Game Theory on Social Networks

John J.E. Smith

April 23, 2017

Research Review

Supervised by Prof. Seth Bullock

Executive Summary

This project aims to contribute to the study of evolutionary cooperation by examining the behaviour of cooperation systems on social networks. A variety of studies exist in this field, examining conditions under which co-operation can evolve and persist in both simple and complex systems.

This project will build upon existing research by applying an existing framework, set out by Santos and Pacheco (2006) [36], to social network graphs generated using the REDS model proposed by Antonioni, Bullock and Tomassini in 2014 [7].

The REDS random graphs were shown in [7] to better resemble the structure of an organic social network than previous random graph models, and so results based upon REDS will be of greater relevance to this field in the setting of social cooperation.

In [9], a study of contagion on REDS networks, it was discovered that 'real' social networks "will tend to be more susceptible to a range of different contagion dynamics than previously thought".

I hope to similarly leverage the REDS model to produce results which are more relevant to true organic social networks; in this case in the setting of evolutionary cooperation.

As such, the aim of this project is to deliver the parametric conditions under which cooperation games can be expected to emerge and persist in both static and dynamic REDS environments, as well as a comprehensive analysis and justification of the results, implications and methodology.

This project will require research in the areas of Evolutionary Game Theory, in particular the field of co-operation and Graph Theory, in particular Random Graph Theory. A degree of Software Development will be necessary to implement simulation of the systems, as well as the statistical analysis of the resultant data. A method for generation of the REDS graph must also be developed.

Contents

Chapter 1

Introduction

The emergence and persistence of cooperation in nature has been a subject of interest since Darwin, and according to [1], "remains identified as one of the biggest challenges of the XXIst century".

This is due to the apparent paradoxical nature of cooperation systems, when reconciled with natural selection. We can observe many biological systems [14][13][12] in which organisms have evolved to cooperate, despite the fact that each individual's genes will be better furthered by defecting.

Intuitively, natural selection should render these systems unstable and therefore transient, but instead we find that "cooperation can be found at all levels of biological organisation" [8].

As such, this problem applies to everything from single-celled organisms to human social interaction, which is the focus of this project.

To properly discuss this topic, it is useful to formally define what we mean by cooperation, and in what kind of systems the problem occurs. This can be done through the use of Game theory as we will see in section 2.1.1. The particular class of game in which we are interested is elegantly summarised as the "Prisoner's Dilemma", a game invented by Merril M. Flood and Melvin Dresher in approximately 1950 [2].

Once we have formulated this idea in a game theoretic sense, through abstraction it becomes clear that this could be expanded and applied to a variety of systems besides evolutionary biology. It is clear that human social behaviour relies to a degree on mutualistic co-operation in a conscious sense. Current research already suggests that understanding the dynamics of social networks has important repercussions in understanding and preventing the spread of disease [3], and possible fields of application could include the modelling of information transmission across human social systems, and explaining the formation of information asymmetries.

Over time research on co-operation has sought to explain first the conditions under which co-operation might emerge in a homogenous system of agents [2] [4], how it might emerge in evenly-spaced lattice structures [5] and most interestingly, how it might emerge when the agents participating are embedded in an uneven (and generally random) graph [36].

For social networks, a random graph is much more realistic than the first two options, (we can immediately observe that people interact with some people more often than others), so the concern is how to generate a graph that most accurately mimics a true social network. Many types of random graphs have been experimented with, including Random Geometric Graphs [6], Erdős-Rényi and Barabási-Albert models [36].

However, most traditional examples of random graphs do not exhibit all of the observed properties of a social network. A paper by Antonioni, Bullock and Tomassini [7] introduces a class of graphs which exhibits all of the most commonly identified properties, and this model is dubbed *REDS*, which is an acronym for the four principle features of the graph's generation: Reach, Energy, Distance and Synergy.

It is the objective of this project to study the conditions under which co-operation is allowed to emerge and reliably persist on REDS graphs, and to examine how the introduction of a dynamic element to the graphs affects these conditions, perhaps by coupling the game's payoffs with a node's 'Energy'.

This project aims to deliver conditions in the form of parametric constraints under which it is possible to expect a static REDS graph to maintain co-operation in its nodes, and the constraints under which extending to a dynamic system can continue to exhibit co-operation. The project should also deliver the analysis required to justify these conclusions, as well as examination of the realistic validity of these parametric constraints in the setting of a social network, and suggest areas to naturally extend this research.

1.1 Objectives

The main aims/objectives of this project are:

- To derive formalised conditions in the form of parametric constraints under which co-operation may be expected to emerge and persist on a static REDS random graph.
- To design justifiable framework under which the REDS graph, in the setting of games, may be extended to have a dynamic element. This may be tied directly to the payoffs of the games being simulated or otherwise.
- To derive the further conditions under which the extended dynamic system can be expected to exhibit co-operation.
- To deliver a detailed analysis and justification of the above results, including an analysis of the implications of the results when considered in the setting of a true organic social network.
- To deliver a full description and justification of the methods and techniques used to generate the results, so that they may be replicated/tested.

1.2 Research Areas

The main areas of research and discussion required for this project to be undertaken are as follows:

- **Game Theory:** A basic clarification of the fundamentals of game theory must be delivered, including definition and light discussion of the concepts of a *Payoff Matrix*, a *Dominant Strategy* and the *Prisoner's Dilemma*.
- **Graph Theory:** Since the scope of this project does not presume knowledge of Graph theory, a brief exploration of some of the fundamental theory shall be included; i.e. the definitions of a *Graph*, a *Vertex*, an *Edge* etc, and possibly the

relation of graphs to linear algebra through the *Adjacency matrix*. This theory may be required to efficiently explain current literature on the evolution of co-operation on graphs.

- **Random Graphs:** Research on the topic of Random Graphs will be required, including the discussion of well-known Random Graph models and how they relate to one another. The level of depth here will depend on the relevance to the aims of this project and the existing literature. Some key models which will likely be studied include the Watts-Strogatz small-world model, the Erdős-Rényi model, the Barabási-Albert scale-free model and Random Geometric Graphs.

- **Evolution of Co-operation:** This will be the most relevant and vital field of research as it will cover the chronology of the field into which this project will contribute. Research will be conducted on the behaviour of co-operation in both simple and complex networks, to lay the contextual foundations for this project. In particular I will focus on the methodology of Santos and Pacheco's 2006 study [36], as I will be adapting their framework for this project.

Chapter 2

Survey

This chapter will summarise the main research and theory relevant to this project.

2.1 Preliminary Research Areas

Before diving straight into the deep end of contemporary research in this discipline, I will give a quick overview of some of the fundamental mathematical principles that are necessary to discuss the topic efficiently. Naturally, first and foremost among these is Game theory.

2.1.1 Game theory

Roger Myerson [10] defines Game theory as "the study of mathematical models of conflict and cooperation between intelligent rational decision-makers". It is a field of study which stretches back to the early 20th century, encompassing a great deal of literature, but for the purpose of this project I need only address a few fundamental principles.

Games can be categorised variously. A game will invariably have some n players, and those players will have access all or part of the information available. For example, Chess is what would be called a 2-player perfect information (or Markov) game, as there are 2 players and each can see the position of all the pieces. On the other hand Poker can have a variable number of players, and is an *imperfect information game*, as players can see only their own hands.

Another way of classifying games is through their degree of randomness. In a *deterministic* game, a player's move applied to a given game state will always have the same effect (e.g. Chess), whereas a non-deterministic (or Bayesian) game will introduce chance in some way (e.g. dice-rolls in Backgammon).

It is also worth noting that a game can be asymmetrical, with each player having a different set of available responses to the same game state.

For this project, we will be considering a class of 2-player, deterministic, symmetrical, imperfect information games, so I will constrain my exploration of this topic to games of this type.

The Normal form and dominant strategies

To describe a game, it is common to use what is known as the *Normal form* which seeks to describe the actions (or strategies) available to each player and equate them to

their outcomes. This form requires the outcomes to be measurable in some way and this measurement is referred to as the *payoff* of an outcome.

The following is based on the elegant definiton found in [11].

Definition 2.1.1 (Normal form)**.** A finite n-player normal form game is a tuple (N, A, u), where:

- N is a finite set of n players, indexed by i;
- $A = A_1 \times \ldots \times A_n$, where A_i is a finite set of *actions* available to player i. Each vector $\mathbf{a} = (a_1, \ldots, a_n) \in A$ is called an *action profile*;
- $\mathbf{u} = (\mathrm{u}_1, \ldots, \mathrm{u}_n)$ where $\mathrm{u}_i : A \to \mathbb{R}$ is a real-valued *payoff* (or utility) function for player i.

In simpler terms, a game can be broken down into three parts; the player space N, the action space A, and a payoff function $\mathrm{u}_i(\mathbf{a})$ for each player (which in a discrete Action space can be represented as a matrix).

Note that in this definition the payoffs of the game are a direct function of the actions taken by the players, which implies that this form relates exclusively to deterministic games.

This presents no problem for us, as we will be operating solely in the domain of deterministic games, but it is very interesting to note that Bayesian games can in fact be reduced to Normal form games (although with a much larger Action space).

The above definition is a general formulation of the problem, and it is worth discussing an example for clarity.

Example 2.1.1 (Rock-Paper-Scissors)**.** If we consider the simple 2-player one-shot game of "Rock-Paper-Scissors", we see that each player can choose between three actions, and is unaware of the action chosen by the other player until they have committed to their own.

This allows us to set $n = 2$, and for each player i,

$$A_i = \{\text{rock}, \text{paper}, \text{scissors}\}$$

Since $n = 2$ the full action space is

$$\begin{aligned} A &= A_1 \times A_2 \\ &= \{\text{rock}, \text{paper}, \text{scissors}\} \times \{\text{rock}, \text{paper}, \text{scissors}\} \end{aligned}$$

So a particular action profile will be of the form

$$\mathbf{a} = (a_1, a_2) \ : \ a_1, a_2 \in \{\text{rock}, \text{paper}, \text{scissors}\}$$

To make the outcomes of the game measurable, we assign a payoff of 1 for a win, 0 for a draw, and -1 for a loss (in keeping with the general consensus that win $>$ draw $>$ lose).

This allows us to determine the payoff functions as a pair of matrices, indexed by the enumeration of A_i:

$$\mathrm{u}_1 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \quad \mathrm{u}_2 = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

Then we can fully describe the game by combining these into the following table of payoffs:

$$\begin{array}{c} \\ \text{Rock} \\ \text{Paper} \\ \text{Scissors} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Rock} & \text{Paper} & \text{Scissors} \end{array} \\ \begin{bmatrix} 0,0 & -1,1 & 1,-1 \\ 1,-1 & 0,0 & -1,1 \\ -1,1 & 1,-1 & 0,0 \end{bmatrix} \end{array}$$

We can see that if one player plays rock and the other plays scissors, then the first player receives payoff 1 whilst the second player receives payoff -1, therefore losing the game.

Definition 2.1.2 (Dominant strategy (2-player))**.** In a 2-player game, a strategy is dominant if and only if

- its payoff is greater than or equal to any other strategy regardless of the strategy of the other player, and;
- its payoff is strictly greater than at least one other strategy regardless of the strategy of the other player.

Note that in Example 2.1.1, there is *no* dominant strategy.

The concept of a dominant strategy leads us very nicely towards one of the most famous games in Game theory; the *Prisoner's dilemma*.

The Prisoner's dilemma

The Prisoner's dilemma was originally formulated by Merril Flood and Melvin Dresher in approximately 1950 [2], but gained its name from the interpretation described by Albert Tucker in the same year.

The interpretation is roughly as follows:

> Two criminals are caught and imprisoned separately, so they cannot speak to each other. The prosecution lacks the evidence to convict the pair for their worst crimes, so they both expect to receive only one year in prison as a result of a minor crimes.
> However, the prosecution makes them both the same offer; if they testify against their partner then they will serve no sentence (whilst their partner serves 3 years). If they both testify however, they will each serve two years.

At first glance, it may seem that the best offer is for neither to testify, both receiving one year (to give the lowest total sentence of all the outcomes). However, this is a game at core, and we are interested in what will happen if each agent is self-interested.

Let us consider the Payoff matrix of this game:

$$\begin{array}{c} \\ \text{Cooperate} \\ \text{Defect} \end{array} \begin{array}{c} \begin{array}{cc} \text{Cooperate} & \text{Defect} \end{array} \\ \begin{bmatrix} -1,-1 & -3,0 \\ 0,-3 & -2,-2 \end{bmatrix} \end{array}$$

If we view this payoff matrix in light of Definition 2.1.2, we see that for both of the opposing player's moves, a player's payoff is higher in defection; if the opponent cooperates, defection gives a payoff 1 higher than cooperation, and if the opponent defects, defection still gives a payoff 1 higher than cooperation. Therefore the *dominant strategy* in this game is to defect.

Consider now that the game is perfectly symmetrical; each player has access to the same strategies and receives the same payoffs. This implies that both players have the same dominant strategy: to defect!

This property is what makes games of this class so interesting, and is also what makes the evolution of cooperative systems in nature so seemingly paradoxical.

Cooperation in Nature: Motivation

It is useful to define cooperation in its game theoretic form for the purposes of analysis, but it is easy to divorce ourselves from reality when discussing the phenomenon in this setting. As such I will now discuss how systems of this form appear in nature, to highlight the relevance of this topic as more than just an abstract discussion.

A great summation of the perpetuation of this form throughout nature is given in West, Griffin and Gardner's 2007 review of the subject; "Evolutionary Explanations for Cooperation" [8], and the following is sourced primarily from that paper.

One specific example of cooperation in nature is the meerkat, which generally live in groups of approx. 30 adults plus their offspring. In such a group, there is generally a dominant male and female who are responsible for the majority of the reproduction of the group, whilst the others simply help the dominant couple raise their offspring [12].

This becomes analogous to the prisoner's dilemma when we consider the subordinate meerkat who chooses to give the food they find to the dominant couple's offspring. That meerkat's genes would be better furthered by keeping the food for itself or its own offspring. One would therefore expect that meerkats who defect from this system will be better selected over those who cooperate, eventually causing the system to evolve into one devoid of cooperation.

Another analogy of the Prisoner's dilemma in nature can be found in human economics, illustrated through the example of the 'tragedy of the commons', which was widely popularised by ecologist Garrett Hardin in 1968 [13].

In the tragedy of the commons, we consider a single (common) piece of land shared equally by a number of independent cattle-farmers. Naturally, additional cattle return additional benefit, but the finite piece of land can only sustain so many cattle before overgrazing occurs.

This becomes a Prisoner's dilemma situation when we consider that the benefit of additional cattle goes exclusively to the owner of those cattle, whilst the damage of overgrazing is shared over the whole group. It is therefore (from an individual economic point of view), a rational decision for one farmer to put more cattle on the commons (therefore defecting).

Naturally, this means that the rational behaviour for *all* farmers is to graze additional cattle, therefore resulting in mass overgrazing. The cooperative state is seemingly not stable, since each individual can always gain by selfishly defecting, but these common grazing systems have long been a stable custom in English villages, and the metaphor can comfortably be applied to resource management on a macro-scale (although one might argue that the cooperative system at the macro level *has* defected).

The theme of cooperative systems continues even down to the cellular level [14]. According to [8], "the tree of life is dominated by single-celled microorganisms that appear to perform a huge range of cooperative behaviours".

We observe that this phenomena is both prevalent and yet seemingly unstable; hence motivating research in this field.

2.1.2 Graph theory

The second preliminary area of study required for this project is that of Graph theory, as our objective is to contribute to contemporary research (see section 2.3) suggesting that embedding systems into (spatial) graphs allows these systems of cooperation to remain stable. This concept is elegantly described as "space scaffolds functional complexity" [15].

I assume a basic knowledge of set theory and linear algebra for this section.

Basic definitions

Most of the graph theory in this section was learnt from material found in [16], but there is a plethora of valid sources which cover these fundamentals.

Definition 2.1.3 (Graph)**.** We define a graph, G, as a set of vertices, V, and a multiset of edges, E, where $E \subseteq V^2$, so if vertices u and v are connected in G then $(u, v) \in E$.

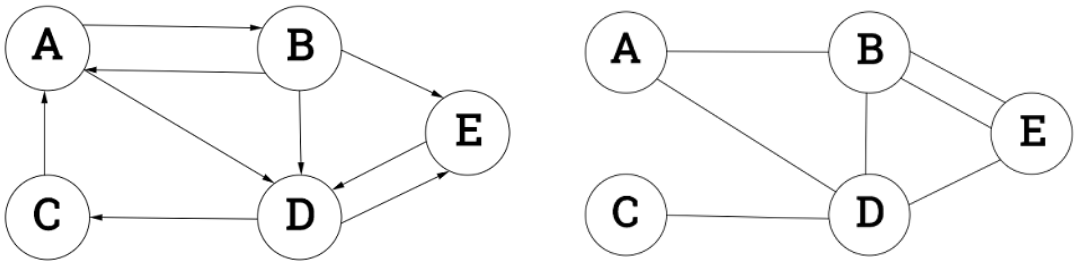


Figure 2.1: Here we see two graphs, one with directed edges, shown by arrows, and the other with undirected edges. Each has an order of 5. Note that since E is a multiset, even in the undirected setting it is still possible to have duplicates of edges (as seen between **B** and **E**).

In simplest terms, a graph is a set of nodes (vertices) which are connected in some configuration (edges). Note that this definition allows edges to be 'one-way', i.e. $(u, v) \in E \implies (v, u) \in E$. This immediately leads to the next definition.

Definition 2.1.4 (Undirected Graph)**.** A graph $G = \{V, E\}$ is *undirected* if and only if for all vertices $u, v \in V$ such that $(u, v) \in E$, we have $(v, u) \in E$.

When discussing undirected graphs, it is normal to view the pair of edges (u, v) and (v, u) as one 'edge', to avoid confusion. The focus of this project will be upon undirected graphs, so from this point on, references to graphs or edges can be taken in the undirected setting, unless otherwise stated.

Similarly, we are interested in graphs for which duplicate edges are not possible, so for the purposes of this review we can assume that E is not a multiset.

We then define the *order* of G to be the cardinality of V, and the *size* of G to be the cardinality of E. If a pair of vertices u and v share an edge $(u, v) \equiv (v, u)$, then we say that u and v are *adjacent*.

Definition 2.1.5 (Adjacency matrix)**.** Let $G = (V, E)$ be a graph, and let $n := |V|$. Then the corresponding adjacency matrix $\mathbf{A} \in \mathbb{Z}^{n \times n}$ is defined by

$$\mathbf{A}_{ij} = \text{no. of edges incident to both } i \text{ and } j$$

Naturally, the adjacency matrix of an undirected graph must be symmetric. Taking the undirected graph from figure 2.1, we can derive the following adjacency matrix:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 & 0 \end{pmatrix}$$

Definition 2.1.6 (Neighbourhood)**.** Let $G = \{V, E\}$ be a graph, and let $u \in V$. Then we define the neighbourhood, $\mathcal{N}$, of u to be

$$\mathcal{N}(u) = \{v \in V : (u, v) \in E\}$$

i.e. the set of all vertices to which u is adjacent. If our graph is a multigraph (allows duplicate edges), then the neighbourhood will be a multiset.

We can then define the *degree* of a vertex u as $\deg(u) = |\mathcal{N}(u)|$.

Definition 2.1.7 (Regular graph)**.** A graph is said to be regular if for every vertex $u \in V$, $\deg(u) = d$ for some constant d.

It is common to refer to a graph with a specific degree d as *d-regular*.

We can construct regular graphs with the same degree in completely different ways, and so two d-regular graphs might behave very differently (see Figure 2.2).

A particular case of a regular graph is a *ring lattice*, in which we take a cycle and add an edge between all vertices who are less than or equal to $k \in \mathbb{N}_{>0}$ steps away, for some k. At $k = 1$ we get the original 2-regular cycle, at $k = 2$ we get a 4-regular graph, and so on (see Figure 2.3).

Definition 2.1.8 (Degree distribution)**.** The *degree distribution* of a graph is the probability distribution of the vertex degrees over the whole graph.

So the degree distribution of a d-regular graph is

$$\mathbb{P}(k) = \mathbb{1}_{\{k=d\}}$$

The degree distribution is an important way of classifying a graph and its properties.

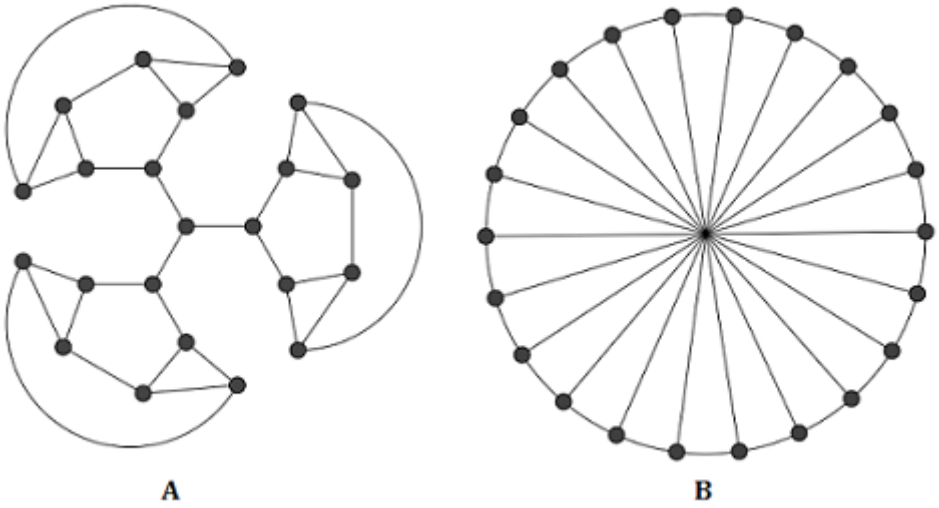


Figure 2.2: 3-regular Graphs: Two graphs with the same number of nodes, each adjacent to three others. Which of these would make a better computer network? Based loosely on an example found in [17].

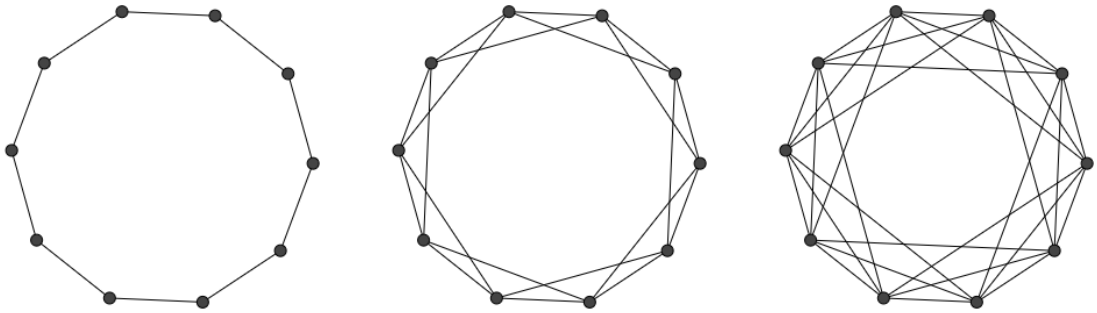


Figure 2.3: On the left we have a cycle, followed by a 4-regular ring lattice ($k = 2$) and a 6-regular ring lattice ($k = 3$).

Graphs in the real-world

Again, it is important to relate the abstract back to reality. In the real-world, graphs can be used to represent a countless number of systems and their analysis has yielded a wide variety of breakthroughs.

An obvious example would be a computer network, in which each node represents a network device such as a server or terminal, and each edge represents a connection. Abstract analysis of graphs can be applied to design networks of arbitrary size effectively, by increasing the connectivity and robustness of the network whilst maintaining a minimal number of (expensive) connections (see Ramanujan graphs in [17]).

Another relateable example of a system that can be analysed as a (directed) graph is the Web, in which each page is represented as a vertex and each link to another page is an edge. Analysis of the Web as a graph is one of the fundamental parts of Google's PageRank algorithm, which was used to rank search results more effectively than the then contemporary search engines.

This algorithm allowed Google to set themselves apart from their competition, and

could be considered as one of the driving factors behind their now gargantuan success. For a more detailed view of PageRank and the theory behind it, I recommend Bianchini et al.'s 2005 paper "Inside PageRank" [18] and Langville and Meyer's companion paper "Deeper Inside PageRank" [19].

Finally, the most relevant example to this project is to consider a social network as a graph. In this setting, we can view an individual as a vertex, and a relationship with another individual as an edge. This is very simple to imagine, but the analysis of the properties that are found in a social network, and the generation of graphs of this type, is a much more complex issue.

To delve into this we will study the topic of random graphs.

2.1.3 Random graphs

The subject of Random Graphs is a young field which marries graph theory and probability theory. A Random graph is described as a random process with which to generate a graph, and so rather than being a graph itself, it is in fact a particular class of graphs which share a probability distribution.

According to Béla Bollobás [20], Random Graph theory "began with some sporadic papers of Erdős in the 1940s and 1950s, in which Erdős used random methods to show the existence of graphs with seemingly contradictory properties".

Paul Erdős, in tandem with Alfréd Rényi, are responsible for the Erdős-Rényi random graph model [21] which can be considered as the canonical Random graph.

Erdős-Rényi model

Paul Erdős and Alfréd Rényi, in tandem with Edgar Gilbert are responsible for the Erdős-Rényi random graph model [21][22]. Here I will describe the variant attributed to Gilbert.

Definition 2.1.9. A $\mathrm{G}(n,p)$ Erdős-Rényi model graph is constructed by taking $n \in \mathbb{N}$ vertices, and adding each possible edge with probability p independent of every other edge.

Theorem 2.1.1. *Let random variable M be the number of edges of a graph generated by a $\mathrm{G}(n,p)$ model. Then M is binomially distributed such that*

$$M \sim \mathrm{Bin}\left(\binom{n}{2}, p\right)$$

Proof. We begin by noting that in a graph of order n, there are a total of $\binom{n}{2}$ possible edges. By Definition 2.1.9 the probability of choosing edge i is drawn from a Bernoulli(p) distribution, and they are chosen independently.

Thus, we can express the number of edges M as

$$M = \sum_{i=0}^{i=\binom{n}{2}} X_i$$

where $X_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. By definition, M is therefore binomially distributed as required. □

Corollary. The expected number of edges of a $G(n,p)$ graph is $\binom{n}{2}p$.

Proof. The result follows immediately from the result of Theorem 2.1.1. □

Lemma 2.1.1. *In $G(n,p)$, every possible graph H with order n and size M has an equal probability of being generated, and this probability is equal to*

$$\mathbb{P}(H|M,n,p) = p^M(1-p)^{\binom{n}{2}-M}$$

Proof. We know that there are $\binom{n}{2}$ edges to draw from. To generate a particular size M graph we must draw M edges, each with probability p and forego a further $\binom{n}{2} - M$ edges, each with probability $1-p$.

Since each draw is independent the probability is simply the product of the probabilities of these draws, i.e.

$$p^M(1-p)^{\binom{n}{2}-M}$$

□

Intuitively, larger values of p are more likely to produce graphs with a greater number of edges, whilst the converse is true of small values of p.

Corollary. For a $G(n, 0.5)$ model, every possible graph of order n will be generated with equal probability.

Proof. For all possible graphs H of size M,

$$\begin{aligned}\mathbb{P}(H|M,n,p=0.5) &= p^M(1-p)^{\binom{n}{2}-M} \\ &= 0.5^M(1-0.5)^{\binom{n}{2}-M} \\ &= 0.5^{\binom{n}{2}}\end{aligned}$$

independent of M. □

Lemma 2.1.2. *The degree distribution of a $G(n,p)$ model graph is*

$$\deg(u) \sim \text{Bin}(n-1,p)$$

or equivalently, the probability that the degree of a vertex is $k \ : \ 0 \leq k \leq n-1$ is

$$\mathbb{P}(k) = \binom{n-1}{k}p^k(1-p)^{n-1-k}$$

Proof. Let u be a vertex in the graph. u has a total of $n-1$ possible edges adjacent to it. Each is chosen independently with probability p, so the total number of edges drawn adjacent to u must be Binomially distributed as required. □

This last result is particularly interesting, as we can apply the Poisson Limit Theorem to assert that as $n \to \infty$ whilst np remains constant, the degree distribution becomes $\text{Poi}(np)$ distributed.

A natural extension of the model is to apply it to an infinite vertex set. If we construct such a graph H, with $0 < p < 1$ then it can be shown (see [24]) that for any $n+m$ elements

$u_1, \ldots, u_n, v_1, \ldots, v_m \in V$, there is almost surely a vertex $c \in V$ such that c is adjacent to all $u_i \ : \ 0 \leq i \leq n$, and is not adjacent to any of $v_j \ : \ 0 \leq j \leq m$.

In a countable vertex space, it has been shown that there is only one such graph (up to isomorphism) [25] which is known as the Rado graph after Richard Rado.

In their 1960 paper [23], Erdős and Rényi remarked on several further interesting properties of the model, including that $\frac{\ln n}{n}$ is the threshold for connectedness of the graph, with values of p below this being disconnected almost surely, and values of p greater than this being almost surely connected. The important pattern to recognise is that Probability theory can be applied to reveal general properties of families of graphs.

Watts-Strogatz small-world model

The Erdős-Rényi model is very effective for determining the features of networks where the topology can be assumed as completely random, just as ring lattices can be effective for examining networks where the topology can be seen as completely regular.

However, as stated by Duncan Watts and Steven Strogatz [26] in the 1998 paper introducing their model, "many biological, technological and social networks lie somewhere between these two extremes".

The *Watts-Strogatz* model therefore attempts to generate networks by finding a middle ground between the random Erdős-Rényi model and the Ring Lattice.

The model aims to construct "small-world" graphs, which are characterised by *local clustering*, where vertices are more likely to be adjacent if they already share a neighbour, and by short average distances between two nodes in the system.

These properties can be clearly seen in social networks. The definition below is based on [26].

Definition 2.1.10 (Watts-Strogatz model)**.** To construct a Watts-Strogatz model small-world graph, we first begin with a k-regular ring lattice $G = (V, E)$ with order n, so that the size of the graph is $\frac{nk}{2}$.

The vertices are labelled u_0 to u_{n-1}, such that there exists an edge (u_i, u_j) if and only if $0 < |i - j| \mod (n - 1 - \frac{k}{2}) \leq \frac{k}{2}$. If we picture the graph as in Figure 2.3, then this is equivalent to numbering the vertices in a clockwise or counter-clockwise fashion.

For each node u_i, we select each edge $(u_i, u_j) \in E$ such that $i \mod n < j$ with probability ρ. Selected edges are 're-wired' by removing them, and then replacing them with an edge (u_i, u_k), where u_k is chosen uniformly from the set $\{v \in V \ : \ v \neq u_i \wedge (u_i, v) \notin E\}$.

In this model, ρ defines the level of randomness of the graph generated, with the special cases being $\rho = 0$, which gives the original ring lattice, and $\rho = 1$ which approaches an Erdős-Rényi $G\left(n, \frac{k}{n-1}\right)$ random graph.

Up to this point, graphs have mainly been characterised by degree distribution, but Watts and Strogatz aimed to classify their graphs by two more features: *characteristic path length* and *clustering coefficient*.

Definition 2.1.11 (Distance)**.** Let $G = (V, E)$ be a graph with vertices $u, v \in V$. The distance between u and v, $\text{dist}(u, v)$ is defined as the number of edges in the shortest path between u and v.

Definition 2.1.12 (Characteristic path length)**.** Let $G = (V, E)$ be a connected graph. Then the characteristic path length, L, is the average distance between distinct pairs of vertices in V.

In general, social networks have been found to have short characteristic path lengths.

Definition 2.1.13 (Clustering coefficient)**.** Let $G = (V, E)$ be a connected graph and let $u \in V$. Then the local clustering coefficient of u is defined as

$$C(u) = \frac{\lambda_G(u)}{\tau_G(u)}$$

where $\lambda_G(u)$ is the number of complete order 3, subgraphs of G containing u, and $\tau_G(u)$ is the number of order 3, size 2 subgraphs or G containing u.

In other words, the clustering coefficient is the number of triangles containing u divided by the number of "possible" triangles containing u.

Definition 2.1.14 (Average clustering coefficient)**.** The average clustering coefficient of graph $G = (V, E)$ is the mean of the local clustering coefficients of all vertices $u \in V$, i.e.

$$C_G = \frac{1}{|V|} \sum_{u \in V} C_G(u)$$

Just as social networks have been found to exhibit low characteristic path lengths, human behaviour causes them to have high clustering coefficients.

In their original paper [26], Watts and Strogatz compared graphs of this model under the conditions of $n \gg k \gg \ln n \gg 1$.

The $k \gg \ln n$ condition ensures connectedness in the $\rho = 1$ case where the model becomes Erdős-Rényi, due to the $p > \frac{\ln n}{n}$ condition derived from [23].

Watts and Strogatz found that for the lattice, the path length $L \sim \frac{n}{2k} \gg 1$, and the clustering coefficient $C_G \sim \frac{3}{4}$., giving a "highly clustered, large world where L grows linearly with n" [26]. The completely random graph at $\rho = 1$ had characteristic path length $L \sim \frac{\ln n}{\ln k}$ and clustering coefficient $C_G \sim \frac{k}{n} \ll 1$, and is therefore poorly clustered but with a short characteristic path length growing only logarithmically with n (for a more in-depth justification of these results, see [27], but note that k is formulated differently in this paper).

As p goes from 0 to 1, however, the characteristic path length drops off much more sharply than the clustering coefficient, giving a 'golden' region in the centre in which the graph features both of the properties of social (and other) networks identified by Watts and Strogatz.

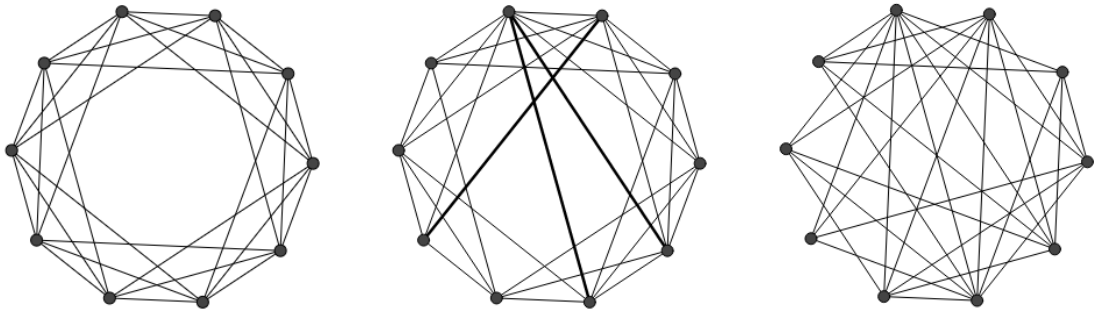
The degree distribution for $0 < \rho < 1$ has been found (see [27]) as

$$\mathbb{P}(\deg(u) = \kappa | u \in V) = \sum_{n=0}^{\mathrm{f}(\kappa,k)} \binom{n}{k/2} (1-\rho)^n \rho^{k/2-n} \frac{\rho^{k/2} \kappa^{-k/2-n}}{(\kappa - k/2 - n)!} e^{-\rho k/2}$$

where $\mathrm{f}(\kappa, k) = \min\left(\kappa - \frac{k}{2}, \frac{k}{2}\right)$ and $\kappa \geq \frac{k}{2}$. This gives a large peak at $\kappa = k$, which decays exponentially to either side.

Barabási-Albert scale-free model

We have seen that the Watts-Strogatz model introduced two important patterns found in social networks, namely clustering and small-world properties, but its graphs lack a feature addressed by Albert-László Barabási and Réka Albert in 1999 [28].



(a) Lattice:	**(b) Small-world:**	**(c) Erdős-Rényi:**
$\rho = 0$	$0 < \rho < 1$	$\rho = 1$
Large distances	Small distances	Small distances
High clustering	High clustering	Low clustering

Figure 2.4: Comparison of the clustering and characteristic path lengths as we interpolate between regular and random graphs. Adapted from example found in [26].

Large networks, including the world-wide web, citation networks and (most importantly) social networks, are found to have degree distribution which follow a *scale-free power law*.

The general concept here is that these networks have a heavy-tailed degree distribution, in which most vertices have a low degree, but a small percentage have an extremely high number of connections.

The findings of Barabási and Albert were that this pattern is caused by a phenomenon dubbed *preferential attachment*, in which vertices are added to the graph over time, and are more likely to connect to a vertex with a higher degree. This kind of system is analogous to the common phrase "the rich get richer".

Definition 2.1.15 (Power law)**.** A functional relationship obeys a power law if a relative change to one quantity results in a (usually inversely) proportional change to the other quantity.

Roughly speaking, a relationship f between two quantities obeys a power law if $\mathrm{f}(x) \propto x^\alpha$ for some constant α.

In preferential attachment networks, we are interested in negative values of α.

Definition 2.1.16 (Scale invariant)**.** A relation f is scale invariant (or homogenous) if and only if scale transformations of the argument result in proportional scaling of the output, i.e. $\forall \lambda \in \mathbb{R}$,

$$\mathrm{f}(\lambda x) = \lambda^k \mathrm{f}(x)$$

for some constant k.

Lemma 2.1.3. *Let* f *be a function obeying a power law. Then* f *is scale invariant.*

Proof. Since f obeys a power law, $\mathrm{f}(x) = \beta x^\alpha$ for some constants α and β. So

$$\begin{aligned} \mathrm{f}(\lambda x) &= \beta(\lambda x)^\alpha \\ &= \lambda^\alpha \beta x^\alpha \\ &= \lambda^\alpha \mathrm{f}(x) \end{aligned}$$

□

Graphs with a power law degree distribution are therefore known generally as being *scale free*.

The method proposed by Barabási and Albert for creating such networks is described below. This definition is based upon the one found in [28].

Definition 2.1.17 (Barabási-Albert model)**.** To construct a Barabási-Albert model scale-free graph, we begin with a connected graph G_0 of some (small) order m_0.

At each time step t we add a new vertex v with $m \leq m_0$ edges, so at time t we have a graph G_t where $|V_t| = m_0 + t$ and $|E_t| = |E_0| + mt \sim mt$.

To choose which vertices these m edges attach to at time step t, we select each other vertex $u \in V_{t-1}$ with probability

$$\mathbb{P}(u) = \frac{\deg(u)}{\sum_{\varphi \in V_t} \deg(\varphi)}$$

This selection process is called *preferential attachement*, as vertices with an already high degree are favoured in future connections.

Barabási and Albert found [28] that graphs constructed in this way would evolve into a scale free network, in which the degree distribution follows a power law with exponent $\alpha = -2.9 \pm 0.1$, leading to the following rule of thumb for the model;

$$\mathbb{P}(\deg(u) = \kappa) \sim \kappa^{-3}$$

The characteristic path length for the model is found in [29] to be

$$L \sim \frac{\ln n}{\ln \ln n}$$

Béla Bollobás [30] showed that the clustering coefficient of these scale free networks also followed a power law;

$$C_G \sim n^{-0.75}$$

It should be pointed out that whilst the Barabási-Albert model exhibits a power law degree distribution and so in this way is advantageous over the Watts-Strogatz small-world model, the small-world clustering coefficient is independent of graph size (unlike the scale-free model), and so neither is strictly better than the other.

RGGs: Random Geometric Graphs

The are a few properties of social networks that have not yet been discussed; positive assortativity, community structure.

Definition 2.1.18 (Assortativity)**.** The (degree) assortativity coefficient of a graph is the correlation of degrees between adjacent vertices in the graph.

Assortativity measures the likelihood that vertices will be joined to vertices with a similar degree, i.e. positive assortativity means that hubs will be connected to hubs.

The second feature, which is often found in biological and social networks [31], is that of *community structure*. Not to be confused with clustering, a graph is said to have community structure if one can describe a graph as a set of densely connected induced subgraphs.

These subgraphs may overlap, meaning that the communities are not necessarily obvious. There are a large number of candidate algorithms for identifying community structure, as no conclusive solution has been found; a paper by Reichardt and Leone in (2008) [32] suggests that there is actually a theshold in which community structure exists but is necessarily undetectable.

The Random Geometric Graph (RGG) model is a model for constructing graphs embedded in space, and it can exhibit both positive assortativity and community structure to some degree.

Definition 2.1.19 (Random Geometric Graph)**.** To construct a RGG, we begin with a set of vertices V embedded in a metric space (M, d), so that $V \subseteq M$.

To add the edges, we take a parameter $g \in \mathbb{R}$ and for every pair of vertices $(u, v) \in V^2$, we set $(u, v) \in E$ iff $\mathrm{d}(u, v) < g$.

In other words, a RGG is a graph in which all vertices less than a threshold distance away from each other are connected, and all others are not. The initial vertex set for the Random Geometric graph model is often drawn from a uniform distribution.

RGGs in 2-dimensional Euclidean space have been shown to have a clustering coefficient approaching $1 - \frac{3\sqrt{3}}{4\pi} \approx 0.59$ [33], and the assortativity has been shown to approach the value of the clustering coefficient for any d-dimensional RGG [34].

The full properties are summed up by Antonioni et al. (2015) [9], "RGG networks thus possess some of the properties associated with social networks (strong clustering, positive assortativity, spatiality) but lack others (short characteristic path length, strong community structure, long-tailed degree distributions".

2.2 Towards an accurate social network

We have seen that there have been many identified properties of a social network, and that there are a variety of different graph families which satisfy some subset of these. In this section I will draw up the full comparison and introduce the recent model proposed by Antonioni, Bullock and Tomassini in 2014 [7]; REDS.

2.2.1 Properties of a social network; existing models

So far, the graph properties we have identified in social networks are as follows:

- High clustering coefficient
- Low characteristic path length (small world)
- Heavy-tailed (scale-free) degree distribution
- Positive assortativity
- Community structure

We have examined some of the most common random graphs, but none exhibit all of these properties:

Model	Clust. coeff.	Path lengths	Deg. dist.	Assort-ativity	Comm. struct.
Lattice	**Strong**	High	$\mathbb{1}_{\kappa=d}$	-	None
Erdős-Rényi	Weak	**Low**	Poi(np)	0	None
Watts-Strogatz	**Strong**	**Low**	Narrow	Low	None
Barabási-Albert	Weak	**Low**	**Scale-free**	0	None
RGG	**Strong**	High	Narrow	**Positive**	**Low**

2.2.2 The *REDS* model

In 2014, a paper by Antonioni, Bullock and Tomassini [7] put forth an extension of the Random Geometric Graph model which demonstrated a complex and non-trivial community structure.

The model is built upon a previous extension of RGGs by Antonioni et al. in 2013 [35] called an Energy-constrained RGG.

The Energy-constrained RGG (EC-RGG) uses an additional constraint to the RGG, in which the sum of the distances of each nodes connections cannot exceed some constant E. Naturally, this allows RGGs with the same vertex space and metric to be constructed in various ways, so the process for construction is to add legal edges at random until no more can be 'afforded'. In this paper, the distance costs are referred to as 'energy' costs.

It was found that, like RGGs, EC-RGGs did not display heavy-tailed degree distributions or a strong community structure.

The REDS model builds upon this and adds a parameter to the model called *Synergy*, which is a direct representation of the likelihood of nodes to share neighbours.

The REDS acronym stands for it's four main construction rules; Reach, Energy, Distance and Synergy, which are described as follows [7]:

- **Reach** an undirected edge (u, v) may exist only if the distance between u and v, $\mathrm{d}(u, v)$, is less than their "social reach" R.
- **Energy** each node u has a finite energy E that can be spent on maintaining edges.
- **Distance** the cost $c(u, v)$ of an edge between u and v is proportional to $\mathrm{d}(u, v)$.
- **Synergy** the cost $\mathrm{c}(u, v)$ varies inversely with the number of shared network neighbours k of u and v, i.e.
$$\mathrm{k}(u, v) = |\mathcal{N}(u) \cap \mathcal{N}(v)|$$
This effect is tuned with parameter $S \in [0, 1]$, so that $S = 0$ gives an EC-RGG.

We can see that Reach alone effectively forms an RGG, and the addition of Distance and Energy gives an EC-RGG. The addition of Synergy allows us to derive our cost function as
$$\mathrm{c}(u, v) = \frac{\mathrm{d}(u, v)}{1 + S\mathrm{k}(u, v)}$$

In [7], the construction process restricts the space to a Euclidean unit square $\Omega \subset \mathbb{R}^2$. A set of n vertices are drawn uniformly from Ω, to form V, and each is given equal initial energy $E_i = E$.

A vertex u_1 is selected uniformly from V, and a second vertex u_2 is drawn uniformly from the set $\{v \in V \ : \ \mathrm{d}(u,v) < R\}$

An undirected edge (u_1, u_2) is added to E if and only if each of the vertices u_1 and u_2 have enough energy to afford the neighbourhood that would result from the new edge.

This is repeated until no more edges can be added.

It is shown that the over all neighbourhood configurations and values of S, the total cost C to vertex u must fall in the interval

$$\sum_{v \in \mathcal{N}(u)} \mathrm{d}(u,v) \leq C \leq \frac{1}{k} \sum_{v \in \mathcal{N}(u)} \mathrm{d}(u,v)$$

In [7], the structure of these graphs was studied over varying E and S (as n and R can be viewed as scale parameters).

The findings were that the results divided sharply into two structural regimes of $E \times S$. In very general terms, systems with high E and S had properties resembling RGGs, whilst those with low E and S became very sparse.

Along the border between the two regimes however, there was found to be a golden zone in which the degree distributions were complex and heavy-tailed, and strong clustering, positive assortativity and community structure prevailed.

The only feature not represented was that of short path lengths, but a solution has been proposed [15] in which the Watts-Strogatz method of re-wiring could be applied to a REDS graph rather than a lattice, to introduce the small-world property to these networks.

In [15], we see that this method (as in the Watts-Strogatz model) allows characteristic path length to drop off before the other features, resulting in a window at around $10^{-2} < p < 10^{-1}$ in which characteristic path length is low, whilst clustering remains high.

2.3 Co-operation in Systems

In this section I will discuss existing literature on studying the evolution of cooperation on networks, which varies both in the types of networks studied and the rules of the 'game'.

Nowak and May, 1992

One paper of note on the subject of evolutionary cooperation systems is Nowak and May's 1992 study "Evolutionary games and spatial chaos" [5].

Nowak and May formulate a prisoner's dilemma game as a 2-player 2-strategy symmetrical game, with the following payoffs:

- $R :=$ mutual cooperation
- $S :=$ cooperate vs. defect
- $T :=$ defect vs. cooperate
- $P :=$ mutual defection

This gives payoff matrix:

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{bmatrix} R & S \\ T & P \end{bmatrix} \end{array}$$

in which $T > R > P > S$.

For standardisation, Nowak and May set $R := 1$ and $S := 0$, and interestingly found that $P > 0$ produced no qualitatively different results to $P = 0$. Therefore they set

$$T = b > 1 \ > \ R = 1 \ > P = S = 0$$

where the variable b represents the temptation to defect.

In this paper, the focus was on simple lattice systems, with the examples given being run on 8-regular graphs. The process that was performed on these graphs was as follows:

- In each cycle, every vertex plays the game exactly once with every neighbouring node. A score or 'measure of individual fitness' is then generated for each node as the sum of the payoffs for that node.
- Each node's strategy is then updated, taking on the strategy of its *highest scoring neighbour* for that round. Note that we allow a node to maintain its strategy if it is higher scoring than its neighbours.
- This cycle is repeated and the evolving dynamics of cooperation are observed.

This strategy updating system is absolute, so each cycle is deterministic. As such, we could view Nowak and May's model as a 'cellular automata'-type system.

In their research, Nowak and May [5] found that systems with $b < 1.8$ resulted in cooperation winning absolutely, whilst the opposite was true of systems in which $b > 2$.

For systems with $1.8 < b < 2$, they found that clusters of cooperation could grow in regions of defection and clusters of defection could grow in regions of cooperation, giving a much more interesting chaotic dynamic.

Asymptotically, they found that the ratio of cooperation to defection fluctuated around 0.32 for most initial configurations, although the actual regions continued to grow and shift erratically.

The results of this paper are very interesting, but it would be fair to view the absolute strategy rule as heavy-handed; entities will likely be more inclined to stay with a strategy than to switch (so the strategy rules should be 'sticky').

Santos and Pacheco, 2006

A much more recent (2006) paper by Francisco Santos and Jorge Pacheco, called "A new route to the evolution of cooperation", suggests an improved model for updating strategies [36].

In this paper, Santos and Pacheco again follow the prisoner's dilemma system of $T = b > R = 1 > P = S = 0$, although they add the new constraint of $b < 2$. Again they run the game on every edge in the network, and calculate an accumulated payoff to represent the individual fitness of a node u, which we will represent as M_u.

To update strategies however, they use a different system, which is a "finite population analogue of replicator dynamics" [37], to which the system will converge in "in the limit of well-mixed populations", i.e. a complete graph.

This evolution system is as follows; to update the strategy of vertex u, we select a vertex v uniformly from $\mathcal{N}(u)$. If $M_u \geq M_v$, then u will retain its existing strategy into the next round. Otherwise, u will adopt the strategy of v with probability

$$p = \frac{M_v - M_u}{\max\left(\deg(u), \deg(v)\right)(T - S)}$$

The denominator here ensures that $0 \leq p \leq 1$. We see that this represents a much more intuitive evolution system, and allows for a degree of unpredictability in results.

Santos and Pacheco obtained results [36] for these simulations on regular lattices, random Erdős-Rényi graphs, Barabási-Albert scale-free networks, and randomised scale-free networks.

Randomised scale free networks are a variant of the Barabási-Albert networks, with a modification to avoid what is known as 'age-correlation' [38]. In this model, a Barabási-Albert network is generated, and then vertices at the ends of pairs of edges are randomly exchanged, therefore removing the age-correlation whilst retaining the degree distribution.

For each type of graph, the starting strategies were randomly and evenly distributed among the vertices. To gain results, the systems were allowed to 'mix' for 10000 generations (to allow time to converge), and the equilibrium frequencies were calculated by averaging over a subsequent 1000 generations.

The results were that for regular graphs, cooperation ceased to persist sharply when $b > 1.1$. This is in contrast to Nowak and May's [5] game, in which Cooperation won for $b < 1.8$. The Erdős-Rényi graphs were better at maintaining cooperation, but this falls slowly as b is increased, and cooperation completely failed to stabilise in systems with $b > 1.6$.

The scale-free networks were much more accommodating for cooperation, falling away very little for large b. At $b = 2$, the random scale-free graphs stabilised with approximately half of the vertices cooperating, and Barabási-Albert model graphs exhibited a stable fraction of cooperation as high as 0.9. See Figure 2.5 for the original results found in [36].

This brings us to the current state of research in the area, and I will now describe the workplan of this project.

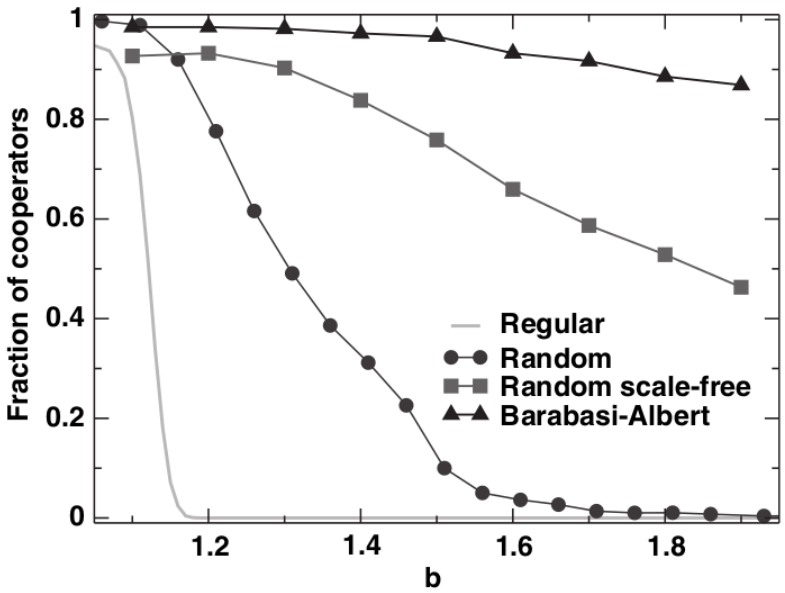


Figure 2.5: Graph taken from [36]. The equilibrium ratio of cooperators to defectors in each graph model tested, over $b \in (1, 2)$.

Chapter 3

Workplan

3.1 Main tasks

3.1.1 Implementing the game model

The first task to undertake is the implementation of a simulation model which matches the standards set by Santos and Pacheco [36].

To do this, I will make use of the Python NetworkX package [39], which provides an existing and standardised framework for storage and manipulation of graph network data structures.

This will allow me to create a framework which can run simulations on various types of graphs, for the cross-checking and testing of results. The NetworkX package includes existing generators for Erdős-Rényi graphs, Watts-Strogatz graphs, Barabási-Albert graphs and Random Geometric graphs.

I will construct a simulation model according to the one set out in [36], and test this on the above types of graphs to replicate and test the results already seen. I will also further test the model on the Random Geometric graph model, to allow for a comprehensive comparison in the results.

3.1.2 Construction of *REDS* networks

A constructor for REDS graph data structures in line with the NetworkX standard must then be implemented. The specification set out in [7] will be followed and adapted for NetworkX.

Once a constructor has been built and tested, I will implement metrics for the social network properties outlined in Section 2.2.1. This will serve both to replicate the results produced in [7], and to test the constructor which I have implemented.

A constructor for small-world REDS graphs will also be produced, extending the initial model.

3.1.3 Games on *REDS* networks

This section will produce the first main area of results. With the REDS graph constructor and simulation model implemented and working, the next stage will be to put the two together.

Following the standard outlined in [36], I will set an initial state of random and equally distributed cooperators and defectors and run the simulations for 10000 iterations to allow for mixing. Results will then be extracted from the average of the subsequent 1000 iterations.

As in [7], the parameters n and R can be set as constant, as they represent 'scale factors' of the network. To test, the simulations will be run on REDS graphs over a range of E, S and b values, to see how cooperation behaves in both regimes of the graph, and in particular in the border area 'sparse' regimes.

This stage of the project should produce results that allow comparison of the behaviour of cooperation between REDS graphs and the other models.

At the conclusion of this stage, the following objectives will be met:

- **Objective 1** To derive formalised conditions in the form of parametric constraints under which co-operation may be expected to emerge and persist on a static REDS random graph.

3.1.4 Games on dynamic *REDS* networks

This stage of the project will be considerably more exploratory. To evaluate the behaviour of cooperation on dynamic REDS graphs, a framework must be designed to allow the structure of a graph to update in a way which intuitively (or at least reasonably) maps to reality.

At this point the main idea is to allow vertices to resign from connections which are consistently costly. This will most likely take the form of a third stage in each cycle, in which nodes are given a probability of evaluating a connection, perhaps over some window of past cycles, and retracting that connection if a condition is met, therefore 'burning a bridge'.

A framework for rewiring should also then be implemented, as both 'freed' vertices will then possess surplus energy.

There are a great many possibilities on how to implement such a system, and thought must be given on how this might affect the convergence rate of the system, and whether to allow graphs to become disconnected.

At this time I believe it would be correct to allow vertices the possibility of reconnecting to vertices they have just 'burnt', because in (spatially) sparse regions of the network, this may be the only choice, and could be thought of as representing a 'small-town' syndrome.

At the conclusion of this stage, the following objectives will be met:

- **Objective 2** A justifiable framework under which the REDS graph, in the setting of games, may be extended to have a dynamic element. This may be tied directly to the payoffs of the games being simulated or otherwise.

- **Objective 3** The further conditions under which the extended dynamic system can be expected to exhibit co-operation.

3.1.5 Results and analysis

This section will involve the communication of the project and its results in the form of a paper. The methods and techniques used must be explicitly and unambiguously stated, and then justified.

The results must then be presented in an intuitive manner, and comparative analysis of the results must be undertaken.

The vast majority of results will be in the form of large datasets (since for each simulation we are returning 1000 iterations of the graph state). As such a comprehensive (but decisively motivated) statistical analysis of the data will be carried out. Due to the size and complexity of the datasets, the analysis will most likely be performed in R.

Finally, the methods and results should be addressed in the context of a social network, to adhere to the motivation of this project.

At the conclusion of this stage, the following objectives will be met:

- **Objective 4** Detailed analysis and justification of the above results, including an analysis of the implications of the results when considered in the setting of a true organic social network.
- **Objective 5** A full description and justification of the methods and techniques used to generate the results, so that they may be replicated/tested.

3.2 Expected Timeline

3.2.1 Breakdown of tasks

To decide on a projected timeline, I will begin by providing a further breakdown of the tasks required for each of the stages detailed above, including estimated time required per task. Note that within stages, it is expected that some overlap will occur, and that some level of writing will be present throughout all stages.

Stage	Tasks	Expected time (days)
Game model	Design and implementation	5
	Testing and analysis	5
REDS constuction	Design and implementation	6
	Testing and analysis	7
Games on REDS	Simulations	10
	Analysis	4
Dynamic REDS	Design and justification	8
	Implementation	6
	Testing	3
	Simulations	10
	Analysis	4
Results	Full comparative analysis	4
Writing	Methods and justification	3
	Results and analysis	3
	Editing	3

Work on the project will fall primarily between 5 June 2017 and 15 September 2017. Based on the estimated times above and accounting for overlaps, particularly in the simulation tasks, the breakdown of tasks over time will be roughly as follows:

Week No.	Week starting	Tasks assigned
1	5 June	Game model
2	12 June	REDS construction
3	19 June	REDS construction
4	26 June	Games on REDS
5	3 July	Games on REDS
6	10 July	Dynamic REDS
7	17 July	Dynamic REDS
8	24 July	Dynamic REDS
9	31 July	Dynamic REDS
10	7 August	Results
11	14 August	Writing
12	21 August	Writing
13	28 August	Writing
14	4 September	*Buffer*
15	11 September	*Buffer*

3.3 Risk Analysis

The main threat to this project is that of unforeseen difficulties causing delays against the timeline. I have attempted to identify the most severe and likely risks below.

Non-replication of previous results

Medium Severity - Low Probability

Since the first two stages of the project involve re-treading the steps of previous studies, to build a standardised framework against which to test, there lies a risk that the results gained from these stages do not match the results found in those studies.

Possible causes for this include that the methodology used in these studies may be ambiguous, or that the implementation of my models carry bugs. Contingency planning is as follows:

In the case of the first cause, a detailed understanding of the underlying principles will afford confidence in the models created, even if they differ parametrically in some way to previous models. In this situation the results attained can still provide insight to the comparative qualitative behaviours of the graph types, but some value added will be lost as direct comparison between studies will be more difficult.

In the case of the second cause, a proportionate amount of time has been allocated for the testing of both implementations of the first stage, and so proper software development process shall be able to deal with this. I estimate a small number of actual lines of code, due to the nature of the problem and the already standardised and documented framework of NetworkX.

Simulation time delays

Low Severity - High Probability

To generate results, long simulations (10^4 cycles) will have to be carried out repeatedly with varying parameters. This will require substantial compute time.

Furthermore, mistakes in the input parameters of a simulation could cause this compute time to be wasted in large chunks.

To account for these issues, I will be dedicating a relatively (to the simulations required) powerful machine solely to this task, separate from the other work required. This will allow tasks to be worked upon and completed in parallel to the simulations being run.

Due to the standardised nature of the simulations and the testing that will be undergone on the models, I intend to lightly automate the sequence of simulations to attempt to negate the possibility for human error when initiating them, and the initial results analysis will be carried out in tandem with further simulations being run.

Loss/corruption of data

High Severity - Low Probability

A severe threat to the project is the loss or corruption of data. In such a scenario, the outcome for the project could be catastrophic and terminal, so rigorous countermeasures shall be put in place.

Simulation data will be backed up each day that simulations are being run to several places; university file space, a secondary machine and the cloud. Storage will then be robust against hard drive failure (as it exists on four separate file systems), as well as physical loss or damage (e.g. fire), as the data will exist in separate geographical locations.

Loss of or damage to equipment required

Low Severity - Low Probability

Events rendering necessary equipment unusable can lead to delay. Thankfully, this project relies primarily on relatively small simulation models, and so the only vital equipment required is a machine on which to run simulations.

The primary machine used will be stationary and dedicated exclusively for this purpose whilst simulations are being run, somewhat reducing the risk of failure. On top of this I have spare components in case of minor failure and two other machines immediately available.

Failure in accurate prediction and analysis of risks

Variable Severity - Very High Probability

It is highly probable that, despite best efforts, potential outcomes will be unforeseen, and their effects will be inaccurately estimated. As such, time estimates for the project have been relatively generous and a two-week buffer period has been allowed to account for delays.

Chapter 4

Conclusion

The aim of this review was to explain the motivation behind this project and to explore the existing state of research and theory relevant to the objectives. This has included comprehensive exploration of the field of evolutionary cooperation systems, including game theoretic interpretations, and the study of how graph structures can more accurately represent (and affect) these systems.

To achieve this, I first covered the motivating issue behind this research (see Motivation section of 2.1.1). To achieve this a basic understanding of Game theory had to be achieved and reformulated in the setting of this problem.

To acquire the background necessary to study evolutionary cooperation system on graphs, an understanding of fundamental graph theory was attained and communicated (2.1.2), and a review of existing research and analysis of random graphs was undertaken (2.1.3).

These two areas were discussed in the context of both actual application and social networks, to relate the relevance and utility of the concepts.

The results of Antonioni, Bullock and Tomassini (2014) [7] were then discussed at length (2.2.2), and the interesting and original nature of the REDS model was ascertained, giving traction to the motivation of this project.

Finally, a summary of the current state of research in the field was provided (2.3), outlining the existing standards by which the evolution of cooperation has been assessed on network systems. The existing results for Regular, Random and Scale-free networks were discussed, highlighting the clear need for study of how the REDS model networks fall into the comparison.

I have outlined my intentions for the execution of this project and look forward to answering the question implicit in this review; how does cooperation evolve on a truly social network?

Bibliography

[1] Santos, F.P., Pacheco, J.M. & Santos, F.C. (2016) "Evolution of cooperation under indirect reciprocity and arbitrary exploration rates", *Scientific Reports*, Nature, **6**

[2] Axelrod, R. (1984) *The Evolution of Cooperation* pp. 216

[3] Read, J.M., Eames, K.T.D & Edmunds, W.J. (2008) "Dynamic social networks and the implications for the spread of infectious disease" *J.R.Soc. Interface* **5** (26)

[4] Di Paolo, E.A. (2000) "Ecological symmetry breaking can favour the evolution of altruism in an action-response game" *Journal of Theoretical Biology* **203** pp. 135-152

[5] Nowak, M. & May, R. (1992) "Evolutionary games and spatial chaos" *Letters to Nature*, Nature, **359** pp. 826-829

[6] Lu, Q., Korniss, G. & Szymanski B.K. (2008) "Naming games in two-dimensional and small-world-connected random geometric networks" *Phys Rev E Stat Nonlin Soft Matter Phys* **1** (2)

[7] Antonioni, A., Bullock, S., & Tomassini, M. (2014) "REDS: An energy-constrained spatial social network model" *Complex Adaptive Systems: Artifical Life* **14** pp. 368-375

[8] West, S. A., Griffin, A. S. and Gardner, A. (2007) "Evolutionary Explanation for Cooperation" *Current Biology* **17** (16) R661-R672

[9] Antonioni, A., Bullock, S., Darabos, C., Giacobini, M., Iotti, B. N., Moore, J. H., & Tomassini, M. (2015) "Contagion on networks with self-organised community structure." *Advances In Artificial Life* pp. 183-190

[10] Myerson, R. B. (1991) *Game Theory: Analysis of Conflict* Harvard University Press, pp. 1

[11] Shoham, Y. and Leyton-Brown, K. (2009) *Multiagent Systems: Algorithmic, Game-Theoretic and Logical Foundations* New York: Cambridge University Press

[12] Griffin, A. S., Pemberton, J. M., Brotherton, P. N. M., McIlrath, G., Gaynor, D., Kansky, R., O'Riain, J. and Clutton-Brock, T.H. (2003) "A genetic analysis of breeding success in the cooperative meerkat (Suricata suricatta)" *Behav. Ecol.* **14** pp. 472-480

[13] Hardin, G. (1968) "The tragedy of the commons" *Science* **162** pp. 1243-1248

[14] West, S. A., Griffin, A. S., Gardner, A. and Diggle, S. P. (2006) "Social evolution theory for microbes" *Nat. Rev. Microbiol.* **4** pp. 597-607

[15] Bullock, S. (2015) "Contagion on networks with a self-organised community structure" Presentation on 9th December 2015, Keele

[16] Bondy, J. A. and Murty, U. S. R. (1976) *Graph Theory with Applications* Elsevier Science Publishing; New York, NY

[17] Krebbs, M. and Shaheen, A. (2011) *Expander Families and Cayley Graphs* Oxford University Press; New York, NY

[18] Bianchini, M., Gori, M. and Scarselli, F. (2005) "Inside PageRank", *ACM Transactions on Internet Technology*, **5** (1), pp. 92-128

[19] Langville, A. N. and Meyer, C. D. (2004) *Deeper Inside PageRank* Tech. rep., North Carolina State University, NC

[20] Bollobas, B. (1998) "Random Graphs" *Modern Graph Theory* Springer; New York, NY pp. 215-252

[21] Erdős, P. and Rényi, A. (1959) "On Random Graphs I" *Publicationes Mathematicae* **6** pp. 290-297

[22] Gilbert, E. N. (1959) "Random Graphs" *Annals of Mathematical Statistics* **30** pp. 1141-1144

[23] Erdős, P. and Rényi, A. (1960) "On the evolution of random graphs" *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **5** pp.17-61

[24] Cameron, P.J. (1997) "The random graph" *The mathematics of Paul Erdős, II* Alg. Comb., **14** pp. 333-351

[25] Erdős, P., Rényi, A. (1963) "Asymmetric graphs" *Acta Mathematica Academiae Scientarium Hungaricae* **14** pp. 295-315

[26] Watts, D.J. & Strogatz, S.H. (1998) "Collective dynamics of 'small-world' networks" *Letters to Nature* **393** pp. 440-442

[27] Barrat, A. & Weigt, M. (2000) "On the properties of small-world network models" *European Physical Journal B.* **13** (3) pp. 547-560

[28] Barabási, AL. & Albert, R. (1999) "Emergence of Scaling in Random Networks" *Science* **286** pp. 509-512

[29] Cohen, R. & Havlin, Shlomo (2003) "Scale-free Networks and Ultrasmall" *Physical Review Letters* **90** (5)

[30] Bollobás, B. (2003) "Mathematical results on scale-free random graphs" *Handbook of Graphs and Networks* pp. 1-37

[31] Girvan, M. & Newman, M.E.J. (2002) "Community structure in social and biological networks" *Proc. Natl. Acad. Sci. USA* **99** (12)

[32] Reichardt, J. & Leone, M. (2008) "(Un)detectable Cluster Structure in Sparse Networks" *Phys. Rev. Lett.* **101** pp. 1-4

[33] Dall, J. and Christensen, M. (2002) "Random Geometric Graphs" *Physical Review E* **66**

[34] Antonioni, A. and Tomassini, M. (2012) "Degree correlations in random geometric graphs" *Physical Review E* **86**

[35] Antonioni, A., Egloff, M. and Tomassini, M. (2013) "An energy-based model for spatial social networks" *Advances in Artificial Life* pp. 192-199

[36] Santos, F.C. and Pacheco, J.M. (2006) "A new route to the evolution of cooperation" *Journal of Evolutionary Biology* **19** (3) pp. 726-733

[37] Gintis, H. (2000) *Game theory evolving: A problem-centered introductionto modelling strategic behaviour* NJ: Princeton University Press

[38] Albert, R. and Barabási, A. (2002) "Statistical mechanics of complex networks" *Rev. Mod. Phys.* **74** pp. 47-97

[39] "NetworkX" https://networkx.github.io/. Accessed 21-04-2017