 9: The number of routed messages until all agents have exploration parameter $\gamma \leq 0.0625$ and a fitted curve

6.4 Experiment evaluation

Looking first at the global distribution over link lengths one can detect similarities to the optimized distributions from 5, in that very short and long links are given low weight and an intermittent pattern begins to appear. The example distribution in figure 10 is perhaps messier than preferable, but the fluctuating pattern begins to appear at low exploration levels and is not a symptom of randomness or inefficiency. Indeed, it is somewhat surprising that it's as evident as it is, given that the intermittent pattern only has fractional impact.

As can be seen in figure 11 the routing performance when agents sample from distributions over link lengths (as defined above) approach, but never outperform, the optimized distributions, as expected. It can also be seen that regret minimizing agents significantly outperform Kleinberg's 1-harmonic distribution, which again conforms to expectation.

For the larger networks the regret minimizing agents begin to outperform even the optimized distributions even though their corresponding distributions over link lengths does not. This is presumably because the agents are not constrained by uniformity and due to the asymmetries that have developed they are able to avoid the mixed symmetric equilibria. However, it should be noted that this slight increase in performance accounts for a majority of the time consumption. As can be deduced from figure 12 the rate at which performance gains are made after the initial stages is very low.

That regret minimizing agents appear to convergence to equilibria similar

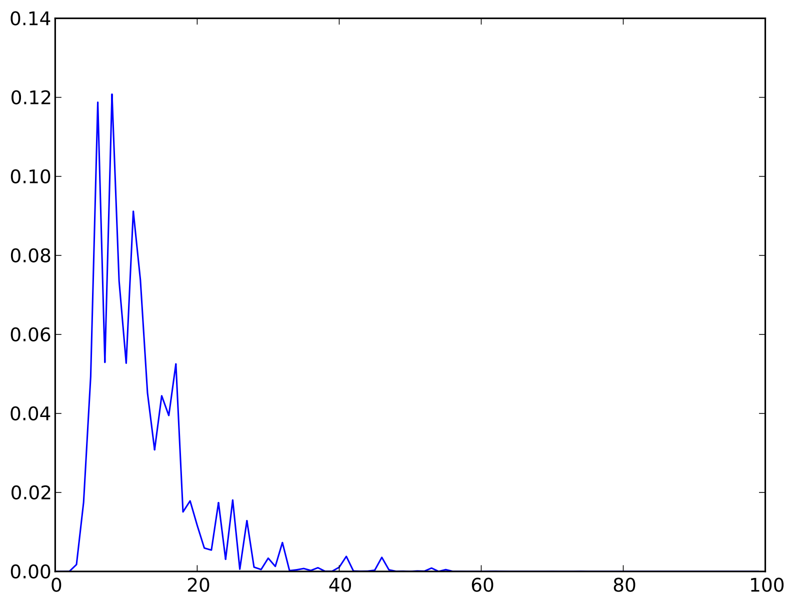

Figure 10: Example distribution over link lengths for n=100

to those of the normal form game studied previously is not a great surprise, as the non-stochastic nature of the problem fits within the adversarial bandit framework and the results in [3] can be extended to include randomized rewards.

However, as a practical model of small world network emergences the adversarial bandit approach seems like a failure, as the time until convergence is too long to be practical (although, to be fair, I can not think of a real world example where a set of nodes quickly form a small world network from scratch). This, I believe, is due to the fact that the assumption of an adversary is too strong, which forces a needlessly defensive approach (and regret is in and of itself a pessimistic approach, or risk averse, as we seek to minimize our error over time, rather than to attempt to find the optimal strategy quickly). Particularly since emergence of small world networks admits a potential function it is often the case that, from the perspective of agent v, what benefits some other agent u tends to benefit v by extension. Of particular note is also how quite efficient networks can be derived from uniform exploration, which suggests the possibility of heuristics with faster convergence if one abandons the theoretical framework of adversarial multi armed bandits.

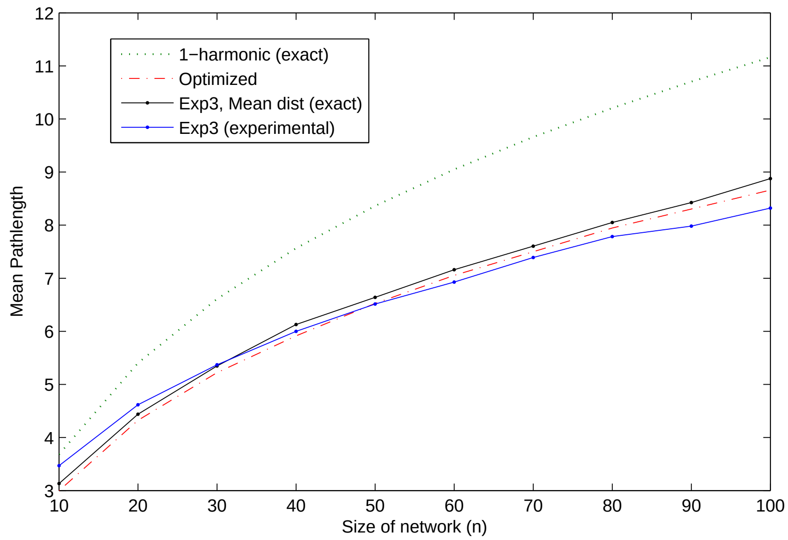

Figure 11: Comparison between regret minimizing agents (experimental measurements and distribution over link lengths), Kleinberg's 1-harmonic distribution, and optimized distributions (previous section)

7 Conclusion

7.1 Result summary

The lower bound on the Price of Anarchy shown in section 4.3 implies that no emergent dynamics that converges to an arbitrary Nash or Correlated equilibria can expect to asymptotically outperform the $\Theta(\log^2 n)$ 1-harmonic distribution, even though there are strategy profiles for which greedy routing performs a full logarithmic factor better. It is indeed the case, as seen in section 6, that networks with regret minimizing agents that are allowed to run for a long period of time not only appears to converge to the value of the corresponding optimized distribution, but the joint distribution begins to show the same intermittent behaviour that is described in §5.

In section 4 some characteristics of the structure of Nash equilibria was investigated. The results were insights into the geometry of the game, such as the existence and optimality of pure Nash equilibria, existence of symmetric mixed Nash equilibria and the price of selfishness. Combined with the experimental investigation in section 6 it paints a fairly comprehensive image, though some questions remain (such complexity results for approximation etc.). See section 7.3 for some of the potentially interesting questions relating to the topic at hand.

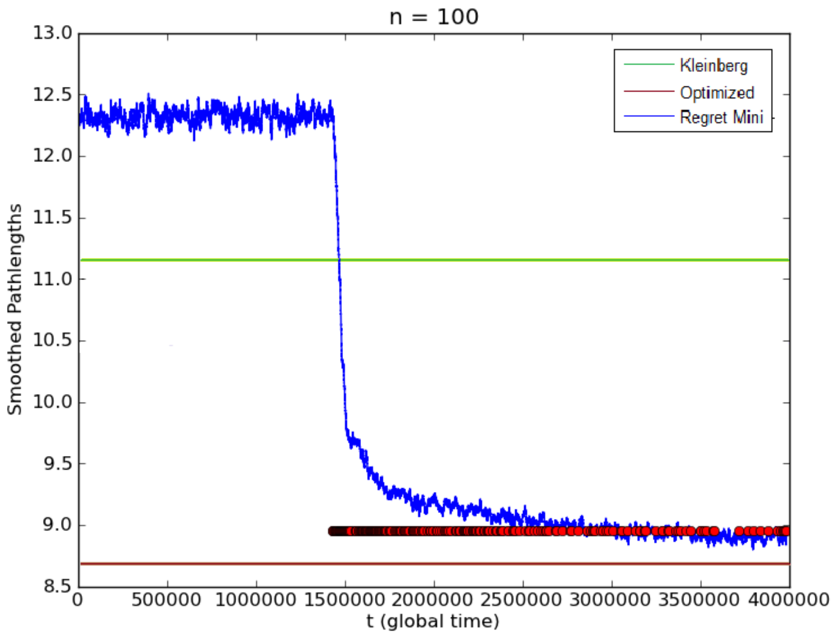

Figure 12: An experimental run with n=100 compared to corresponding mean pathlengths of the 1-harmonic distribution and optimized distribution. The red dots are a visual aid that signals when an agent is sufficiently confident to decrease it's exploration. Note the initial uniform exploration phase and the subsequent gain when agents begin employing their learned strategies.

7.2 Relation to research

In the literature there has been some attempts at defining network creation games with the purpose of studying the creation of Small World Networks, however, to my knowledge none has accurately captured the small world phenomenon. The game theoretic generalisation of Kleinberg's algorithmic approach studied in this report suggests that when distributed agents strive to partake in short paths, the resulting game has benign properties that facilitate emergence. Furthermore, the symmetric equilibrium shown in 4 stands in support of Kleinberg's model, as well as the body of research that followed, since it implies that symmetric behaviour is both theoretically valid and experimentally probable.

7.3 Future work

7.3.1 Congestion game

Congestion games have been widely studied in Algorithmic Game Theory. It's known that every congestion game is a potential game [16], and [14] shows that every potential game is a congestion game. Thus there exists an isomorphic congestion game to the Small World Game. Studying the corresponding congestion game might be interesting on it's own, but most promising is to draw on some of the results for congestion games.

In particular, Replicator Dynamics from Evolutionary Game Theory are known to converge in congestion games. This is interesting not only because small world networks are prevalent in biological systems, but it might cast further light on which members of the set of equilibria various dynamics converges to, and which members of the set of equilibria are stable in an evolutionary sense (able to withstand mutation of a subset of the strategies).

7.3.2 Computing a symmetric equilibrium

As established the game has a symmetric equilibrium, which is of particular note since it could be said to be both the expected outcome due to the identical nature of the agents, as well as the experimental observed outcome when agents employ regret minimizing strategies. The obvious question is then to ask how difficult it is to compute this equilibrium, so as to gain insight in to how feasible it is to exploit the geometry of the game in practice.

7.3.3 Alternate model

As has been seen, the difficulty involved in the emergence of small world networks is due largely to the chaotic nature of the evolving network under dynamical conditions, when all nodes simultaneously try to learn their no regret strategies. If, from the perspective of an agent, the network can be assumed to be stationary then the problem reduces to the much easier classical multi armed bandit where every pure strategy can be modelled as an independent Markov process.

An interesting model to consider is thus an expanding network, where nodes arrive over time. If the network is already of small world type then this could potentially facilitate the expansion process. Is it easier to incrementally grow a network to size n than it is to learn an efficient shortcut configuration for n distributed nodes?

It is, for example, interesting to consider Google's impact on the world wide webs link structure. Any new website would presumably use Google to learn of the authoritative sites on topics related to it's purpose, and subsequently link to them. Perhaps this can explain the effect observed by Kleinberg in [11] where websites with a high degree of incoming links accrue new links at a higher rate than other nodes. Or, put another way, can an expanding model explain preferential attachment as nodes that are easier to find or learn.

7.3.4 More outgoing links

Can the conjectured existence of a symmetric mixed Nash equilibria be carried over to networks where agents have more than one outgoing link? Using existing bounds in the literature would suggest a of Anarchy of $\frac{\Omega(\log^2 n/\log n)}{\Omega(\log n/\log\log n)} = \Omega(\log\log n)$ when nodes have $\log n$ outgoing links.

7.3.5 Other questions

- The game could be defined for other routing protocols, either greedy protocols with higher information content (such as augmenting the routed message with a list of the nodes which it has hit, thereby increasing the information about the network configuration available to distributed agents) or other protocols entirely (to my knowledge it is an open question if greedy routing is inherently inefficient)

- Can one find a correlated equilibria with asymptotically lower routing time, through central coordination of only a few agents?

- Lower bound the payoff gain in best response dynamics for arbitrary strategy profiles?

References

[1] I. Abraham, D. Malkhi, and G.S. Manku. Papillon: Greedy routing in rings.

[2] J. Aspnes, Z. Diamadi, and G. Shah. Fault-tolerant routing in peer-to-peer systems. In *Proceedings of the twenty-first annual symposium on Principles of distributed computing*, pages 223–232. ACM, 2002.

[3] P. Auer, N. Cesa-Bianchi, Y. Freund, and R.E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1):48–77, 2003.

[4] L. Barrière, P. Fraigniaud, E. Kranakis, and D. Krizanc. Efficient routing in networks with long range contacts. *Distributed Computing*, pages 270–284, 2001.

[5] A. Blum, M.T. Hajiaghayi, K. Ligett, and A. Roth. Regret minimization and the price of total anarchy. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 373–382. ACM, 2008.

[6] M. Dietzfelbinger and P. Woelfel. Tight lower bounds for greedy routing in uniform small world rings. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 591–600. ACM, 2009.

[7] M. Flammini, L. Moscardelli, A. Navarra, and S. Perennes. Asymptotically optimal solutions for small world graphs. *Distributed Computing*, pages 414–428, 2005.

[8] Y. Freund and R. Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *Computational learning theory*, pages 23–37. Springer, 1995.

[9] George Giakkoupis and Vassos Hadzilacos. On the complexity of greedy routing in ring-based peer-to-peer networks. In *Proceedings of the twenty-sixth annual ACM symposium on Principles of distributed computing*, PODC '07, pages 99–108, New York, NY, USA, 2007. ACM.

[10] J. Kleinberg. The small-world phenomenon: an algorithm perspective. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 163–170. ACM, 2000.

[11] J. Kleinberg and S. Lawrence. The structure of the web. *Science*, 294(5548):1849–1850, 2001.

[12] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Proceedings of the 16th annual conference on Theoretical aspects of computer science*, pages 404–413. Springer-Verlag, 1999.

[13] C. Martel and V. Nguyen. Analyzing kleinberg's (and other) small-world models. In *Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing*, pages 179–188. ACM, 2004.

[14] D. Monderer and L.S. Shapley. Potential games. *Games and economic behavior*, 14:124–143, 1996.

[15] J. Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295, 1951.

[16] R.W. Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

[17] D.J. Watts and S.H. Strogatz. Collective dynamics of small-worldnetworks. *nature*, 393(6684):440–442, 1998.

[18] H. Zhuge and X. Sun. A virtual ring method for building small-world structured p2p overlays. *Knowledge and Data Engineering, IEEE Transactions on*, 20(12):1712–1725, 2008.