UNIVERSITY OF SOUTHAMPTON

Dynamics and stability of small social networks

by

Elisabeth zu Erbach-Schoenberg

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Business and Law
School of Management

June 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF BUSINESS AND LAW
SCHOOL OF MANAGEMENT

Doctor of Philosophy

by Elisabeth zu Erbach-Schoenberg

The choices and behaviours of individuals in social systems combine in unpredictable ways to create complex, often surprising, social outcomes. The structure of these behaviours, or interactions between individuals, can be represented as a social network. These networks are not static but vary over time as connections are made and broken or change in intensity. Generally these changes are gradual, but in some cases individuals disagree and as a result "fall out" with each other, i.e. , actively end their relationship by ceasing all contact. These "fallouts" have been shown to be capable of fragmenting the social network into disconnected parts. Fragmentation can impair the functioning of social networks and it is thus important to better understand the social processes that have such consequences.

In this thesis we investigate the question of how networks fragment: what mechanism drives the changes that ultimately result in fragmentation? To do so, we also aim to understand the necessary conditions for fragmentation to be possible and identify the connections that are most important for the cohesion of the network. To answer these questions, we need a model of social network dynamics that is stable enough such that fragmentation does not occur spontaneously, but is simultaneously dynamic enough to allow the system to react to perturbations (i.e. , disagreements). We present such a model and show that it is able to grow and maintain networks exhibiting the characteristic properties of social networks, and does so using local behavioural rules inspired by sociological theory.

We then provide a detailed investigation of fragmentation and confirm basic intuitions on the importance of bridges for network cohesion. Furthermore, we show that this topological feature alone does not explain which points of the network are most vulnerable to fragmentation. Rather, we find that dependencies between edges are crucial for understanding subtle differences between stable and vulnerable bridges. This understanding of the vulnerability of different network components is likely to be valuable for preventing fragmentation and limiting the impact of social fallout.[1]

[1]This work was supported by an EPSRC Doctoral Training Centre grant (EP/G03690X/1).

Contents

List of Figures

List of Tables

Declaration Of Authorship

I, Elisabeth zu Erbach-Schoenberg, declare that the thesis entitled *Dynamics and stability of small social networks* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as zu Erbach-Schoenberg et al. (2014).

Signed: ....

Date ....

Acknowledgements

A large number of people have helped me in so many different ways and I would like to thank all of them for their help and support.

My supervisors Prof Seth Bullock and Prof Sally Brailsford for their guidance and continued interest in my topic, especially in times of modelling dead ends where their enthusiasm has helped to motivate me to try out yet another model variant. The Modelling Social Systems group for providing the supportive but critical interdisciplinary community that is so hard to find. Nicki for always sorting out any bumps on the Phd road and helping create the environment in the DTC that makes all phases of the Phd a less solitary experience.

My family and godparents for supporting me in many ways, especially my parents for believing in me during the less productive phases of my work where progress is hard and slow.

My friends who have helped me with this thesis through discussions and feedback and celebrating small and large successes that are so easily ignored when there is always more to do.

And to Rob, who has taught me many things about good science and provided so much support in many ways. For listening to many thesis problems, small and large, theoretical and technical, and always finding a way to help me solve them: through discussion, questions, examples of good and bad science or simply listening.

To Opa,

who taught me that maths is a

powerful tool that should be used
in engineering and cake making alike.

Chapter 1

Introduction

> "Social networks are dynamic by nature. Ties are established, they may flourish and perhaps evolve into close relationships, and they can also dissolve quietly, or suddenly turn sour and go with a bang."
> —Snijders et al. (2010)

The world we live in is becoming more and more complex. Or rather, our perception of it is becoming more and more complex since we have reached the boundaries of knowledge acquisition through reductionist science. Focus in many fields of science is shifting from the study of system components in isolation to studying interactions and relationships.

In social systems, such as groups of humans or animals, organisations or countries, behaviour of the system as a whole is shaped by individual actions and the aggregated effect arising from the interplay of the actions. Thus, the system behaviour is the result of a non-linear aggregation of individual actions, rendering it a complex system (Newman, 2011). The complicated set of interactions between humans or social entities can be described as a social network (Wasserman and Faust, 1994). In a network, nodes represent the system components and edges between them the relationships or interactions between the components. In a social network the nodes are most commonly individuals, however the nodes can also represent groups of people, such as organisations or governments. The nature of the interactions being represented depends on the setting considered. Network connections might represent mere proximity between individuals, or a history of explicit interactions of some type, e.g. sexual contacts (Liljeros et al., 2001), face-to-face meetings (Cattuto et al., 2010), phone calls (Onnela et al., 2007b) or online interactions (Grabowski, 2007; Leskovec and Horvitz, 2008; Szell and Thurner, 2010). Alternatively, connections can represent interactions with a specific function, such as one individual seeking advice from another. Thus, social networks can have a variety of functions such as communication, information spread and support. In other cases the function of a social network is less obvious or less clear cut as edges might fulfill different functions at different times, such as support at one time and information transfer another. Some networks that fall under the term social network are purely epiphenomenal, such as the network of co-starring actors. While this type of network represents data on real-world interactions, the connections are not perceived to have any function from the perspective of the actors involved.

When trying to understand systems shaped by the actions of many individuals, the way these actions interact to form a system level outcome is important (Strogatz, 2001). It has been shown that the interaction structure strongly influences the outcomes in many contexts, ranging from disease transmission (Keeling, 2005) to movement decisions in animals (Bode et al., 2012).

In many contexts, information and knowledge are transferred along the edges of social networks. In organisations a formal, often hierarchical network is present but in addition, an informal network might develop (Albrecht and Adelman, 1984; Albrecht and B.W., 1997; Albrecht and Hall, 1991; Contractor et al., 2012; Huberman and Adamic, 2004) and this informal network could be crucial for the performance of the organisation (Borgatti and Cross, 2003; Burt, 1992; Eisenberg and Monge, 1987; Geard and Bullock, 2010; Krackhardt, 1994; Papa, 1990) as it influences information diffusion within the organisation, knowledge transfer and the quality as well as speed of organisational decisions (Carley, 1999; Davis and Greve, 1997). Despite the importance of informal networks on organisational performance and the fact that a lot of empirical research exists on this, according to Monge and Contractor (2001) almost no work on theories of mechanisms driving small-group network formation has been published in the previous 20 years.

Well-connected networks, in which multiple paths between pairs of nodes exist, have been shown to have an important influence on team performance (Sinclair et al., 2012). The influence of interactions and information flow on team performance has been studied in a range of settings, such as medical teams in hospitals (Leykum et al., 2012) and football teams in the English Premier League (Grund, 2012). In some contexts lives depend on the cohesion of the group and intact communication structures such as military operation or research expeditions; one example that has been studied in this context being South Pole research station staff during the winter months where the team can not leave the station (Johnson et al., 2003). Even in cases where social fragmentation would not be life-threatening to anyone in the network it can significantly impact individuals' well-being and happiness (Salzinger, 1982). A well known fictitious example describing the potential effects of group fragmentation is given in the novel “Lord of the Flies” by William Golding.

Thus, it is pivotal in many contexts for communication networks to stay well connected and not fragment into isolated components. As with many complex systems, social systems cannot be easily steered – as examples from organisations have shown: here, an informal communication network frequently forms in addition to a formal one. It is therefore important to understand the factors that contribute to the formation of stable and well-connected networks, and which of these factors could yield points of influence.

In cases where network cohesion is weak, it has been observed that the social network fragments into isolated components, completely cut off from each other (Zachary, 1977). Even though the problem of network fragmentation has been mentioned in the literature (Borgatti and Halgin, 2011) and most of us will have anecdotal evidence of it occurring in our own social networks, not much work has been done in investigating the processes responsible for network fragmentation. Note that we do not always want to prevent fragmentation: in some case we might want to influence the network to actually make it more fragile. This has been discussed in relation to disease transmission (e.g. fragmentation of a HIV transmission network, Rothenberg et al., 1996) and we could imagine interest in this in the context of terrorist networks (Krebs, 2002) or companies and political parties (regarding competitors and opposing parties). In any case, understanding the mechanisms of fragmentation is crucial for achieving any chance of influencing network stability.

Although network stability (robustness to node or edge removals) has been studied in detail when considering technological networks (e.g. electricity grid, Crucitti et al., 2004b), processes on social networks are quite different from current flow on electrical networks. We will discuss why findings regarding the stability of technological networks cannot be directly transferred to social networks later in this thesis. In social networks, fragmentation has been studied intensively using coevolutionary opinion dynamic models (Holme and Newman, 2006). In these dynamic models individuals are connected by a network. They hold opinions and change these based on the people they are connected to. This means that the network topology influences the information flow between people. In addition to these dynamics on the network, there is a dynamic *of* the network: individuals break connections to others that hold an opinion that is too dissimilar from their own. Depending on the rates or timescales of the two processes (opinion change and edge breaking), consensus or fragmentation can be observed (Gil and Zanette, 2006; zu Erbach-Schoenberg et al., 2011).

In this thesis we investigate the phenomenon observed in Zachary (1977): a single disagreement causes the network to fragment. In this example of a university Karate club, the club president and the instructor had a disagreement about lesson prices that eventually led to fragmentation of the social network between club members. Fragmentation of a network following a single disagreement can result from two different processes (or an interplay of those): individuals taking sides and interacting preferentially with individuals in the same faction; and individuals being drawn to one side of the conflict without explicitly choosing sides. The first process is simple and relatively easy to understand intuitively, even though the complexities introduced by the networked nature of the system make it hard to make intuitive predictions about sizes and exact membership of the resulting disconnected components. This mechanism is well described by models of coevolutionary opinion dynamics, as introduced above, and has been well-studied. However, most of us have experienced or heard of situations in which individuals did not explicitly choose a side and nevertheless fragmentation occurred. This scenario is clearly distinct from the first scenario, and the mechanisms underlying fragmentation in these conditions are not yet well understood.

In this thesis, we will present a mechanistic explanation for a previously stable network fragmenting as a consequence of a single disagreement, even when the two disagreeing individuals are the only ones to change their behaviour. To study the mechanisms that lead to fragmentation, we explicitly simulate network dynamics. To demonstrate that the fragmentation is caused by the perturbation (i.e. a disagreement) the network must have reached a stable topology before any perturbation is introduced. In order for the system to be able to react to a perturbation, it is at the same time necessary for the system to exhibit ongoing dynamics on the micro-level. Here we will introduce a dynamic model satisfying both of these requirements. In the model presented here, local rules grow and maintain a network which exhibits the properties characteristic of social networks. These topological characteristics remain stable over time while on the micro-level relationships continue to fluctuate, thus exhibiting ongoing dynamics. Using this model we show that in cases where individuals do not explicitly take sides, dependencies between edges play a crucial role in the fragmentation process.

1.1 Methodology: agent-based models

In this thesis, we use an agent-based model to simulate the dynamics of network fragmentation.

In an agent-based model individual components interact and the behaviour of these individuals or agents is specified by local rules. Agent-based models (Bonabeau, 2002) are one of the main techniques used for modelling complex systems. They are well suited for modelling complex systems as the macro-level behaviour emerges as a result of non-linear interactions between micro-level components (Bedau, 2002). Social systems are complex systems since the social behaviour does not follow linearly from individual behaviour (Moss, 2008). Thus, the use of agent-based models has been popular in the social sciences (Gilbert and Troitzsch, 2005; Miller and Page, 2009) and many believe that agent-based models play a crucial role in bridging the macro-micro gap between macro-level observations and theory on the micro-level (Raub et al., 2011). Their use ranges from investigating whether hypothesised behavioural mechanisms are able to generate certain system level outcomes, to models incorporating large amounts of data with the purpose of making predictions about system-level behaviour. Agent-based models can integrate data from various sources such as qualitative and experimental data, as well as empirically validated theoretical knowledge (Alam and Geller, 2012).

For the problem investigated in this thesis, an agent-based modelling approach is highly appropriate: we want to explore behavioural rules and constraints that allow growth and maintenance of a social network as well as creating fragmentation following a fall-out (i.e disagreement between two individuals). Using such an approach enables us to examine, at a conceptual level, mechanisms and processes that are entangled in fragmenting social networks. We want to examine whether it is even possible conceptually to create the outcome of fragmentation under the conditions described, as no mechanism has been proposed for this so far and there is a lack of time-resolved data that has a fine enough resolution to allow us to investigate potential mechanisms. Purely mathematical approaches, on the other hand, do not deal well with heterogeneities in the system, an aspect that is key for understanding how fragmentation could result from a single behavioural change in a specific location.

1.2 Methodology: networks of agents

Agent-based modelling in the social sciences is often combined with networks to represent interaction structures (Alam and Geller, 2012). Instead of assuming that agents randomly select interaction partners, potential interactions are represented by a social network.

Networks are often assumed to be binary, meaning that a connection between two individuals is either present or absent. However, it is often appropriate to attribute a strength or weight to each network connection, since some social interactions may be stronger or take place more frequently than others (Barrat et al., 2004a, 2005). When representing such weighted interactions as a binary network, important information may be lost when inferring network topology from interaction patterns as it is necessary to apply some form of thresholding to decide whether a certain frequency of interaction should be represented as an edge or not. How this threshold is chosen influences the resulting topology (Franks et al., 2010). Additionally, when simulating the dynamics of networks, using binary connections restricts the edge dynamics to make and break dynamics, making it impossible to model the gradual changes in interactions that can be observed for real-world relationships.

To study group fragmentation in the context of coevolutionary opinion dynamics binary networks are sufficient since these models assume individuals to make the decision to break a connection to individuals with an opposing opinion. If we want to study how a disagreement can lead to other edges disappearing (ultimately leading to fragmentation) without individuals explicitly deciding to break edges, it is obvious that more gradual dynamics are needed. Therefore, we will use a weighted network to model associations between the individuals, meaning we assign a weight to each edge. The dynamics will change these edge weights. If an edge weight becomes zero we assume the edge to be broken. If an edge weight increases from zero to a non-zero value, a new edge has been made.

Above we have established the suitability of the agent-based modelling paradigm and the importance of using weighted network dynamics for studying non-strategic fragmentation. In the remainder of this chapter we provide an outline for how the thesis is organised.

1.3 Structure of the thesis

In Chapter 2 we discuss the background of the research presented in this thesis and introduce related literature. We establish that there is a gap in the literature for a model of weighted social network dynamics that is able to both grow and maintain a network using only localised, non-strategic behavioural rules.

In Chapter 3 we outline such a model, and discuss assumptions made and their relation to evidence about human behaviour. We describe constraints on the interactions between agents and how these constraints are parametrised in the model.

In Chapter 4 we study the topological features of the resulting networks and show that they exhibit features characteristic of real-world social networks. We show that the system reaches a quasi-stable state in which the defining topological features reach equilibrium while the edge weights, quantifying the strength of the relationships, continue to fluctuate. We discuss the influence of parameters and initial conditions on the resulting topology and identify regimes within the parameter space for which the generated networks resemble real social networks. We also explore how the system scales when increasing the number of agents.

In Chapter 5 we investigate the effects of the perturbation in terms of topological changes and edge weight changes and discuss the mechanism that results in the impact. Through this investigation, we demonstrate that our model can exhibit fragmentation as a result of a disagreement perturbation, and identify some of the key features of the network topology for this to occur.

In Chapter 6 we discuss shortcomings of existing models of network dynamics. We list limitations of our model and put it into context with empirical literature on social networks. We present suggestions for further work, including how our model could be integrated with other existing models to form a more mature model of social network dynamics. We discuss implications of our findings for real-world systems and how measures could be applied.

1.4 Key contributions

A social network model with ongoing micro-scale dynamics. In Chapter 3 we introduce a model of weighted network dynamics. This model fills the gap that we identified in Chapter 2, and puts us in a position to study the process of fragmentation as a consequence of small perturbations. This is a substantial contribution that enables the investigations in subsequent chapters: the model must balance the requirements of exhibiting some fluctuations in order to react to a perturbation, but it cannot be so fragile that it falls apart unperturbed. Our model is capable of growing and maintaining a network in a quasi-stable equilibrium (Bryden et al., 2011) without the need for global balancing processes.

Chapter 4 shows that our model does exhibit the dynamical behaviour required for studying non-strategic fragmentation: the topology of the generated networks reaches quasi-stable equilibrium, while on the micro-level edge weights exhibit ongoing dynamics. Furthermore, the resulting networks exhibit the topological features characteristic of social networks and we show that this is stable across a range of parameter settings.

Understanding mechanisms that result in fragmentation after small changes in interaction structure. In Chapter 5 we show that fragmentation can happen following a disagreement, even though almost all individuals continue to behave as before. We demonstrate that this fragmentation is a result of the perturbation and not due to noise through the use of unperturbed control runs. We investigate edge characteristics, and identify characteristics that indicate which of the edges must remain unperturbed in order for the network to stay connected. We examine the underlying mechanism by which a single perturbation can lead to breaking of several edges, ultimately leading to fragmentation.

Chapter 2

Background

2.1 Networks for complex systems

A network or graph describes components of a system and the relations or interactions between them. Thus, it simultaneously describes parts that comprise the system, called *vertices* and the structure of relationships between the components, called *edges*. Different notations exist in different fields. In computer science the elements are called nodes and the connections links, in the context of sociology they are referred to as actors and ties and yet another notation exists in the physics literature, sites and bonds (Newman, 2010). Despite these different names, the aim in studying networks is the same in all disciplines: formalising and understanding the structure of interactions between the elements that form a system. The study of complex systems is focused on these interactions between a system's elements and how they give rise to global-scale behaviour. Therefore, networks, together with agent-based models, cellular automata and differential equations models form the basis for describing and modelling complex systems (Holme et al., 2007).

The nodes and edges in a network can represent different entities, depending on the real-world system that is being modelled. The interactions between people can be modelled as a social network. We will describe social networks in detail in Section 2.2. Other examples include technological networks, ranging from the Internet as a network of routers, the telephone network and transportation networks, to delivery and distribution networks such as the power grid or water supply networks (Newman, 2010). In biology, systems commonly described as networks include neural networks, ecological networks or food webs (Solé and Montoya, 2001), biochemical networks such as metabolic networks (Jeong et al., 2000; Ravasz et al., 2002), protein-protein interaction networks and genetic regulatory networks (Newman, 2010). Documents that reference other documents can form a network of references such as the world-wide-web or a citation network (Boccaletti et al., 2006).

2.1.1 Research questions in network research

The questions asked by scientists about the behaviour of systems represented as networks are diverse and depend on the target system. Questions studied include "which flight connections should we offer for maximum profit?" for a flight connection network (Cancho and Solé, 2003; Gastner and Newman, 2006), "what is the effect of removing nodes of the network" for distribution networks such as the electricity grid (Crucitti et al., 2004a) or "how does the topology of a network influence the spread of an infectious disease?" (Keeling, 2005; Keeling and Eames, 2005; Newman, 2003b).

If we compare the real-world questions asked, it is obvious that many of them address similar problems and processes. One problem concerns which nodes and connections are most crucial for the system in fulfilling its function. In technological systems such as computer networks or distribution networks this is a very important problem as nodes and connections are subject to failure or targeted attack. Closely related to this is the question of robustness and resilience, concerned with how a system will cope with failures of components and connections, especially if one failure can trigger another leading to a failure cascade. This is especially important in supply networks, such as the power grid, but its relevance for social systems forms the basis of the research question in this thesis. Another frequently posed question is how to build or change a system to optimise certain properties such as cost, efficiency, robustness or all three, given certain constraints. In air traffic networks the tradeoff is between the length of a journey (how often does a passenger need to change planes, which should be as low as possible) and the number of edges present, which should be as low as possible as each edge means a flight connection that needs to be supplied which is costly (Cancho and Solé, 2003). Yet another line of research focuses on the role of networks in spreading processes such as the spread of information (Borgatti and Cross, 2003; Daley and Gani, 2000; Friedkin, 1982; Noble et al., 2004; Wu et al., 2004), opinions (Deffuant et al., 2000; Franks et al., 2008; Hegselmann and Krause, 2002), behaviours (Christakis and Fowler, 2007, 2008; Coviello et al., 2014; Rosenquist et al., 2010), infectious diseases (Barthélemy et al., 2005b; Brockmann and Helbing, 2013; Pastor-Satorras and Vespignani, 2001) or computer viruses (Chakrabarti et al., 2008).

The details of the system studied need to be taken into account for answering more specific questions especially if predictions are to be made. For spreading processes different constraints apply depending on what spreads on the network. Information can be transferred in the context of a phone call whereas most diseases require at least some proximity between the individuals. For modelling the transmission of a particular disease it is important to distinguish between diseases where infection and subsequent recovery lead to immunity and diseases for which reinfection is possible. However, certain network topologies are more likely to allow for spread (of disease or information) to reach a larger proportion of the network than others, which is why results from epidemic spread can be of interest to marketing executives wanting to spread product information. Frequently, such general findings hold for networks from different research areas. This realisation that some of the questions asked and processes studied with network models are similar in very different areas of research has led to the development of a shared multidisciplinary terminology and theory base, now generally called network science. Based on the well-established mathematical field of graph theory (Bollobás, 1998, 2001; Harary, 1994; West, 2000) and borrowing statistical measures from sociology (Wasserman and Faust, 1994), network science tries to distill out properties and behaviours common to networks in different fields. For a detailed introduction to network science see the book by Newman (2010) or an earlier review article by the same author (Newman, 2003a). Focusing on the developments in the area of network science are the reviews by Lewis (2011), Albert and Barabási (2002) and Dorogovtsev and J.F.F. (2002).

2.1.2 Complex networks

In the last 10 years, network research has undergone a transition to the study of "complex networks" (Barrat et al., 2008; Boccaletti et al., 2006; Cohen and Havlin, 2010; da Fontoura Costa et al., 2011). The focus has shifted from describing topologies and static properties to the study of dynamical processes on the network as well as the processes that shape the "evolution"[1] of networks. Using networks to describe the interaction structure of a dynamic system (Strogatz, 2001) has become a widely used approach in many areas of research that before had tended to assume random mixing for simplicity. Examples for this are the study of spreading processes through a population, synchronisation processes and games on networks (Jackson, 2005).

In addition, network models have started to become more complicated by taking into account more features of real-world networks. One example for this shift is considering weighted networks instead of binary networks. In a binary network, an edge is either present or absent. In most real-world systems, connections can differ in strength or capacity (Barrat et al., 2004b; Barthélemy et al., 2005a; Granovetter, 1973; Yook et al., 2001). In a social network we might like some people very much and others less so. In a distribution network, different edges might be able to accommodate different flows. To be able to represent this fact in a network, we can add a weight to an edge, denoting the strength of the connection. In social networks, adding weights to the edges allows distinguishing between friends and acquaintances, a distinction that is very important in social networks, since these two types of relationship play different roles (Fingerman, 2009; Granovetter, 1973, 1983). Furthermore, weights on edges allow for a more gradual change of edges compared to rewiring of links. Rewiring is used in many network models of network change but is arguably not a realistic abstraction of human behaviour.

[1]The term evolution in network science is used to describe the change of network topology over time and should not be confused with the biological meaning of the word.

Another step towards complex network models can be made by allowing heterogeneity of nodes. The heterogeneity might be structural, such as how many neighbours a certain individual has which in turn could be a result of heterogeneity in behaviour. For example, some people prefer to interact with only a small circle of close friends while others might prefer contact with a range of more casual contacts (Backstrom et al., 2011). In addition, other factors such as age or type of job can influence the number of contacts an individual has (Danon et al., 2012; Roberts et al., 2008) This heterogeneity has important implications for a range of systems. In sexually transmitted diseases (STDs) most nodes have few contacts and therefore only a small number of edges compared to other diseases such a influenza. However, a few individuals with many contacts can sustain a STD epidemic in the population (Jones and Handcock, 2003; Potterat et al., 2002). The effects of network topology on disease dynamics has been intensively studied for various diseases (Bansal et al., 2007; Bearman et al., 2004; Colizza et al., 2006; Jeger et al., 2007; Keeling, 2005; Moreno et al., 2002; Newman, 2003b; Pastor-Satorras and Vespignani, 2001; Salathé and Jones, 2010; Shirley and Rushton, 2005; Volz, 2008; Volz and Meyers, 2007; Wu and Liu, 2008).

2.1.3 Coevolutionary networks

Recently, the interplay of processes on the network and change of topology has been studied extensively and several review articles and books on this topic exist (Blasius and Gross, 2009; Gross and Blasius, 2008; Gross and Sayama, 2009; Mukherjee et al., 2013). In an adaptive or coevolutionary network there is mutual feedback between a dynamic process happening on the network and the network topology. Nodes interact with each other and these interactions are mediated by the topology. These encounters then lead to changes in topology, closing the circle. The study of social coevolutionary networks can be divided into three main areas depending on the process constrained by the network: opinion dynamics, games on networks and disease spread on networks.

In the context of epidemics on networks, susceptible individuals can rewire away from infected individuals, adding feedback from dynamics to topology which alters the disease spread patterns (Gross et al., 2006; Shaw and Schwartz, 2010; Van Segbroeck et al., 2010; Zanette and Risau-Gusmán, 2008). Related are models of adaptive networks using opinion or cultural dynamics as the process on the network (Castellano et al., 2009; Jalili, 2013). Some of these use binary opinions (Kacperski and Hołyst, 2000; Nardini et al., 2008; Vazquez et al., 2008; Zanette and Gil, 2006) while others use continuous opinions (Dittmer, 2001; Kozma and Barrat, 2008; Krause, 2000; Prettejohn and McDonnell, 2011). Rewiring can be selective (e.g. to someone with the same state) (Holme and Newman, 2006; Vazquez et al., 2008) or random (Kozma and Barrat, 2008; Nardini et al., 2008) or may take into account local rewiring using triangle closure (Iñiguez et al., 2009). The interplay between dynamics and network topology change has also been studied using games on networks where the outcome of games played with neighbours influences which edges are kept and which ones are dropped or rewired (Biely et al., 2007; Demirel et al., 2011; Ebel and Bornholdt, 2002; Pacheco et al., 2006; Poncela et al., 2009; Santos et al., 2006; Skyrms and Pemantle, 2000; Van Segbroeck et al., 2011, 2009; Zimmermann et al., 2004). While the areas of coevolutionary opinion dynamics and coevolutionary game theory models are related to the topic of this thesis and are important topics in their own right, we will only discuss a few relevant models here as a comprehensive review of these topics is beyond the scope of this thesis. We will discuss the differences in assumptions between our model and coevolutionary network and opinion dynamics models in detail in Chapter 6.

2.1.4 Networks and space

Another important feature of real-world networks is that almost all of them are embedded in physical space. In social networks, face-to-face meetings are very important for maintaining relationships between individuals and these meetings require the individuals to be in the same physical location. Another example of real-world systems where space is crucial is the optimisation of transport networks (Barrat et al., 2004a; Caschili and De Montis, 2013; Kaluza et al., 2010) where the balance of edge-count-distance versus spatial distance generates very different topologies (Gastner and Newman, 2006). If the per-edge cost dominates, the resulting networks often have a hub or star structure such as commonly found in air transport networks. If the actual spatial distance is the more important factor, then the resulting optimal networks are significantly more branched, as in the case of road networks. The simplest model of a spatially embedded network is a random geometric graph where nodes are distributed uniformly in space. Two nodes are connected if their distance is below a certain threshold (Dall and Christensen, 2002; Penrose, 2003). Compared to non-spatial random networks, these networks show higher clustering (meaning that it is likely that two neighbours of a node are themselves connected) and assortativity (positive correlation in degree for network neighbours) (Herrmann et al., 2003; Wong et al., 2006). Both of these are qualities found in many real-world networks such as social networks. If the assumption of homogeneous node distribution is relaxed, even higher levels of assortativity can be observed (Barnett et al., 2007; Bullock et al., 2010). However, network models that are only constrained by space produce networks that lack other important aspects of social networks such as communities (Barnett et al., 2007; Illenberger et al., 2012). We will encounter more examples of spatially embedded networks in the following sections.

2.2 Social networks

The interaction structures of people can be represented as a network. In such a *social network* the nodes represent individual people and the edges represent interactions, relations or sentiments. The term social network in a broader sense is used for networks such as Facebook or collaboration networks, which represent information about peoples' sentiments or interactions but do so on a higher level. For example, consider a social network generated by asking university students who they spend most of their time with. This network constitutes an abstract representation of the real student friendship network. The Facebook network between the same people will depend strongly on the friendship network, but give us a more time-aggregated, higher-level view. Many people will still be Facebook friends with people from their school class even though many of these contacts are no longer active or maintained. However, removal of edges is possible in Facebook in contrast to another type of network studied: collaboration networks amongst academics. Collaborations provide a fully time-aggregated mapping of an underlying collaboration network time series. Each collaboration resulting in a publication before the snapshot was taken will be represented by an edge in the network. It is important to make the distinction between snapshots of a network and aggregated snapshots (Holme et al., 2007). While time-aggregated social networks are worth studying to render insights into specific questions, it is important to note that some of the features (such as scale-free degree distribution) of these higher level social network abstractions do not generalise to non-aggregated friendship networks (Amaral et al., 2000). By contrast, snapshots also introduce biases since, by nature, they are also collected over a period of time, even if this period of time is short.

Data on the interaction structure of people in certain settings has been collected by social scientists for a long time (Wellman, 1926) and therefore an extensive body of case studies, techniques and theoretical work exists (Scott, 1988; Wasserman and Faust, 1994). Sociology is concerned with both describing the topology of networks and the implication of this topology for processes constrained by the network, as well as the implications of an individual's position in the network for that individual. Many of these measures have been integrated and developed further within the framework of network science and we will describe the most important ones for the work presented in this thesis later. In addition to the descriptive approach of characterising network topology, many theories exist about how individuals behave and make decisions that shape the social network they are embedded in.

2.2.1 Different questions generate different networks

At the heart of social network research in the social sciences is data collection to formulate new theories and validate or refute existing theories. Therefore, an extensive base of data exists. Data generally shows the interactions between individuals but the setting and type of these interactions span a large range. Even in the same setting or social group the resulting network depends on the questions asked to determine strength or presence of an edge (Borgatti and Halgin, 2011; Caldarelli et al., 2004). Examples for social networks range from small groups with fairly well-defined boundaries[2] such as members of a university sports club (Zachary, 1977), novitiates in a monastery (Sampson, 1968), the network of a terrorist cell (Krebs, 2002) or even the interactions between dolphins and killer whales (Foster et al., 2012; Lusseau et al., 2003) to the study of large groups such as a university (Kossinets and Watts, 2006), a large online gaming community (Grabowski and Kruszewska, 2007; Leskovec et al., 2010) or even a whole country (Blondel et al., 2008; Onnela et al., 2007a). The types of interactions considered are also diverse (Borgatti et al., 2009) and include communication and interaction networks (Contractor et al., 2012), friendship or acquaintance networks (Bernard et al., 1988; Fararo and Sunshine, 1964), advice networks, co-citation or co-starring networks (Watts, 1999a), kinship and marriage networks (Padgett and Ansell, 1993) as well as marking someone as a friend (or foe) in an online social network or game (Leskovec et al., 2010; Szell and Thurner, 2010).

A distinction can be made between *state-type ties* such as like/dislike, family ties or friendship ties and *event-type ties* (sending email, meeting someone, making a sale) (Borgatti and Halgin, 2011). A common problem for the collection of data on state-type ties is that individuals have different scales on which they rate interactions. Some people might consider everyone who they talk to regularly as a friend, while others would only see their closest contacts as belonging to their circle of friends. Individuals also tend to forget to list people they are acquainted with leading to errors in sampling (Wang et al., 2012). Therefore, even when focusing on the study of friendship networks, contact or communication networks are frequently used as a less biased alternative. Individuals tend to have very biased observations of who they interact with and the duration of these interactions (Bernard and Killworth, 1977; Bernard et al., 1984; Killworth and Bernard, 1976; Marsden, 1990). Therefore, recent studies on contact networks often aim to collect data directly instead of asking individuals for their perception of the events. Interaction data can be collected by asking participants to fill out contact diaries (Danon et al., 2012; Read et al., 2008), collecting or analysing existing online interaction data or by measuring contacts using wearable sensors (see Section 2.3.1). More and more data sets on online interactions have become available for study. This data can come from social networking services (Ahn et al., 2007; Buscarino et al., 2012; Grabowski, 2007; Panzarasa et al., 2009), massively multiplayer online role playing games (Grabowski and Kruszewska, 2007; Szell and Thurner, 2010), mobile phone call records (Onnela et al., 2007b; Wesolowski et al., 2012) or email and instant messenger logs (Cole et al.,

[2]See Laumann et al. (1989) for a discussion on why boundaries of a social system are often hard to define and why it is important to define them well and explicitly specify where the boundary of a certain group was drawn.

2005; Ebel et al., 2002b; Klimt and Yang, 2004; Leskovec and Horvitz, 2008). Contact networks are especially important for the study of disease spread. Depending on the type of disease, a contact can be anything from being in spatial proximity (such as being in the same room or bus), touching the same object (shopping trolley handles come to mind) to physical contact such as handshaking or as the case for many sexually transmitted diseases, sexual contact or needle sharing between drug users. In contrast, the spread of information might take place via many channels of which face-to-face conversation is just one. Others include phone calls, text messages, emails or letters, online chats and instant messages, forums or feeds such as on twitter or Facebook.

2.2.2 Measures of topological properties

In social sciences and biology *social network analysis (SNA)* denotes the data-driven study of real-world networks. SNA has developed many measures to characterise properties of networks which have since then become more universally used due to the multi-disciplinary efforts in network research. The measures used in SNA broadly fall into two categories: measures to describe characteristics of the network topology and measures to describe the importance or influence of an actor (Borgatti and Foster, 2003). These network measures are used to study the topology of a static snapshot of a social network. However, they are generally used to understand the constraints placed on dynamic processes happening on the network. Although using a set of numbers to describe a network is a very strong abstraction, the measures used by social scientists convey information about the potential dynamics of the system. Knowing several network measures allows us to make a rough inference about the dynamics. As an example, if we consider information spread on a social network, then knowing a network is sparsely connected tells us that information flow on this network will most likely be slow. If we know that a network consists of several disconnected parts, then we expect that a piece of information discovered in one part of the network will never reach the individuals in the other components. Therefore, even though the study of network dynamics should extend beyond pure analysis of static network snapshots, using statistical measures allows us to describe and characterise networks. We introduce measures as they are used in this thesis but for more detailed descriptions and formal definitions we refer the interested reader to Newman (2010).

2.2.3 The distinct features of social networks

Several properties have been recently discovered to be omnipresent in real-world networks, differing significantly from the simplifying assumption of a regular, lattice like topology used in early work on networks (Boccaletti et al., 2006). Most real-world networks (social networks in particular), exhibit a remarkably short average path length (Watts and Strogatz, 1998). The average shortest path length in social networks is much lower than the path length we would expect to see in a regular network, such as a lattice, but larger than the path lengths observed in random networks. The existence of shortcuts through the network has been illustrated by Milgram (1967) in his famous experiment in which letters were handed out to participants, telling them the name, city of residence and occupation of the intended receiver but nothing else. The task was to pass the letter to one of their friends or acquaintances which they thought would be closer to the target than themselves. Surprisingly, most letters reached the target after only 6 steps. A more recent version of this experiment is described in Dodds et al. (2003) using email instead of paper mail. The experiment confirmed the number 6 for the average path length (the paths were on average 5-7 hops long) using a large number of people. Another example of this "small-world" phenomenon is the Bacon number for film actors. In the film actor network, two actors are connected by an edge if they have co-starred in a film. The Bacon number of an actor is the number of hops from the actor to Kevin Bacon. The highest defined number found is surprisingly low with 8 hops (Watts, 1999a). The same effect can also be observed in collaboration networks (Newman, 2001).

We all experience the phenomenon described above every now and then when we find out that two of our acquaintances from different friendship circles know each other or share a common friend unknown to us. This is surprising, since social networks tend to be very sparse networks. This effect can be explained as a result of social networks both having high levels of clustering and low global separation (Watts, 1999b). The low global separation is a result of *bridges* connecting otherwise disconnected clusters consisting of internally densely connected nodes. This combination of densely connected areas with high clustering and connecting bridges or *weak ties* to other groups of well-connected nodes is called community structure. Community structure is another characteristic property found in (but not limited to) social networks (Blondel et al., 2008; Fortunato, 2010; Girvan and Newman, 2002; Newman, 2006). Communities result from heterogeneity in the population such as grouping by location, family ties, occupation or interests, age, gender, race, religion or even explicit group membership (Newman, 2003b). This assortativity, associating preferentially with individuals similar to oneself gives rise well-connected clusters of individuals sharing certain properties. Assortativity according to degree is another important characteristic of social networks (Newman and Park, 2003; Toivonen et al., 2007). Social networks display positive values for the assortativity coefficient, indicating that higher degree nodes tend to be directly connected more often than would be expected to occur by chance, whereas other types of networks have negative values indicating disassortative mixing (Newman, 2002).

Most social networks also display broad degree distributions, with some individuals having many connections and others only very few. The shape of this distribution has been the subject of an ongoing debate, with some authors arguing that social networks exhibit scale-free degree distributions (Barabási and Albert, 1999) while others argue that degree distributions are Poisson distributions (Bearman et al., 2004). However, a distinction should be made of the type and size of social network considered. In certain types of social networks such as actor networks or Facebook friend networks it is possible for the degree distribution to be scale-free and for the degree of nodes to reach very high maximum values since the contacts do not require maintenance. In the actor network, two actors can star in the same film without ever meeting. For Facebook networks, individuals in urban areas come into contact with many people they can connect to. However, they can still only maintain a certain number of those potential relationships. So even though someone might have met 2000 people and added them as friends on Facebook, he or she will not be able to keep up regular contact with all of them. In contrast, in networks where the edges require regular maintenance, such as in a friendship network (Bernard et al., 1988) the degree distribution tends to have a characteristic scale or display a sharp cutoff at higher numbers (Amaral et al., 2000).

To summarise the main identifying features of social networks are high levels of transitivity or clustering, community structure (densely connected areas with fewer connections between them), short average path length, assortative mixing according to degree and right-skewed, broad degree distributions (Boguñá et al., 2004; Hamill and Gilbert, 2010; Jackson and Rogers, 2007; Newman and Park, 2003; Toivonen et al., 2007; Wong et al., 2006).

2.2.4 Theories of behaviour

Many sociological theories about human behaviour in social settings have been proposed and empirically tested over the years. Here, we will give an overview of a small subset of theories important for the type of network that is the focus of this thesis. A more extended review can be found in various works on communication theories in an organisational context for example in Monge and Contractor (2001) on theories and their connection to processes and outcomes in organisational settings or in Contractor et al. (2006, 2012) on the use of these theories for statistical modelling of networks.

One theory based on the cognitive abilities of humans suggests that friendship networks generally do not exceed 150 acquaintances (where a name and other information can be associated with the person and contact is maintained to some degree) (Dunbar, 2008). This limit has been confirmed by empirical evidence (Hill and Dunbar, 2003). This limit has been taken into account by some companies by making sure that individual units stay within this limit. One example is GoreTex where no factory unit has more than 150 employees (Gladwell, 2000). As discussed in Section 2.2.3 degree distributions in certain types of social networks might have higher cutoffs depending on the type of network considered. However, when considering friendship networks of actively maintained relationships or organisational communication networks the number of friends is generally in the range of hundreds.

Another important theory for social networks is balance theory. Balance theory postulates that people are more likely to connect to friends of their friends. This leads to increased transitivity in the network by closing open triangles, a process commonly called *triangle closure*. This effect can be measured as it leads to larger clustering coefficients. In addition, studies looking at longitudinal network data have observed open triangles closing over time, thus providing support for balance theory (Hallinan and Hutchins, 1980). In signed networks (networks with positive and negative links) balance theory predicts subgroups to be stable if all cycles have a positive sign, thus the network can be split into groups with positive links within the group and negative links in between (Cartwright and Harary, 1956). Not many networks are explicitly signed. However, an example of a study of a signed network can be found in Leskovec et al. (2010) where the authors study signed online networks.

Another theory focusing on which connections are likely to be made and maintained in a network is the theory of homophily. According to this theory, individuals who are close in demographic space are more likely to interact with each other (Contractor et al., 2012). Gender has been shown to be particularly important in this context, for example in organisational networks (Brass, 1985), classrooms (Smith-Lovin and McPherson, 1993) or voluntary associations (Popielarz, 1999).

McPherson et al. (2001) distinguish between two dimensions of assortment: *induced homophily* and *choice homophily*. Induced homophily refers to the limited options individuals have to form ties to individuals with radically different demographic attributes due to the social context they are embedded in. As an example, consider the setting of a school. If most girls had only other girls as their friends, this could be due to induced homophily for example in cases where the school is an all-girls school. In this case the available pool of potential friends would be strongly biased towards making friends with girls. On the other hand, in a class with exactly the same number of boys and girls homophily could be the result of the explicit choices the children make and therefore considered to be a case of choice homophily. Kossinets and Watts (2009) investigate the contribution of both types of homophily on the observed total homophilous effect in a university email network and conclude that both play an important role in the system studied.

Many social interactions take place in the context of *activity foci*. This includes groups with explicit membership, such as religious groups, clubs or classes as well as friendship groups or workplaces. It has been argued that foci present one of the most important contributions to clustering in social networks (Feld, 1981). This effect has been shown to be particularly significant in student populations (van Duijn et al., 2003). Another important factor of induced homophily is spatial embedding, which limits the pool of potential friends to within a certain vicinity or reach (Wong et al., 2006). This theory is referred to as *proximity theory* (Contractor et al., 2012). This effect has been shown empirically by Backstrom et al. (2010) and Illenberger et al. (2012).

Often several of these aspects of homophily are considered as separate hypotheses in modelling social networks, for example in Contractor et al. (2012) where the influence of gender homophily, the effect of shared activity foci as well as spatial proximity are considered separately. Another example investigates homophily effects regarding individuals belonging to the same organisation type (government organisation or private company) (Contractor et al., 2006).

A particularly well known theory of social networks is the theory of weak ties (Granovetter, 1973). Granovetter observed that most job seekers discovered new job opportunities through contacts they classified as acquaintances, not close friends. He proposed that this stems from the fact that groups of close friends are more transitive and clustered, while acquaintances provide bridges between these clusters. The high clustering of strong edges means that it is likely that you will have the same information as your friends. Therefore, information within clusters is redundant, so that your friends are unlikely to have new information for you. Acquaintances share fewer friends with us, thus they provide a bridge to the information their friends have, which will most likely be novel to us. The existence of these bridges and the fact that they are generally weak ties have been empirically confirmed in studies by Onnela et al. (2007a); Toivonen et al. (2007), analysing a mobile phone call network. They find that weak ties indeed do not share many neighbouring ties, but have high betweenness centrality, meaning they feature in the shortest paths between many pairs of nodes. They also observe that removing random weak links leads to faster fragmentation of the network than removal of strong edges, suggesting that components are indeed frequently connected by weak ties. Another study highlighting the importance of weak ties is by Dodds et al. (2003). In their modern version of the Milgram experiment using email instead of letters the authors found that successful chains often went through medium or weak edges, thus confirming the theory.

Describing the same phenomenon but taking a different angle is the theory of structural holes (Burt, 1992). Burt proposes that individuals have a potential advantage depending on to their position in the network. Individuals acting as bridges have access to information from two clusters that otherwise do not exchange information. This means they span a structural hole in the network. In cases where new information is valuable, spanning structural holes can be advantageous for individuals.

The two theories have in common that they both emphasise the importance of bridges or non-redundant ties. They are important, because they are more likely to give an individuals access to novel information. For more detail on these theories and their different angles, we refer the reader to Borgatti and Halgin (2011). Both original theories focus on the benefits for individuals in the network. Considering the network as a whole, weak ties often act as bridges between densely connected clusters. Therefore, they are important for the structural cohesion of the network as a whole as they provide the only channel for information flow or communication between groups.

2.3 Network dynamics

Most of network research until recently has focused on static snapshots of networks. Even though many have argued for the study of network dynamics (see for example Bansal et al. (2010); Doreian and Stokman (1996, 2003); Palla et al. (2007); Snijders et al. (2010)) a large proportion of research on social networks has considered static network snapshots. This might be due to limited availability of time-resolved data or the extension of models to dynamics models not being trivial. In addition, statistical measures are only defined for static networks and some authors suggest that, at least for growing networks, the properties measured (such as diameter (denoting the maximum distance in the graph), clustering coefficient and average degree) change significantly (Barabási et al., 2002; van Duijn et al., 2003), thus requiring redefinition or extension of measures.

In this section we will discuss the recent availability of time-resolved data and then present statistical and simulation models considering the dynamics of networks, demonstrating that dynamic models are now feasible.

2.3.1 Data on dynamics of social networks

In the social sciences the importance of collecting longitudinal data has been recognised and attempts to collect longitudinal data to foster understanding of mechanisms of network change have been made, but these are of course subject to limitation of manual data collection. Examples of longitudinal data collected by hand include a study by Morgan et al. (1996) on the social networks of recent widows studied over the course of a year, a study by Sampson (1968) on the interactions of young men in a monastery documented at different timesteps, the study of a changing social network of members of a university Karate club by Zachary (1977), the change in topology of a drug-user needle sharing network (Rothenberg et al., 1998) and the social network of sociology freshmen at the university of Groeningen (van Duijn et al., 2003). It is important to note that two of these studies have reported fragmentation of the observed small groups (Sampson, 1968; Zachary, 1977), indicating that this is not a rare phenomenon. Limitations of these studies are the small number of snapshots for each network and that they often focus on ego-centered sub-parts of larger networks and therefore do not allow propositions about dynamics of the whole network.

Digital data storage and the web have led to vast quantities of data being recorded, stored and made available. Examples for online interactions between individuals are blog comments, Facebook messages or wall posts, diggs, upvotes or likes and tweets or retweets (De Choudhury et al., 2010). Popular data sources are social networks such as Facebook, where users can mark other users as friends and send messages (Holme et al., 2004; Leskovec et al., 2008; Panzarasa et al., 2009). Another important digital source of data on human communication are email networks, where it is recorded who sends emails to whom and with what frequency. Recording the frequency also makes it possible to assign weights to the edges and therefore to distinguish between strong and weak or frequent and infrequent interactions (Kossinets et al., 2008; Kossinets and Watts, 2006). The inclusion of weights is particularly important if we are interested in potential routes of information flow as a longer path of very frequent contacts could lead to faster information transmission than a short path going through infrequent contacts.

Another source of interaction data is data on mobile phone calls (Blondel et al., 2008; Onnela et al., 2007b). These data sets are particularly interesting as some also provide approximate locations of the callers and therefore allow studying correlations between phone calls and proximity of the callers. One study found that people often call each other just before they meet face-to-face for coordination and therefore mobile phone call networks are correlated more strongly with face-to-face meetings than one might expect (Calabrese et al., 2011).

Massively multiplayer online games build a virtual world in which players can interact with each other. In addition, many games allow sending messages to other players and forming allegiances with them. Data on individuals interacting in these games has been collected (Leskovec et al., 2010; Szell and Thurner, 2010) and games have even been designed explicitly to study the social behaviour (Spraragen et al., 2013).

To measure face-to-face contacts (which is especially important for studying the spread of contagious diseases) several novel data collection methods have been proposed. One of them is measuring proximity via bluetooth equipped phones (Eagle and Pentland, 2006; Eagle et al., 2009). A different approach using dedicated sensors based on RFID technology has been deployed mainly in conference settings (Cattuto et al., 2010; Chaintreau et al., 2007; Isella et al., 2011). Use of a different type of dedicated sensor is described in (Fischbach et al., 2009). In this study, in addition to the sensors, email communications were also recorded for the same period. This allowed for comparison between face-to-face communication and email communication which were found to show different structures as people that were spatially close only very rarely emailed each other. This is an important fact to take into consideration when inferring communication networks from a single source such as email contact data. Further studies are needed to investigate the relationship and interactions of communication networks over different media (such as face-to-face meetings, email, calls).

2.3.2 Statistical modelling of social network dynamics

Available data can be used to build statistical models. Statistical models conceive of each social tie in a network as a random variable and try to assess the relative importance of various processes working on node attributes and creating characteristic local structures (such as triangles or reciprocal dyads in directed networks) in creating the global topology observed. The stochasticity of these models is sometimes argued to be necessary because of the stochasticity of the system itself, or it could reflect lack of knowledge of processes that are fundamentally deterministic. By exploring how parameters must be tuned in order to generate (maximum likelihood) networks that agree with empirical properties of interest, these models can answer questions regarding the relative contributions made by different social processes to key features of network structure. These models are called *ERGMs* (exponential random graph models) or p^* models. We refer the interested reader to the introductions found in Robins et al. (2007a); Snijders (1996) and examples in Contractor et al. (2006); Pattison and Robins (2004); Robins et al. (2007b).

These models have been extended to models of longitudinal data using discrete (Hanneke and Xing, 2010; Robins and Pattison, 2001) or continuous time (Snijders et al., 2010). Apart from modelling network dynamics, the continuous models are suitable for modelling out-of-equilibrium networks (Snijders et al., 2010). They use an actor-orientated approach, assuming that the network changes due to the choices of individual actors. As these models use continuous-time Markov chains they can no longer be solved analytically and therefore have to be solved by simulation. An implementation of this framework named *SIENA* is presented by Snijders (Snijders, 1996, 2002; Snijders and van Duijn, 1997). Example applications can be found in Contractor et al. (2012) where longitudinal data on communication within an organisation is modelled and van Duijn et al. (2003) where the framework is applied to data on university freshmen friendships. More recent models are also capable of considering coevolutionary dynamics of network change and behaviour change of the actors (Burk et al., 2007; Steglich et al., 2006).

Statistical models provide a powerful way of modelling network dynamics. However, they are not yet able to consider weights on edges and are limited to fairly small networks of no more than 1000 nodes (Burk et al., 2007; Snijders et al., 2010).

2.3.3 Simulation modelling of social network dynamics

Although statistical models of the type described in the last section and agent-based models are both types of simulation that take an actor based view on the system, the two approaches are somewhat different (Snijders et al., 2010). While the former (which we will refer to as statistical actor models) tend to employ more sophisticated statistical analysis, the latter (which we will refer to as simulation models) concentrate on explicit representations of social interactions, allowing for weighted dynamics and heterogeneity of nodes.

Modellers using statistical models have argued that the two modelling methods complement each other as simulation allows for hypothesis generation and conducting thought experiments whereas statistical models are concerned with explaining observed empirical data and making predictions about which processes have generated the observed topology. Authors have argued that simulation models can fill the macro-micro gap between descriptive statistical models on the macro-level (which do make predictions about network processes but only based on observation of their resulting topological substructures) and the sociological theories about individual behaviour on the micro level of actors or dyads and triads (Robins et al., 2007a). Simulation is able to bridge this gap by relating micro-level behaviour proposed by sociological theory and validated by empirical research to observed macro-level network topology or network evolution.[3] In addition to linking certain micro-level conditions to macro-level outcomes (Raub et al., 2011), simulation models allow for the direct observation of the processes driving the network evolution and therefore can provide insights into the modelled phenomena (Alam and Geller, 2012; Jackson, 2005).

In the following sections, we will give an overview on existing simulation models of network evolution. Some of the models discussed here are compared to each other (and ERGMs) in Toivonen et al. (2009). This review is however focused on the physics and statistics community of network science and we will add to it by discussing agent-based and coevolutionary models of network change.

2.3.3.1 Growing network evolution models

One class of models for modelling dynamic change of social systems is titled *growing network evolution models* or *growing NEMs* by Toivonen et al. (2009) as they model network growth. This means nodes are added sequentially until a certain number of nodes has been reached. This type of network evolution model is a good abstraction for growing systems such as the Internet or collaboration networks. However, in many other real world systems, such as most social settings, the continued making and breaking of edges after a short initial growth phase is what dominates the dynamics of the topology. For such cases these network formation models are less appropriate models. Many NEMs are explicitly not aimed at modelling network evolution but instead at efficiently generating a static network with certain properties that will be used for some other modelling context (Toivonen et al., 2006).

One example is presented in Holme and Kim (2002) where the traditional preferential attachment mechanism (Barabási and Albert, 1999) is extended by adding a triad forming mechanism to increase clustering in the resulting network. The first edge for each node added is made according to preferential attachment. The following edges connecting the new node to the existing network are made either by preferential attachment

[3]Stokman and Doreian (1996) argue that *network dynamics* should denote the description of network change (quantitative or qualitative) while *network evolution* should be reserved to contexts where we also have a theoretical understanding of the processes generating the changes in addition to the description of those, which is the aim of most simulation models.

(with probability p) or connect the new node to neighbours of the node selected in the first step (with probability $1 - p$).

Another model extending the preferential attachment model creates a nested community structure (Park et al., 2006). The network is grown by preferential attachment for a number of steps. Then the two vertices with highest degree are picked to create a group. Each of the remaining nodes gets assigned to one of these groups depending on which of the two selected nodes it is closer to. Now the groups grow in size by preferential attachment within the group. Depending on the the nestedness desired in the final network, the splitting of groups can be repeated to split the groups into smaller subgroups before continuing the preferential attachment.

Other models are based on random attachment with subsequent linking to neighbours to increase clustering. Vázquez (2003) add a new node to the network with probability $1 - p$ and connect it to a random node j. They then add potential edges from the new node to all of j's neighbours. These edges are not made, they are just listed as possible edges that can be created at a later stage. With probability p a random edge from the list of potential edges is created. A similar model can be found in Toivonen et al. (2006). Here each new node is connected to either one (probability p) or two (probability $1 - p$) initial contacts. For the first initial contact j, a number m_{sec} $U[0, k]$ is drawn which denotes how many edges to j's neighbours the new node will attempt to make. The new node will then connect to m_{sec} randomly chosen neighbours of j (if they exist). If a second initial contact was chosen in the first step, the same procedure is repeated for the second contact. An extension of this is presented in Ivanova and Iordanov (2012) using two populations of nodes to represent violent and non-violent nodes in a model of terrorist networks. This model allows for different edge probabilities for nodes of the same group compared to edges between one violent and one non-violent node.

To model the growth of ego-centered Facebook friendship networks, Buscarino et al. (2012) attach new nodes to randomly chosen nodes with probability p. With probability $1 - p$ the new node selects neighbours randomly from a subset, corresponding to a community. To determine the subset, communities in the network are calculated using a community detection algorithm and one of the resulting communities is chosen to act as a subset for the new node to link into. While this allows us to add community structure to the generated network, the mechanism obviously does not to attempt to model the process by which social networks form and organise into communities.

2.3.3.2 Spatial network models

As discussed in Section 2.1.4 many systems that can be described as networks involve spatial arrangement of their elements. However, many models of social networks ignore the location of nodes for simplicity's sake. While placing the nodes in space adds complexity to models, it has also been shown that important features of social networks come “for free” if we embed social networks in space and place some restriction on interactions depending on the distance between nodes. This is a valid assumption to make according to proximity theory, which has proved that in most cases individuals located in close proximity have a higher probability of interacting than individuals further apart. We will refer to models that take spatial arrangement into account when forming links as *spatial network models*. In Toivonen et al. (2009) these are denoted *nodal attribute models* or *NAMs*.

Several spatial extensions of the traditional preferential attachment model exist (Barrat et al., 2005; Barthélemy, 2003; Manna and Sen, 2002; Xulvi-Brunet and Sokolov, 2002). Most models make the simplifying assumption of homogeneously distributed nodes. This is an unrealistic assumption in many cases, one example being the Internet where nodes typically are clustered in space due to denser population in urban areas. Models concerning the growth of the Internet take into account these biases by either placing new nodes preferentially close to existing nodes (Xulvi-Brunet and Sokolov, 2007) or by using population density data for the calculation of location probabilities (Yook et al., 2002). Spatial embedding of the nodes has also been added as a factor in utility-based models of social network formation (Illenberger et al., 2012; Johnson and Gilles, 2000).

As an extension to random geometric graphs, which have a distance cutoff for interactions, Boguñá et al. (2004) distribute N nodes uniformly in 1D space and then link the nodes with probability $p = \frac{1}{1+(d/b)^\alpha}$ (where d is distance between the nodes, while b defines a characteristic length scale that is used to control the average degree and α represent degree of homophily). They also discuss extensions of their model to higher-dimensional space. Another model with probabilistic, distance based edges is presented in Wong et al. (2006). Here nodes are distributed homogeneously in space. Two nodes are connected subject to the following rules: if their distance d is smaller or equal than a certain threshold H ($d \leq H$), connect the nodes with prob $p + p_b$. If the distance between the nodes exceeds the threshold H, connect them with a different probability, $p - \Delta$ (where Δ is chosen so that a specific total fraction of possible links is created).

A model with finite distance cutoff is presented in Hamill and Gilbert (2009). Nodes are randomly distributed in space and connected if they are less then a certain distance apart. This is implemented as follows. Every node has a *social reach* r defining a circle around the node of radius r. Individuals connect if and only if their social reaches overlap at some point. Extending from this random geometric graph model, the authors investigate the effect of heterogeneity in distance threshold between different individuals and find that the addition of heterogeneity in r adds a fat tail to the degree distribution. A similar model is presented in Antonioni et al. (2014, 2013) where, in addition to the interaction threshold, there are costs associated with forming an edge, depending on the length of the edge.

In most of the models, the concept of space is not limited to geometric space but could be extended to the notion of *social space*. Social space includes factors and dimensions that lead to some actors being closer to each other than others. Possible dimensions in social space include geometric arrangement but extend to potentially include demographic attributes such as age or gender, interests and opinions or shared affiliations (Pattison and Robins, 2004).

While the models presented in this section are important models for social networks, they do not include network evolution. A first step in this direction is made in Hamill and Gilbert (2010) where agents are allowed to move in space which leads to changes in the network structure due to altered distances.

2.3.3.3 Dynamic network evolution models

Another model where movement of the agents in space drives ongoing network change is presented in González et al. (2006). In the model presented in the paper, nodes are deleted and added to maintain a certain edge density in the network. The nodes or agents have a radius of size r and are not fixed in space but move around. A collision between two agents leads to the creation of a link, after which the agents change direction. The space in this model is not meant to represent physical space but rather social space. While this is a reasonable abstraction to make, the assumption that individuals change direction in the journey of the social space after a collision without any correlation between the two individuals now connected by an edge is not explained by the authors or linked to any sociological theory or real-world observation. The model is able to match the levels of community structure and assortative mixing as well as average shortest path length and clustering coefficient of a real-world school friendship network.

Movement of agents in space is one of the mechanisms driving network change but not the only one. A simple model for the dynamics of a social network is described in Davidsen et al. (2002) and Ebel et al. (2002a). The model works as follows. A node i is selected at random and if it has less than two neighbours it is connected to another randomly chosen node. Otherwise, two neighbours of i are picked and connected, closing an open triangle. With probability p a randomly chosen node is removed from the network, together with all its links and a new node is added to the network. This removal probability p controls the clustering and density of the network. Here, the resulting networks exhibit in a power-law degree distribution and high clustering. Since most real-world social networks do not show scale-free degree distributions, this model is more appropriate for simulating the dynamics of online networks where unmaintained links can persist.

A similar model is presented in Marsili et al. (2004). Here a randomly selected node i connects to another randomly selected node with probability η. With probability ξ, it chooses the new contact from the set of its neighbours' neighbours (again leading to triangle closure). In this model, however, nodes are not removed to keep the network from fully connecting as in previous models. Instead, with probability λ for each timestep, a random edge is deleted.

The models described so far all assume binary relations, meaning an edge can be either present or absent. This is a common abstraction to make, but if we want to model the dynamics in continuous time we need to allow for gradual changes of edges, which requires weights to be assigned to the edges. Kumpula et al. (2007) and Kumpula et al. (2009) present a model that simulates the dynamics of a weighted network. Each step a randomly selected node i connects to random nodes or friends of friends, similar to the models described previously. The node selects a friend of a friend k by weighted search along the present edges. Node i is therefore more likely to meet a friend of a good friend than a friend of a casual acquaintance, which is a realistic assumption to make. Once a suitable contact k is found, i connects to k with probability p (with initial weight w_0) if the two individuals are not already connected. If they are connected, the edge weight w_{ik} is increased by δ. In addition, the edges visited during the search also receive a weight increase of δ. Independently, i connects to a randomly selected node with probability p_r. As in models discussed previously, with probability p_r a node is removed with all its edges and a new node added to the network.

This models allows for gradually changing edges and models the mutual feedback of weights and dynamics. This is very important as we have to remember that the social network is just a visualisation of a very complex social system. In the social interaction system, the actors have a memory of past interactions (which are represented as the social network) which influences their current actions. For models using binary edges, this memory can distinguish between two states, denoting whether the other individual has been encountered before or not. Adding weights to the edges allows for a much finer distinction between frequent and infrequent contacts and opens the possibility to studying phenomena like the emergence of weak ties as bridges (Granovetter, 1973). In Kumpula's model and a subsequent extension by Jo et al. (2011) (to model burstiness observed in human interaction patterns) weights represent a history of past meetings as they are increased with every meeting.

Even though the models reviewed here have taken the first steps to modelling the dynamics of social networks (with the models by Kumpula even moving towards gradually changing, weighted edges) they require a constant in- and outflow of nodes or random deletion of edges to reach a dynamic equilibrium. While random removal of edges is a good way for an abstract model to keep the network from filling up with edges, there is no sociological process that could be abstracted as random edge removal. A throughput of individuals can be a realistic assumption in many contexts if the system is studied at the right time scale. In friendship networks individuals move away and new individuals are added. In an organisational context, people get hired and people leave for new jobs or are fired. However, when studying network fragmentation, the timescales are often much shorter. It is therefore safe to assume that the number and identity of the actors stays the same over the course of the simulation, making the models presented unsuitable as a base model for the investigating the topic of this thesis. We will now review some models that do not require replacing nodes.

One such model, where the number and identity of nodes is fixed and edges are weighted is described in Jin et al. (2001). In this model meetings between individuals happen with a certain probability p_{ij} for the pair i and j. This probability increases with the number of mutual friends and decreases with the number of friends the two individuals considered already have. This could relate to the real-world fact that if you have many friends you are probably less likely to make many new friends, due to limitations on time available for maintaining social connections. When two individuals meet, their edge weight is set to 1 and is then subject to decay, leading to a weighted network. We find this way of simulating edge weight dynamics counterintuitive, as we would expect a connection to get progressively stronger with subsequent meetings. In the model by Jin et al. (2001) no matter how frequently two individuals meet, the weight of the edge between them will be 1. Furthermore, there is no positive feedback of individuals prioritising strong ties over weaker ones when allocating time. Therefore, while able to generate realistic-looking social networks, this model is not suitable for studying the weighted dynamics of social networks.

2.3.4 Groups and network evolution

Another view on network dynamics views groups as crucial for the evolution of social networks. This is based on theories of homophily and interaction foci. In models of group dynamics, groups and community structure emerge in the social network as a result of the individuals' interactions and fragmentation might be observed.

One model is presented in Grönlund and Holme (2004). The theory behind this model differs from other models of group evolution in that the authors assume that people want to be different from others, but not too different. In this model the least central node i of three randomly selected vertices is picked. Then another random vertex is selected and its edge rewired to i and its neighbours. The model produces networks with clear community structure; however, the mechanism does not mimic human interactions. Another model explains the emergence of groups as a result of biasing communication towards local communication (Rosvall and Sneppen, 2009). Here individuals communicate with other individuals and build a map of the network based on a memory of these encounters.

Other models model groups explicitly and study the coevolution of groups and network topology. In Geard and Bullock (2008) a social network with a fixed number of N nodes and M edges is used. Each node has an associated trait vector, modelling individual heterogeneity. Each iteration, G individuals are selected to each start a new group. Every group has a dimension of the trait vector associated to it. One setting where this might be a realistic assumption are university clubs, where each club is centered around an activity or opinion. Each of the group starters then invites neighbours who are not yet associated with a group to join their group. Each invited individual now compares the value at the relevant position of their trait vector with the group's value and joins the group with a probability proportional to the distance. Thus, the more closely aligned an individuals' relevant attribute is to the groups attribute, the more likely it is for the individual to join the group. If the individual chooses to join the group, it itself can recruit new members from its friends by the same process. Once every group member has tried to recruit all of their neighbours, each individual that is in a group will adjust their connections by rewiring an edge from their most dissimilar friend to the most similar group-comember (that they are not already connected to). Now a new round of group formation starts by dissolving all groups and selecting G new individuals to start new groups. The authors find that group growth is in some cases self-limiting by generating very pronounced community structure, limiting the flow of information (invitations) which, in extreme cases, can lead to fragmentation of the network into disconnected parts.

As in reality groups do not dissolve periodically, the authors develop the model to allow for more persistent groups (Geard and Bullock, 2009). Individuals in this model have a single attribute assigned to them which is a number between 0 and 1. Individuals create new groups with probability g and, contrary to the model discussed before, they can also leave groups. The probability for leaving a group increases with the number of friends not in this group (with a particular weight on friends who used to be members but are not any more) and also with the number of other groups the individual is associated with (imposing time constraints). The probability of leaving decreases for every friend who is currently in the group. The probabilities of joining a group after being invited by a friend depend on the same factors but with reversed effects. These factors are in line with sociological theory and empirical work (McPherson et al., 1992). The network evolves by rewiring randomly selected edges. An edge is either rewired to a group co-member of one of the groups that an individual belongs to or to a randomly selected individual from the whole population. If the distance in attributes between the two new neighbours is larger than between the two old neighbours, the connection might be rejected with a certain probability.

A model similar to this, but with slightly less sophisticated joining and leaving probabilities, is presented in Geard and Bullock (2010). It studies the influence of time constraints on the resulting network, by assigning each individual a fixed capacity for group investment (TAE). If an individual is invited to join a group and joining that group would exceed the its TAE, it either refuses to join or it has to drop a current association to join the new group. Groups more distant to the individual in social space are dropped with a higher probability. Note that the models discussed here again only allow for binary edges. Therefore, while they provide a good abstract model for group dynamics, the mechanisms driving topology change are rather simple. The models highlight how the interplay/positive feedback between choices (choice homophily) and social environment (induced homophily) can generate high levels of homophily and pronounced community structure. When individuals have larger time budget available, community structure is less visible and an increase in group size and lifetime is observed. In contrast, for low values of the group investment TAE, the networks are more prone to fragmentation. In the case where membership in only one group is allowed, this model is very similar to opinion dynamics models, which we will discuss next.

2.3.5 Communities and fragmentation as a result of opinion dynamics

In opinion dynamics models group formation and fragmentation is frequently observed. We will discuss some models of opinion dynamics here. However, we want to stress that while the mechanisms and individual behaviours implemented in models of opinion dynamics provide an explanation for network fragmentation, they are limited to cases where the individuals explicitly break edges. In contrast to this is the question that is the central point of this thesis, which is how edges can erode away even though the individuals involved do not explicitly make a decision to break or cease to maintain them. Following from the previous section we will first discuss a model where groups emerge as a result of opinion dynamics (Bryden et al., 2011). In the model presented, four different processes operate on the network. Two are state change processes: state spread, where one node adopts the state of a neighbour and innovation where a random node is assigned a novel state that is not present in the network yet. The two other processes are rewiring processes, either random or homophilous, to connect two individuals with the same state. Depending on the relative rates of these processes, three main regimes can be observed: random networks, networks displaying community structure, and fractured networks, in the case when rewiring dominates the dynamics. This model is similar to models studying the co-evolution of opinions and network topology. One model of opinion dynamics coupled with homophilous rewiring is presented in Holme and Newman (2006). In this model a randomly selected node either rewires one of its edges to a node with the same opinion (probability ϕ) or adopts a neighbour's opinion with probability $1 - \phi$. The authors show that for a critical value ϕ_c the system transitions from all opinions persisting in disconnected clusters to a giant component with one dominating opinion. Durrett et al. (2012) study a variant of this model where rewiring is random instead of homophilous. They find that this change leads to a continuous transition around the critical value, thus, counterintuitively, leading to a less rapid fragmentation into same-opinion components compared to the original model.

A model considering homophily as well as heterophily (tendency to connect with individuals dissimilar to oneself) is studied in Kimura and Hayakawa (2008), extending the model by Holme and Newman (2006). In this model, nodes can also rewire to a randomly chosen node with a different opinion (heterophilous rewiring). This leads to bridges between clusters of same opinion individuals being maintained. This model shows how certain tendencies can counteract the fragmentation of the network into opinion-homogeneous components. A model explicitly studying the fragmentation of the network based on differing opinions is discussed in Henry et al. (2011). Opinions are static in this model and edges are randomly rewired with a probability proportional to the neighbours' opinion difference and an aversion bias p quantifying the level of tolerance the agents have towards differing opinions.

These models describe the emergence of groups through opinion dynamics as well as fragmentation as a result of explicit disagreement. However, the processes are very abstract and gradual changes can not be modelled due to edge dynamics being simulated by rewiring and not gradual edge weight changes.

2.4 Disruptive change

We will now discuss models of cascades and fragmentation as well as factors that make networks vulnerable to fragmentation. The stability of networks is a very important topic in many areas of network research. One area investigating network stability is the study of network percolation (Callaway et al., 2000). Edges or nodes are removed until the giant component of the network disappears. This is relevant for distribution networks where we want the network to remain connected and functioning even when edges or nodes fail. For epidemics, fragmentation is desired, so we want to know how many nodes need to be vaccinated to stop the spread of an epidemic by fragmenting the network into components between which transmission can not occur.

Node or edge removal in these models can be random or targeted. Random removal corresponds to random technical failures while targeted removal could represent targeted attack by terrorists or hackers. Targeted attacks are simulated by removing high degree vertices, high betweenness vertices or vertices identified as important by some weighted measure such as strength for weighted networks (DallAsta et al., 2006). Different topologies react differently to random and targeted attacks. Scale-free networks can tolerate a higher number of randomly chosen nodes being removed than random graphs but are more vulnerable to targeted removal of high degree nodes (Albert et al., 2000; Cohen et al., 2000, 2001; Crucitti et al., 2003, 2004b).

Percolation theory studies independent failures, but one failing part could also start a cascade of failures (Crucitti et al., 2004b; Motter and Lai, 2002). This has been observed frequently in electric power grids. Here a failure of a line leads to higher load on neighbouring lines, potentially leading to failure of those if they were already close to being overloaded (Crucitti et al., 2004a). These cascading failures are not constrained to a single network but can cross to another co-dependent network. One example of this is the blackout in Italy in 2003 where a power network and the Internet failed in a series of cross-network cascades as the Internet dependent on the power for functioning and the power network was regulated via the internet (Buldyrev et al., 2010; Vespignani, 2010). Such interdependent networks show a breakdown behaviour very different to failures in a single network, displaying a faster decay and sharp transition to fragmentation. Both cases have been studied extensively. Another interesting study has been presented in Holme (2002) where the authors show that growth by preferential attachment can potentially lead to the network fragmenting itself as it grows if edges are sensitive to overloading, so that the network itself triggers the cascade.

The models reviewed here do not extend to fragmentation of social networks. The processes at work here are very different from overload of edges. It has been shown that cascades in technological networks propagate further in networks with little community structure (Wu et al., 2006). In contrast, in social networks community structure leads to a higher risk of triggering fragmenting of the network as we will discuss later. As described before, we need to distinguish between two cases of fragmentation dynamics in social networks. In the first, fragmentation is due to individuals making an explicit decision to take sides or disconnect from certain parts of the network. This case is modelled very well by models of opinion dynamics as described at the end of Section 2.3.3.3. The other case, which we will focus on here, can not be modelled by opinion dynamics. In this case a single edge failure due to a disagreement leads to a cascade of rearrangements in the network. In contrast to cascading failures in technological networks this is not a result of other edges failing through overloading. Instead the same social processes that coordinated meetings between individuals before now lead to fragmentation of the network. Individuals' behaviour rules remain the same and individuals do not take sides with one of the individuals involved in the fall-out. However, the network that individuals base their decisions on has changed and this leads to further changes.

Certain topologies are more prone to fragmentation than others. As discussed before, social networks often exhibit community structure with densely connected clusters connected by a few, often weak, edges. Percolation studies of weighted social networks have shown that the removal of weak edges leads to faster fragmentation, thus confirming that components are often connected by weak bridges (Onnela et al., 2007b; Toivonen et al., 2007). The structuring of a network into communities is obvious for many social networks, such as friendship networks of students (Eagle et al., 2009; Zachary, 1977), mobile phone calls networks (Blondel et al., 2008), networks of drug-users (Weeks et al., 2002) and even dolphin social networks (Lusseau et al., 2003).[4]

[4] A picture of this network can be found in Fortunato (2010).

The observed community structure can often be traced back to some external constraints. In the mobile phone call network the data is from Belgium and the observed communities correlate with the French and Dutch language communities (Blondel et al., 2008). In the social network of a Karate club observed by Zachary (1977) (which has become a benchmark for community detection since a split was actually observed later; Fortunato (2010)) the club members are grouped by loyalty to the trainer or the president of the club. Another example is presented in Lazer et al. (2009) in a citation network between political blogs where each of the two very distinct clusters corresponds to an American political party. Clustering can also be due to spatial location of the nodes as shown in an email network studied by Huberman and Adamic (2004).

Networks can either fragment or become more densely connected. Both cases have been observed by Potterat et al. (1999). Fragmentation has been described by Zachary (1977), where the university club split into two clusters according to the community structure[5]. Another example of fragmentation of small groups is presented in the study by Sampson (1968) who observed the interactions of young men in a monastery over a period of time during which many members left the monastery, either resigning or being expelled.

In cases such as the university club, the monastery, or in settings such as companies, fragmentation can lead to splitting of the organisation or even complete dissolution. In these cases, we are interested in ways to make networks more resistant to fragmentation. There are other cases, though, where we are interested in how to aid fragmentation of networks such as terrorist networks or disease transmission networks (Bearman et al., 2004; Weeks et al., 2002).

Factors increasing group stability have been studied in the setting of groups with explicit membership (Geard and Bullock, 2008, 2010) and theoretically in sociology (White and Harary, 2001). On the other hand fragmentation without explicit groups but as a result of differing opinions has been studied in detail (Durrett et al., 2012; Henry et al., 2011). Fragmentation of small groups without explicit group memberships where the fragmentation is not a result of explicitly taking sides has been mentioned as an important problem in the literature (Borgatti and Halgin, 2011), however there is a lack of both data and explanatory models. In this thesis we will present a model investigating this phenomenon and the potential mechanisms that can lead to group fragmentation.

[5] With the exception of one individual, who was in the community centered around the club president but associated with the trainer in order to finish his black-belt training.

Chapter 3

Simulating the dynamics of social networks

3.1 Introduction

In the previous chapter we have reviewed models that are able to grow networks mimicking social network structure by exhibiting the characteristic topological features observed in real-world social networks e.g. (Holme and Kim, 2002; Toivonen et al., 2006; Vázquez, 2003). While these models grow the network in a dynamic fashion by adding links sequentially, the growth process is typically a means to achieve some class of final network structure rather than an attempt to model the human social behaviours that continuously generate and maintain real social networks. Other models explicitly model the dynamic change in topology (Davidsen et al., 2002; Ebel et al., 2002a; Kumpula et al., 2007; Marsili et al., 2004) but rely on periodic exogenous removal of randomly chosen edges or nodes (including all their edges) in order to maintain the system in a steady state. Models of *coevolutionary* or *adaptive networks* focus on the mutual feedback between processes constrained by the network and topological change as a result of individual behaviour (Gross and Blasius, 2008; Skyrms and Pemantle, 2000) but rely on a specific process taking place on the network such as opinion dynamics or pairwise games (Eguíluz et al., 2005; Holme and Newman, 2006; Takács et al., 2008; Van Segbroeck et al., 2009) restricting their applicability to the study of strategic interactions Thus, we have identified a gap in current work on network dynamics for a general model of network evolution, where individual behaviour drives network evolution and the network reaches a dynamic steady state which can be maintained without artificial, global balancing processes. Here, we propose such a model in which the network topology is created as a result of ongoing changes in interaction patterns. The interaction patterns emerge as a result of the nodes' decisions and can be formally described as a weighted network; we will investigate the topology of networks generated by the model in the next chapter.

In this chapter we will discuss the choice and implementation of behavioural rules on the individual level that make up our model. The decisions on which components of human behaviour to include in the model were guided by sociological theory and empirical evidence. We will elaborate on the assumptions and design choices made for the various model components and how they relate to the literature on theory and data of social networks/interactions.

3.2 Modelling social network dynamics

3.2.1 Problems modelling the dynamics of weighted networks

Before arriving at the model presented in this chapter, we explored several other mechanisms. In this section we will discuss these approaches and why they were insufficient to model the dynamics of weighted social networks. Not many models of weighted social networks exist and we postulate that this is due to the fact that several issues arise when simulating network dynamics, that can be avoided in binary networks but not in weighted networks. Here we will discuss these problems before presenting a model that is able to overcome these difficulties.

The main issue when simulating network dynamics (binary or weighted) is that for a network to reach a steady-state the number of edges made must balance the number of edges deleted. In many models this problem is avoided by using rewiring dynamics (Demirel et al., 2011; Geard and Bullock, 2009; Holme and Newman, 2006). When using rewiring for modelling network dynamics, the number of edges is kept constant as edges are disconnected from one node and rewired to another, but never created or removed. Thus, achieving a stable network state with a certain density is guaranteed by the rewiring mechanism. In cases where edges can be created, an edge removal process must exist to balance the increase in edge number (unless the focus is on network growth which is not the focus of the model presented here). This is generally implemented as network growth with an added periodic removal of edges (Marsili et al., 2004) or of nodes with all their edges (Ebel et al., 2002a; Kumpula et al., 2009). Thus, these models utilise a globally set removal rate that needs to be tuned for obtaining stable steady-state networks. While this is an effective mechanism to achieve stable networks, edges or nodes to be removed are chosen randomly at the network level, rather than removal being a local process reflecting a real-world process.

One of the few models using weighted dynamics, which is presented in Kumpula et al. (2009), employs this technique of periodic node removals to achieve a steady state. While it is able to produce weighted networks, this model does not provide any dynamics of those weights. Similar to growing models being limited to models of network growth, this model is limited to modelling weight strengthening: edges are only ever strengthened until they are removed (when a node they are connected to is removed).

When simulating weight dynamics, balancing is again the main issue. For a network to reach a steady-state the total increase in weights must be balanced by a decrease in weights. There are several options to implement a balancing process. We could implement a global decay process that decreases each edge weight by a certain predetermined amount each time step. However, in preliminary studies we found this method very difficult to parameterise as the increase in edge weight is highly heterogeneous depending on the edge's embedding in the network. We would frequently observe fully connected or completely empty networks.

Another option is to implement a local process scaling all edges attached to a particular node once the node strength reaches a predetermined value. We found that this resulted in very homogeneous networks and often lead to spontaneous fragmentation of the network into disconnected clusters during the initialisation phase, a feature not desired for the baseline model. Furthermore, this means that each edge is adjusted by the two nodes it connects independently, which might be a reasonable assumption for sentiment networks but is not a good assumption for modelling contact networks, since the contact frequencies represented by the edges are a result of the aggregated decisions of both individuals connected by the edge.

These unsuccessful initial modelling approaches prompted us to build a more refined model through studying assumptions made in sociological and organisational literature and deriving the behavioural rules of the individuals from these assumptions. Through this approach we have arrived at a more elegant way of balancing edge weights that arises naturally from the dynamics. We will describe this method in Section 3.3.1.4, after discussing the assumptions that have influenced these rules.

3.2.2 Assumptions

Following Snijders et al. (2010) our selection of model mechanisms was "guided by theory, subject-matter knowledge, and common sense". We will first discuss the assumptions we make about human behaviour that inform our model and link them to sociological theory and empirical observations.

3.2.2.1 Decay and time allocation

Most importantly, we assume that social ties have to be actively maintained, otherwise they decay in intensity and importance and eventually vanish (Burt, 2000; Cummings et al., 2006). As ties need to be maintained by social interactions, individuals have to devote resources to maintain a tie (such as time, money, emotional energy and memory space) and these resources are limited (Contractor et al., 2012; Kossinets and Watts, 2009; Miritello et al., 2013; Sutcliffe et al., 2012). Thus, individuals have to decide on how to allocate their time to the maintenance of contacts.

It has been shown that people are biased to interact with individuals they have interacted with previously (Caldarelli et al., 2004) as the contact has already been established and the individuals have established a certain level of trust and familiarity. In addition, exposure of two individuals to each other seems to increase the likelihood of people liking each other more (Reis et al., 2011) which might be another reason for them to seek to maintain the relationship. Time allocation in real-world interactions is also influenced by many other factors and how well two individuals get on also depends strongly on their character traits, opinions and interests. The fact that individuals in general prefer to interact with others who are like them in some respect is called homophily. Homophily has been shown to be an important predictor for edge existence (Kossinets and Watts, 2006) and friendship formation (van Duijn et al., 2003).

3.2.2.2 Spatial constraints

Direct contact between two individuals requires physical proximity which necessitates travel, costing time and money. People therefore have a limit on the distance they are willing to travel for social interactions (Illenberger et al., 2012) and only connections between individuals located close enough can thus be maintained through regular face-to-face meetings. In addition, it has been shown that individuals located close to each other have a higher probability for casual meetings leading to contact which maintains the connection (Conrath, 1973; Festinger et al., 1963; Monge et al., 1985; Van den Bulte and Moenaert, 1998; Zahn, 1991). Real-world social networks are embedded in space and many of the properties (such as high clustering and assortativity) that define social networks are linked to spatial embedding (Barnett et al., 2007; Bullock et al., 2010; Herrmann et al., 2003).

3.2.2.3 Triangle closure

So far we have mainly discussed the assumptions regarding the maintenance of existing relationships. Connections in social networks are often made when an individual brings two of their friends into contact, thus facilitating a new connection. This process is called triangle closure and has been shown to play an important role in real-world social networks (Kossinets and Watts, 2006; van Duijn et al., 2003). It is used in several of the models discussed in the previous chapter to achieve strong clustering, meaning connected nodes are likely to have network neighbours in common (Newman, 2010).

3.2.3 Modelling the dynamics of relationships

In a social network the edges represent some kind of real-world relationship or interaction between individuals. The most important distinction to be made here is between contexts where edges represent quantified sentiments and contexts where they stand for some type of contact. In this thesis we investigate the phenomenon that despite sentiments between people remaining unchanged relationships between them might dissolve as a result of changes to the local environment in which the relationship is embedded. In the model presented here we therefore measure contact frequencies and represent these as edges. This means an edge does not represent how much two individuals like each other but rather how often they come into contact. We represent this contact frequency of a pair of individuals as an undirected edge (representing a symmetric relationship) between them.

Unlike in other models, we simulate the interaction between individuals explicitly and the edge weights are a representation of this simulated interaction, just as edges in an observed contact network represent real world interactions between people. Therefore, an edge's existence is not under the control of one of the actors as is the case in many models of strategic interaction. This means that individuals do not actively change edge weights or break edges. Rather, edge weights change as a result of contacts which in turn are a result of the individual decisions of both nodes involved in a tie. Edges disappear or are broken when individuals decide to not maintain the contact anymore in which case the edge weight gradually decays to zero.

3.2.4 Desired properties for generated networks

As discussed in Chapter 2, real-world social networks exhibit several defining properties. Models of social networks aim to create networks that match these properties. Therefore, we will revisit the list of properties characteristic for social networks and later assess the networks created by our model using this specification.

Real-world social networks exhibit high levels of transitivity or clustering, meaning that two individuals who have mutual friends are likely to be connected themselves. Another characteristic feature is short average path length, meaning that on average a pair of nodes is connected by a relatively short path. Short here refers to being shorter than would be expected from a regular network, such as a lattice, while potentially being longer than the very short average paths in random networks.

Social networks are generally sparse, with only a subset of potential connections being present. Networks that are both sparse and show high levels of clustering exhibit another defining topological features of real-world social networks: community structure. Community structure describes the existence of parts of the networks that are densely connected internally, but have few connections between them. Often the existence of community structure in a social network is assessed simply by visual inspection. Communities show a high edge density within the community with fewer edges existing between communities. More formally community is defined as sets of nodes that have more connections between each other than to nodes outside of the group.

Another feature that is used frequently to distinguish the topology of social networks and other networks is degree correlation. In social networks degree correlation between nodes connected by an edge is high and positive, which is referred to as assortative mixing. In contrast, in other networks (e.g. technological networks) negative degree correlations are frequently observed.

Regarding the degree distribution of the network, there is some debate on the exact nature of the distribution and it is likely that the exact shape depends on the type and size of the social network. However, a general consensus is that social networks should exhibit broad degree distributions with a certain degree of skew.

3.3 The *RASH* Model

In this section we will give a brief overview on the key mechanisms of the model and the corresponding parameters, before discussing how we have translated the assumptions listed above into model mechanisms. We will go into detail on how the constraints on interactions discussed in Section 3.2.2 are formalised into decision rules and how the constraining processes are parametrised. After that we will describe the algorithm in more detail.

3.3.1 Mechanisms

The core of the model is very simple. Each timestep, every individual invites all of their network neighbours to a gathering where connections between individuals are strengthened and new connections formed. Whether an individual accepts an invitation depends on several constraining processes which model the social processes described in Section 3.2.2. We will now introduce the parameters related to these processes before providing a detailed description of the implementation of each process.

1. Interactions are constrained by a limited interaction *range*: individuals separated by a spatial distance greater than R cannot accept each others' invitations.
2. Interactions are constrained by *affinity*: with probability A there is enough affinity between a pair of individuals to interact. With probability $1 - A$ there is a lack of affinity and the individuals will not accept each others' invitations. They can however, meet at a mutual friend's gathering.

3. Interactions are constrained by a limited number of time *slots* available: each individual can accept at most S invitations per timestep.

4. Interactions are constrained by *history*: invitations tend to be accepted when the individuals involved have a strong history of interaction over the preceding H timesteps.

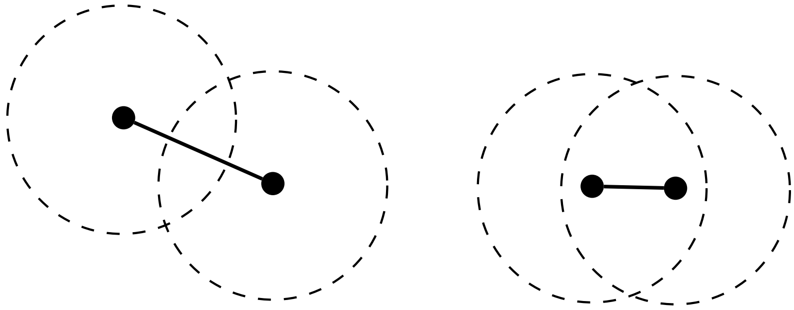
3.3.1.1 Spatial constraints

To model spatial embedding of the nodes and the resulting constraints on interaction we assign each of the N nodes a random position (x_i, y_i) (where x_i and y_i are chosen uniformly from the interval $[0, M]$) in a two-dimensional bounded square arena of size M and area M^2. The probability of interaction depends on the distance between two individuals; this is modelling the constraint that geographic embedding has on interactions. We model this by restricting interactions to pairs of nodes that are no further than R apart. The distance between two nodes is measured as the Euclidean distance between their positions. This is similar to the spatial restriction used by Hamill and Gilbert (2009, 2010) and following them we will refer to it as the *social reach* or *reach* of an individual. There is a small difference in the way the reach of an individual is described between their model and ours. In their model, two individuals interact if their reaches overlap therefore leading to an interaction distance of up to $2R$, meaning that nodes further apart than $2R$ can not interact. In contrast, in our model individuals must be located within each other's reach of radius R, so that the parameter R describes the threshold distance for interaction. We illustrate this difference in the definition of reach in Figure 3.1.

This constraint only applies to direct interactions where one individual attends a friend's gathering. In addition to these directly maintained interactions, individuals in our model can interact with friends of their friends, even if these individuals are located beyond their own direct social reach, as they can meet at a mutual friend's gathering. This relaxes the spatial limitations on triangle closure present in Hamill's social reach model.

3.3.1.2 Homophily and affinity

In addition to spatial distance, pairs of individuals may be separated by differences of personality or belief. To reflect this, each pair of individuals is assigned an affinity value describing how well they get along. This value is mutual and therefore symmetrical, and, once assigned, it remains constant for the duration of a run. For simplicity we only distinguish between two cases. With probability A, a pair of nodes have sufficient affinity to send invites to each other. With probability $1 - A$, the pair's affinity is not sufficient for invitations to be exchanged. This incorporates homophily into our model



(A) Illustration of interaction restriction used in Hamill and Gilbert (2009).

(B) Illustration of interaction restriction used here.

FIGURE 3.1: Defining the interaction range of an individual as a circle of radius R around each individual, we illustrate the differences between the model by Hamill and Gilbert (2009) and the model presented here. In the model presented by Hamill and Gilbert two individuals are within reach if their interaction circles overlap (resulting in maximum interaction distance of $2R$). In the model presented here, individuals must lie within each others interaction circle, resulting in a maximum interaction distance of R. Thus the parameter R in our model denotes the threshold distance for active maintenance of a contact between two individuals.

in an abstract way. Note that similarly to the spatial distance, a connection between people who would not actively maintain contact (due to being to far apart in space or out of disinterest due to low affinity) can be maintained indirectly through a mutual friend.

3.3.1.3 Time constraints

As discussed above, social interactions are limited by the time budgets of the individuals involved. In general, individuals have more opportunities for social interactions than time available to pursue these opportunities. A model that explicitly models these limited time budgets is presented in Geard and Bullock (2010) where individuals have a fixed capacity for investing in group membership. We impose a similar capacity by restricting individuals to attend at most S gatherings per time step, chosen from amongst the (typically greater than S) invites that they receive. They choose based on affinity as well as relative familiarity (based on the previous encounters reflected in the edge weight), similar to the mechanism of choosing interaction partners described in the "Friends" models in Skyrms and Pemantle (2000). This mechanism leads to positive reinforcements as familiar interaction partners are preferred, resulting in more familiarity. As discussed in Section 3.2.2.1 this effect has been observed empirically.

3.3.1.4 Decay and maintenance of edges

Finally, real social connections must typically be actively maintained or they erode and eventually disappear. Thus, we must include some form of decay if we want to model the dynamics of a social network.

A simple mechanism to achieve this is to use a global decay process which reduces all edge weights every time step. We found this almost impossible to calibrate, often leading to networks either fragmenting completely if decay is too strong, or connecting fully if decay is too weak. This is a major problem that bottom-up models of weighted social networks face, as stochasticity together with heterogeneity in degree and positive feedback make it difficult to predict low-level outcomes. We will discuss this in more detail in Chapter 6.

Some models avoid this problem through occasionally removing nodes with all their edges. However, since our ultimate aim for this model is to understand how edges disappear as a result of changed context, edge weight decay has to be a linked to individual decisions.

Consequently, rather than implementing global decay or turnover processes, we have chosen a localised mechanism inspired by the fact that friendships need to be actively maintained. We therefore assume that there is a baseline contact frequency that is required for an edge to continue to exist. This means that if a contact is not maintained regularly enough, individuals will "forget" each other.

To achieve this, we model the interactions between individuals explicitly, meaning that the edge weights denote relative frequency of contact between two individuals (rather than sentiments, as discussed in Section 3.2.3). This means that the intensity of a relationship is defined as the history of meetings within a certain time frame. In detail, each individual has a memory recording their meetings with other individuals for each of the previous H time steps. The number of encounters between two individuals i and j at timestep t is denoted h_{ij}^t and can range from 0 to $S+1$ occasions per timestep (the maximum value corresponding to the situation where both individuals attend each others' gatherings and also attend the same $S-1$ gatherings hosted by their other shared neighbours).

This means that H defines the time period within two individuals need to at least have one contact for the edge between them to continue to exist. Thus, if two individuals have had no meetings in the last H timesteps, there will be no edge between them. Otherwise an edge exists with the weight depending on the number of encounters within the last H timesteps.

Formally, the weight of an edge between individuals i and j, is calculated as

$$w_{ij} = \frac{1}{H} \sum_{t=1}^{H} h_{ij}^{t} \tag{3.1}$$

summing the total number of meetings in the pair's recorded interaction history h_{ij} and dividing this number by the history length H (thus yielding the average number of encounters per time step). This gives us a value that is meaningful and allows for comparison between runs with different memory lengths.

3.3.2 Algorithm

With the mechanics described, we can specify the model more formally as follows:

1. Each of N nodes is independently assigned a location, (x_i, y_i), selected uniformly at random on a map of size M (only one parameter is needed as we use a square map of area M^2).
2. Every pair of nodes, (i, j), where $i \neq j$, is assigned a symmetric affinity, $a_{ij} = a_{ji}$, set equal to one with probability A, and zero otherwise.
3. For every pair of nodes, (i, j), where $i \neq j$ and the distance separating them, d_{ij}, is not more than R, a connection is established by adding a single meeting at a randomly chosen point, $0 < t < H$ in their interaction histories, h_{ij}^t and h_{ji}^t, initialising the remainder of the history with 0 meetings.[1]
4. For each simulated timestep $t < t_{max}$:
   (a) For each edge, (i, j), update the interaction history by shuffling more recent values one step down the list and overwriting the oldest value h^H, i.e., $\forall t \in \{0, 1, \ldots, H-1\} : h_{ij}^{t+1} \leftarrow h_{ij}^t$. As the edge weight w_{ij} is defined as the mean number of meetings per timestep this update results in an update of the edge weight.
   (b) For each pair of nodes, (i, j), invite j to i's gathering (and vice versa) if $w_{ij} = w_{ji} > 0$ meaning j is a current network neighbour of i, and $a_{ij} = 1$, and $d_{ij} \leq R$.
   (c) For each node, i, accept at most S invitations, each selected with probability proportionate to the edge weight w_{ij}.

[1] Note that the resulting steady-state networks are not sensitive to the placement of the initial single meeting slot. We chose to initialise in a random slot as this means that during the initialisation phase $t = 0$ to $t = H$ edges that were initialised but then never maintained decay to 0 at different points and not all at the same time leading to a smoother transition to equilibrium phase.

(d) For each node, i, consider each pair of individuals (j, k) attending i's gathering (including i themselves), where $j \neq k$, and increment their interaction history by one meeting in the most recent time slot, i.e. increment h^0_{jk} and h^0_{kj} by one.

3.3.3 Algorithm illustration

We will describe one timestep of the algorithm using a hand-generated example network. For clarity, we will focus on one individual and its neighbourhood. The individual we will focus on here is individual 0, highlighted with a grey circle (Figure 3.2a).

At each step invitations are sent between network neighbours. They are only sent if the individuals are no further than distance R apart. In Figure 3.2b we visualise the social reach of node 0 by the grey circle. Node 0 can only send invitations to and receive invitations from individuals within this circle, i.e. nodes 2, 3 and 4. Furthermore, invitations will only be sent between pairs of nodes that have sufficient affinity. Let us assume that in this example the affinity value associated with each pair of nodes is 1, except for $(0, 3)$ (highlighted in Figure 3.2c by the edge between them being coloured orange).

Given these constraints, individual 0 sends out invitations to (and receives invitations from) all individuals within reach that have sufficient affinity, in this example nodes 2 and 4 (Figure 3.2d). This means that in the example node 0 will only send and receive two invitations. If we assume that each individual can select up to 2 meetings here ($S = 2$) then no selection is required here as the number of invitations is the same as the number of available time slots. For most parameter settings this is generally not the case and nodes will on average receive more invitations than they have timeslots available. To exemplify this, let us consider another node, individual 2. Individual 2 has received invitations from all its neighbours that are within reach: 0, 1, 3, 6 and 7, indicated by green edges in Figure 3.2e. Node 2 now has to select two of these invitations to accept. The likelihood of selection is proportionate to the edges weights, which is the same approach as used in Skyrms and Pemantle (2000). Thus, the most likely outcome here would be to select invitations from nodes 0, 3 and 7 (corresponding to the edges with the highest weights).

Node 2 has decided to attend node 0's gathering. We will now describe what happens at this gathering. If we assume individual 4 has accepted the invitation as well, then there are 3 individuals present at node 0's gathering: the two guests, 2 and 4, and the host, 0 (Figure 3.2f). Edges that already exist are strengthened (or maintained) which here applies to the edges $(0, 2)$ and $(0, 4)$. In addition, new connections can be formed between individuals. Here nodes 2 and 4 are introduced to each other and form a new, initially weak edge. This edge closes a previously open triangle (Figure 3.2g).

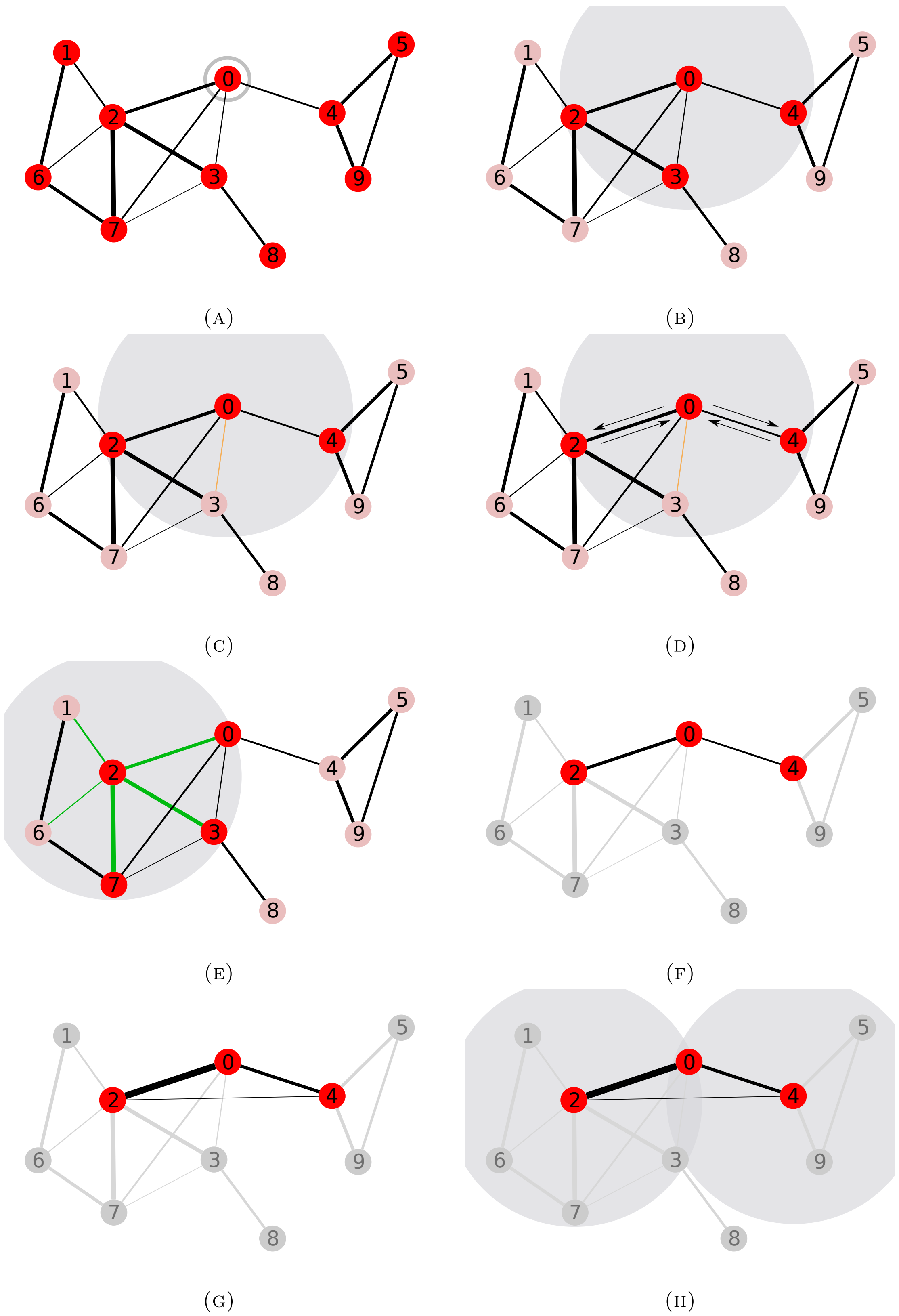


Figure 3.2: Illustration of the algorithms using a small example network. A detailed description of these figures is given in the main text.

Note that 2 and 4 are outside of each others reach (Figure 3.2h). Therefore, the edge between them can only be maintained indirectly by their mutual friend, node 0, who is located within the intersection region of their reaches. This means that both nodes can attend node 0's gatherings and maintain a connection between them by doing so. In the same way, the edge between node 0 and 3 can be maintained, even though the individuals do not attend each others gatherings due to lack of affinity.

The process demonstrated for one individual here happens for all nodes, meaning that each timestep N gatherings take place. New edges are formed, existing edges are maintained or strengthened and unmaintained edges decay in weight. Edges that have not been maintained in the last H timesteps will disappear.

3.3.4 Implementation details

3.3.4.1 Proportionate selection of invitations

Due to time constraints, each timestep every individual has to select S invitations from a typically larger set of invitations received. These are selected with probabilities proportionate to the edge weight. We implement this as a roulette-wheel selection.

For this an individual i assigns an invitation from its neighbour j the following score: $w_{ij} * a_{ij} * r_{ij}$ where w_{ij} is the weight of the connection between i and j, a_{ij} is the affinity score and $r(i,j)$ is a function of the distance between i and j. In the current version of the model we use a distance cutoff for interactions, meaning that

$$r(i,j) = \begin{cases} 1, & \text{if } d_{ij} \leq R \\ 0, & \text{otherwise} \end{cases}. \tag{3.2}$$

Similarly the affinity score a_{ij} is either 1 denoting sufficient affinity to attend a neighbours gathering, or 0 in which case the affinity is not sufficient. This means that being further away than R or having a affinity value of 0 is equivalent to an invitation not having been sent as it is always ignored. This means that effectively only invitations from neighbours within range are considered and only if the pair of nodes has sufficient affinity. From the set of individuals to which these two conditions apply, selection is proportionate to the edge weight.

3.3.4.2 Initialisation of the meeting history

When a new edge is made during the run (which happens when two individuals are introduced by a mutual friend), it is initialised with an empty history and at the end of the time step updated like all other edges. This means that the most recent slot is set to the number of meetings that occurred during the current timeslot. This number can be larger than one if two individuals get introduced to each other at one party and then immediately meet again at another party in the same timestep. This would be equivalent to meeting someone new one week and meeting them again at another party the same week.

This procedure means that there is no need for a parameter that determines with which weight a connection is initialised as the initial weight follows naturally from edge weights being calculated from observed meetings.

Chapter 4

Characteristics of generated networks

In this chapter we will discuss the characteristics of the network structures generated by the model presented in the previous chapter. We will discuss how the interplay of the constraining processes leads to networks that are more than simple random geometric graphs, even though the spatial restrictions are crucial for generating the desired topological structures. We show that the generated networks reach equilibrium on the macro-level, while changes on the micro-level continue to happen and why this is crucial for studying the type of network fragmentation that is the topic of this thesis.

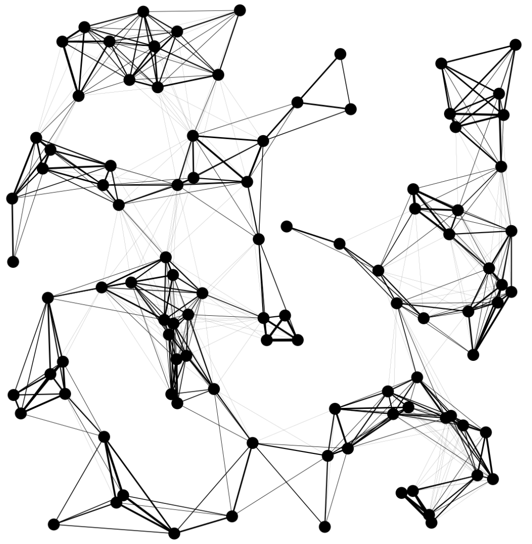
4.1 Topology of resulting networks

4.1.1 Topological measures

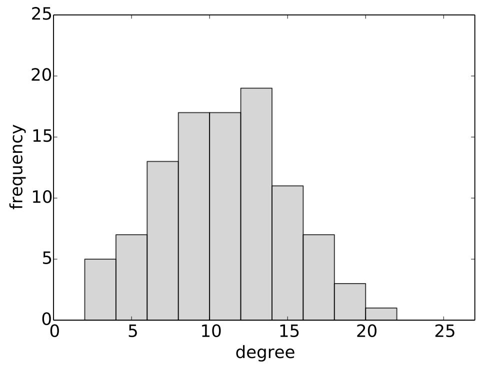
Social networks are typically sparse and they exhibit community structure as well as high clustering, positive assortativity with respect to degree and short characteristic path length. We will first investigate whether the structures resulting from running our model exhibit these properties. We use $R = 30$, $A = 0.75$, $S = 3$, $H = 50$ and $M = 200$ as the standard parameter setting (also listed in Table 4.1). We will discuss the influence of the parameters on the resulting networks in detail in Section 4.4. Setting $R = 30$ means that nodes can invite (and thus directly interact) with other nodes no further away than 30. The map the individuals are located on is of size 200 and area $200^2 = 40000$. Individuals can accept up to 3 invitations every time step, limiting the edge weight to a maximum of $w_{ij}^{max} = S + 1 = 4$. Around 75% of all pairs of nodes will have sufficient affinity to interact directly, while the remaining 25% will not invite each other. This standard parameter setting is used for all experiments unless stated otherwise.

Parameter	value
R	30
A	0.75
S	3
H	50
M	200

TABLE 4.1: Standard parameter values used if not stated otherwise.



(A) Network



(B) Degree distribution

FIGURE 4.1: An example network at $t = 10,000$ for the standard parameter setting. The network exhibits community structure with densely connected components, which are linked together by weaker edges. The degree distribution of the network is shown in Figure 4.1a. The distribution is broad, spanning an order of magnitude and limited for higher values, as observed in real social networks that require edge maintenance, however, it lacks the skewedness observed in many real-world networks.

An example network generated with the standard parameter setting is shown in Figure 4.1a. Regarding the desired characteristics defined in Section 3.2.4, the obtained networks are sparse, exhibiting both densely connected clusters and weaker bridges between clusters. Thus, the resulting networks exhibit community structure. In Table 4.2 we list network measures calculated for the networks resulting from running the model for $t = 10000$ timesteps with the standard parameter settings. Each value shown is the rounded mean taken over 100 runs, with the standard deviation shown in brackets. The values obtained for the statistical measures match those of real-world social networks with high clustering and positive assortativity with respect to degree. The average shortest path length is low, however, looking at a larger system size of $N = 1000$, the average shortest path length does not scale logarithmically as observed in many real-world networks. This is due to the strong limitations on contacts in our model, imposed by the interaction threshold R. We will discuss the scaling behaviour of the system in detail in Section 4.3.

	N=100		N=1000	
nodes*	98.49	(± 1.23)	992.06	(± 3.41)
edges	535.56	(± 56.76)	6153.7	(± 163.57)
clustering	0.75	(± 0.04)	0.71	(± 0.01)
clustering weighted	0.25	(± 0.03)	0.2	(± 0.01)
transitivity	0.69	(± 0.04)	0.66	(± 0.01)
assortativity	0.49	(± 0.11)	0.55	(± 0.05)
shortest path	3.49	(± 0.75)	12.16	(± 1.78)
diameter	8.49	(± 2.23)	31.69	(± 5.89)
degree	10.88	(± 1.17)	12.41	(± 0.34)
strength	12.22	(± 0.37)	12.67	(± 0.09)
components	2.88	(± 1.39)	8.37	(± 2.96)

Table 4.2: Measures characterising the topology of the resulting networks for network sizes $N = 100$ and $N = 1000$, showing mean (over 100 runs) and standard deviation for each. For networks consisting of more than one component, the average shortest path length and diameter were measured for the largest component. When increasing the system size, the number of edges increases proportionately with the number of nodes (by a factor of 10 in this case). Diameter, average shortest path length and the number of components increase by a factor of $\sqrt{10}$. This results from scaling the grid to ensure the same density of nodes when increasing the number of nodes. We will discuss how the different measures scale with N in detail in Section 4.3. (*When calculating these measures, singletons are ignored, meaning only nodes with at least one edge are considered. Listed here is the number of nodes that were not singletons and therefore included in the calculation of the measures shown.)

The degree distribution is relatively broad, spanning an order of magnitude, and centred around a characteristic value, as has been observed for real-world social networks where links require regular contact to be maintained (Amaral et al., 2000). The maximum degree has a cut-off, which has been suggested as another important characteristic of interaction networks (Figure 4.1b) (Hamill and Gilbert, 2010). However, the distribution does not exhibit an obvious skew, while skewed distributions have been observed for many real-world networks.

4.1.2 Resulting networks differ from random geometric graphs

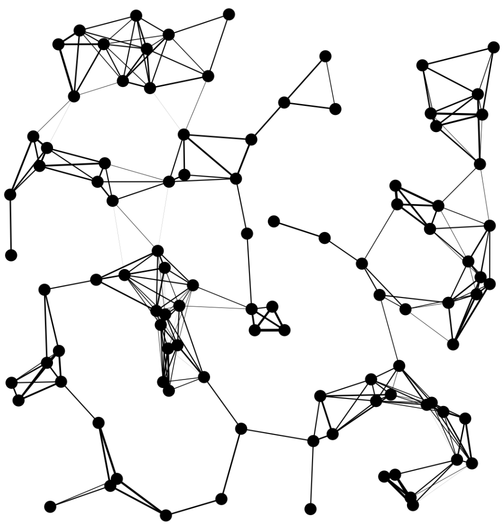
The spatial constraints imposed by the finite interaction reach encourage the high clustering and community structure typical of a social network. However, the networks are not just random geometric graphs with threshold distance R. Clustering is higher than for a random geometric graph (Dall and Christensen, 2002) as gatherings bring about longer range edges with a length up to $2R$. However, these longer range edges can only be created if they close a triangle. They are able to persist only as long as connections to the shared neighbour that facilitated the edge creation exist. Because of this dependence on other edges we will refer to these edges as *secondary edges* and to the edges with a length of less than R (which can be directly maintained by mutual invitation) as *primary edges*. Since edges with length between R and $2R$ have to be maintained by a mutual friend, many nodes separated by such a distance are not connected (see Figure 4.2). Nodes at that distance might be lacking a mutual friend for several reasons. There might be simply no node placed in the space between them, meaning no potential edge facilitator exists. Or there might be a lack of affinity between one of the nodes and a potential mutual friend. All of these effects can lead to potential secondary edges not being created. The influence of affinity results in not all possible primary edges being present either as some individuals will not invite each other. Furthermore, the stochasticity of the meeting process in combination with the time constraint can lead to a primary edge being lost in which case it can only be recovered if the pair of nodes have at least one common neighbour.

Both the additional process of triangle closure bringing about secondary edges as well as the added constraints imposed by the restrictions on affinity and time available for maintaining connections lead to networks that are more complex than random geometric graphs with connection threshold R. Furthermore, the triangle closure mechanism creating secondary edges leads to dependencies between edges. We will show later that these dependencies are crucial for modelling the social processes that lead to fragmentation in social networks.

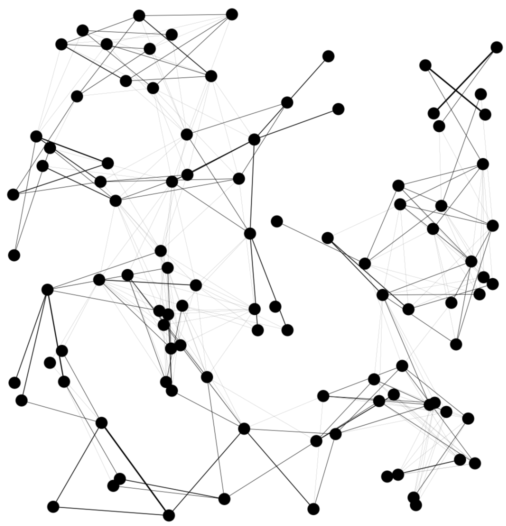
4.1.3 Variability between different realisations of the same spatial arrangement

The stochasticity of the initialisation and the meeting process allows for some variability between networks generated with the same node locations but different initial seeds. Figure 4.3 shows four networks resulting from running the model with the same spatial arrangement of nodes but varying the seed initialising the random number generator. This means that the histories of the initial edges are initialised differently, affinity values might vary between runs and the selection of invitations to accept is changed (but still biased in the same way towards edges with higher weights). The resulting network topologies look similar but also exhibit some obvious differences. For example, the networks on the left hand side have a higher number of components as nodes on the far left form a separate component. This is due to the stochasticity in the meeting process and the fact that edges, once broken, can only be re-formed if a mutual friend is present.

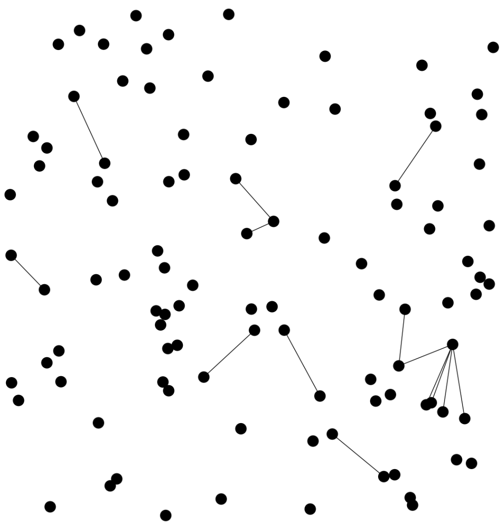
To study the variability between different runs on the same spatial arrangement we ran 1000 repeats of the model with identical node positions. For each edge observed in at least one of the resulting networks we calculated in how many of the 1000 repeats this edge was observed. An example run of this is shown in Figure 4.4. Primary edges are shown in grey and secondary edges are shown in cyan. This shows that many of the primary edges are present in almost all of the 1000 repeats, whereas secondary edges typically only occur in some of the realisations.



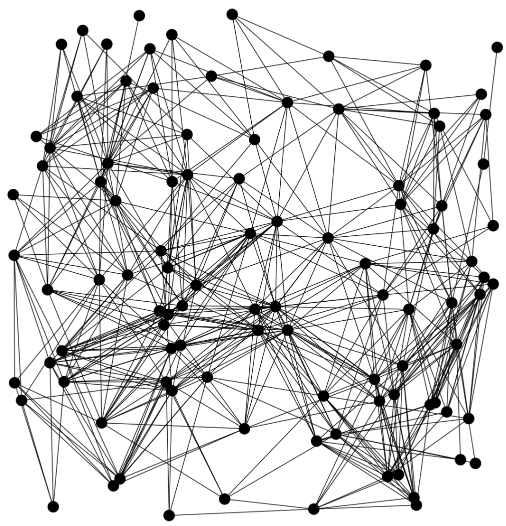
(A) Primary edges present



(B) Secondary edges present



(C) Primary edges missing



(D) Secondary edges missing

FIGURE 4.2: Classes of edges within the example network shown in Figure 4.1a. Figures 4.2a and 4.2b show the edges that are present in the network, whereas Figures 4.2c and 4.2d show potential edges that are not observed. The edges are divided into two groups, primary edges and secondary edges. Primary edges are between nodes close enough to invite each other to gatherings (i.e. $d_{ij} \leq R$). Secondary edges are between nodes that may only interact at the gatherings of mutual neighbours (i.e. $R < d_{ij} \leq 2R$). A typical network features many but not all of the possible primary edges and a much smaller fraction of the possible secondary edges. We discuss this in more detail in Section 4.4.2.

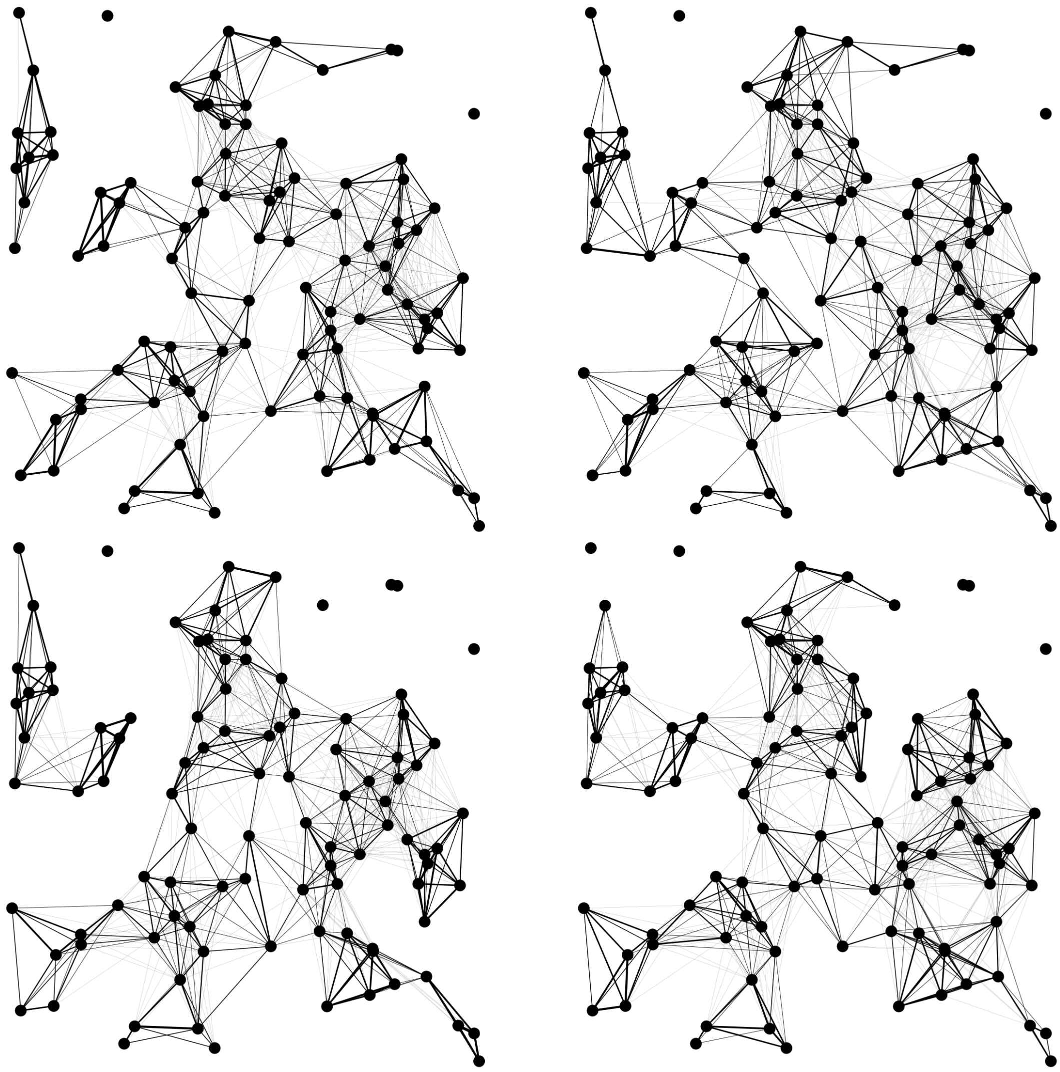


FIGURE 4.3: Four networks resulting from running the model with the same spatial arrangement of nodes but different random number generator seeds. The standard parameter setting was used. While exhibiting similar topology the networks also have some differences. Note for example that nodes on the far left form a separate component in the two networks on the left hand side.

We then average over 20 different spatial arrangements, running the model with 1000 different seeds for each arrangement. This means that again any particular edge can be present in up to 1000 repeats (if it is found in every single repeat). The results are shown separately for primary and secondary edges in Figure 4.5. On the x-axis the number of repeats that an edge can be present in is shown and the y-axis shows the percentage of primary or secondary edges respectively that are present in that number of repeats. This shows that almost all primary edges are present in all repeats for a particular spatial arrangement, i.e. the presence of primary edges seems to be only minorly influenced by the particular chain of events as determined by a particular random generator seed. In contrast, there is a much higher variability across secondary edges. Most secondary edges are present in more than half of all repeats on one node placement, but the distribution is broad.

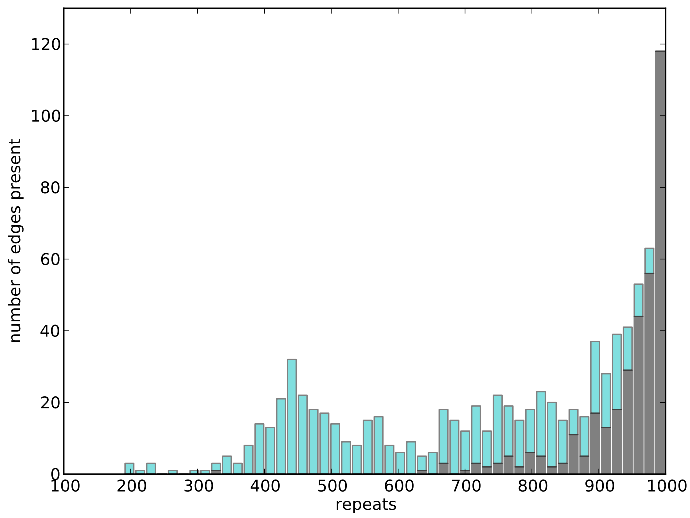


FIGURE 4.4: Number of edges that are present in a certain number of network realisations generated on the same spatial arrangement of nodes. Primary edges (shown in grey) seem to be fairly stable across different repeats whereas secondary edges (cyan) show more variability. This figure shows the results of 1000 realisations on exactly the same spatial arrangement of the nodes.

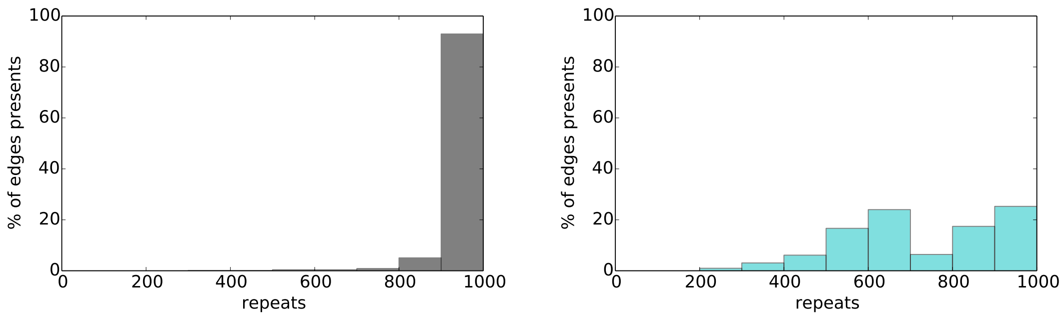


FIGURE 4.5: These figures were created by generating 20 different spatial arrangements. For each of the arrangements, 1000 repeats were run, each with a different random seed (as for Figure 4.4). We then calculated in how many of these 1000 repeats each edge (that was found in at least one repeat) was present, giving us a measure of variability between runs. We calculated these counts for all edges observed and for each of the 20 spatial arrangements. Using these counts we derive the numbers of edges that are present in a certain number of repeats (out of 1000 total repeats). From this we calculate the percentages of edges that were present in a certain number of repeats; these are the values shown in the figures. The left hand figure shows that almost all (more than 90%) of all primary edges are present in all repeats for a particular spatial arrangement. The right hand figure displays the percentages for secondary edges, which tend to only be present in some of the repeats.

4.2 Dynamical behaviour

4.2.1 Macro-level

Regarding the dynamics of the system, the topological measures reach equilibrium after H timesteps as shown in Figure 4.6 (the standard parameter setting was used here with $H = 50$). After H steps the initialisation effects have been removed. At this point all H slots of the history contain entries that reflect simulated meetings as all entries that were set during initialisation phase have been replaced by newer entries. Once this has happened the topological measures are remarkably stable. In the initialisation phase new edges are formed, resulting in an increase in mean degree and strength of the nodes as well as a decrease in diameter and average shortest path length.

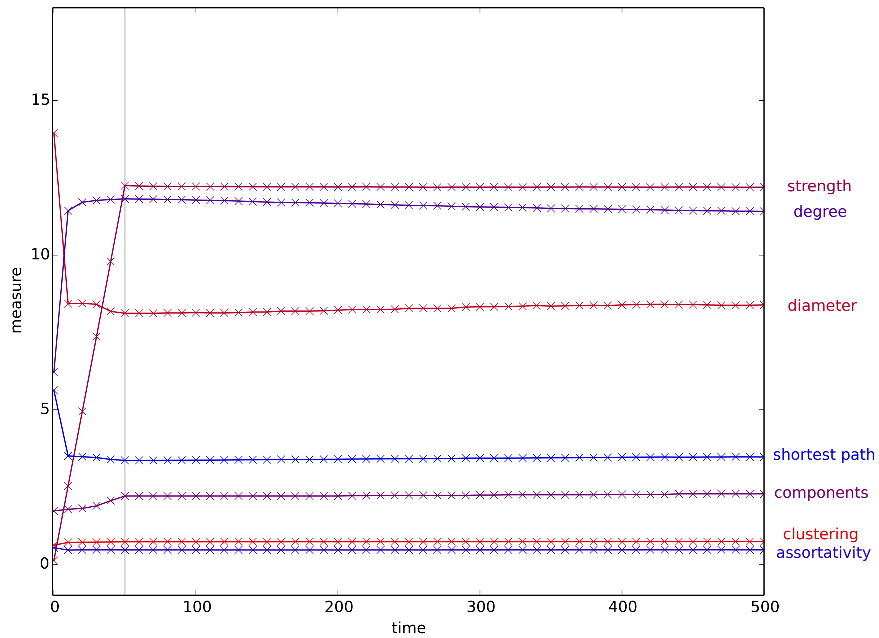
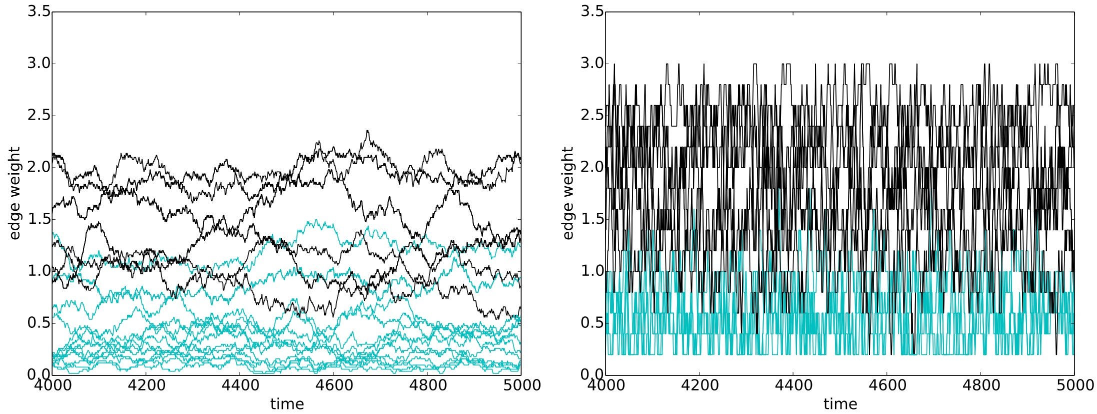


FIGURE 4.6: Topological measures equilibrate rapidly, within $H = 50$ timesteps. After $t = 50$ all initialisation effects have been removed and the system reaches a quasi-stable equilibrium. All data points shown are means taken over 100 runs. The standard parameter setting $H = 50$ is marked by the grey vertical line.

4.2.2 Micro-level

While the values of measures characterising the topology are stable once all history slots contain values reflecting actual meetings, the exact topology described by edges and their weights is subject to random fluctuations. This is exemplified in Figure 4.7 where the change in edges weights is shown for all edges attached to a randomly selected node (for a single run). The edge weights fluctuate over time due to the stochasticity in the meeting process generated by limiting the number of invites accepted to S (and the random selection of these) and having only a finite memory of length H. The magnitude of the fluctuations can be controlled by changing H. A higher value (longer memory) means that the edge weights are calculated over a larger window of time, leading to less volatility (Figure 4.7).



(A) $H = 50$ (B) $H = 5$

FIGURE 4.7: Edge weights over time for all edges attached to a randomly chosen node, shown for the interval $t = 4000$ to $t = 5000$. The system has reached global equilibrium at this point and topological measures are stable, while individual weights continue to fluctuate. Primary edges are plotted in black, secondary edges in a cyan. To highlight the effect of the history parameter H on the stochastic fluctuations the settings $H = 5$ and $H = 50$ are shown, displaying a much larger magnitude and frequency in fluctuations for shorter history dynamics. Note also that secondary edges on average have a lower edge weight, as they rely on indirect maintenance through mutual friends and the increased distance between them ensures that they share fewer of these neighbours compared to nodes closer together.

This means that the system reaches a quasi-stable equilibrium in which the statistical properties of the topology do not change, whereas edge weights (and therefore the exact topology) continue to change. This seems to be a property of social systems in equilibrium and has been demonstrated for other models/systems as well (Bryden et al., 2011). Related is Snijders's definition of a dynamic equilibrium as "stochastic fluctuations without a systematic trend" (Snijders et al., 2010), which we observe in our model as fluctuating edge weights without these fluctuations propagating to the macro-level.

4.3 Scaling behaviour

The network's topological properties discussed in Section 4.1 scale appropriately with increasing system size. For instance, for a larger system consisting of $N = 1000$ nodes (the size of the grid was increased appropriately to ensure the density of nodes remained the same for $N = 100$ and $N = 1000$), we observe the same structures as before, leading to comparable values of clustering, assortativity and average degree (see Table 4.2 for values and Figure 4.8 for an example network). Since these measures are linked strongly to the connectivity, they are mainly constrained by the spatial restrictions placed on the interactions. Therefore, the increased number of nodes does not result in growth of the existing structures but in the presence of more of these local structure that were previously observed. They are linked together to form a larger network, leading to increased diameter, shortest path length and number of components (see Figure 4.9). The increase in those measures is due to the fact that long range links are not possible in this model because of the strict restrictions on social reach, thus prohibiting the existence of edges bridging long distances.

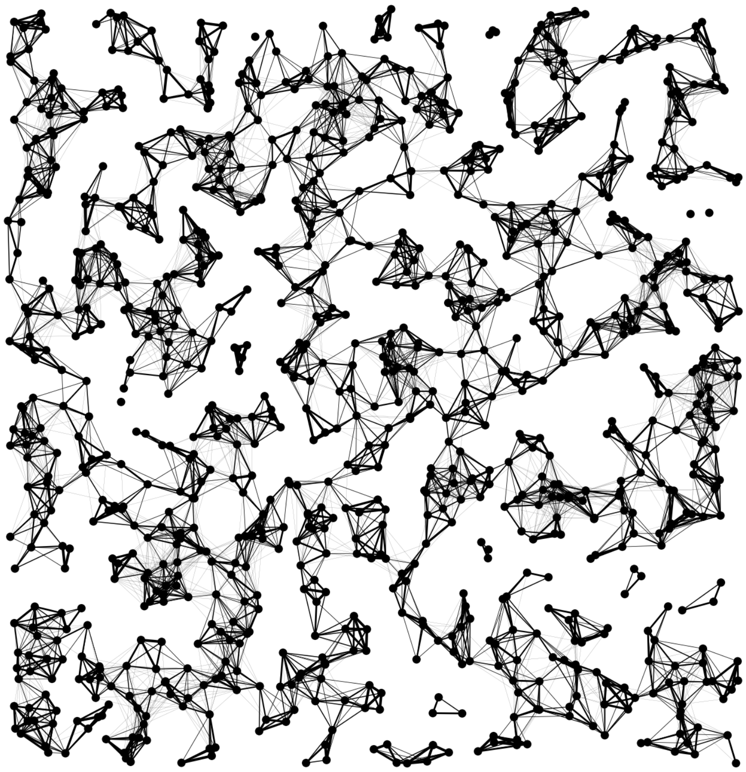


Figure 4.8: An example network with many nodes. Parameters are the same as in Figure 4.1a except for the increased number of nodes $N = 1000$ and the size of the map being enlarged to $M = 632$ to ensure the same node density as for the run with fewer nodes. The resulting network exhibits the same characteristic local structures as for $N = 100$ but inevitably has a larger diameter and longer shortest path length due to longer range edges being prohibited by the spatial constraint (see Table 4.2 for a comparison of the topological measures).

To demonstrate that the increase in diameter, shortest path length and number of components results from components of the same scale being connected together we will now compare sub-networks of the same spatial size cut from the networks with different sizes. For each value of N we select an area of the same size (140 by 140) from the middle of each spatial map (of total size $M = \sqrt{400 \cdot N}$) and only consider nodes and edges within the selected area for calculating the topological measures (see Figure 4.10 for illustration). Thus, we obtain a subnetwork for each network which contains only the nodes located within a certain area. This area is chosen to be the same size for each of the networks to be compared, therefore allowing us to compare the topological structures found on a local level across different system sizes.

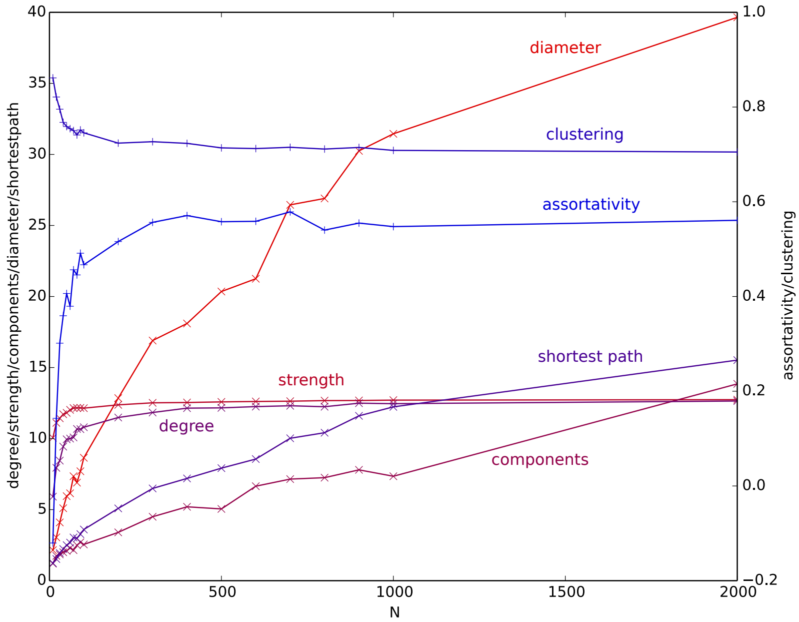


FIGURE 4.9: Topological measures for different network sizes N. Clustering, strength and average degree as well as assortativity do not scale with N (apart from for small values of N) since the social reach R creates a characteristic scale. In contrast, diameter, average shortest path length and the number of components increase with increasing N. All values shown are averages over 20 runs. The map was enlarged with increasing N to keep the node density constant across different values of N according to $M = \sqrt{400 \cdot N}$, all other parameters were kept constant to the standard parameter setting.

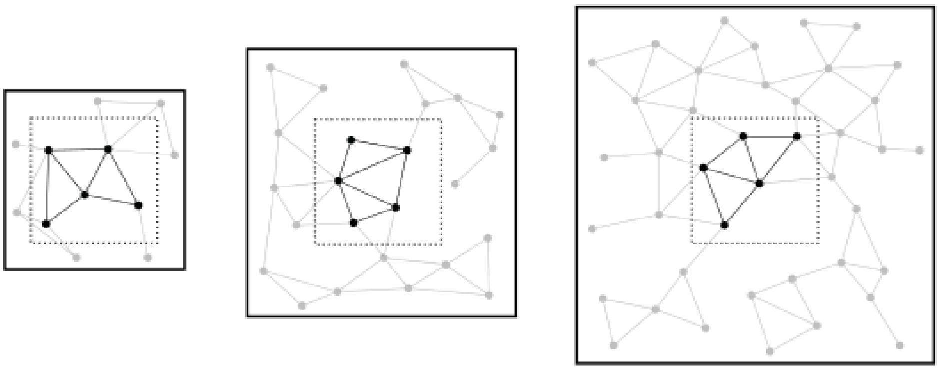


FIGURE 4.10: To show that structures on the local level have the same properties irrespective of the system size, we compare sub-networks of the same size from each network. An area of size 140 by 140 is selected from the middle of the map and all nodes within this area and all edges between them are selected.

Figure 4.11 shows the topological measures calculated for different system sizes N. The values calculated from the selected area (with same size across each network size) are shown in colour and for comparison the measures for the whole network are plotted in grey (as shown in Figure 4.9). If we only consider the selected sub-networks, we observe that diameter, shortest path length, and number of components do not scale with N. This means that for all system sizes, the structures observed on a local level are the same. This is due to R placing a strong limit on interactions, thereby generating a characteristic scale independent of the system size N.

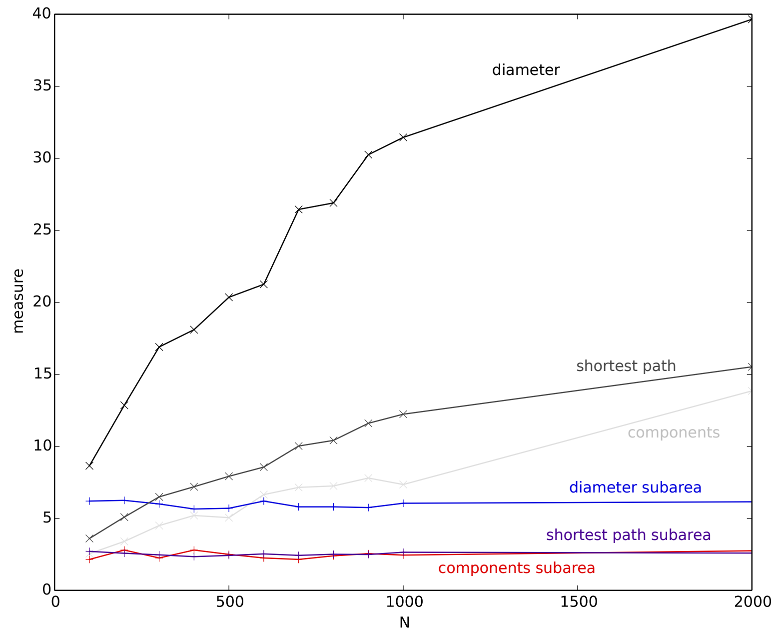


FIGURE 4.11: Measures for different network sizes N calculated for an area of size 140 by 140 from the middle of each spatial map, as illustrated in Figure 4.10. Only nodes and edges within that area were taken into account for calculating the measures. The fact that diameter, shortest path length and components do not scale with N if focusing on an area of particular size shows that the effects observed for the whole system (shown in grey here) are due to the system scaling by connecting components of the same scale to form a larger network. Each data point shown is the mean taken over 20 runs.

4.4 Influence of parameters

The model has four main parameters that control network dynamics and, consequently, influence topology: the social reach, R; the probability of affinity, A; the maximum number of invitations accepted per time step, S; and the length of the interactions histories, H. In this section we explore the effect that varying these parameters has and determine the portion of parameter space for which the model produces networks with desired characteristics.

4.4.1 Equivalence of varying M and R

We omit the parameter M from our analysis as we can show that varying M is equivalent to varying R and it is therefore sufficient to investigate the effect of varying R. Each node can only interact with a subset of the total number of nodes in the system. Since the N nodes are distributed approximately evenly over the map of size M^2, the number of nodes within reach of a node can be approximated by the fraction of the whole map (of size M^2) that is covered by the circle of radius R around the node. Formally, we can write this as:

$$\frac{N}{M^2}\pi R^2 \tag{4.1}$$

Decreasing the size of the map (by halving M) is equivalent to increasing the reach (by 2):

$$\frac{N}{(\frac{1}{2}M)^2}\pi R^2 = \frac{N}{(\frac{1}{2})^2 M^2}\pi R^2 = \frac{N}{M^2}\pi 2^2 R^2 = \frac{N}{M^2}\pi (2R)^2 \tag{4.2}$$

We illustrate this effect in Figure 4.12. Note that the same effect can also be achieved by increasing the number of nodes (without increasing the map size M accordingly as was done in the experiments described in Section 4.3), although here the equivalence would only hold on a local level as discussed in the previous section.

4.4.2 Limits on connectivity imposed by parameters

The social reach R places a hard limit on the number of interactions. R restricts the number of nodes to which a node can make connections and is thus the most important factor influencing the connectivity of the network. Direct interaction is not possible for nodes further apart than R, therefore the other processes governed by the parameters A, S and H can only operate within the limit imposed by R and reduce the set of possible connections further, but not extend it. We will investigate this upper bound on connectivity set by the parameter social reach and how the number of connection opportunities changes with (increasing) R.

For a sufficiently large system we would expect the average number of nodes within an individual's reach to depend on the proportion of the map covered by the interaction radius of R as described by equation 4.1. How this value scales with increasing values of R is shown in Figure 4.13 by the dashed black line. We can easily calculate the number of nodes within reach of a particular node simply by counting the number of nodes around it that are no further away than R (i.e. lie within a circle of radius R around the individual of interest). In Figure 4.13 we show the number of nodes within reach averaged over all nodes in a network. To cancel out any noise introduced by the specific placement of nodes for a particular run we take the mean over 280 spatial arrangements. The observed values are shown as a solid black line.

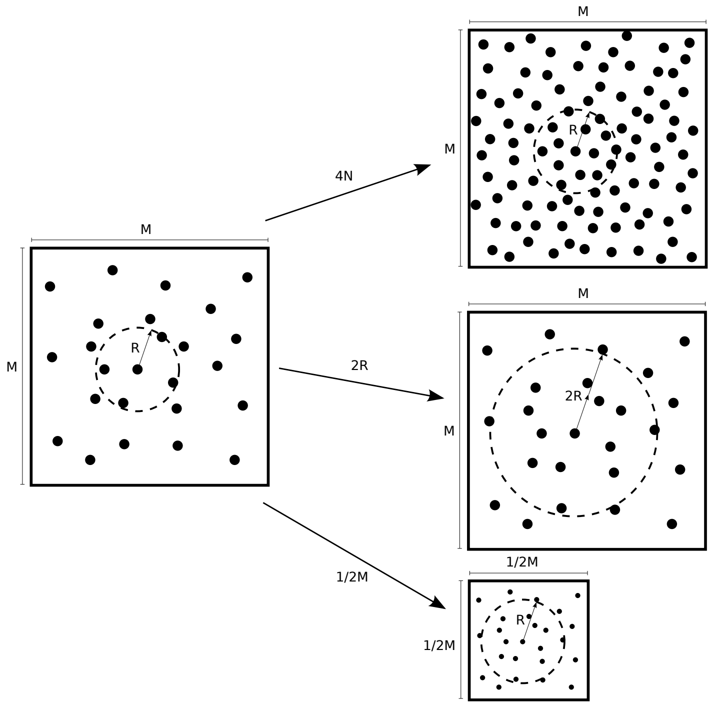


FIGURE 4.12: Figure illustrating how varying R, M and N can lead to comparable results. Increasing the interaction radius R leads to a larger proportion of the map being within the interaction radius of each individual. The same effect results from shrinking the map on which the nodes are distributed without changing R. Increasing N increases the node density and therefore leads to more nodes being located in an individual's interaction radius, thus increasing the number of potential interaction partners. However, in contrast to the equivalence in varying R and M, increasing N will lead to an increase in certain topological values, such as diameter, as discussed before.

If we compare observed and expected values for the number of nodes that are on average within an individual's reach, we see an obvious discrepancy here. This is due to boundary effects as the map has a finite size. Thus, nodes positioned closer than R to any of the borders have less potential interaction partners, as part of their interaction radius lies outside of the map [1] (illustrated in Figure 4.14). This effect can be quantified by measuring the mean number of nodes within reach separately for nodes within the central region of the map ($R \leq x \leq M - R$ and $R \leq y \leq M - R$, shown in green) and the boundary region (shown in red) (Figure 4.13). Nodes located in the central region of the map on average have the expected number of neighbours whereas nodes located in the boundary regions of the map have a significantly lower number of potential interaction partners. As R increases, the number of potential interaction partners increases in both regions. However, as the size of the central region shrinks with R (as nodes need to be further and further away from the border of the map to have no part of their interaction radius outside of the map) fewer and fewer nodes lie within that region. At $R = \frac{M}{2}$ the border region spans the whole network and no central nodes are observed any more.

[1] Note that we do not use periodic boundary conditions in the model.

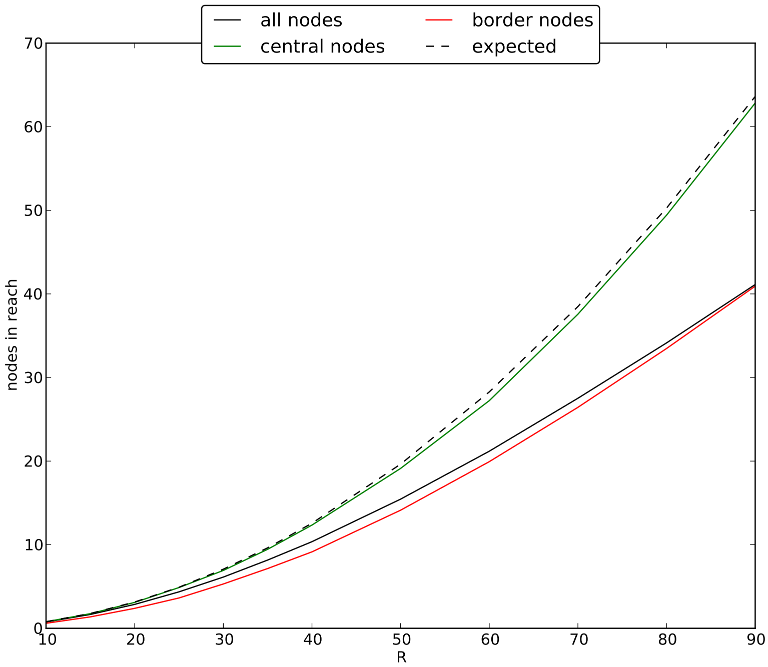


FIGURE 4.13: Number of nodes within reach of an individual. Shown in dashed is the expected number of nodes within reach as calculated in equation 4.1. The observed number is shown as a solid black line. Note that the observed number is much lower than the expected number due to the finite size of the map and resulting boundary effects. To highlight this, the number of interaction neighbours was calculated separately for nodes within the central region (green) and nodes located in the boundary region (red line) where generally less interaction partners are available (see Figure 4.14 for illustration). Each data point was calculated as the mean over 280 runs. The map size used was $M^2 = 200^2$, as in the standard parameter setting.

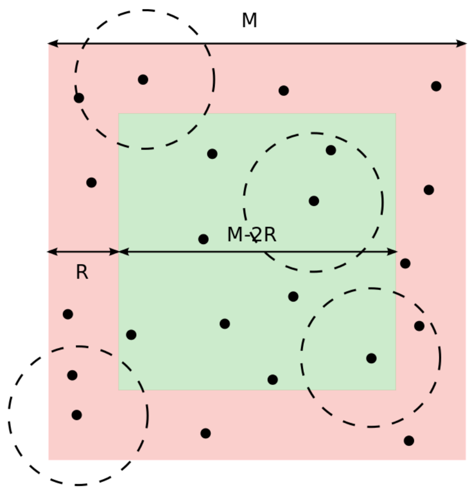


FIGURE 4.14: Illustration of boundary effect observed in Figure 4.13. Nodes located in the boundary region of the map (shown in red) have less potential interaction partners since part of their interaction circle defined by R lies outside of the map.

To summarise, R limits the number of edges that can be made to a set of potential edges. We will now investigate how many of these potential edges are made in a typical run of our model and how this number depends on the parameters R, A, S and H.

The number of edges that could be made if there were no restrictions on connections is given by:

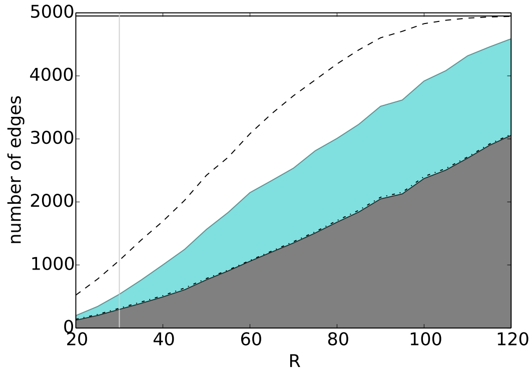
$$\frac{N(N-1)}{2} \tag{4.3}$$

This corresponds to a fully connected network, so for the standard parameter setting with $N = 100$ the network would contain 4950 edges. Analogously to Section 4.4.2 we can calculate the upper limit of edges possible in our model, given the spatial restrictions imposed by R. The upper limit here includes all potential primary edges (edges that can be maintained directly since the two nodes i and j are less than R apart - $d(i,j) \leq R$) as well as all potential secondary edges (calculated as the number of pairs i, j where $R < d(i,j) \leq 2R$).

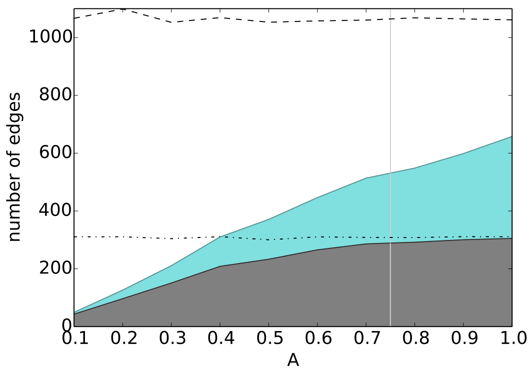
In Figure 4.15a we show these upper limits together with the observed numbers of edges. Shown as a solid black line is the maximum number of edges for a spatially unrestricted network, which corresponds to a fully connected graph. The number of edges possible (primary and secondary) given the restrictions imposed by R is shown as a dashed line. As discussed in Section 4.4.2 the number of possible edges increases with R and reaches maximum connectivity (fully connected) for sufficiently large R - when all nodes are within reach of each other, either directly or indirectly. For reference, we show the number of possible primary edges (as shown by a solid line in Figure 4.13) as a dot-dashed line.

We will now take a look at the number of edges present and how this number relates to the number of possible edges. In Figure 4.15a we show the number of observed primary edges in grey (with the area under the line coloured in grey) and the number of observed total edges (primary and secondary) in cyan. The area between the cyan and grey line is coloured cyan and corresponds to the number of secondary edges observed. As the parameter R increases, the social reach of each node gets larger leading to more potential direct interactions. As we can see in Figure 4.15a, the number of observed primary edges (shown in grey) is very close to the number of potential primary edges (shown by a dot-dashed line), meaning that very few primary edges are missing.

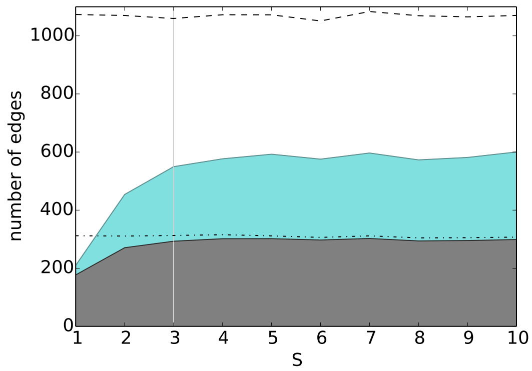
Thus, the number of actual interactions increases with the number of possible interactions. The number of primary edges for a run with the standard parameter setting can be approximated by the number of node pairs where the nodes are no further than R apart. With an increasing number of primary edges being present for larger R, the number of secondary edges present increases as well. However, the total number of edges does not reach the maximum number possible under spatial restrictions (shown in dashed). We calculated this number as the number of pairs that could technically



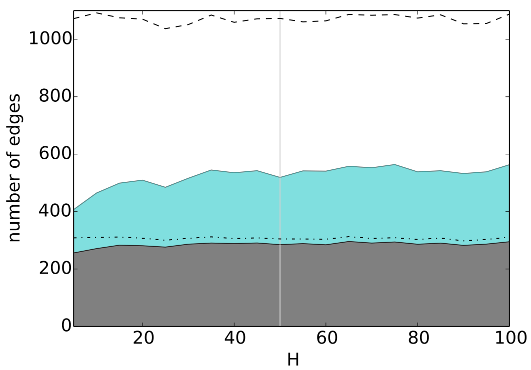
(A) number edges over R



(B) number edges over A



(C) number edges over S



(D) number edges over H

FIGURE 4.15: Changes in connectivity of the resulting networks when varying the 4 main parameters R, A, S and H. The parameter varied is shown on the x-axis, all other parameters were set to the standard parameter setting. R has the largest impact on connectivity and sets a hard upper limit on the number of edges possible within which the other parameters can reduce the number of edges further, but not increase it. With increasing R, more and more edges are present. Decreasing H reduces the number of edges due increased noise in the edge weights, but only for smaller values. S has a similar effect, but here the standard parameter setting of $S = 3$ is in the regime where the parameter has a limiting effect on the connectivity. The parameter A directly influences the number of primary edges that can be formed. Increasing A therefore leads to more pairs of nodes having sufficient affinity to maintain a direct connection. Note that in Figure 4.15a a larger range is used for the y-axis. For each parameter settings we ran 20 repeats. The standard value for the parameter studied/varied is indicated by a grey vertical line.

form a connection, either direct or indirect. In contrast to primary edges (which can form and be maintained directly if nodes are no further than R apart) secondary edges ($R < d(i,j) \leq 2R$) can not be maintained directly and therefore rely on the two nodes having a mutual friend. Thus, many potential secondary edges are not present because the two individuals between which an edge could be formed do not share any mutual friends.

We will now examine how the remaining parameters A, S and H influence network connectivity. We show the obtained results in Figures 4.15d to 4.15b. Since $R = 30$ (standard parameter setting) for these runs, the number of edges possible under the spatial restrictions imposed by this value of R is much lower than for the highest values shown in Figure 4.15a. Therefore, to visualise the effect of the other parameters on connectivity remaining subgraphs are shown on a smaller scale.

The parameter H (length of the meeting history) defines the number of past timesteps that are considered in calculating the strength of a connection. H therefore influences to what extend the stochastic fluctuations in meeting patterns translate into fluctuations of edge weights. For higher values, H has little influence on the number of edges present and long meeting histories correspond to low levels of noise. For small values ($H < 20$) H decreases the number of edges. This is because the increase in noise leads to edges being broken and some of these broken edges can not be re-formed due to a lack of mutual friends.

Next, we will discuss the influence of the affinity probability A. Increasing A leads to a larger number of primary edges as more direct interactions are possible because more pairs of individuals have enough affinity to accept each others invitations[2]. As secondary edges depend on the presence of primary edges, the number of secondary edges increases with the number of primary edges. For $A = 1.0$ all possible primary edges are present, which can be seen in Figure 4.15b as the number of primary edges (grey line) reaching the number of potential primary edges (shown as dot-dashed line).

Lastly, we will focus on the influence of the parameter S which specifies the number of interaction opportunities per timestep. Changing the value of S only has a visible effect on the network connectivity for small values of S, mainly because H is chosen fairly large here $H = 50$ so that even with a relatively small number of available time slots many interactions can be maintained. However, for the standard parameter setting of $S = 3$ the number of edges present is still somewhat lower than the number of possible edges. This means that under the standard parameter setting both R, A and S influence the connectivity, with R setting an upper limit within which A and S can reduce the connectivity further.

This effect is visualised in Figure 4.16. This figure is similar to Figure 4.15a, but instead of the standard parameter setting A and S are set to higher values. The affinity probability A is set to the maximum value of 1.0, meaning all pairs of nodes have sufficient affinity to interact directly. The number of meeting slots S is set to $S = 50$. In contrast to Figure 4.15a we observe fully connected networks for $R = 120$. This shows that even though the imposed spatial restrictions are most crucial for creating sparse networks, the other limiting processes (time constraints and affinity) influence the connectivity as well, within the limits set by R.

[2] Note that edges between individuals without sufficient affinity can still exist, as they can be maintained through a mutual friend, similar to secondary edges. We will discuss this in more detail in Chapter 5.

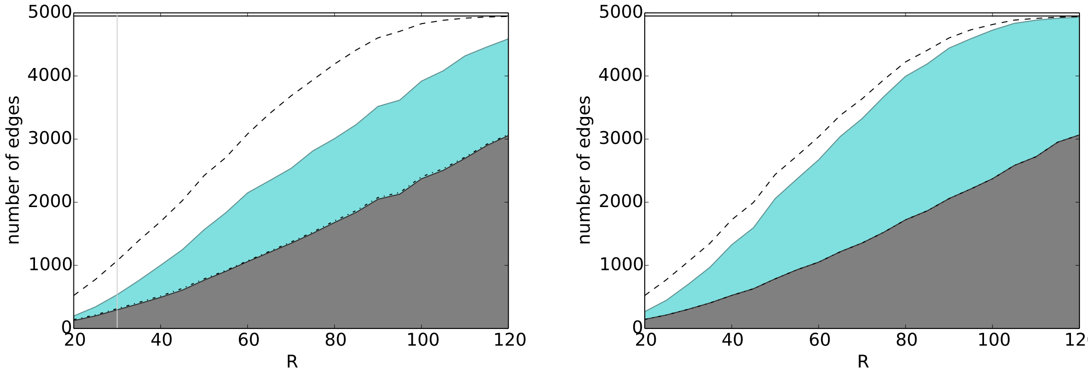


FIGURE 4.16: This figure shows the results from Figure 4.15a on the left hand side to allow comparison with the same analysis for the parameter setting A=1.0, S=50 and H=50 (shown on the right hand side). It highlights the limiting effect of S and A on connectivity as, in contrast to the case shown on the left hand side, the network connects fully for high values of R with the observed number of edges reaching the maximum number of edges possible. For lower values of S and A (as shown on the left side here) this value is not reached, as lower value of S and A were keeping some edges from forming.

4.4.3 Influence of parameters on other aspects of the network topology

In the previous section we have discussed the influence of the parameters R, A, S and H on the connectivity of the networks produced by the model. We will now investigate the effects of varying these parameters on other topological measures, such as the number of components, assortativity and average node strength (which is the weighted equivalent of degree, calculated by summing the weight of all edges attached to a particular node).

In Figure 4.17 we show example networks resulting for different combinations of R and S. For all runs the same underlying spatial placement of nodes was used to allow for easier visual comparison. R increases from left to right and S from top to bottom.

We can see clearly here that as R increases, edges get longer as they can span larger distances and therefore the number of possible and observed connections increases, as discussed in the previous section. For low R, networks are sparse and fragmented into many components as each node has only a limited number of potential interaction partners (see leftmost column of Figure 4.17). As R increases, components connect until the network consists of one component. This point is reached at around $R = 40$. At $R = 30$ (the value for R in the standard parameter setting) the majority of nodes belongs to a single connected component. However, some singleton nodes might still be present. At this point the diameter and characteristic path length are at their maximum as shown in Figure 4.18.

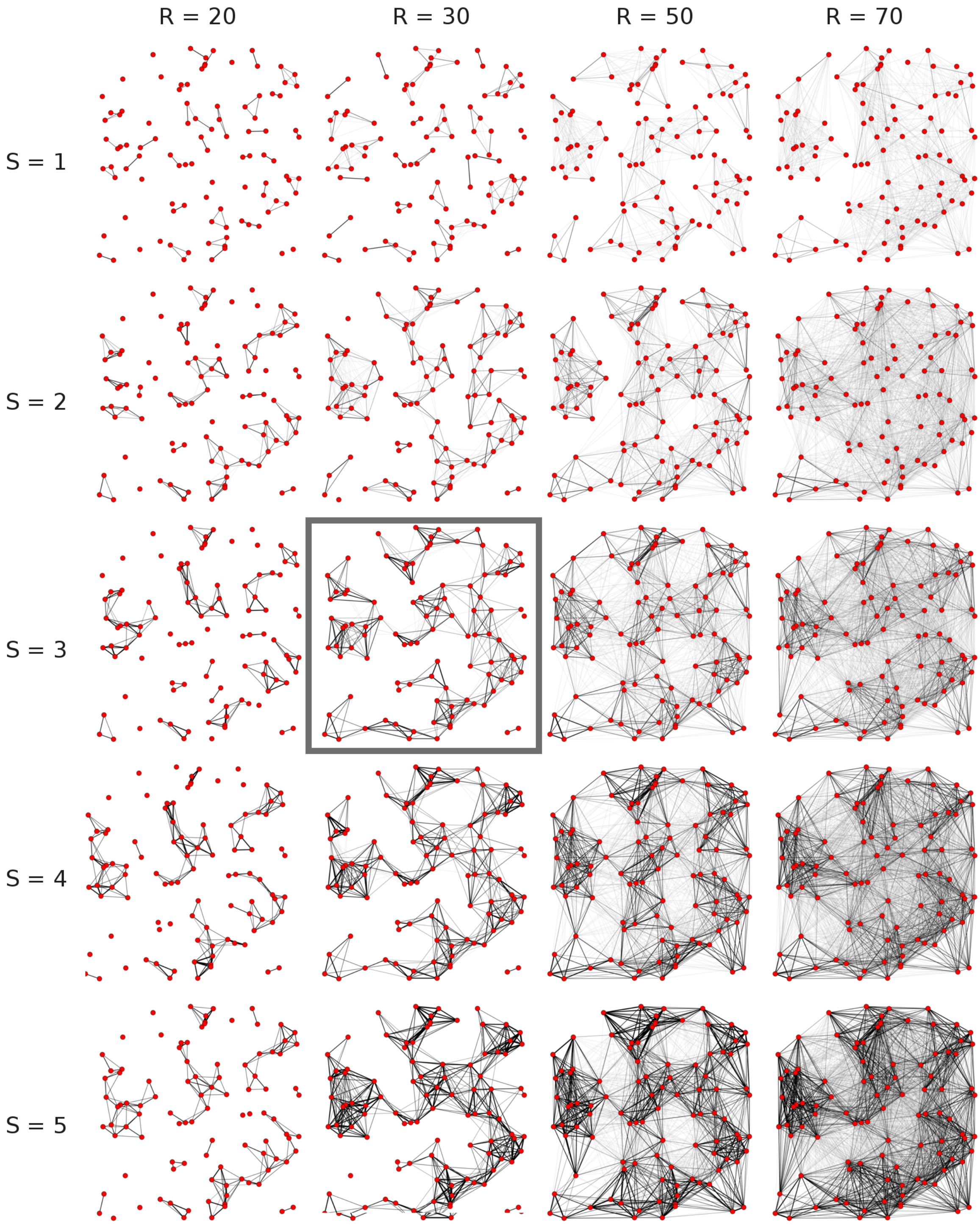


FIGURE 4.17: Resulting networks after $t = 10000$ steps for $H = 50$, $A = 0.75$ and varying R (columns, increasing from left to right) and S (rows, increasing from top to bottom), using the same placement of nodes for all runs. The network generated with the standard parameter setting ($R = 30$, $S = 3$) is highlighted by a grey box.

At this point the largest component spans the whole network but the social reach is still too small to allow direct connections between pairs of nodes on opposite sides of the space.[3]

With increasing R clustering increases. However, around $R = 40$ a small dip in clustering can be observed as previously isolated components connect without all triangles being able to close. Note that clustering stays fairly high in the system across all parameter settings shown, due to the triangle closure mechanism present.

Increasing the value of R further weakens the influence of spatial proximity until the networks lose the characteristics of social networks and eventually transition to random networks. As the model makes this transition, assortativity decreases until it becomes negative as shown in Figure 4.18 as the red dashed line. The value of R at which the network transitions to a random network decreases with increasing values of the history length H (see Figure 4.19).

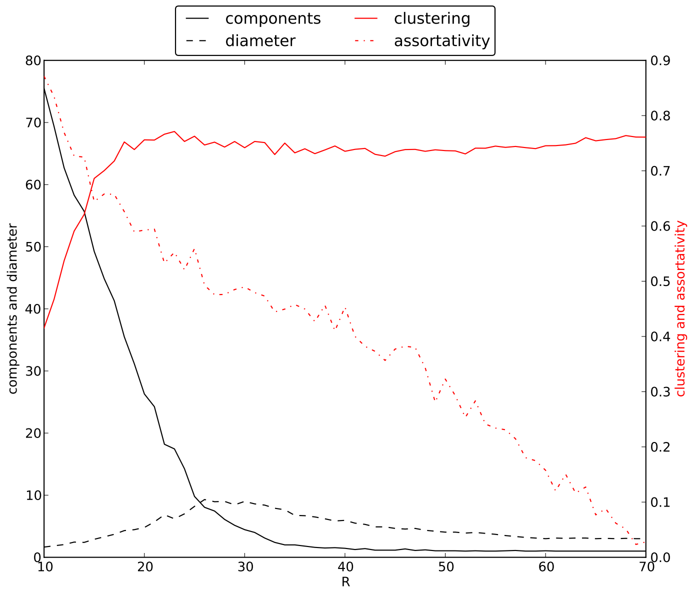


Figure 4.18: Topological measures changing in reaction to varying the social reach R (R ranging from 10 to 70, 20 repeats per data point). Assortativity and the number of components decrease with increasing R, whereas clustering and diameter increase until reaching a peak at around $R = 25$. At this point the reach is large enough for a connected network to form, but not large enough for a significant number of triangles and shortcuts to be created. Note that in contrast to Table 4.2, the calculation of the number of components does include singletons as we are considering parameter settings for which significant numbers of singletons exist. Other parameters were set as in the standard parameter setting.

Low values of H lead to more noise and greater fluctuation in the edge weights, increasing

[3]For lower R the network consists of several components meaning that the diameter and average shortest path length are not defined for the whole network. We therefore resort to calculating these measures for the largest component. As R increases components will merge, resulting in the largest components being bigger than before, thus leading to an increase in diameter and average shortest path length.

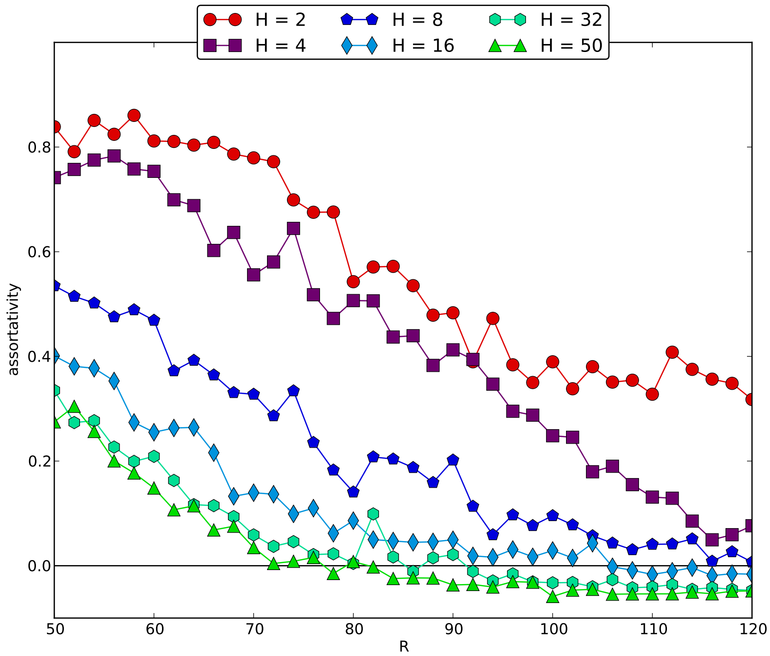
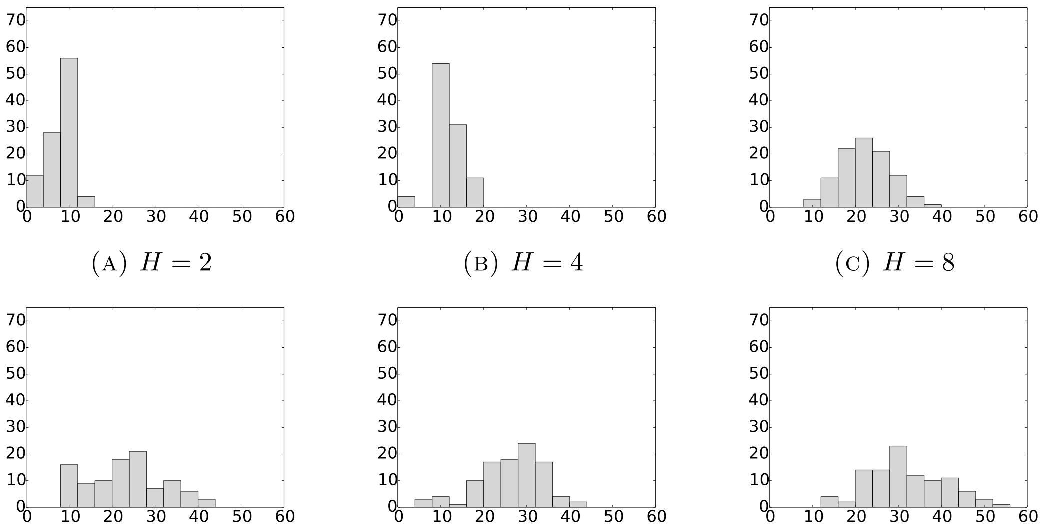


FIGURE 4.19: Mean of assortativity (calculated over 20 runs per data point) showing that the value of H influences the point at which the network transitions towards a random network, crossing to negative values of assortativity. R ranging from 50 to 120, shown for different values of H.

the probability for an edge being lost due to noise. Increasing H reduces fluctuations and therefore lowers this risk. In addition, increasing H allows each node to keep track of an increased number of relationships, as the minimum number of meetings required to maintain an edge is lower. This leads to a denser network relative to the a network with the same value of R (but lower H), which is why increasing H shifts the point of transition to a random network towards lower values of R. This increase in connectivity caused by H can also be observed as an increase in the maximum and average degree, resulting in a flattening of the degree distribution (as shown in Figure 4.20).

This means that for the standard value of $H = 50$ realistic-looking social networks with positive assortativity are obtained for intermediate values of the parameter R of approximately $20 < R < 70$, with the exact range depending on the values of the other parameters as well as the desired density of the resulting social network.

We have shown the effect of H on fluctuations in edge weight in Figure 4.7 (showing the dynamics of the edge weights of all edges attached to a particular node for one run) showing an example for $H = 50$ and $H = 5$. The higher value of H shows substantially smaller fluctuations. More fluctuation in the edge weights leads to more fragmentation events where a component is split in two. Since components cannot fuse, decreasing H leads to an increase in the number of components (see Figure 4.21). For very low values ($H = 2$) the network is fragmented into many components and displays almost maximal clustering since only fully connected cliques can persist in the presence of high



(A) $H = 2$ (B) $H = 4$ (C) $H = 8$

(D) $H = 16$ (E) $H = 32$ (F) $H = 50$

FIGURE 4.20: Degree distributions for different values of $H = 2, 4, 8, 16, 32, 50$. For all other parameters the standard parameter setting was used. The increase in connectivity for higher values of H results in an increase in the maximum and average degree, resulting in a flattening of the degree distribution.

noise. Note that for the standard parameter setting of $H = 50$ edge weights are fairly stable and accidental fragmentation is unlikely to occur. We have chosen a fairly large value of H to ensure that any fragmentation effects observed are not caused by edge weight fluctuations. This will be crucial when we investigate fragmentation processes of networks.

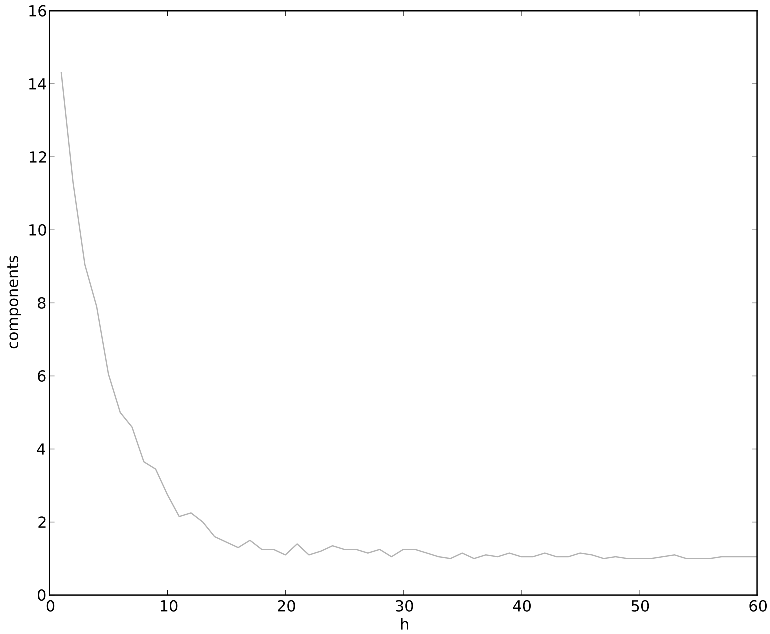


FIGURE 4.21: Number of components decreasing for increasing values of H (memory length). Note that in contrast to Table 4.2, the calculation of the number of components does include singletons. Each data point represents the mean taken over 20 data points.

4.4.4 Balance between degree and strength

Increasing R or H increases the number of edges in the network[4] and therefore results in a higher average degree. In contrast, average node strength is primarily constrained by the parameter S. Figure 4.22 shows that increasing R only increases the average node strength up to a certain point at which it is limited by S.

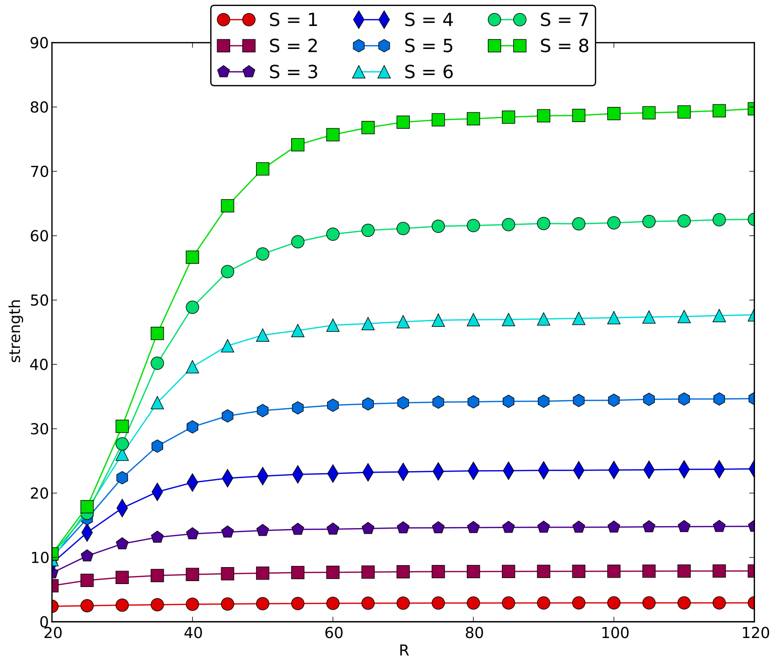


FIGURE 4.22: Strength in reaction to varying R for different values of S. The average strength increases for increasing R, with the maximum value reached being determined by the number of slots available per timestep (S). Each data point shown is the mean over 20 repeats.

As discussed before, the maximum weight of each edge is $S + 1$, therefore increasing S increases the possible strength of the nodes. This effect can be seen in the bottom row of Figure 4.17 where the edges are plotted thicker in proportion to their increasing weights. For the standard parameter value of $R = 30$ increasing S from 1 to 2 and then 3 first leads to an increase in degree. This increase in connectivity also leads to a decrease in the number of components. Increasing S further to 4, 5 and 6 has a much smaller effect on the degree as the number of nodes within reach is limited by $R = 30$. Therefore, the increase in "socialising time" leads to an increase of the average edge weight[5] as the extra time can not be spent on creating new relationships, so instead it is spent strengthening existing ones.

This interplay between increasing socialising options (controlled by R) and opportunities (limited by S) on existence and weight of edges is highlighted in Figure 4.23. On the left-hand side the number of edges present in the network is shown (mean over 20 repeats

[4]As discussed in the previous section, R has a much larger influence on this than H.
[5]The increase in average edge weight can also be observed as an increase in node strength.

with the same parameter setting). The number of edges increases with R, and within the limits set by R, S controls how many connections can be maintained by imposing time constraints (as discussed in Section 4.4.2). The number of edges increases with both R and S. On the right side the average weight of an edge for different combinations of R and S is shown and here we observe that the average weight is highest for smaller values of R. In this case there are only few socialising options but many opportunities, meaning that individuals have a very limited set of individuals to interact with and therefore increase in socialising time can only lead to stronger connections but not more.

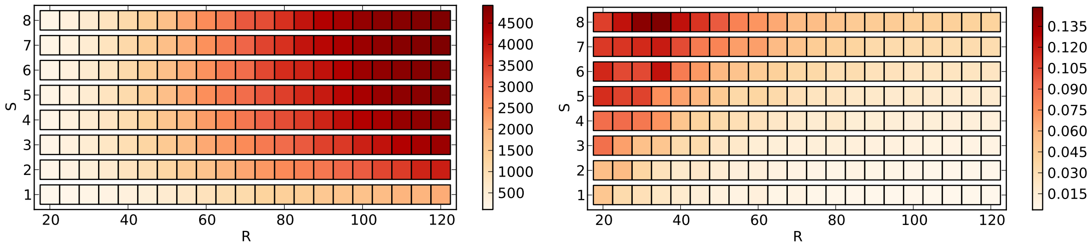


FIGURE 4.23: Number of edges (left) and average edge weight (right) for varying values of R and S. Higher values are shown in deeper shades of red. The number of edges increases for increasing R and S, while the average edge weight is highest for large S but small R. This is because in this case each individual is only connected to a few other individuals but has a large time budget for socialising. Due to the limited socialising options, socialising time is generally spent on the same individuals, leading to high edge weights.

4.4.5 Conclusion

In this chapter we have shown that the model introduced in Chapter 3 produces networks with the topological properties defining real social networks: community structure, high clustering, positive assortativity with respect to degree and short characteristic path length. The spatial restrictions imposed by only allowing interactions within a finite social reach are crucial for generating these structures. Together with the triangle closure mechanism present in the model this leads to networks that share some properties with random geometric graphs. We have noted that the generated networks are more complex than random geometric graphs as triangle closure allows for some edges to form that depend on the presence of other edges. As we will discuss in the next chapter, the presence of these dependencies between edges is crucial for understanding the dynamics that lead to the type of network fragmentation that is the focus of this thesis.

Furthermore we have shown that the system reaches a quasi-stable equilibrium with macro-level topological features reaching an equilibrium while on the micro-level the exact topology defined by the edge weights continues to fluctuate. This means that the system reaches a stable state but can still react to perturbations as a result of ongoing dynamics on the micro-level. This is very important if we want to model fragmentation as a result of a small perturbation. If the system would not reach equilibrium, we could not attribute fragmentation to an applied perturbation. If the system reached a stable equilibrium that was not accompanied by ongoing low-level edge dynamics, then it would not be able to react to perturbations, rendering it useless for investigating the phenomenon that we are considering here.

In the next chapter we will discuss the nature of the perturbation used and the reaction of the system to it.

Chapter 5

The impact of social fallout on topology and weight structure

5.1 Introduction

In this chapter we focus on fragmentation of social networks and will show how our model can shed light on this phenomenon. It has been observed in the real-world that a disagreement between a pair of nodes can have consequences that lead to a previously connected social network fragmenting into two disconnected parts (Zachary, 1977). While described as a potential problem for some networks and a potential opportunity in others, not many attempts have been made to propose an underlying mechanism responsible for this phenomenon. Here, we propose such a mechanism and show that it can lead to fragmentation.

In the previous chapter we have shown that the model presented in Chapter 3 has the characteristics required to study fragmentation dynamics of social networks: the system reaches a quasi-stable equilibrium, where the characteristic features of the generated networks remain stable, while on the lower level edge weights continue to change. We have discussed why these features are crucial for modelling fragmentation dynamics.

In this chapter we will show that a disagreement between two people can be a sufficient perturbation to fracture the network. This fragmentation happens even though all nodes (apart from the two nodes having the disagreement) continue to act according to the same behavioural rules as before. In cases where the network does not fracture, we can nevertheless observe a significant impact within a confined neighbourhood and we will introduce measures that are able to quantify this impact. Furthermore, we discuss edge measures that could be predictors for identifying the edges most critical for maintaining cohesion of the social network and how well different measures can predict the observed impact.

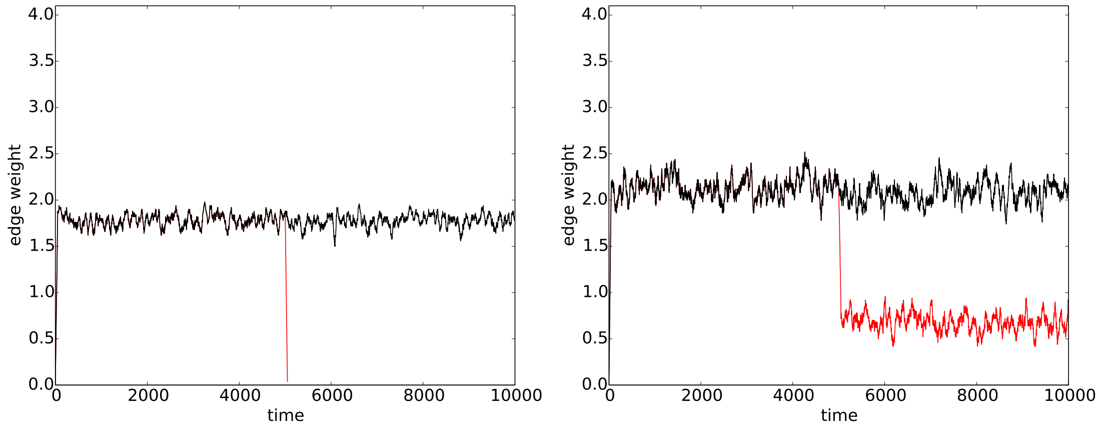
5.2 Implementation of social fallout

The aim of the model presented here is to allow investigation of the impact that a disagreement (or *fallout*) between two people can have on network topology. We model a fallout event between two individuals by running the base model presented in Chapter 3 until it has reached equilibrium and then introducing a perturbation by changing the affinity value associated with a pair of nodes to 0. This means that the two individuals cease to maintain their connection. We do not directly remove any edges or change the behavioural rules of any of the individuals. Any edges broken or changing in weight are a result of the perturbed dynamics that the change of one affinity value introduces into the system. It is important to allow the system to settle into a quasi-stable state before introducing the perturbation, as in this state major topological changes due to noise are extremely rare. Therefore, we know that any major topological changes are a result of the perturbation.

The effect of the fallout on the system is not trivial. Considering the weight of the edge between the two individuals involved in the fallout highlights the complexities in the system. We would expect nodes to break their connection after a disagreement, but, while in some cases this is indeed the case, in others the edge between the nodes that fall out weakens, but persists. This can be explained by the fact that the relationship is embedded within the context of the social network. Even though the two individuals in question no longer actively maintain their relationship, they might still be forced into contact through mutual friends. Note that in the case of a strong disagreement, people might actively avoid each other to an extent where they will refuse to attend any gathering that brings them into contact. While this certainly does happen in the real world, we do not model this here. We want to show that fragmentation is possible even though individuals do not change their behaviour, thus it is essential that the behavioural rules remain unchanged. To what extent this new avoiding behaviour would influence dynamics would be a separate question.

Examples for the two possible cases (the fallout edge disappears or it persists with a lower weight) are shown in Figure 5.1 where the edge weight of the fallout edge is depicted over time. The fallout perturbation is introduced at $t = 5000$ and the seed used to initialise the random number generator is the same for both fallout and base run, therefore their trajectories are identical up to the fallout point. We can see that in the case shown in Figure 5.1a the edge weight of the fallout edge rapidly drops to 0, meaning the edge is broken as a result of the fallout. In contrast, in Figure 5.1b the edge weight drops steeply as well but to a non-zero value as the edge continues to be maintained indirectly through mutual neighbours of the nodes as a consequence of both the directly involved nodes attending the same gatherings hosted by mutual friends.

The examples shown here consist of a single run. If we want to ensure that the effects observed are due to the introduction of a perturbation and not noise, we need to quantify



(A) Edge weight of fallout edge drops to zero following fallout (B) Edge weight of fallout edge drops to non-zero value following fallout

FIGURE 5.1: The edge weight of the fallout edge over time. In black, we show the edge weight for the base case without fallout, in red we show the edge weight for the fallout case.

the level of noise present in equivalent runs in order to decide whether an observed effect is due to chance or an underlying systematic trend. As discussed in Section 4.1.3, fluctuations due to noise have a strong effect during the initial phase, leading to variability in the resulting equilibrium topology. So while it is necessary to assess any fallout impact across different sequences of events, it would not be very useful to compare runs with completely independent sequences of random numbers as the same edge might not be present in different runs and affinity values would be distributed differently. We can, however, make use of the fact that once the system has reached the quasi-stable equilibrium, the effect of the randomness in edge fluctuations is less pronounced with regard to its effect on the overall topology. On the local level, the edge weights will differ over time but the same edges will be present in general, allowing us to assess whether an edge disappearing is due to noise or can be attributed to the fallout.

Therefore all results are ensembles of runs with the same initial seed up to the fallout point $t_{\text{fo}} = 5000$. This ensures that all runs follow the same trajectory initially, resulting in exactly the same topology at $t = 5000$, the point where the fallout is introduced. Note that for $H = 50$, as per the standard parameter setting, the system has been in equilibrium for a long time. Running the model for a long period allows for any spontaneous fragmentation due to unstable topological features to occur before we introduce the fallout. At $t_{\text{fo}} = 5000$, the seed of each instance within the ensemble runs is reset to a new, individual value, allowing the runs to diverge due to stochastic effects. In Figure 5.2 we show the same information as in Figure 5.1 but for an ensemble of 10 runs for both the base case and fallout case. The base case runs are shown in shades of grey and the fallout runs in orange-red. We can see here that the individual runs diverge somewhat after being reseeded with different seeds at $t = 5000$. However, base case and fallout trajectories remain clearly separated, meaning we can attribute the gross difference in

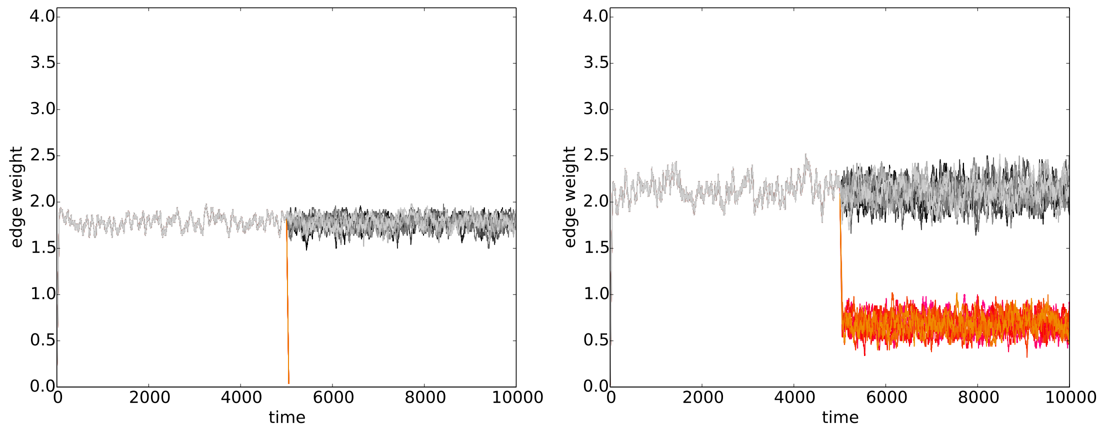


FIGURE 5.2: Analogous to Figure 5.1 this figure shows the weight of the fallout edge over time. For both the base case and the fallout case 10 different runs are shown. The effect observed in Figure 5.1 is still visible, with the edge decaying to 0 on the left hand side and to a non-zero value on the right side.

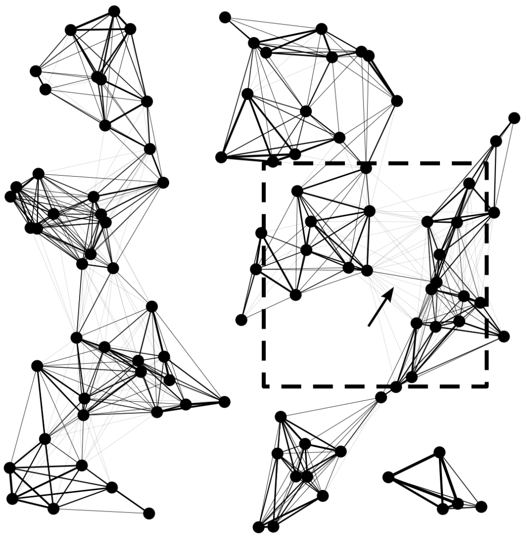
edge weight to the perturbation. Averaging over 25 different networks we observe that in 12% of the cases the fallout edge is always lost (in each of the 50 runs), in 85% it is never lost and in the remaining 3% it is lost in a proportion of runs within an ensemble.

As we can see from these examples, edge weights change rapidly to adjust to the new conditions. To measure the impact of the fallout, we consider the networks resulting at $t = 5500$, which is 500 timesteps after the fallout. We use the standard parameter setting of $R = 30$, $A = 0.75$, $S = 3$ and $H = 50$, meaning that this timespan is equivalent to 10 history lengths.

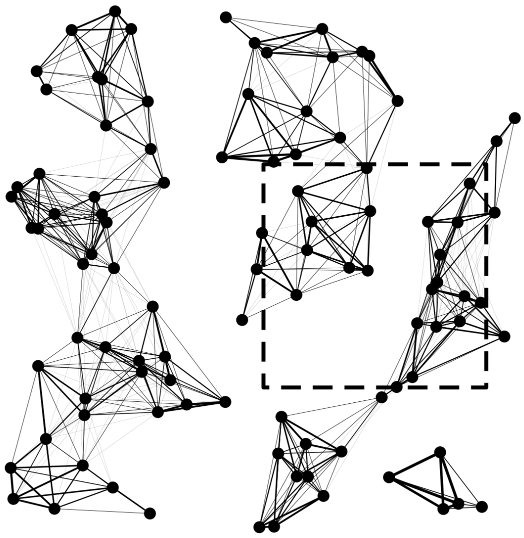
5.3 Effects of fallout

5.3.1 Network fragmentation

We will now look at the effects the fallout on the network topology. First, and most importantly, we will show that this process can indeed lead to fragmentation, which confirms our main hypothesis: that previously stable social networks can fracture even though only one relationship is directly perturbed and all other individuals continue to behave as before. Networks do fracture as a result of the fallout, implemented as described above. An example is shown in Figure 5.3. In this figure we show the network at $t = 5500$ for a run with fallout (Figure 5.3b) and for a base run where no perturbation is introduced (Figure 5.3a). To reiterate, both runs are identical up to the point where the fallout occurs ($t_{\text{fo}} = 5000$) and each run (in both cases) is reseeded with a new seed at that point. As we can see here, one of the two large components in the network has fragmented into two components in the fallout case (but not in the base case). In the fallout case the edge connecting the two disagreeing individuals (marked by an arrow in the left subfigure) has disappeared together with edges surrounding it. We will discuss the mechanisms that leads to this later in this chapter.



(A) base case at $t = 5500$



(B) fallout case at $t = 5500$

FIGURE 5.3: An example for network fragmentation. On right hand side we see the network at $t = 5500$, 500 steps after the fallout. The fallout edge is indicated by the black arrow. One of the two components is split into two for the fallout case. For comparison the network for a base case run without the fallout is shown on the left hand side, where no fragmentation occurs. The area where the fragmentation occurs is marked by a black box.

Whether the network fragments or not depends on which edge is the fallout edge. Later in this chapter we will discuss measures that allow us to predict which edges are crucial for the network to stay connected. For each network we simulate a fallout for each edge that is present in the network at $t = 5000$. We run each fallout independently, meaning we run a separate iteration of our model for each fallout simulated. For each fallout edge we run 50 reseeds with fallout and 50 base cases without. Fragmentation following fallout is generally only observed for some edges in the network. For these edges, fragmentation will often be observed in all 50 repeats. However, for some edges fragmentation only occurs in a subset of the fallout runs. Note that we almost never observe fragmentation in any of the base cases, meaning we can attribute the fragmentation to the fallout and not to noise in the system. A single network for which spontaneous fragmentation was observed in the base case has been excluded from our analysis to make sure any fragmentation observed is due to the fallout.

5.3.2 Networks can fracture at different points

Generally networks have more than one point that is vulnerable to fragmentation. This means that for most networks we can observe several fragmentation patterns, depending on the pair of nodes between which the fallout occurs. In Figure 5.4 we show an example for this. In Figure 5.4a the network is shown for $t = 4999$, just before the fallout occurs. In the remaining subfigures, different fragmentation patterns found for this network are shown at $t = 5500$. As we can see, the already existing components can split further at different points, depending on which edge is the fallout edge. The component on the left has several potential fracturing points and a small component at the bottom left has two fracturing points. Note that in this example the largest component is stable and has no fragile points.

The sizes of the resulting components vary. In Figure 5.3 we observed a component splitting into two components of roughly equal size. In other cases, such as shown in Figures 5.4e and 5.4f, small fringe components containing only one or a few nodes split off. Often these are located in the periphery of the network where socialising options are sparse. This shows that there is variety in the effect of a fallout. How we rate the impact of a particular effect depends on our focus. If we are concerned with impact on the information flow in the network, then the largest component splitting in two removes many communication paths and therefore has a severe impact on the function the network serves. In contrast, a small fringe component splitting off has only a minor effect on information flow across the whole network. However, if we are concerned with the impact on individuals, then a small group or a single individual splitting off isolates them whereas in the case of a larger component splitting into two still large components breaks some edges but all nodes involved retain some contacts.

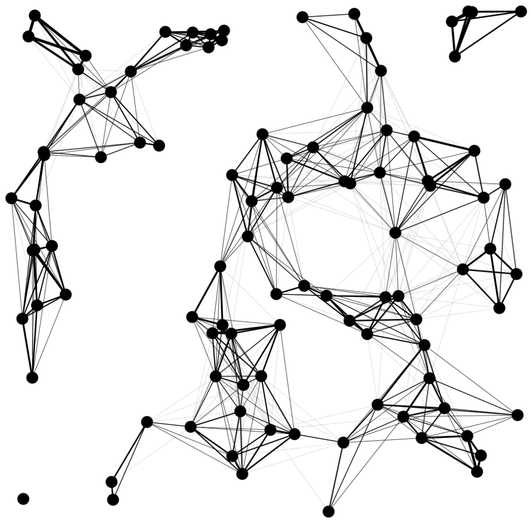
5.4 Local fragmentation

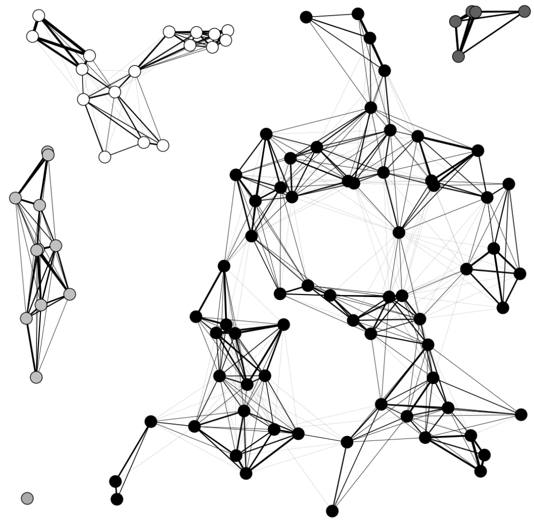
Fragmentation is the most obvious and drastic result of a fallout for the network topology. However, even in cases where fragmentation does not occur, the fallout can have an impact on the local topology. It might lead to edges within a component being broken even though redundant paths keep the component connected. This could lead to local fragmentation where only a much longer path maintains cohesion of the module. An example is shown in Figure 5.5 where local fragmentation occurs in the lower part of the network. However, the number of components stays the same since redundant, albeit much longer, paths exist.

5.5 Non-fragmentation impact

In cases where a perturbation results in neither global nor local fragmentation, it can still lead to significant rearrangement within an area of the network. Even in the case where no edges are broken significant adjustment of edge weights can follow the fallout.

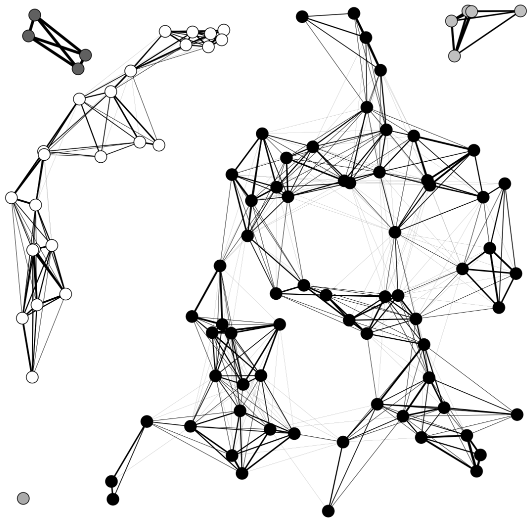
To capture these effects and measure the effect of fallout on the edge weights, we compare

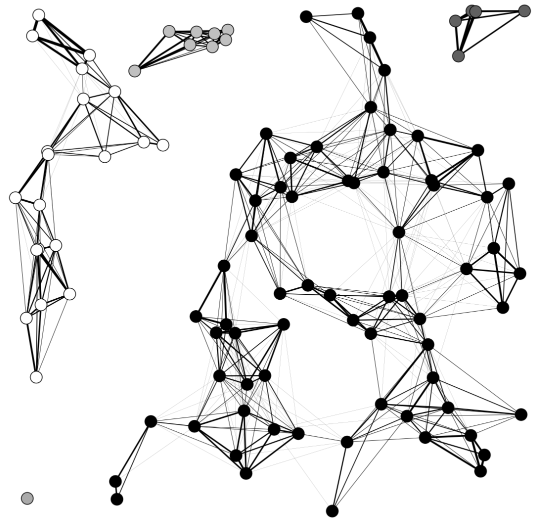




(A) Network before fallout at $t = 4999$

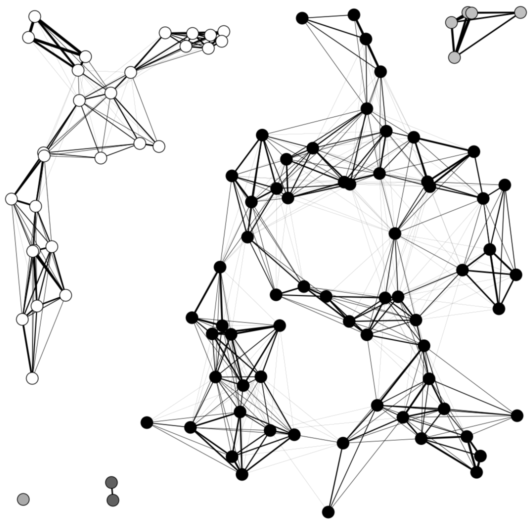
(B) Left component fractures in the middle

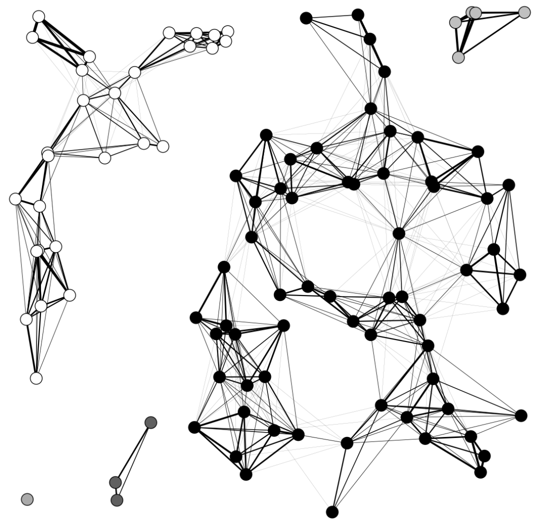




(C) 4-clique on the left splits off

(D) Left component splits at the top

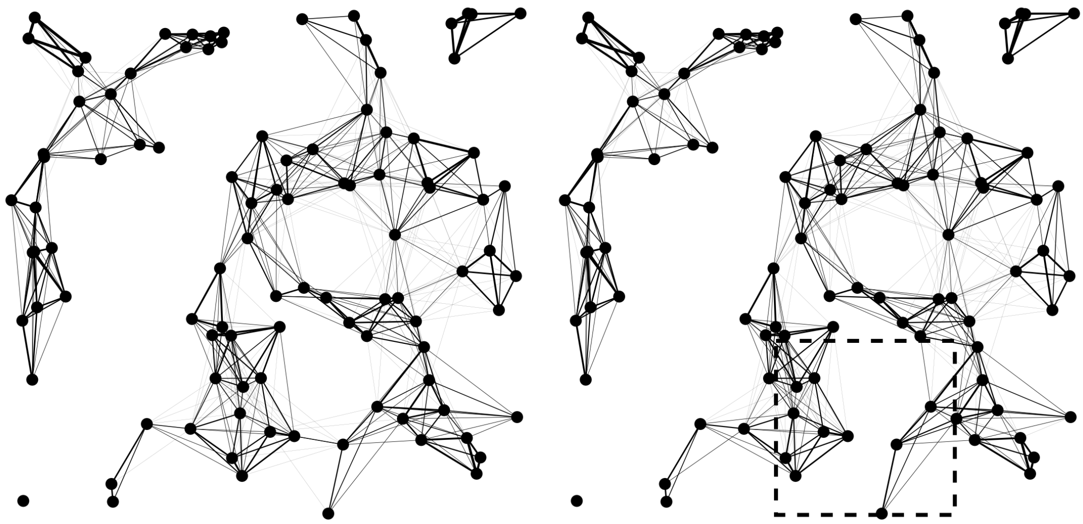




(E) Small component consisting of two nodes at the bottom left is formed.

(F) Small component consisting of three nodes is formed.

FIGURE 5.4: Depending on which pair of nodes falls out, the network can fracture in different places. Here we show the initial network for comparison together with five different observed fragmentation patterns.



(A) network at $t = 4999$ (B) network following fallout at $t = 5500$

Figure 5.5: For the resulting network presented on the right hand side the fallout does not result in an increase in the number of components. However, we observe that on the local level two previously connected modules of the largest component are now only indirectly connected, leading to local fragmentation. On the left hand side we show the network before the fallout for comparison.

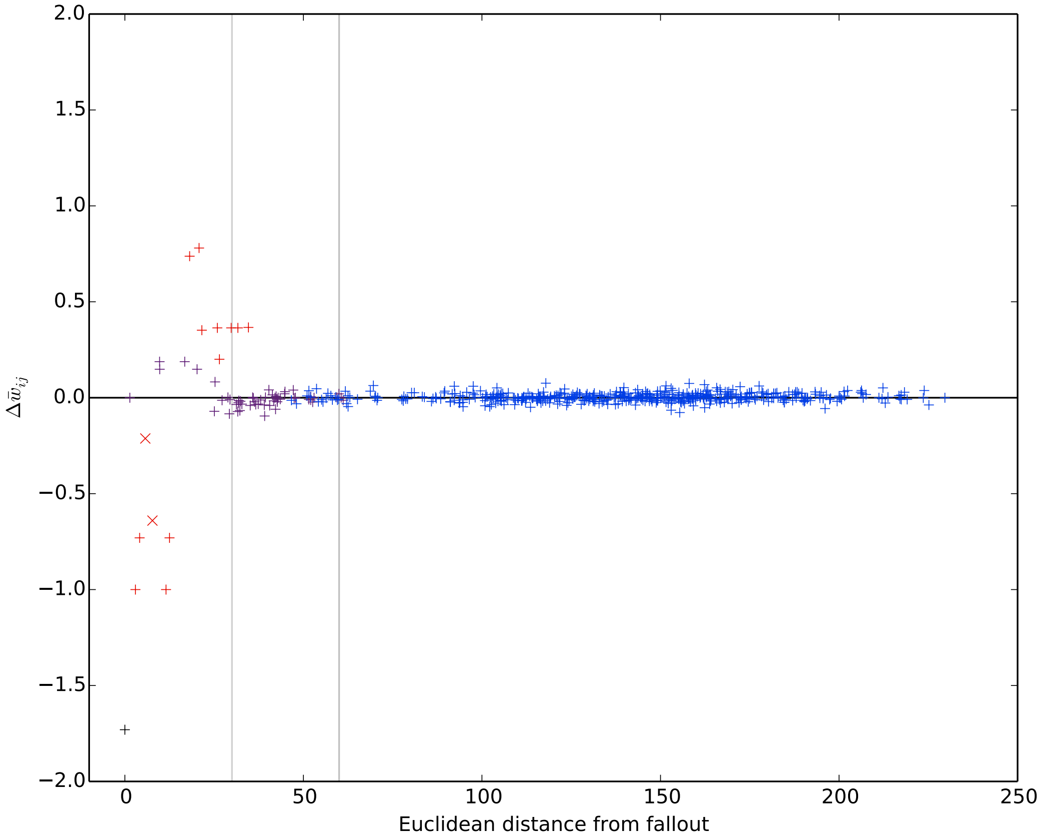
edge weights in the resulting network for the base case and the fallout case. To do so we calculate the difference in weight at $t = 5500$ between fallout and base cases for each edge that is present at the time of fallout. Due to the stochasticity in the model, edge weights fluctuate. The attenuate the noise introduced by this we run ensembles of runs by reseeding at the point of fallout as described in Section 5.2. For both fallout and base case we run 50 repeats with different random seeds for reseeding, meaning the runs are the same up to the fallout point when their trajectories diverge.

We then calculate the difference in average weight between fallout and base case for each edge fallout. To do so we take the mean over all 50 repeats for the fallout case ($\bar{w}^f_{ij}$) and subtract the mean weight calculated for the 50 base reseeds ($\bar{w}^b_{ij}$):

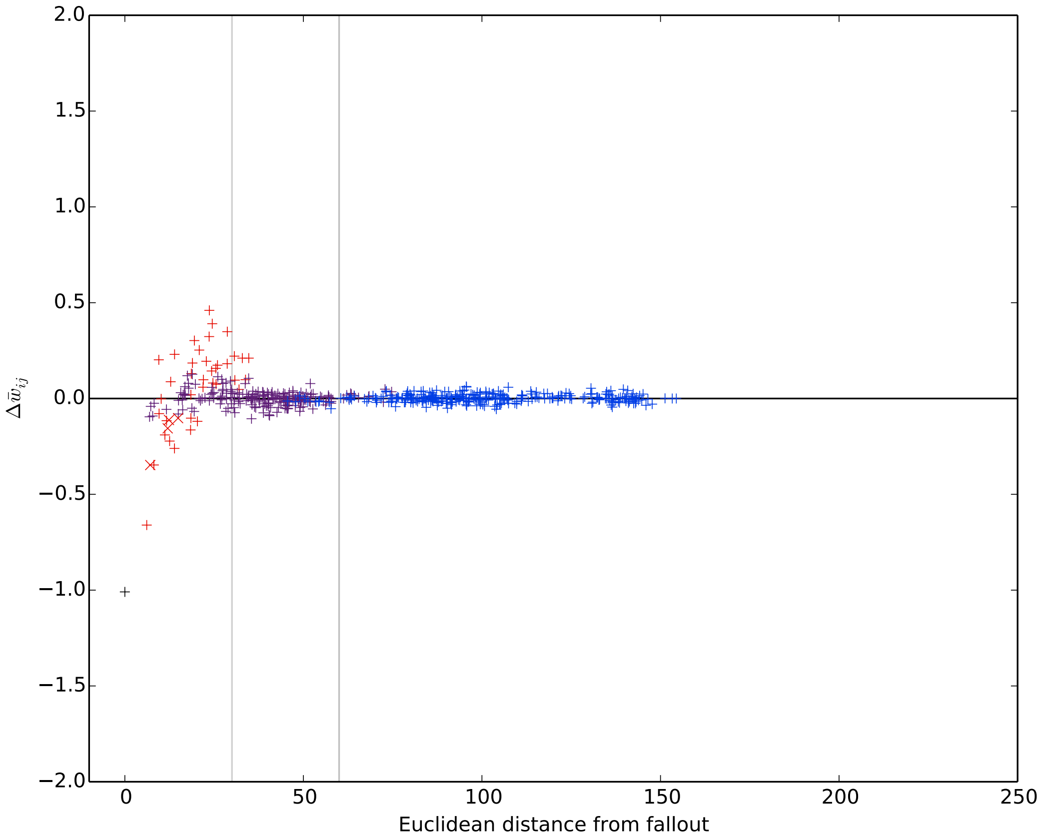
$$\Delta \bar{w}_{ij} = \bar{w}^f_{ij} - \bar{w}^b_{ij}. \tag{5.1}$$

Positive values indicate that an edge's weight has increased as a result of the fallout whereas negative values indicate a decrease. For cases where an edge was broken in the fallout case (in one or more repeats) we assumed its weight to be zero for any repeats where the edge was not present.

In Figure 5.6, we show the differences in weight calculated as described in Equation 5.1 for all edges that were present before the fallout at $t = 4999$. Each point shown corresponds to an edge in the graph. Each figure shows the effect for a different fallout edge (edge $(24, 28)$ in Figure 5.7a and edge $(58, 72)$ in Figure 5.7b). The y-axis represents the magnitude of edge weight differences and while the x-axis corresponds to the distance



(A) edge $(24, 28)$



(B) edge $(58, 72)$

FIGURE 5.6: Edge weight differences resulting from two fallouts. Edges that were present in the base cases but not in any of the fallout repeats are shown by the symbol "×". See text for details and explanation of the colour coding.

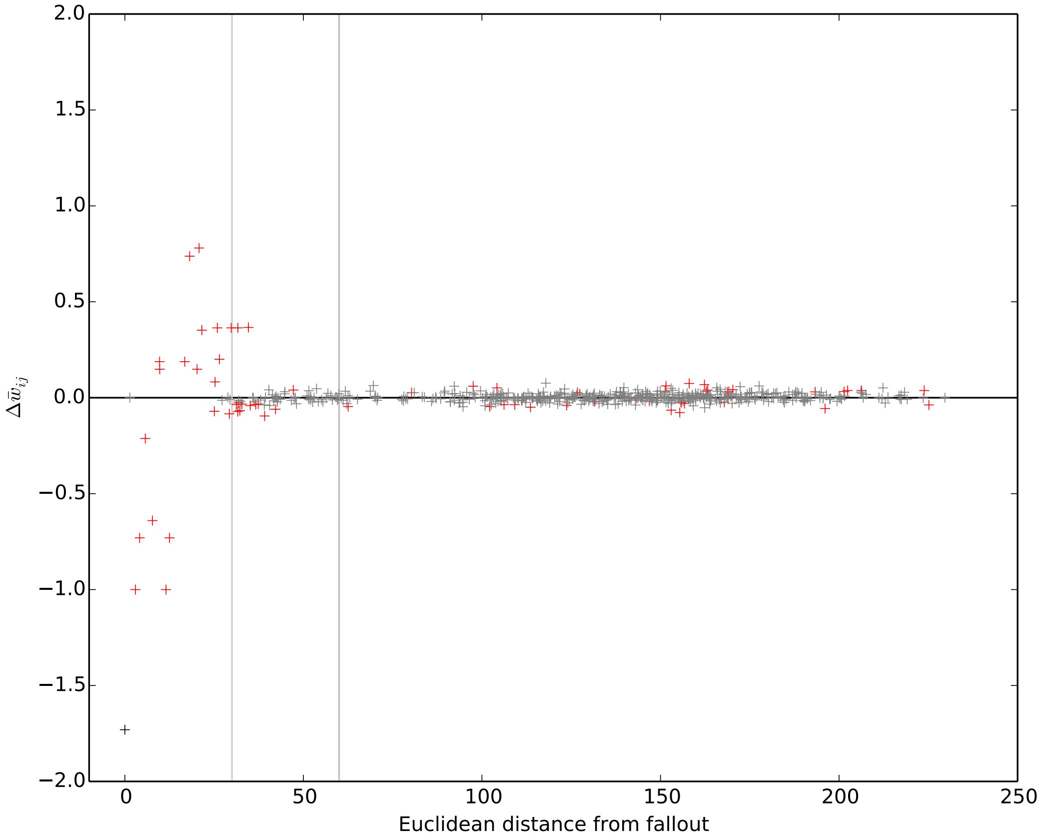
of an edge to the fallout edge (meaning the edge at $x = 0$ is the fallout edge itself). We calculate the distance between edges as the Euclidean distance between the midpoints of the edges. The distances R and $2R$ (corresponding to the limits for direct and indirect interactions) are shown as grey vertical lines to give an indication of the scale.

In a spatially embedded network, edge-distances can be measured in two ways: distance in space as discussed above or geodesic distance on the network by measuring the number of edges on the shortest path from one edge to another. Here, we will refer to this as hop distance. Edges that share an endpoint are one hop away from each other. Edges whose shortest path goes through an edge that is one hop away have hop distance 2 and so forth. Hop distance is correlated with spatial distance in our model due to the spatial restriction imposed. However, depending on the topology, edges can be close to each other in space but several hops apart in the network. Even for edges located very close together on the map there might be no path at all between them if they are in different components. In Figure 5.6, we visualise the hop distance of an edge to the fallout edge as the colour of the corresponding data point: the fallout edge itself (with hop distance 0) is shown in black, edges with hop distance 1 to the fallout edge (meaning they share a node with the fallout edge) are shown in red, edges with hop distance 2 in purple and any remaining edges (with a hop distance of more than 2) are shown in blue.

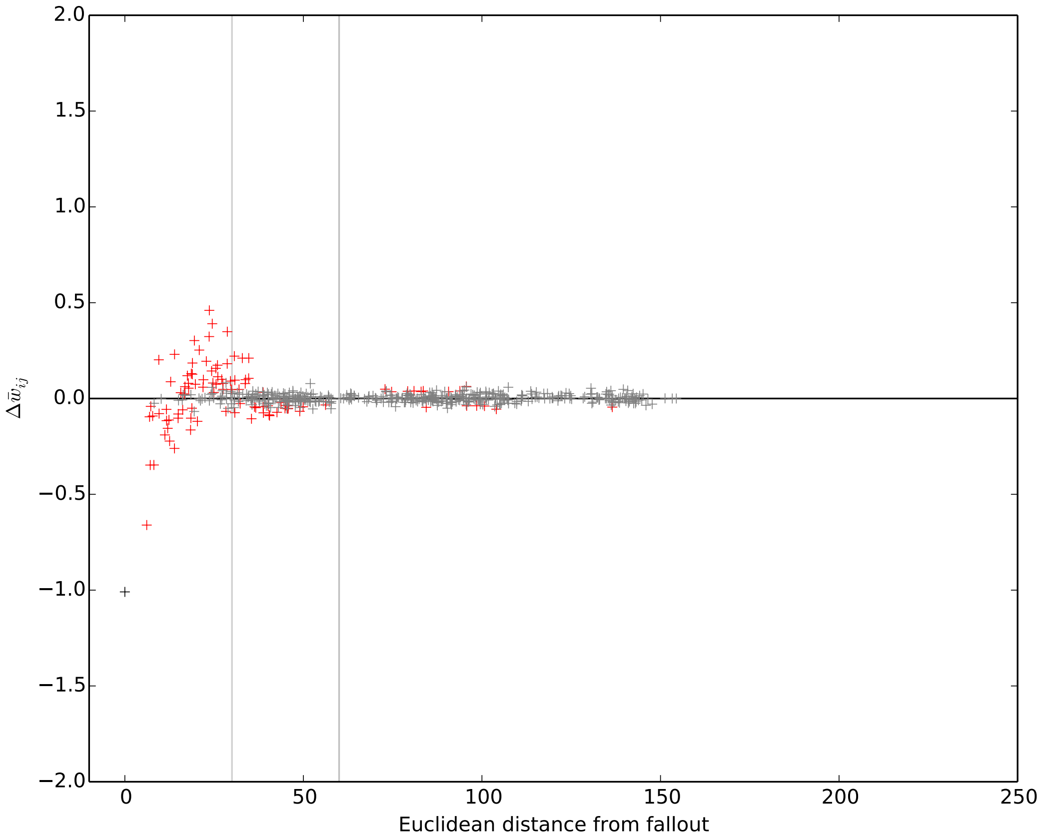
Edges that were present in the base reseeds but not in any of the fallout repeats are shown using the marker "×" all other edges are shown using a "+". Since these edges are not present in any of the fallout repeats but can be found in the base repeats, their disappearance is very likely to be a result of the fallout. The number of edges broken as a result of the fallout is generally low, but the distribution has a long tail (see Figure 5.8). Note that we do not observe the reverse: edge formation as a result of the fallout.

In Figure 5.7 we show the same data used in Figure 5.6, with points coloured according to whether the weight difference between fallout and base case is statistically significant. To determine whether this is the case, we pair each of the 50 observations of an edge weight in the base case with one of 50 observation in the fallout case. We calculate the mean of this distribution of weight differences (as done previously) and the confidence intervals for this mean. Weight differences for which the confidence interval does not include 0 are shown in red, all others in grey. We observe that the majority of edges that exhibit a significant weight difference between fallout and base case are found in the vicinity of the fallout edge.

If we consider the change in edge weights as plotted in Figure 5.6, then we can see that larger changes in edge weight are limited to edges that are located within distance $2R$ of the fallout edge. Moreover, the edges that exhibit the largest change in weight are edges that share a node with the fallout edge (shown in red). This means that both spatial proximity on the map as well as geodesic proximity influence the effect the fallout has on a particular edge. This means that the effect of a fallout is localised to an area



(A) edge $(24, 28)$



(B) edge $(58, 72)$

FIGURE 5.7: Edge weight differences resulting from two fallouts. Points shown in red show a significant difference between fallout and base case. Edges exhibiting a significant weight change as a result of the fallout are mostly located in the vicinity of the fallout.

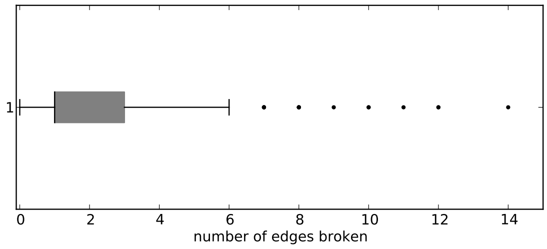


FIGURE 5.8: Number of edges broken following the fallout. Distribution shown over all edges for a particular network. The number of edges broken as a result of the fallout is generally low, but the distribution has a long tail.

around the fallout. We will discuss this later in more detail. The fallout edge itself (at distance 0) shows the largest change in weight. The fallout leads to a negative change in average edge weight for the edges closest to the fallout edge, close in terms of both distance measures, map distance and hop distance. For edges further away, we observe positive changes in edge weight, meaning the average edge weight is higher in the fallout case than in the base case. These positive changes in edge weights highlight how the system adjusts and re-balances itself following a fallout.

5.6 Rearrangement of edge weights following perturbation

Changing the affinity value of a pair of individuals from 1 to 0 means the nodes in question will no longer actively maintain the connection between themselves. This leads to a decrease in edge weight, either to 0 or a non-zero value lower than the previous weight, for both the fallout edge itself (as discussed in Section 5.2) as well as edges that were indirectly maintained through the fallout edge. This decrease in edge weight corresponds to data points with negative values in Figure 5.6.

The two nodes involved in the fallout now have additional time available since they no longer invest any time in maintaining the connection between them. The time freed up can now be allocated to the maintenance of other contacts, leading to an increase in the weights of edges to other contacts. This effect is visible in Figure 5.6 as positive values. This balancing effect becomes obvious when visualising the magnitude of edge change by colouring the edges of the network according to the magnitude of edge weight change observed. In Figure 5.9 the local neighbourhood of the fallout is shown and edges are coloured according to the magnitude and direction of weight difference between fallout and base cases. Edges with a decrease in edge weight are shown in red (darker shades indicate a larger change) and edges for which an increase in edge weight is observed are shown in shades of blue.

The fallout edge $(24, 28)$ (nodes 24 and 28 shown in red), shows the largest difference in weight (decrease in weight, therefore coloured red). The edges from node 28 to mutual

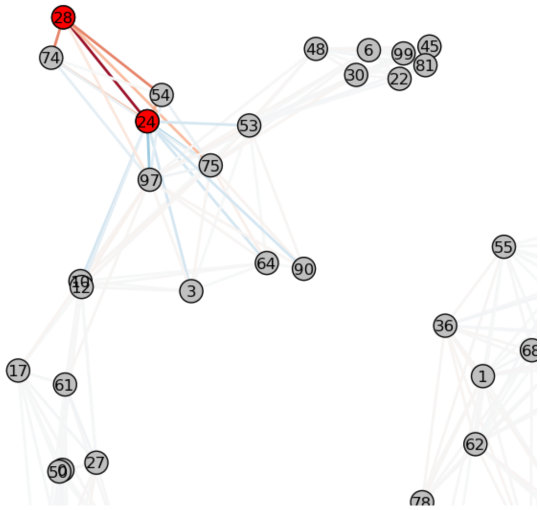
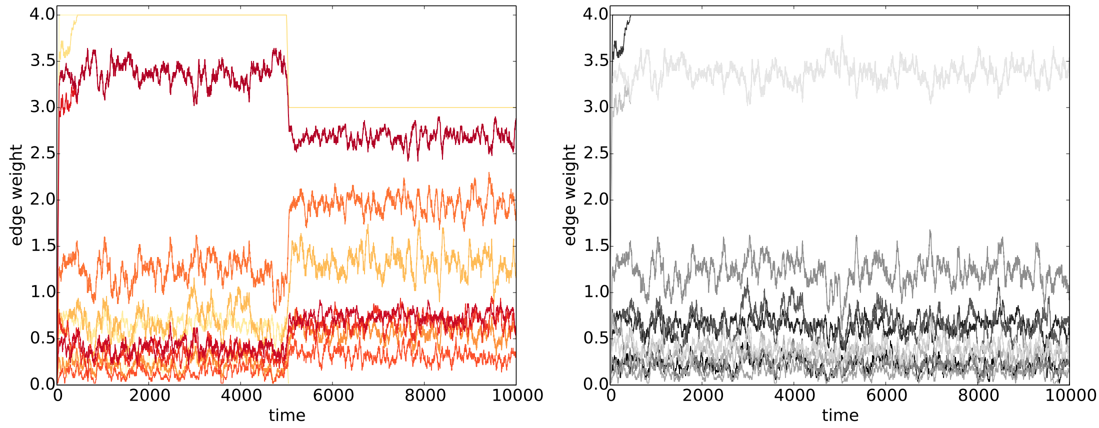


Figure 5.9: Magnitude and direction of edge weight changes for neighbourhood of the fallout. The fallout nodes (24 and 28) are shown in red here. Negative changes indicated by red edge colour, edges with increased edge weight as a result of the fallout are coloured blue. Darker shades indicate larger magnitude of edge weight changes. Some of the edges decrease in edge weight, most visibly the fallout edge itself, coloured in dark red as it shows the largest magnitude of weight change. This negative change of edges decreasing in strength results in social capacities being freed up which leads to other edges increasing in weight.

friends of 24 and 28 also exhibit a decrease in weight, as they loose the indirect maintenance of meeting node 28 at node 24's gatherings. Other edges experience an increase in edge weight (indicated by blue colour). Following the fallout, node 24 has additional time to invest in strengthening some of its connections since it no longer requires part of its time budget to maintain the edge to node 28. Note that by investing time to strengthen an existing link by more frequently attending an individual's gathering, connections to mutual friends are strengthened too, through meeting at the same gathering. Therefore, negative changes in weights are balanced out by positive changes. The net effect of the fallout is generally small and can be positive or negative. This depends on the topology surrounding the fallout edge and therefore highlights the non-linear nature of the system studied.

We can observe this balancing effect of some edges decreasing in weight while others increase by showing the change of edge weights over time for selected edges. In Figure 5.10 we show all edges that are attached to either of the fallout nodes. The network, fallout edge and local neighbourhood used are the same as shown in Figure 5.9. On the left hand side of Figure 5.10 we can see that following the fallout at $t = 5000$ the system rapidly adjusts to a new quasi-stable steady state. This readjustment consists of some edges decreasing in weight, while others increase. For comparison, we show the weights for the same edges in the base case, where no perturbation is introduced and we can see that no rearrangement happens.



(A) edge (B) base

FIGURE 5.10: Weights over time for all edges attached to one of the two nodes involved in the fallout. On the left side the edge weights are shown in the fallout case. The fallout occurs at $t = 5000$ and we observe subsequent changes in edge weights for the edges that share a node with the fallout edge. Note that the change can be positive (increase of edge weight) or negative (decrease in weight). On the right side the base case is shown for comparison.

5.7 From fallout to fragmentation

In the previous section we have discussed how the fallout results in edges decreasing in weight and how this, in turn, leads to other edges being strengthened due to reallocation of individuals' time budgets. We will now investigate the mechanism by which the fallout of a single pair of individuals can fragment the network.

The key to this is the existence of dependencies between edges. These dependencies are created through the spatial constraints imposed on the connections as well as the restrictions through the affinity value structure. As discussed in Section 4.1.2, edges between nodes that are located more than R apart in space (but less than $2R$) cannot be directly maintained, as the individuals do not attend each others gatherings. They can, however, meet at a mutual friends gathering and thus indirectly maintain their connection.

In Chapter 4 we introduced the distinction between primary and secondary edges based solely on distance. Some of the edges that are classified as primary edges by distance are nevertheless not directly maintained due to lack of affinity. This means that these edges are effectively secondary edges as they can only be maintained through mutual friends. Here, we will further distinguish between true primary edges and secondary edges by affinity which we will simply call *affinity secondary edges*. We classify all edges as affinity secondary edges that are no more than R in length (and are therefore primary edges by distance) but do not have sufficient affinity to maintain the connection directly.

In Figure 5.11 the network previously shown in Figures 5.5 and 5.4a, is represented with the edges drawn according to the three classes of edges: primary edges are shown in

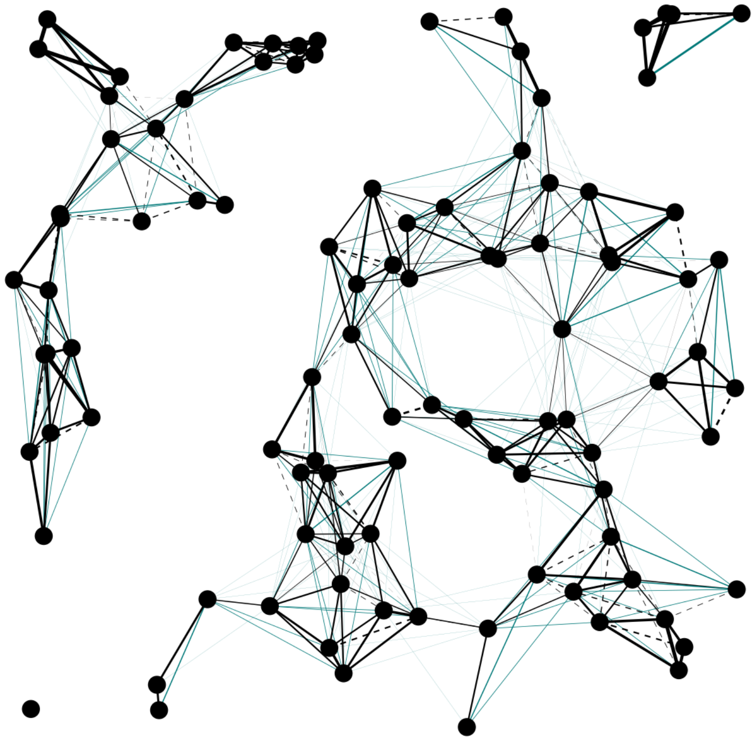


FIGURE 5.11: Network used in Figures 5.5 and 5.4a with edges classified into primary edges (black solid), affinity secondary edges (black dashed) and secondary edges (cyan). This allows us insight into the dependency structure of the edges.

black, affinity secondary edges are shown as black dashed lines and secondary edges are shown in cyan.

As discussed in Section 4.4.2, secondary edges depend on the presence of at least one mutual friend in order to be formed. Since secondary edges (both secondary by distance as well as affinity secondary edges) cannot be maintained directly, they also rely on mutual friends for maintenance and, thus, existence. A secondary edge that is facilitated through a mutual friend requires two edges to continue to exist: the edges from both of its endpoints to the mutual friend. Note that for this individual to be able to facilitate the connection, both edges to it need to be primary edges so that the two nodes can both attend the mutual friend's gathering.

This introduces dependencies into the system. Secondary edges depend on the existence of other edges. Based on the exact network structure, a secondary edge might be maintained through several mutual friends whereas in other cases it is fully dependent on the existence of connections to a single mutual friend.

These dependencies are crucial for understanding how a single fallout can lead to network fragmentation. Created by the combination of indirect strengthening of connections which individuals did not directly choose to reinforce and the constraints on which connections can be maintained directly, these dependencies are the main reason that the system is a non-linear and therefore complex system. Without these dependencies, the system could never show fragmentation as a result of a single fallout. If no dependencies exist between edges, two nodes changing their behaviour cannot have any effect on the rest of the network unless other nodes change their behaviour too.

The dependencies of secondary edges on the existence of a particular primary edge lead to the effects observed. When a primary edge disappears, all secondary edges that depend on that edge will decrease in weight. In the case where the edge was exclusively maintained through this connection, its weight will decrease to zero resulting in the edge being broken. In other cases, where more than one mutual friend (and therefore more paths for indirect maintenance) exists, the edge decreases in weight but continues to exist as it is still maintained through the remaining connections.

Fallout in primary edges can lead to the effects discussed above. It effectively turns a primary edge into an affinity secondary edge. Therefore, fallout in secondary edges has no effect since the nodes never were able to maintain their connection directly. This becomes clear if we consider the mechanism by which fallout is implemented. When introducing the fallout perturbation we set the affinity value of a pair of individuals to 0. If they previously had an affinity value of 1, this changes the dynamics of edge maintenance. If, however, the affinity value was 0 to begin with (which is what defines an affinity secondary edge) then nothing has changed for which edges are maintained directly and which are not. We will highlight this effect in the following section.

5.8 Quantifying impact

We have discussed how the fallout of primary edges can lead to changes in edge weights and edges breaking as well as network fragmentation. We will now move on to studying these effects systematically over sets of networks. In Figure 5.6 we observed that the effect of a fallout is localised to a certain region in the vicinity of the fallout edge. However, we also observe a non-zero change in edge weights further away even though no systematic edge weight change should be observed there. These changes are due to random fluctuation as we observe similar levels of noise within the base case samples (see Figure 5.12). Here, instead of comparing the average weights of the edges between fallout and base cases, we calculate the difference between the weight averages taken over one half of the 50 base runs to the average over the other half. Runs are randomly assigned to the halves.

From visual inspection we can confirm that the levels of noise observed in the tail in Figure 5.6 (at distance $> 2R$) from the fallout are similar to the levels of noise observed within the base runs as shown in Figure 5.12. To assess which edges in the vicinity of the fallout edge experience significant weight changes we need to quantify the baseline levels of noise.

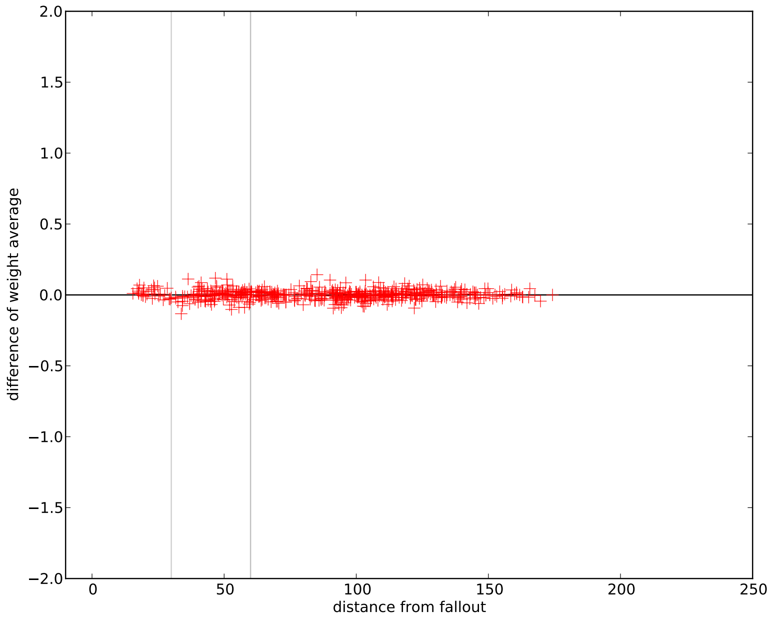


FIGURE 5.12: Differences in weight averages between two ensembles of 25 baseline runs. This highlights that due to noise and divergence of trajectories after reseeding, we should expect to observe some changes in edge weights even when no perturbation is introduced.

To do so, we consider the distribution of the edge weight differences between fallout and base case (as shown in Figure 5.6). To establish the levels of noise in the tail of the spatial plot, we calculate the standard deviation in the tail σ_{tail} over all edges that are further than $2R$ away from the fallout edge.

This σ_{tail} is around 0.2 for the scenarios shown in Figure 5.6 and the mean of σ_{tail} observed over all fallouts for that particular network is 0.022. We use this technique to distinguish between significant and non-significant edge weight changes when assessing the number of edges that experience weight changes as a result of the fallout. To do so we use a fairly conservative threshold of $5\sigma_{\text{tail}}$. We count the number of edges where $|\Delta \bar{w}_{ij}| > 5\sigma_{\text{tail}}$ (meaning the absolute edge weight change for that edge is larger than the threshold). Since each fallout is scaled to its individual tail noise level we can compare the number of edges affected across different fallouts (for one network) and across different spatial arrangements and networks.

In Figure 5.13 we show the distribution of the number of edges with above threshold weight difference over all edge fallouts for one particular network. The number of edges affected varies depending on which pair of nodes falls out and large numbers can be observed for some cases. We also see that there are no significant effects of fallout where the fallout edges are secondary edges (both secondary and affinity secondary), as discussed in Section 5.7. Thus, we will only consider primary edges for all following experiments.

So far we have focused on the study of single network examples to illustrate the effects

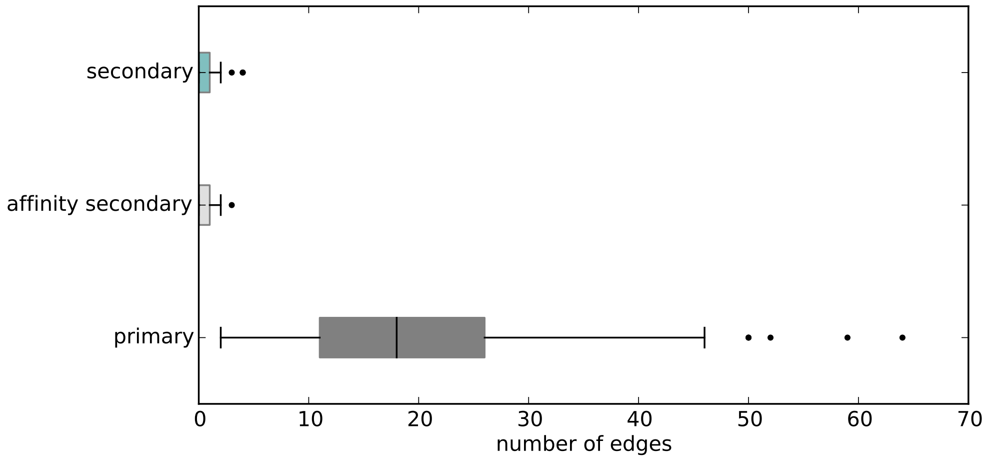


FIGURE 5.13: Distribution of number of edges with $|\Delta\bar{w}_{ij}| > 5\sigma_{\text{tail}}$ over all possible fallouts for one network. The distribution is split up by type of the fallout edge. Only for primary edges we observe a median that is larger than 0, as expected. For most cases reasonable number of edges is affected, but for some cases we observe as many as 64 edges that have changed in weight as a result of the fallout.

observed. We will now move on to assessing the impact of fallout on both topology as well as weights across multiple network realisations.

As discussed above, fallout can result in different types of impact on the network. The most drastic and obvious form of impact is fragmentation, with a component of the network splitting into smaller components. On a more local level, edges might be broken, but without fragmentation occurring. Even when no edges are broken, we might observe changes in edge weights, both negative and positive as discussed in Section 5.6. Here, we will measure these forms of impact in the following ways.

To quantify fragmentation, for each fallout edge we calculate the fraction of fallout repeats for which fragmentation is observed (i.e. cases for which we observe an increase in number of components). We here use 50 repeats for each individual fallout. To assess the impact in terms of edges broken, we count the number of edges that are present in the base runs but not in any of the fallout runs. To assess the impact on edge weights, we count the number of edges that exhibit a difference in average weight between fallout and base runs that is greater than the previously discussed threshold in magnitude, $|\Delta\bar{w}_{ij}| > 5\sigma_{\text{tail}}$.

We will now try to assess what characterises edges that are likely to lead to a certain type of impact. We would expect fallout in edges that are connected to nodes with high degrees to have a stronger impact than fallout for edges in isolated areas of the network. For node removal a common measure is to simply take the degree of the node as a measure of its importance. Extending from this, we here use a measure we call *degree sum*, which is simply the sum of the degrees of the two nodes that are connected by the fallout edge. Following from Section 5.6, we would expect fallout in edges with high degrees sum to lead to above-threshold changes in weight for many of the adjoining

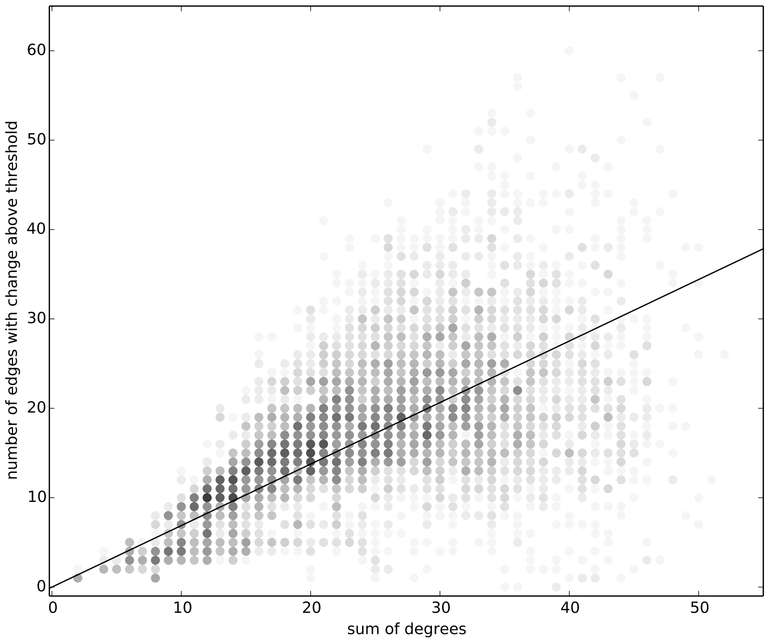


FIGURE 5.14: Number of edges where an above-threshold change in edge weight is observed versus the degree sum for the fallout edge. As we can see here, the more neighbours the nodes of the fallout edges have, the more nodes exhibit an above-threshold change in weight in the fallout case compared to the base case. We observe a linear correlation, the regression line with slope 0.688126 is shown as a black line.

edges, since the two fallout nodes will distribute the newly available time no longer needed for maintaining the fallout edge across other neighbours. Thus, our hypothesis is that there is a linear correlation between the degree sum and the number of edges that show an above threshold change in weight following the fallout.

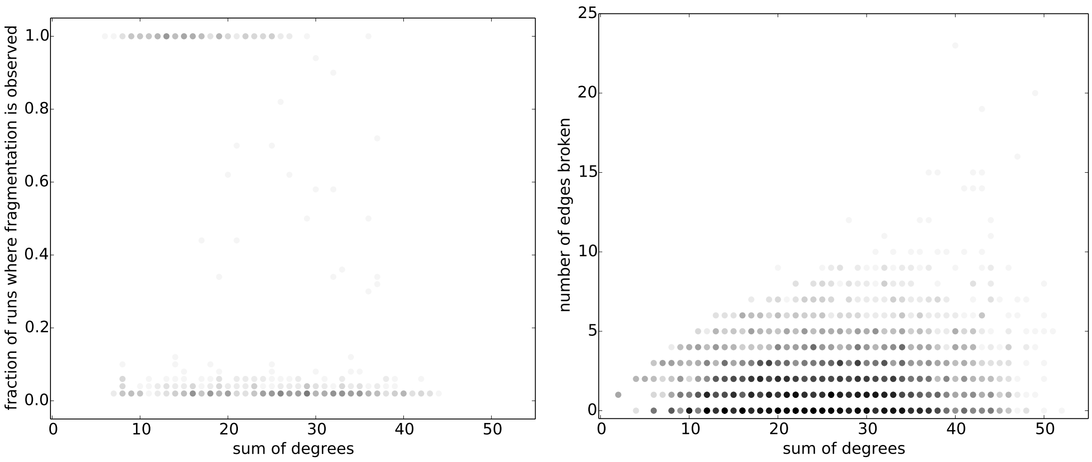
In Figure 5.14 we show the number of edges that exhibit an above-threshold change plotted against the degree sum of the fallout edge. In this figure we show the aggregated results for 24 networks.[1] For each network a fallout scenario was run for each primary edge and in each case 50 base and fallout reseeds were run to calculate edge weight changes as described in Equation 5.1. Each data point aggregates the impact on all edges in the network for a particular fallout edge and represents the point cloud shown in each of the subfigures of Figure 5.6. The data points are plotted with low opacity to allow us to identify regions where many points are located in the same space.

We observe that the number of edges that exhibit an above threshold change in edge weight as a result of the fallout is linearly correlated with the degree sum of the fallout edge. Number of edges with above threshold change can be predicted from the sum of degrees by the following formula: $y = 0.688 \cdot x, R^2 = 0.862$ (forced through 0). This means that the more neighbours the fallout nodes have, the more individuals (or relationships to be exact) will be subject to some change following the fallout. Thus,

[1]We initially used 25 different networks, but further analysis revealed that one network needed to be excluded as it exhibited spontaneous fragmentation in a base case.

our hypothesis is confirmed in this case.

For fragmentation and the number of edges broken, however, the degree sum is not a good predictor (Figure 5.15). A binomial logistic regression of sum of degrees against fragmentation shows a low odds ratio of -0.106 and a linear regression for number of edges broken versus sum of degree found a weak correlation between $y = 0.032 \cdot x + 0.989, R^2 = 0.021$. While both of these relationships were significant ($p < 10^{-10}$), this is largely due to the large sample sizes involved. Fragmentation seems to be rare for higher values of degree sum, presumably because there is a higher likelihood for redundant connections to exist in more densely connected areas. Regarding the number of edges that are broken as a result of the fallout, we can see that the maximum number observed does increase with the degree sum, but only weakly. More edges can potentially be broken, since more edges around the fallout edge exist. Thus, we can conclude that degree sum, while a useful measure in assessing the potential weight change impact of a fallout, is not a good predictor for the number of edges breaking.



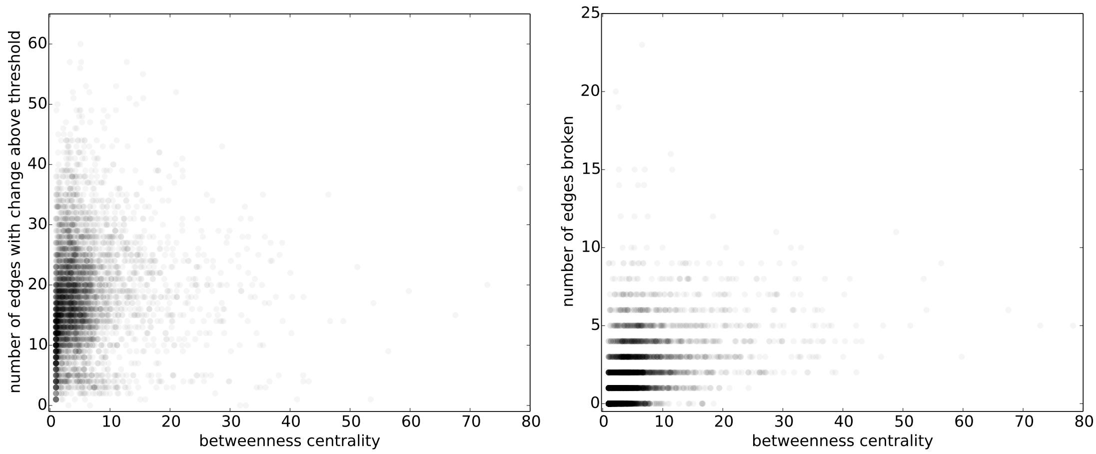
(A) Fraction fragmentation (B) Number of edges broken

FIGURE 5.15: Fraction fragmentation and number of edges broken for degree sum of the fallout edge.

For network cohesion, edges with high betweenness centrality are often thought to be particularly important, as many shortest paths run through them. Thus, removing edges with high betweenness centrality often leads to paths between nodes increasing in length. This is especially important for systems where transmission occurs along the edges. Effects of the removal of high-betweenness edges for percolation have been investigated in Onnela et al. (2007a) and we will discuss this in more detail in Chapter 6. Here we will investigate whether high centrality edges are the most crucial edges for network cohesion or whether other edge measures better characterise the set of edges most crucial for cohesion. Our hypothesis is that edges with high betweenness centrality are likely to lead to fragmentation of the network when broken.

We calculate edge betweenness centrality using networkX, taking into account the weights of the edges. The algorithm used by the networkX package is described in Brandes (2001).

In Figure 5.16 we show the edge level measures of fallout impact versus the edge betweenness centrality of the fallout edge. These show no interesting relationship, although we can observe a decrease in the maximum number of edges broken with increasing centrality and a weak but significant ($p < 10^{-10}$) linear correlation is observed (according to the formula $y = 0.149 \cdot x + 16.991, R^2 = 0.011$ for centrality and number of edges with above threshold change and $y = 0.151 \cdot x + 0.966, R^2 = 0.203$ for number of edges broken predicted by centrality). Since this is simply a result of edges with higher centrality being located in sparser neighbourhoods, this effect is already captured by the degree sum (edges in a more densely connected area will exhibit a higher degree sum as well as lower centrality as more redundant paths exist).



(A) Number of edges above threshold (B) Number of edges broken

FIGURE 5.16: Edge with weight change above threshold and number of edges broken against betweenness centrality of the fallout edge.

In Figure 5.17 we show the fraction of fragmentation for fallout edges with a certain scores of betweenness centrality. From this we can conclude that betweenness centrality alone is not a good predictor for the effect of the fallout on the global topology (a binomial logistic regression indicates a weak relationship with a low odds ratio of 0.099, $p < 10^{-10}$). However, our hypothesis is confirmed for edges with very high levels of centrality are important for network cohesion is confirmed since in cases where betweenness centrality of the fallout edge is high, the network is likely to fragment. In the 24 networks considered there are only a few of edges with very high betweenness centrality. The issue with using betweenness centrality as a measure to predict fragmentation is that, while it correctly identifies edges that are important connectors between parts of the network (and this includes bridges), it is a global measure and thus takes into account the context of the edge. When assessing edge importance for network cohesion, it results in high values for bridges that connect large components. In contrast, an edge that connects relatively small modules to the rest of a component might be more

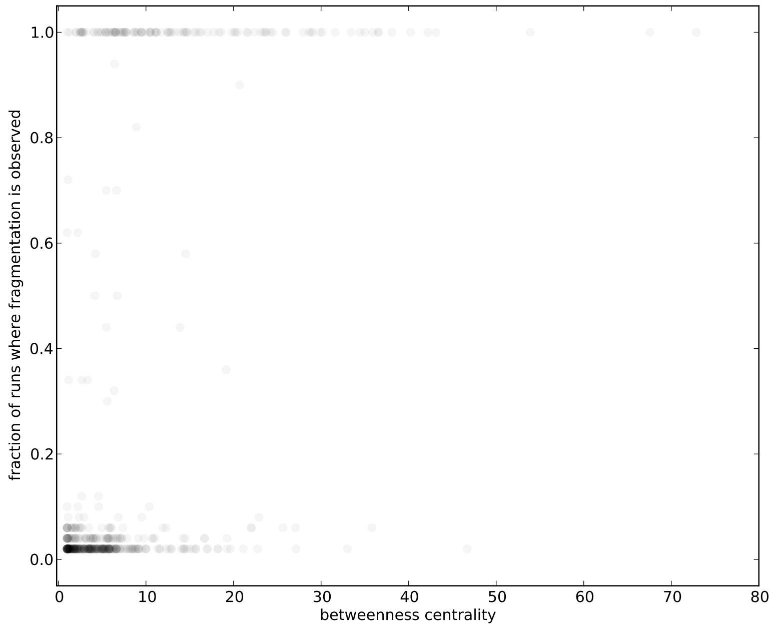


FIGURE 5.17: Fraction fragmentation versus edge betweenness centrality

important for the integrity of that component, but the edge will still have a relatively low score of betweenness centrality since there are only a limited number of shortest paths going through that edge, due to the small size of one of the sub components. Thus, edge betweenness centrality might be a good measure if we are mostly interested in the stability of large components. If we are interested in fragmentation of smaller subgroups, edge betweenness centrality fails to identify the edges crucial for connecting small modules as important.

A measure developed to describe the "bridginess" of an edge is overlap (Onnela et al., 2007a). As it is a local measure, it is not biased by the size of the subcomponents on either side of the bridging edge. Furthermore, due to its localised nature, it can be efficiently calculated for very large networks. The overlap score of an edge between two nodes i and j is calculated according to the following equation:

$$O_{ij} = \frac{m_{ij}}{(k_i - 1) + (k_j - 1) - m_{ij}} \tag{5.2}$$

k_i and k_j denote the degrees of the nodes i and j respectively. m_{ij} denotes the number of mutual neighbours of i and j. Overlap is therefore defined as the fraction of mutual friends with respect to all neighbours of i and j and ranges from 0 to 1. High overlap means that to a large proportion of i's friends are also j's friends (and vice versa), thus corresponding to a low level of "bridginess" for the edge. Here we therefore use $1.0 - O_{ij}$ which we will refer to as complement overlap. Thus, a high complement overlap score for an edge indicates that the two nodes it connects share few neighbours, therefore the edge is likely to be a bridge.

In Figure 5.18 we show the fraction of fragmentation against the complement overlap.

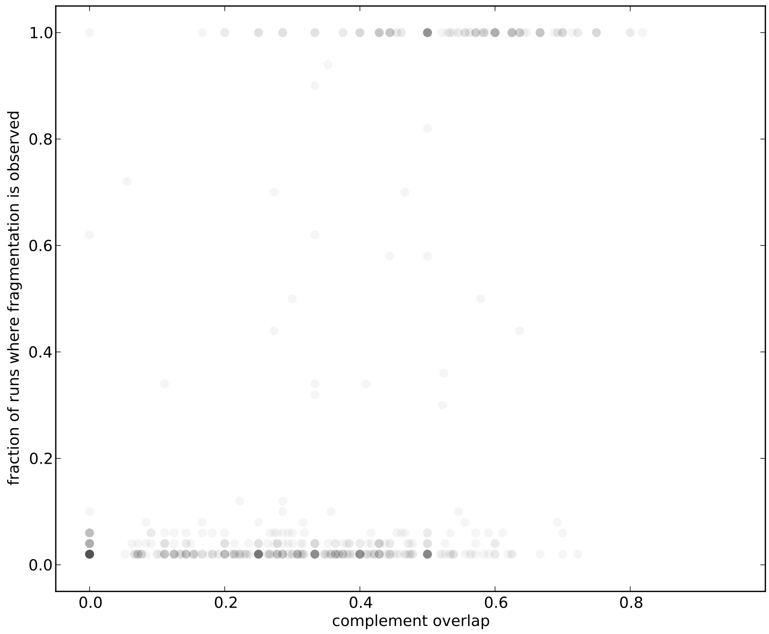


FIGURE 5.18: Fraction fragmentation versus complement overlap of the fallout edge. For low levels of complement overlap, fragmentation is generally not observed in any of the reseeds. For fallout edges with high levels of complement overlap fragmentation is likely, most often in all 50 reseeds. For intermediate values, either case might be observed; sometimes even cases where fragmentation occurs only for some of the repeats.

For low values, fragmentation is unlikely and generally occurs in none of the 50 fallout reseeds. For edges with intermediate values of around 0.5, fragmentation can occur but does not always. For most edges the network either fragments in all 50 reseeds or in none, however, for some fallout edges fragmentation is observed only for some of the 50 repeats. This is most likely for edges with medium levels of complement overlap (see Figure 5.18). Note that we only measure global fragmentation here. In some cases, fallout for edges with intermediate to high levels of complement overlap might lead to local fragmentation but this would not register as global fragmentation. For edges with high scores of complement overlap, fragmentation is very likely. A logistic regression indicates a strong relationship with a high odds ratio of 7.587 that is statistically significant ($p < 10^{-10}$). This means that overlap is a good predictor for which edges will lead to fragmentation in case of fallout.

The maximum number of edges that break as a result of the fallout increases with complement overlap (Figure 5.19). A positive, linear correlation exists between the number of edges broken and complement overlap, according to the following formula: $y = 5.603 \cdot x + 0.077, R^2 = 0.245$, with $p < 10^{-10}$. The increase in the maximum number of edges breaking is due to the fact that it is more likely that edges exist that fully depend on the fallout edge as a bridge, as bridges span locally sparse regions of the network. Thus, we observe that for very high values of complement overlap there is a decrease in the maximum of broken edges as in the case where the two nodes of the fallout edge share no mutual friends the fallout edge will not have any edges that depend

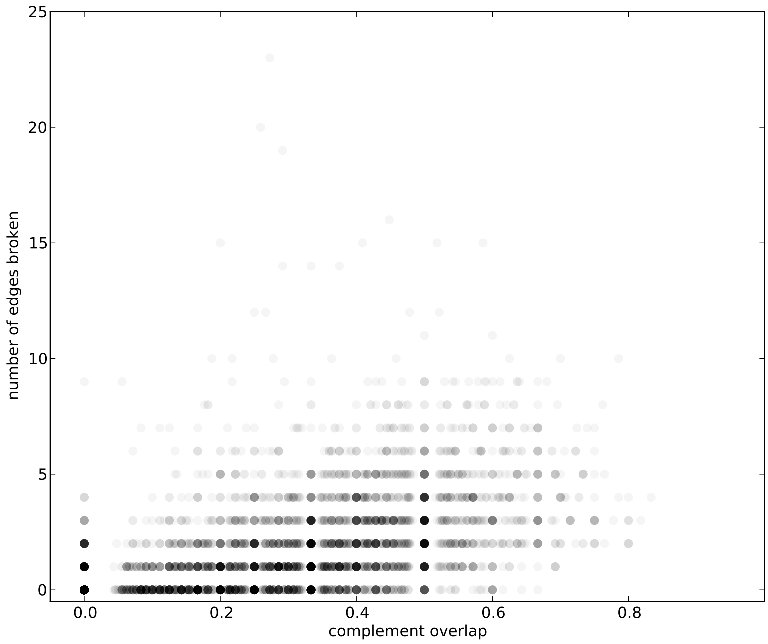


FIGURE 5.19: Number of edges broken versus complement overlap of the fallout edge. We observe an increase in maximum number of edges broken for increasing values of the complement overlap and a decrease for very high value. Note that the banding in the figure is due to overlap being calculated as a fraction.

on it.

Next, we will introduce two measures that explicitly take into account the dependencies between edges and therefore are able to better predict the number of edges broken as a result of the fallout. As discussed in Section 5.7, edges are broken as a result of a supporting primary edge being turned into a secondary edge through fallout.

The first measure we introduce here simply counts the number of secondary edges that are maintained exclusively through a particular primary edge. Note that each secondary edge requires two primary edges to be present to close the triangle, thus both will receive an increase in score of 1 for that edge. Since the score consist of a count of edges it is an integer. We call this measure the *dependency score*. Edges with a high dependency score maintain many secondary edges that could not exist without the edge itself existing. Thus, the dependency score provides a lower bound for the number of edges that will disappear following a fallout in the edge. We observe this effect in Figure 5.20. As expected, all edges that depend on the fallout edge break (corresponding to the dependency score) and often the fallout edge itself decays to zero, leading to the number of edges broken being the dependency score increased by one. In some cases we observe more edges disappearing. One example for this is the case where a pair of nodes is connected by a weak primary edge, that relies mainly on indirect maintenance. In other cases edge weight perturbations following weight rearrangement might lead to edges breaking.

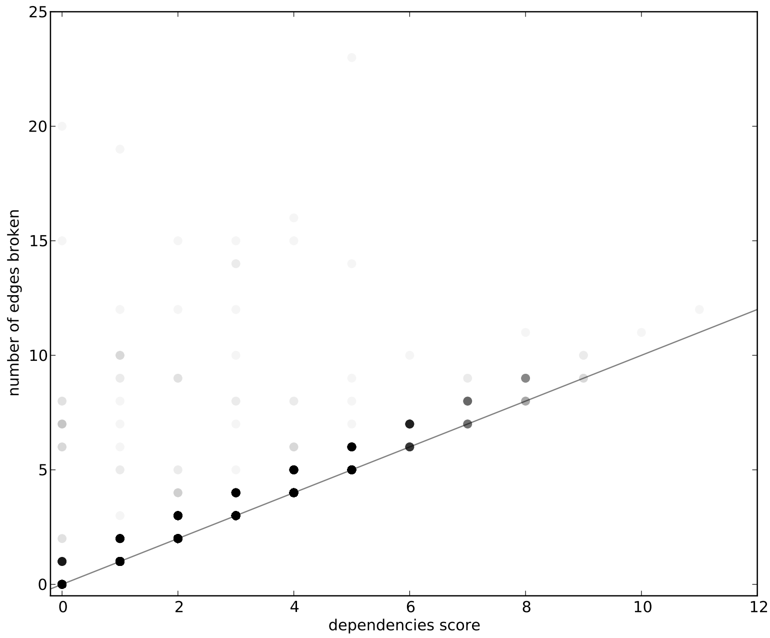


FIGURE 5.20: Number of edges broken versus dependencies score of the fallout edge. We show the diagonal $y = x$ for reference as a black line. Points on this diagonal correspond to cases where the number of edges broken as a result of the fallout is exactly the number of edges that depend strongly on the fallout edge being present. In some cases the fallout edge itself disappears, resulting in a second diagonal of data points. Additionally, we observe cases in which additional edges are broken following the fallout although this is less common.

We observe a high correlation between the dependency score and the number of edges broken: $y = 1.080 \cdot z + 0.092, R^2 = 0.797, p < 1^{-10}$. The dependency score is able to predict the number of edges broken with high accuracy as it utilises perfect information on the dependency structure. For real-world systems we are unlikely to be able to extract information like this as it is unlikely that dependency structures are simple enough to assume exclusive dependencies. We will therefore introduce a measure that takes into account dependencies without having to rely on an edge being part of the only path through which a particular secondary edge can be maintained. Even though this is still a very abstract measure that might be impossible to calculate in real-world setups, we will discuss in Chapter 6 how similar measures to be used for network analysis could be derived.

We will refer to this measure as an edge's *facilitation score*. It is similar to the dependency score, however, secondary edges that can be maintained through several paths are taken into account for this measure, whereas for the dependency score only strong dependencies (which are maintained through only one path) are considered. If we assume that a certain secondary edge can be maintained through x mutual friends, all edges on the paths through each mutual friend receive an increase in facilitation score of $\frac{1}{x}$. If we compare this measure to the dependency score, we observe the same effect, but in addition we observe values underneath the diagonal (see Figure 5.21, the corre-

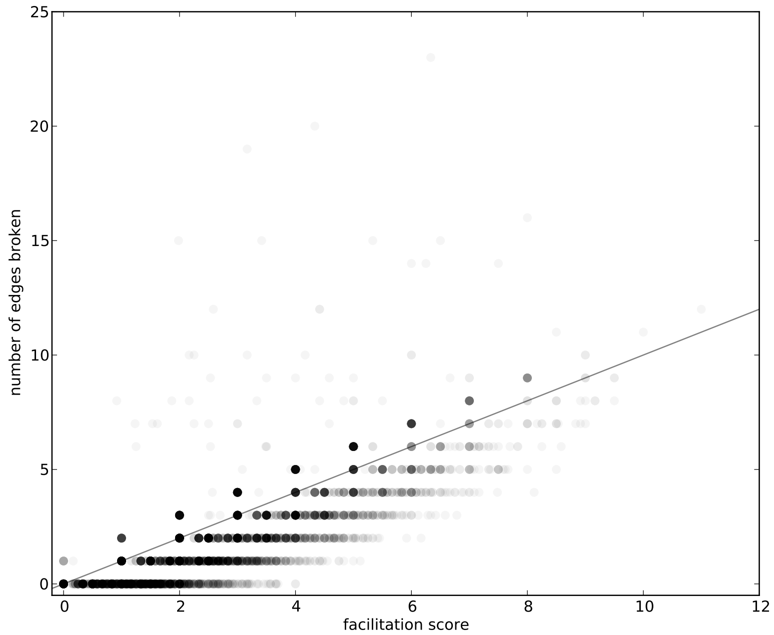


FIGURE 5.21: Number of edges broken versus facilitation score. As for Figure 5.20, the number of edges broken is correlated with the score. However, here we observe values below the diagonal since the definition of the facilitation score includes secondary edges that can be maintained through several paths and thus are not always broken following a fallout.

lation in this case is weaker with: $y = 0.943 \cdot -0.795, R^2 = 0.653$, but still statistically significant with $p < 10^{-10}$). This corresponds to edges that have a higher facilitation than dependency score. The edges that depend on these fallout edges do not depend on them exclusively. Therefore, many of these edges will not disappear following the fallout as redundant paths for maintenance exist. They will, however, still be affected by a decrease in edge weight since one of the maintenance paths has disappeared. This is also reflected in an increase of both minimum and average number of edges that exhibit above threshold edge changes with increasing facilitation (Figure 5.22, observed linear correlation is relatively low (but significant with $p < 10^{-10}$) according to the formula $y = 1.682 \cdot x + 13.190, R^2 = 0.116$). Thus, for a more holistic assessment of both global and local impact, the facilitation score provides a better measure than the dependency score.

5.9 Conclusion

In this chapter we have discussed how fallout in one edge affects the network both on a global and local level. We have shown that changing the behaviour of a single pair of nodes leads to weight changes in the edges surrounding the fallout edge, even though all other individuals behave in the same way as before. On a global level we have confirmed that fragmentation is possible as a result of the fallout and we have discussed

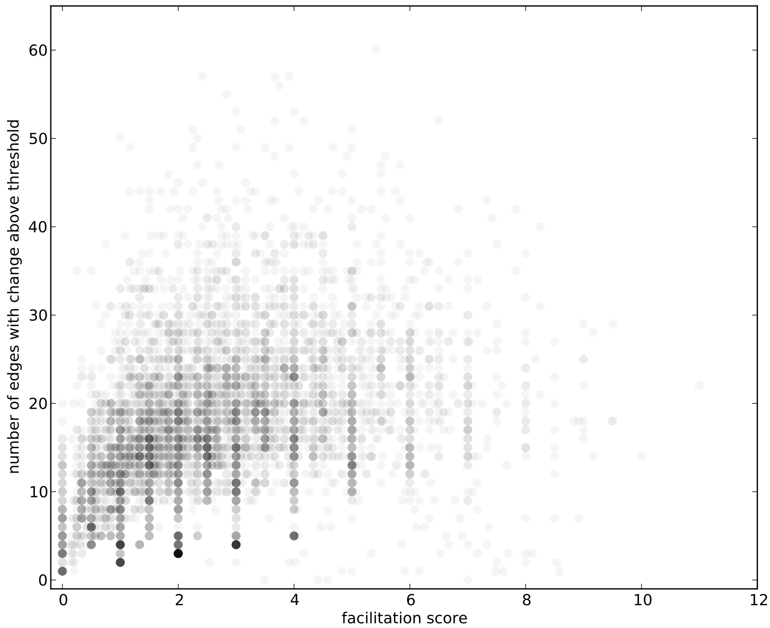


FIGURE 5.22: Number of edges with weight change above threshold against facilitation score. Both minimum and average increase with increasing facilitation score of the fallout edge.

how dependencies between edges are crucial for this to happen.

We have investigated different edge measures to characterise fallouts that impact the network in a certain way. We find that the sum of degrees of the two nodes connected by the fallout is correlated with the number of edges that exhibit significant weight changes following the fallout. We have shown that betweenness centrality is a good predictor for some types of fragmentation impact and discussed its limitations and reasons for these limitations. Furthermore, we have shown that complement overlap, a measure derived from overlap (as presented in Onnela et al., 2007a) which describes the “bridginess” of an edge, is not biased in the same way as it is a local measure and is thus able to provide a good prediction for fragmentation impact in general. We have introduced two measures to quantify the dependency structure of the edges and have shown that the number of edges broken as a result of the fallout is strongly correlated with these measures.

In the following chapter we will discuss the limitations, but also the possible applications of these measures.

Chapter 6

Discussion

6.1 Connections to existing theoretical and empirical work

6.1.1 Shortcomings of other modelling approaches

Chapter 2 introduced several models of social network dynamics. Here, we revisit this topic and discuss in more detail why the modelling approaches presented fall short in modelling the dynamics of social networks.

Many models of social networks are mathematical models. The advantage of these models is that they can be solved analytically and a large theory base exists. The issue with these models is that many modellers stay within the bounds of analytically solvable models and as a result choose problems based on whether they can be solved by a particular tool. Thus, a vast proportion of network research is concentrated around problems for which analytical solutions exist and mean-field approaches can be applied. While significant progress has been made using mathematical network models, the mathematically minded branch of network science seems to centre its efforts around a certain set of models as a result of being unwilling to move out of the area of simple, mathematically tractable models.

A related bias can be observed in the agent-based modelling community. In the context of this community, richer models regarding node behaviour can be found, as well as coevolutionary models, studying the interplay of topology dynamics with processes on the network. However, similar to the mathematical network science community, a narrow focus has developed on incremental work. A majority of models centre around three main areas: disease spread, opinion dynamics and games on networks. The public health implications of urbanisation and globalisation explain the interest in, and funding for, the study of network-mediated disease spread. The others owe their popularity mainly to the fact that there are simply not many other coevolutionary models of human behaviour.

When starting work on the models presented in this thesis, we were looking for a base model for coevolutionary social network dynamics. We were surprised to find that no general model of social network dynamics seemed to be publicly recognised as a starting point for modelling non-strategic human behaviour. While many models exist that are useful abstractions of certain settings, such as (Ahrweiler et al., 2011; Kaufmann et al., 2009; Noble et al., 2012), for general and abstract models of network dynamics the modelling landscape is dominated by game theoretic models and opinion dynamics models. These models, however are not as general as they claim to be, due to the specific nature of the processes.

Game theory focuses on modelling strategic interactions between pairs of individuals and is thus well suited for situations where individuals might rationally assess different options of interaction behaviour and strategically decide on the behaviour that is likely to give the individual the largest benefit. Thus game theoretic network models play an important role for modelling a certain subset of problems where individuals interact and for each interaction choose their behaviour from a range of possible behaviours. For simulating casual encounters between humans, they are less useful, since a lot of this behaviour is the result of non-strategic decisions, limited options or simply past choices (for example when individuals stick with the same behaviour since it has become a habit). Opinion dynamics can provide valuable insights into network fragmentation driven by strategic decision making. Thus, models of coevolutionary opinion dynamics provide the counterpoint to the approach taken in our model. Again, this assumes strategic decisions and, more importantly, explicit breaking of edges. While explicit breaking of connections certainly occurs in real world social networks, in many cases edges disappear unceremoniously by decaying due to not being maintained. Therefore, human behaviour is more complicated than assumed by game theoretic models and opinion dynamics models and involves a non-rational component. We thus argue that behavioural rules that are not based on strategic decisions should be considered more frequently to complement the approaches taken by existing models.

Another problem is that most models of dynamic networks do not use weighted networks, limiting the edge dynamics to making and breaking of edges. Even models that do take into account edge weights (Kumpula et al., 2007; Skyrms and Pemantle, 2000) do not implement any self-balancing mechanisms for the edge weights. Instead, edges or nodes are periodically removed to ensure that the network density remains within a certain regime.

While the reasons behind this design choice are generally not discussed, we suspect that this global mechanism of node or edge removal is chosen because implementing a self-balancing mechanism of network dynamics is rather hard. There are several reasons for this and they are consequences of the complex nature of social networks. Stochasticity of the dynamics, heterogeneity in both topology and node behaviour, and positive feedback all play a major role in this. For example, in most network models we observe topological heterogeneity as nodes vary in their degree. Since real social networks exhibit broad degree distributions this is a desired feature and should therefore not be omitted from social network models for the sake of simplicity. Heterogeneities in the network are often a result of initial differences being magnified by the stochasticity of the system. This makes it very hard to calculate information needed for balancing mechanisms, such as the average increase in weight per timestep. This in turn leads to difficulties for the design of an appropriate decay mechanism for balancing edge weights.

We tried several other approaches before settling on the approach presented here, which is to make edge weights an explicit representation/record of past meetings. While it is desirable to have the edges reflect contact frequencies, as it potentially allows for comparison with empirical contact data, we have also found that it is crucial to implement edge decay this way as it is very difficult to predict the global increase in weights per step, due to the complexity of the systems as discussed above. Local balancing mechanisms that rely on a fixed decay rate fail for our model since the magnitude of weight increase depends on the aggregated decisions of all individuals due to the fact that connections can be reinforced indirectly at friend's gatherings. This is a specific problem of non-binary weight changes, which might be one of the reasons why not many social network models incorporate weight dynamics. In addition, since edges in networks represent relationships between two people, we believe that edge weights should be influenced by the actions of both individuals and thus should not be regulated by the individuals independently. In addition, allowing individuals to independently increase and decrease edge weights to limit node strength can lead to less well connected nodes getting disconnected even though they invest time in the maintenance of particular contacts. This does not happen for the mechanism presented in this thesis, as unreciprocated relationships are possible (an individual can attend another individuals party even if the reverse never happens). With the mechanism presented here edge weights are bounded as a result of external constraints and they are a result of the action of both of the individuals that are connected by the edge.

6.1.2 Limitations of the modelling approach presented here

Due to the dearth of models of non-strategic network behaviour we could not build on existing models when creating the model presented in this thesis, as they were unsuitable for our purpose in one way or another. While inspired by mechanisms present in other models, our model combines these mechanisms in new ways and additionally incorporates mechanisms that are novel in modelling social network dynamics. This means that our model is just a first step in the direction of modelling the dynamics of weighted social networks and therefore fairly simple. While it is a good starting point and sufficient to answer the main question posed in this thesis, there are many ways in which it could be extended to match human behaviour better. In this section we discuss limitations of the model and possible extensions.

Referring back to the desired properties for artificial social networks discussed in Section 3.2.4, we will first discuss to what extent the networks generated as a result of our model match the desired properties. As discussed in Section 4.1, the generated networks are sparse and exhibit high levels of clustering and positive assortativity. They exhibit visually obvious community structure, which is a characteristic topological feature of social networks. With regard to these characteristics our model is able to match the desired features.

Regarding average shortest path length, our model shows larger values than observed in real social networks for larger values of N. This is due to the fact that our model implements a strict cutoff for edges above a certain length, therefore prohibiting shortcuts between more distant nodes. However, there are several mechanisms one could think of to introduce additional long range connections to the existing model. Allowing random connections, not constrained by spatial placement is used in many models to generate long range connection as well as local connections that do not close triangles. Whether this is a realistic assumption to make depends on the system to be modelled. Another option would be to allow nodes to move in space and retain some of the previously made edges, even if they span distances larger than the reach R.

The degree distributions observed are broad, but do not exhibit any obvious skew. Real-world social networks have right-skewed degree distributions and our model fails to recreate this feature. Preliminary work by Hamill and Gilbert (2009) suggests that heterogeneity in social reach R might be able to address this issue and introduce a degree skew to the distribution.

In the model presented here, nodes are located in space. As discussed, this is an important factor for the generation of realistic-looking social networks. Spatial placement in our model is according to a uniform distribution, whereas in many real-world systems we observe nodes to be clustered both with respect to locations in space as well as distributions of traits or opinions. Thus, extending the model to allow for different spatial distributions is a possible next step. Moving towards non-uniform node distributions could also influence the likelihood of sparse local areas, thus leading to more bridges and increased probability of fragmentation. The initialisation of networks in our model as a random geometric graph is a further simplification. Some preliminary studies show that starting with a slightly less dense network does not produce obvious differences in topology, given there is some variability between individual runs with different seeds due to stochasticity in any case. However, if initialising with an entirely different topology we might find that this affects the topology of the resulting networks in more obvious ways. In any case the initialisation will be an abstraction of how social networks form from scratch, but this is a different problem and beyond the scope of the work presented here.

In addition to the simplifications regarding spatial placement of nodes, we can also imagine possible extensions using continuous mechanisms for implementing a spatial influence on possible connections. In our model, the reach R sets a discrete threshold distance for direct interactions, whereas in the real world we would assume a more probabilistic mechanism, with connection probability decreasing with increasing distance. Furthermore, we would expect this decrease in probability to be influenced by other factors, such as the similarity of the individuals in question, as well as the number of socialising opportunities in the closer neighbourhood. This last factor has been studied empirically and we will come back to this in the next section where we discuss our model in the context of empirical findings on social networks.

Using a threshold allows us to divide edges into primary and secondary edges which makes it easier to understand the mechanism that causes fragmentation following a fallout. We would expect this distinction to be less clear cut in the real world, meaning that dependency structures would also be more complicated. For example, if we assume probabilistic interactions based on distance we could still observe dependencies but they would no longer be a binary distinction between primary and secondary edges. Instead of an individual never accepting an invitation, it would simply be highly unlikely for that individual to accept. We would not expect this extension to change the dynamics drastically, but nevertheless it would be interesting to study fragmentation in the case where this slightly forced distinction between primary and secondary edges does not exist and investigate to what extent exactly this influences fragmentation dynamics.

Another simplification that should be addressed in further work is the fact that currently nodes are homogeneous with respect to attributes and behaviour. We have not yet investigated the influence of heterogeneity for any of the parameters. Hamill and Gilbert (2009) find that in their model heterogeneity in reaches leads to a longer tail in the degree distribution. Their model is not dynamic, thus it would be interesting to investigate whether a similar result can be obtained in our dynamic model. The effect of heterogeneity in individual's reaches on the actual dynamics is a particularly interesting possible extension as this would again affect the dependency structures, which influences the fallout dynamics.

Similarly, introducing heterogeneity in the time budgets (parameter S in our model) could represent that some individuals spend more of their time socialising than others, or the simple fact that some individuals might have less time available. This heterogeneity should have implications for the resulting topology. Heterogeneity in history length (H) could reflect the fact that some people are better at keeping in contact with old friends whereas others show more variability in their set of their friends over time, which has the potential to affect the dynamics of network change and stability of the created networks.

An important step for the future would be to add richer homophily dynamics to replace the affinity values. We could add several dimensions of traits and values, and individuals could assess whether to maintain a connection based on differences in these values. The spatial placement of nodes in our model can be viewed as distinct from the social value space but could also represent location in an abstract social space. If we take this view, spatial embedding and affinity values could be replaced with a higher dimension trait space with likelihood of direct interaction being a distance function between the individuals' locations in this space.

Another type of heterogeneity that most models, including our own, fail to incorporate is that different types of connections can exist with entirely different dynamics. One example are romantic relationships which, at least in general, undergo discrete transitions in type and are exclusive. Romantic ties are also more likely to be actively broken, which is why it would be interesting to add such a dynamic to our model especially to include fallout as an internal process instead of an external perturbation. Another type of ties are family ties, which persist over long periods and are generally more resilient to being broken, even when not maintained for a long time. While our model does not capture any of these heterogeneities, it could be extended to do so.

In the following section we review findings of empirical work on social network dynamics and discuss how they relate to the model presented here. In this context we discuss features observed in empirical work that emerge in our model and we present empirical findings about human behaviour that are currently missing but could be included in future iterations of the model (depending on desired applications).

6.1.3 Empirical context of the work presented here

Regarding persistence and decay of relationships, Burt (2000) notes that embedding edges in a dense network with shared friends results in a higher probability of edge survival. This can be observed in our model, as edges that connect individuals with mutual friends are less likely to decay as multiple paths of maintenance exist. This only holds for non-fallout situations though, which is in accordance with the findings described by Burt (2000), namely that the probability of decay is increased for embedded edges where the embedding is disrupted. However the author also observes that strong relationships and edges that have existed for long periods of time show lower rates of decay, a fact that is not represented in our model.

Using data on the dynamic change of an email communication network, Kossinets and Watts (2006) study factors influencing triadic closure. They find that tie strength strongly influences the probability of triadic closure. This effect is an emergent property of our model as stronger ties lead to a higher probability of two individuals attending each other's gatherings and thus meeting each other's friends, resulting in triangle closure. Thus, triadic closure is more likely for a pair of nodes with strong connections to the mutual friend. We can also observe a certain degree of heterogeneity in individuals' socialising behaviour, similar to what is observed by Backstrom et al. (2011): some individuals interact with only a small circle of close friends intensively, others have more friends but invest less time in each of the relationships. We can observe this in the model presented here; it is a result of the spatial embedding of the nodes and the resulting heterogeneity in the number of contact opportunities. In the real-world, the type of time allocation across contacts (focused on few individuals or spread across many) seems to be stable for an individual over time and is thus unlikely to be the result of contact limitations. Our model does not capture intrinsic differences in time-allocation in its current form, even though this might be an important factor for extending the model towards heterogeneous behaviour.

Regarding spatial effects, Illenberger et al. (2012) study data on leisure contacts and the geographic distance between the individuals and find that the probability of forming a connection decreases with distance, as modelled in our model by the spatial constraint. They also observe that the number of contacts seems to be independent from the spatial location, meaning that individuals located in sparse areas compensate for this fact by accepting longer range connections, a fact that is not captured in the current state of our model. Furthermore, the constraint on edge length is very rigid in our model prohibiting any long range links. As discussed in Chapter 4, the absence of long range links results in diameter and average shortest path length not scaling logarithmically, as observed in real-world social networks. Therefore, relaxing this constraint to allow some longer range edges might be important for the future as well as moving from a threshold of interaction to a probabilistic approach.

van Duijn et al. (2003) confirmed that spatial proximity, similarity and network opportunity all influence the probability of friendship formation. These factors are implemented in our models as the spatial constraints (parameter R), the affinity value structure (parameter A) and the triangle closure mechanism. In real-world friendship formation the importance of these factors changes over time, a fact that cannot be represented in our model in its current form. The authors show that proximity is important in the first stages of friendship formation, whereas "invisible similarity" only becomes important later on as the friendship develops (as it can not be assessed at the beginning when individuals do not yet have any information about each other). It could be interesting to include richer mechanisms of friendship formation in future models of the dynamics of friendship networks.

Regarding the development of relationships over time, Reis et al. (2011) find that two individuals being exposed to each other increases two individuals' affinity to each other. While this effect is indirectly captured in our model by preference for attending meetings hosted by individuals to which a strong edge exists (indicating frequent previous contact), we do not model this fact explicitly by modelling the dynamics of affinity values.

6.1.4 Fragmentation and percolation

Not much empirical work studying network fragmentation exists. This might be due to the fact that this requires collection of time-resolved data, but with the advent of new collection techniques for time-resolved contact and communication data (see Section 2.3.1) we will hopefully see progress in this area in the future.

In Onnela et al. (2007a) the authors study the percolation behaviour of a real world mobile phone call network. They observe that many weak links seem to have low overlap but high betweenness centrality, confirming their role as bridges as suggested by Granovetter (1973). Thus, they are important for the cohesion of the network. In our model, edges that are bridges by default have a lower weight as they are only directly maintained and do not receive additional indirect maintenance through mutual friends. Our model therefore recreates this structure, now thought to be a defining feature of social networks. Our experiments on fallouts confirm that edges with high levels of betweenness centrality and low levels of overlap (high levels of complement overlap) are more likely to lead to fragmentation of the network. This also fits with the definition of group cohesion in the social sciences literature, where groups are considered to be more cohesive if more edges need to be removed to split the group (White and Harary, 2001).

In models of group evolution, spontaneous fragmentation has been observed when community structure of the network becomes so pronounced that groups no longer have access to potential new members and thus can grow no further, leading to network fragmentation (Geard and Bullock, 2008). In a later model, Geard and Bullock (2010) observe that increasing individuals' social time budgets results in more robust networks that are less likely to fragment, as they show less distinctive community structure.

As discussed previously, fragmentation has also been studied in the context of opinion dynamics. In coevolutionary opinion dynamics, the timescales of the processes involved (opinion update and edge rewiring) play a major role in determining whether the system develops towards fragmentation or connected consensus (zu Erbach-Schoenberg et al., 2011). The process that leads to fragmentation in these models (individuals explicitly break edges to individuals with a differing opinion) thus can not explain the questions that are the focus of this thesis: how networks can fragment without the individuals taking explicit action towards breaking edges. Nevertheless, since human behaviour in social conflict probably lies between explicitly taking sides and breaking edges and not taking sides at all, these models should be considered for future studies of fragmentation dynamics.

Cascading failures are observed in many technological systems (Holme, 2002; Motter and Lai, 2002). Whether edge failure cascades are possible in social networks as well has not been answered by empirical studies yet. The model we present in this thesis does not exhibit cascades as described in other works. The effect of the fallout we observe in our model is localised. Only edges located close to the fallout edge in space and which have a short hop distance to the edge are affected. This means that cascades, as observed in other systems, are not possible in this model. However, the breadth of the impact is only limited by the degree of the fallout edge, so fallout of one edge can still lead to changes in edge weights for many edges. And even though the impact is localised, fragmentation does occur as shown in Chapter 5.

6.2 Steps towards richer models of social network dynamics

Many of the models discussed in Chapter 2 are able to accurately model certain aspects of human behaviour. Here we identify several key mechanisms of social network dynamics that have been the focus of these models and suggest how they could be combined to construct richer models of social network dynamics.

One important aspect of social network dynamics is triangle closure. Triangle closure is used in many models of social networks to achieve high levels of clustering. In an empirical study of a dynamics network Kossinets and Watts (2006) claim that social network evolution is indeed driven by triangle closure and additionally, homophily. Homophily is generally divided into two dimensions. The first is baseline homophily (also called induced homophily), which is a result of limited interaction opportunities. One example for this is that students at a university are generally young adults, and are therefore more likely to interact with other young individuals. Similarly, meeting someone in the UK it is likely that both of you will speak English. While this homophily is influenced by past choices, it involves no current bias on choosing interaction partners. The additional bias of choosing to interact with a similar individual instead of another is called choice homophily. It describes the bias to preferentially interact with individuals similar to oneself. In our model, both types of homophily are represented in an abstract way: baseline homophily is present due to the spatial location of the individuals and interactions are further biased by the affinity values, introducing a choice homophily mechanism.

An important aspect of social network dynamics we have not discussed so far is focal closure. Focal closure refers to the fact that individuals meet through activity foci. Activity foci are associated with some form of activity such as work, a hobby or a religious group. Activity foci bring individuals together for particular activities and this enables the formation of new connections. In contrast to triangle closure, activity foci can lead to the formation of connections between individuals that do not share mutual friends and is therefore less constrained by the network. In real-world social networks both network mediated new contacts as well as contacts not directly mediated through the network are important for driving network evolution. The model presented in this thesis does not capture non-network-mediated contacts, which is why components, once fragmented, cannot reconnect. This process is abstracted as random encounters of individuals in many models. While this can be a reasonable abstraction to make, we believe that in many cases the fact that activity foci are defined through some shared values or interests (and thus lead to baseline homophily) is important and should be considered even for abstract models of network dynamics.

In Kossinets and Watts (2009) the authors suggest the construction of a simulation model in which choice homophily, focal closure, and cyclic closure could be varied to study their effects and the model presented in Chapter 3 could serve as a basis for this. However, our model lacks the possibility for connections to be made between nodes that do not share a mutual friend. As just discussed, non network mediated connections can happen though shared activity foci or groups. Some previously introduced models study the interplay of networks and group dynamics (Geard and Bullock, 2008, 2010). The main focus of these models is group dynamics. Therefore the network dynamics side is rather simple and does not include weighted dynamics. Combining these models with our model using weighted edges could be a promising step towards a more mature model of social network dynamics.

6.3 Implications for real-world social networks

As discussed above, the study of social network dynamics is still in its infancy. Therefore, claims about mechanisms of fragmentation in real-world systems are still further down the line. Nevertheless, we have provided insights into mechanisms that can lead to fragmentation and will discuss implications of these findings for further study of the cohesion of social networks.

To draw conclusions about real-world friendships or organisational networks a more specific model would be necessary. However, we have provided a basis for such a model. In Chapter 5 we confirmed previous results: edges that are bridges are important for the cohesion of a social network and these edges are characteristically edges with low overlap and high betweenness centrality.

Neither of these measures are able to capture the dependency structures present in many networks, which are the focus of our explanation for fragmentation following a single disagreement. Our work highlights the importance of these hidden dependencies. We could try to get an idea of these dependencies in two different ways. First, it would be possible to question individuals about their sentiments to find out which connections are directly maintained and which ones are not. This might be feasible for smaller networks. For larger networks this approach would, however, suffer from the same problem as other manual data collections: high costs. Furthermore, for organisational networks, we might not be interested in the sentiments themselves but rather in identifying who the facilitators of certain contacts are. For this it does not matter whether two individual would like to actively maintain a work connection, if the organisational structure makes this impossible in any case. Thus, observation of the facilitation structure would be more useful here.

The key fact for indirectly maintained relationships (secondary edges) is that they are always reinforced together with the maintaining edge. Instead of purely thinking in terms of dyads, it would be beneficial to consider groups in this context. This approach is not limited to face-to-face meetings, but could also be used for online interactions such as emails. If two individuals never send direct emails to each other but frequently occur together in email conversations within a larger group of individuals then this would be a strong indicator that the two individuals are connected by an secondary edge. Finding out which other individual (or group of individuals) is the facilitator of this edge might be more difficult than identifying the secondary edge, especially as this situation might not be as clear-cut as in the model.

6.4 Summary

In this chapter we have discussed the shortcomings of other coevolutionary models of social networks and the bias towards models of strategic interactions such as games on networks or opinion dynamics. The field of agent-based social network models seems to centre mainly around models where actors are assumed to make rational decisions and models of non-strategic behaviour are rare. Real human behaviour surely lies somewhere between these extremes for most settings and we have thus argued that non-strategic interactions should be explored more frequently.

In this thesis we have presented such a model that uses non-strategic behavioural rules to build a model of social network formation and maintenance and explore the phenomenon of network fragmentation. Due to not being able to build on an extensive body of previous work in this area, since there is a lack of models with similar assumptions, our work is only a first step and thus oversimplifies human behaviour in the opposite direction to strategic models: it assumes that interactions are entirely nonstrategic. We have discussed these limitations in this chapter and have pointed out parts of the model that could benefit from richer mechanisms. In addition we have discussed how results and mechanism of our model fit in with existing work on theory and empirical work on social networks.

Lastly, we have shared some ideas about how our model complements other modelling approaches for network dynamics, such as opinion dynamics, models of homophily, and group dynamics, and how combining these could lead to richer, more advanced models of coevolutionary social network dynamics.

Chapter 7

Conclusion

In this thesis, we have presented a dynamic network model that is able to grow and maintain networks. These networks exhibit the defining features of social networks, such as sparseness, high levels of clustering, and positive assortativity. In contrast to other models, the interactions between individuals are non-strategic and thus our model complements existing approaches of modelling network dynamics using strategic interactions. Strategic models use explicit make-and-break decisions to drive network change, whereas in our model the network is shaped through positive feedback (reinforcing connections that were reinforced in the past) as well as a range of constraints on interactions inspired by sociological theory.

We have discussed limitations of existing models and highlighted the difficulties present in building models of self-limiting network growth. The edge weight dynamics in our model are self-limiting and thus, in contrast to other models, do not require a global balancing process. While the model presented here is quite simple, we have shown how it could be combined with existing models and incorporate additional processes suggested by empirical evidence to build a richer, more mature model of social network dynamics. Crucially, we have shown that our model exhibits the dynamical features that are essential for studying non-strategic fragmentation dynamics: it reaches stability on the macro-level while exhibiting continuous dynamics on the micro-level.

One aim of this thesis was to show that fragmentation of social networks can occur even when individuals do not consciously choose to break connections to other individuals. We have shown that this is possible and have discussed which edges are crucial for maintaining a cohesive network topology. Unsurprisingly, these edges fulfil a bridging function in the network and we have shown that a measure derived from overlap is well suited to identify such edges. However, this topological feature is not sufficient to fully explain fragmentation. For bridges that consist of several edges, fragmentation will only occur if all of these edges are broken. This can happen as a result of dependencies between edges. These dependencies are a result of the constraints on connections implemented in our model which results in some edges depending on mutual friends for maintenance. It is through these dependencies that a single disagreement can lead to a number of other edges disappearing. This will lead to fragmentation in cases where a bridge consists of one primary edge and several dependent secondary edges. We have proposed some preliminary measures to quantify these dependencies and have discussed how they could be approximated in real-world settings.

We have contributed to the understanding of an important social process that can significantly impact the functioning of social networks with respect to information transfer. As we have discussed, cohesion of communication networks is essential for many settings, such as organisational and team performance. Thus understanding the dynamics of fragmentation is important to identify points at which we might be able to circumvent (or otherwise influence) its onset.

Bibliography

Ahn, Y., Han, S., Kwak, H., Moon, S., and Jeong, H. (2007). Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th international conference on World Wide Web*, pages 835–844. ACM.

Ahrweiler, P., Pyka, A., and Gilbert, N. (2011). A new model for university-industry links in knowledge-based economies*. *Journal of Product Innovation Management*, 28(2):218–235.

Alam, S. and Geller, A. (2012). Networks in agent-based social simulation. In *Agent-based models of geographical systems*, pages 199–216. Springer.

Albert, R. and Barabási, A. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47–97.

Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382.

Albrecht, T. and Adelman, M. (1984). Social support and life stress. *Human Communication Research*, 11(1):3–32.

Albrecht, T. and B.W., B. (1997). *Communication in complex organizations: A relational approach*. Harcour Brace NY.

Albrecht, T. and Hall, B. (1991). Relational and content differences between elites and outsiders in innovation networks. *Human Communication Research*, 17(4):535–561.

Amaral, L., Scala, A., Barthélémy, M., and Stanley, H. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21):11149–11152.

Antonioni, A., Bullock, S., and Tomassini, M. (2014). Reds: an energy-constrained spatial social network model. In Lipson, H., Sayama, H., Rieffel, J., Risi, S., and Doursat, R., editors, *ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems*. MIT Press.

Antonioni, A., Egloff, M., and Tomassini, M. (2013). An energy-based model for spatial social networks. In *Advances in Artificial Life, ECAL*, volume 12, pages 226–231.

Backstrom, L., Bakshy, E., Kleinberg, J., Lento, T., and Rosenn, I. (2011). Center of attention: How facebook users allocate attention across friends. In *Proc. 5th International Conference on Weblogs and Social Media*.

Backstrom, L., Sun, E., and Marlow, C. (2010). Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th international conference on World wide web*, pages 61–70. ACM.

Bansal, S., Grenfell, B., and Meyers, L. (2007). When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, 4(16):879–891.

Bansal, S., Read, J., Pourbohloul, B., and Meyers, L. (2010). The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics*, 4(5):478–489.

Barabási, A. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286:509–512.

Barabási, A., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., and Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3):590–614.

Barnett, L., Di Paolo, E., and Bullock, S. (2007). Spatially embedded random networks. *Physical Review E*, 76(5):056115.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. (2004a). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752.

Barrat, A., Barthélemy, M., and Vespignani, A. (2004b). Weighted evolving networks: coupling topology and weight dynamics. *Physical review letters*, 92(22):228701.

Barrat, A., Barthélemy, M., and Vespignani, A. (2005). The effects of spatial constraints on the evolution of weighted complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2005:P05003.

Barrat, A., Barthelemy, M., and Vespignani, A. (2008). *Dynamical processes on complex networks*, volume 574. Cambridge University Press, Cambridge.

Barthélemy, M. (2003). Crossover from scale-free to spatial networks. *EPL (Europhysics Letters)*, 63:915.

Barthélemy, M., Barrat, A., Pastor-Satorras, R., and Vespignani, A. (2005a). Characterization and modeling of weighted networks. *Physica A*, 346(1):34–43.

Barthélemy, M., Barrat, A., Pastor-Satorras, R., and Vespignani, A. (2005b). Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of theoretical biology*, 235(2):275–288.

Bearman, P., Moody, J., and Stovel, K. (2004). Chains of affection: The structure of adolescent romantic and sexual networks1. *American Journal of Sociology*, 110(1):44–91.

Bedau, M. (2002). Downward causation and the autonomy of weak emergence. *Principia*, 6(1):5–50.

Bernard, H. and Killworth, P. (1977). Informant accuracy in social network data ii. *Human Communication Research*, 4(1):3–18.

Bernard, H., Kilworth, P., Evans, M., McCarty, C., and Shelley, G. (1988). Studying social relations cross-culturally. *Ethnology*, 27(2):155–179.

Bernard, H., Kilworth, P., Kronenfeld, D., and Sailer, L. (1984). The problem of informant accuracy: The validity of retrospective data. *Annual Review of Anthropology*, 13:495–517.

Biely, C., Dragosits, K., and Thurner, S. (2007). The prisoner's dilemma on co-evolving networks under perfect rationality. *Physica D: Nonlinear Phenomena*, 228(1):40–48.

Blasius, B. and Gross, T. (2009). *Reviews of Nonlinear Dynamics and Complexity, Volume 2*, chapter Dynamic and topological interplay in adaptive networks, pages 63–106. Wiley-VCH.

Blondel, V., Guillaume, J., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*, 2008(10):P10008.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308.

Bode, N. W., Franks, D. W., Wood, A. J., Piercy, J. J., Croft, D. P., and Codling, E. A. (2012). Distinguishing social from nonsocial navigation in moving animal groups. *The American Naturalist*, 179(5):621–632.

Boguñá, M., Pastor-Satorras, R., Díaz-Guilera, A., and Arenas, A. (2004). Models of social networks based on social distance attachment. *Phys. Rev. E*, 70:056122.

Bollobás, B. (1998). *Modern graph theory*. Springer Verlag.

Bollobás, B. (2001). *Random graphs*. Cambridge University Press.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280–7287.

Borgatti, S. and Foster, P. (2003). The network paradigm in organizational research: A review and typology. *Journal of management*, 29(6):991–1013.

Borgatti, S. and Halgin, D. (2011). On network theory. *Organization Science*, 22(5):1168–1181.

Borgatti, S., Mehra, A., Brass, D., and Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916):892–895.

Borgatti, S. P. and Cross, R. (2003). A relational view of information seeking and learning in social networks. *Management science*, 49(4):432–445.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177.

Brass, D. (1985). Men's and women's networks: A study of interaction patterns and influence in organizations. *Academy of Management Journal*, 28:327–343.

Brockmann, D. and Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342(6164):1337–1342.

Bryden, J., Funk, S., Geard, N., Bullock, S., and Jansen, V. (2011). Stability in flux: community structure in dynamic networks. *Journal of The Royal Society Interface*, 8(60):1031–1040.

Buldyrev, S., Parshani, R., Paul, G., Stanley, H., and Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028.

Bullock, S., Barnett, L., and Di Paolo, E. (2010). Spatial embedding and the structure of complex networks. *Complexity*, 16(2):20–28.

Burk, W., Steglich, C., and Snijders, T. (2007). Beyond dyadic interdependence: Actor-oriented models for co-evolving social networks and individual behaviors. *International journal of behavioral development*, 31(4):397–404.

Burt, R. (1992). Structural holes: The social structure of competition. *Harvard University Press.*

Burt, R. (2000). Decay functions. *Social networks*, 22(1):1–28.

Buscarino, A., Frasca, M., Fortuna, L., and Fiore, A. (2012). A new model for growing social networks. *Systems Journal, IEEE*, 6(3):531–538.

Calabrese, F., Smoreda, Z., Blondel, V., and Ratti, C. (2011). Interplay between telecommunications and face-to-face interactions: A study using mobile phone data. *PloS one*, 6(7):e20814.

Caldarelli, G., Coccetti, F., and De Los Rios, P. (2004). Preferential exchange: Strengthening connections in complex networks. *Phys. Rev. E*, 70:027102.

Callaway, D. S., Newman, M. E., Strogatz, S. H., and Watts, D. J. (2000). Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85(25):5468–5471.

Cancho, R. and Solé, R. (2003). *Optimization in complex networks*, volume 625, chapter Optimization in complex networks, pages 114–126. Springer.

Carley, K. (1999). On the evolution of social and organizational networks. *Research in the Sociology of Organizations*, 16:3–30.

Cartwright, D. and Harary, F. (1956). Structural balance: A generalization of heider's theory. *Psychological Review*, 63(5):277–293.

Caschili, S. and De Montis, A. (2013). Accessibility and complex network analysis of the us commuting system. *Cities*, 30:4–17.

Castellano, C., Fortunato, S., and Loreto, V. (2009). Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81(2):591–646.

Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J., and Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed rfid sensor networks. *PloS one*, 5(7):e11596.

Chaintreau, A., Hui, P., Crowcroft, J., Diot, C., Gass, R., and Scott, J. (2007). Impact of human mobility on opportunistic forwarding algorithms. *Mobile Computing, IEEE Transactions on*, 6(6):606–620.

Chakrabarti, D., Wang, Y., Wang, C., Leskovec, J., and Faloutsos, C. (2008). Epidemic thresholds in real networks. *ACM Transactions on Information and System Security (TISSEC)*, 10(4):1.

Christakis, N. and Fowler, J. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4):370–379.

Christakis, N. A. and Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, 358(21):2249–2258.

Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. (2000). Resilience of the internet to random breakdowns. *Physical Review Letters*, 85(21):4626–4628.

Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. (2001). Breakdown of the internet under intentional attack. *Physical Review Letters*, 86(16):3682–3685.

Cohen, R. and Havlin, S. (2010). *Complex networks*, volume 1. Cambridge University Press.

Cole, J., Colledge, M., Megaw, T., Powls, M., Bullock, S., and Keen, J. (2005). The implementation of electronic services: planned or organic growth? *Informatics in primary care*, 13(3):187–194.

Colizza, V., Barrat, A., Barthlemy, M., and Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7):2015–2020.

Conrath, D. (1973). Communications environment and its relationship to organizational structure. *Management Science*, 20:586–603.

Contractor, N., Wasserman, S., and Faust, K. (2006). Testing multitheoretical, multilevel hypotheses about organizational networks: An analytic framework and empirical example. *Academy of Management Review*, 31(3):681–703.

Contractor, N., Whitbred, R., Fonti, F., and Steglich, C. (2012). Understanding the ties that bind: A longitudinal investigation of the evolution of a communication network. *Western Journal of Communication*, 76(4):333–357.

Coviello, L., Sohn, Y., Kramer, A., Marlow, C., Franceschetti, M., Christakis, N. A., and Fowler, J. H. (2014). Detecting emotional contagion in massive social networks. *PloS one*, 9(3):e90315.

Crucitti, P., Latora, V., and Marchiori, M. (2004a). A topological analysis of the italian electric power grid. *Physica A: Statistical Mechanics and its Applications*, 338:92–97.

Crucitti, P., Latora, V., Marchiori, M., and Rapisarda, A. (2003). Efficiency of scale-free networks: error and attack tolerance. *Physica A: Statistical Mechanics and its Applications*, 320:622–642.

Crucitti, P., Latora, V., Marchiori, M., and Rapisarda, A. (2004b). Error and attack tolerance of complex networks. *Physica A: Statistical Mechanics and its Applications*, 340(1):388–394.

Cummings, J., Lee, J., and Kraut, R. (2006). *Computers, phones, and the Internet: Domesticating information technology*, chapter Communication technology and friendship during the transition from high school to college, pages 265–278. Oxford University Press.

da Fontoura Costa, L., Oliveira Jr, O. N., Travieso, G., Rodrigues, F., Boas, P., Antiqueira, L., Viana, M., and Rocha, L. (2011). Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics*, 60(3):329–412.

Daley, D. J. and Gani, J. (2000). *Epidemic Modelling.* Cambridge University Press.

Dall, J. and Christensen, M. (2002). Random geometric graphs. *Phys. Rev. E*, 66:016121.

DallAsta, L., Barrat, A., Barthélemy, M., and Vespignani, A. (2006). Vulnerability of weighted networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(04):P04006.

Danon, L., House, T., Read, J., and Keeling, M. (2012). Social encounter networks: collective properties and disease transmission. *Journal of The Royal Society Interface*, 9(76):2826–2833.

Davidsen, J., Ebel, H., and Bornholdt, S. (2002). Emergence of a small world from local interactions: Modeling acquaintance networks. *Physical Review Letters*, 88(12):128701.

Davis, G. and Greve, H. (1997). Corporate elite networks and governance changes in the 1980s. *The American Journal of Sociology*, 103(1):1–37.

De Choudhury, M., Sundaram, H., John, A., and D.D., S. (2010). *Handbook of Social Network Technologies and Applications*, chapter Analyzing the Dynamics of Communication in Online Social Networks, pages 59–94. Springer.

Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(87):87–98.

Demirel, G., Prizak, R., Reddy, P. N., and Gross, T. (2011). Cyclic dominance in adaptive networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 84:541–548.

Dittmer, J. C. (2001). Consensus formation under bounded confidence. *Nonlinear Analysis: Theory, Methods & Applications*, 47(7):4615–4621.

Dodds, P., Muhamad, R., and Watts, D. (2003). An experimental study of search in global social networks. *Science*, 301(5634):827–829.

Doreian, P. and Stokman, F. (1996). Introduction. *The Journal of Mathematical Sociology*, 21:1–2.

Doreian, P. and Stokman, F. (2003). Introduction. *The Journal of Mathematical Sociology*, 27:85–87.

Dorogovtsev, S. and J.F.F., M. (2002). Evolution of networks. *Adv. Phys.*, 51:1079–1187.

Dunbar, R. (2008). Cognitive constraints on the structure and dynamics of social networks. *Group Dynamics: Theory, Research, and Practice*, 12(1):7.

Durrett, R., Gleeson, J., Lloyd, A., Mucha, P., Shi, F., Sivakoff, D., Socolar, J., and Varghese, C. (2012). Graph fission in an evolving voter model. *Proceedings of the National Academy of Sciences*, 109(10):3682–3687.

Eagle, N. and Pentland, A. (2006). Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268.

Eagle, N., Pentland, A., and Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106(36):15274–15278.

Ebel, H. and Bornholdt, S. (2002). Coevolutionary games on networks. *Physical Review E*, 66(5):056118.

Ebel, H., Davidsen, J., and Bornholdt, S. (2002a). Dynamics of social networks. *Complexity*, 8(2):24–27.

Ebel, H., Mielsch, L., and Bornholdt, S. (2002b). Scale-free topology of e-mail networks. *Physical Review E*, 66(3):35103.

Eguíluz, V., Zimmermann, M., Cela-Conde, C., and San Miguel, M. (2005). Cooperation and the emergence of role differentiation in the dynamics of social networks1. *American Journal of Sociology*, 110(4):977–1008.

Eisenberg, E. and Monge, P. (1987). *Handbook of organizational communication*, chapter Emergent communication networks, pages 304–342. Sage Publications.

Fararo, T. and Sunshine, M. (1964). A study of a biased friendship net. *Youth Development Center, Syracuse University.*

Feld, S. (1981). The focused organization of social ties. *American journal of sociology*, 86(5):1015–1035.

Festinger, L., Schachter, S., and Back, K. (1963). *Social pressures in informal groups: A study of a housing project.* Stanford University Press.

Fingerman, K. (2009). Consequential strangers and peripheral ties: The importance of unimportant relationships. *Journal of Family Theory & Review*, 1(2):69–86.

Fischbach, K., Gloor, P., and Schoder, D. (2009). Analysis of informal communication networks–a case study. *Business & Information Systems Engineering*, 1(2):140–149.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3-5):75–174.

Foster, E. A., Franks, D. W., Morrell, L. J., Balcomb, K., Parsons, K., van Ginneken, A., and Croft, D. P. (2012). Social network correlates of food availability in an endangered population of killer whales, *Orcinus orca*. *Animal Behaviour*, 83:731–736.

Franks, D., Noble, J., Kaufmann, P., and Stagl, S. (2008). Extremism propagation in social networks with hubs. *Adaptive Behavior*, 16(4):264–274.

Franks, D. W., Ruxton, G. D., and James, R. (2010). Sampling animal association networks with the gambit of the group. *Behavioral ecology and sociobiology*, 64(3):493–503.

Friedkin, N. E. (1982). Information flow through strong and weak ties in intraorganizational social networks. *Social networks*, 3(4):273–285.

Gastner, M. and Newman, M. (2006). Optimal design of spatial distribution networks. *Physical Review E*, 74(1):016117.

Geard, N. and Bullock, S. (2008). Group formation and social evolution: a computational model. In *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, pages 197–203. MIT Press, Cambridge, MA.

Geard, N. and Bullock, S. (2009). Homophily and competition: a model of group affiliation. In Edmonds, B. and Gilbert, N., editors, *Conference proceedings: The 6th Conference of the European Social Simulation Association*, page 7.

Geard, N. and Bullock, S. (2010). Competition and the dynamics of group affiliation. *Advances in Complex Systems*, 13(4):501–517.

Gil, S. and Zanette, D. (2006). Coevolution of agents and networks: Opinion spreading and community disconnection. *Physics Letters A*, 356(2):89–94.

Gilbert, N. and Troitzsch, K. (2005). *Simulation for the social scientist.* McGraw-Hill International.

Girvan, M. and Newman, M. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.

Gladwell, M. (2000). *The Tipping Point: How Little Things Can Make a Big Difference.* Little Brown.

González, M., Lind, P., and Herrmann, H. (2006). System of mobile agents to model social networks. *Physical Review Letters*, 96(8):088702.

Grabowski, A. (2007). Interpersonal interactions and human dynamics in a large social network. *Physica A*, 385(1):363–369.

Grabowski, A. and Kruszewska, N. (2007). Experimental study of the structure of a social network and human dynamics in a virtual society. *International Journal of Modern Physics C*, 18(10):1527–1535.

Granovetter, M. (1973). The strength of weak ties. *American journal of sociology*, 78:1360–1380.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological theory*, 1(1):201–233.

Grönlund, A. and Holme, P. (2004). Networking the seceder model: Group formation in social and economic systems. *Physical Review E*, 70(3):036108.

Gross, T. and Blasius, B. (2008). Adaptive coevolutionary networks: a review. *Journal of The Royal Society Interface*, 5(20):259–271.

Gross, T., D'Lima, C. D., and Blasius, B. (2006). Epidemic dynamics on an adaptive network. *Physical Review Letters*, 96(20):208701.

Gross, T. and Sayama, H. (2009). *Adaptive networks: Theory, Models and Applications.* Springer, New York.

Grund, T. (2012). Network structure and team performance: The case of english premier league soccer teams. *Social Networks*, 34(4):682–690.

Hallinan, M. and Hutchins, E. (1980). Structural effects on dyadic change. *Social Forces*, 59(1):225–245.

Hamill, L. and Gilbert, N. (2009). Social circles: A simple structure for agent-based social network models. *Journal of Artificial Societies and Social Simulation*, 12(2):3.

Hamill, L. and Gilbert, N. (2010). Simulating large social networks in agent-based models: A social circle model. *Emergence: Complexity and Organization*, 12(4):78–94.

Hanneke, S.and Fu, W. and Xing, E. (2010). Discrete temporal models of social networks. *Electronic Journal of Statistics*, 4:585–605.

Harary, F. (1994). *Graph Theory.* Westview Press.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3):paper 2.

Henry, A., Prałat, P., and Zhang, C. (2011). Emergence of segregation in evolving social networks. *Proceedings of the National Academy of Sciences*, 108(21):8605–8610.

Herrmann, C., Barthelemy, M., and Provero, P. (2003). Connectivity distribution of spatial networks. *Physical Review E*, 68(2):026128.

Hill, R. A. and Dunbar, R. I. (2003). Social network size in humans. *Human Nature*, 14(1):53–72.

Holme, P. (2002). Edge overload breakdown in evolving networks. *Physical Review E*, 66(3):036119.

Holme, P., Edling, C., and Liljeros, F. (2004). Structure and time evolution of an internet dating community. *Social Networks*, 26(2):155–174.

Holme, P. and Kim, B. (2002). Growing scale-free networks with tunable clustering. *Physical Review E*, 65(2):026107.

Holme, P., Min Park, S., Kim, B., and Edling, C. (2007). Korean university life in a network perspective: Dynamics of a large affiliation network. *Physica A*, 373:821–830.

Holme, P. and Newman, M. (2006). Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E*, 74(5):056108.

Huberman, B. and Adamic, L. (2004). Information dynamics in the networked world. *Complex networks*, pages 371–398.

Illenberger, J., Nagel, K., and Flötteröd, G. (2012). The role of spatial interaction in social networks. *Networks and Spatial Economics*, pages 1–28.

Iñiguez, G., Kertész, J., Kaski, K., and Barrio, R. (2009). Opinion and community formation in coevolving networks. *Physical Review E*, 80(6):066119.

Isella, L., Stehlé, J., Barrat, A., Cattuto, C., Pinton, J., and Van den Broeck, W. (2011). What's in a crowd? analysis of face-to-face behavioral networks. *Journal of theoretical biology*, 271(1):166–180.

Ivanova, K. and Iordanov, I. (2012). Two-population dynamics in a growing network model. *Physica A: Statistical Mechanics and its Applications*, 391(4):1811–1821.

Jackson, M. (2005). A survey of network formation models: stability and efficiency. *Group formation in economics: Networks, clubs and coalitions*, pages 11–57.

Jackson, M. and Rogers, B. (2007). Meeting strangers and friends of friends: How random are social networks? *The American Economic Review*, 97:890–915.

Jalili, M. (2013). Social power and opinion formation in complex networks. *Physica A: Statistical Mechanics and its Applications*, 392(4):959–966.

Jeger, M., Pautasso, M., Holdenrieder, O., and Shaw, M. (2007). Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist*, 174(2):279–297.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z., and Barabási, A. (2000). The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654.

Jin, E., Girvan, M., and Newman, M. (2001). Structure of growing social networks. *Physical Review E*, 64(4):046132.

Jo, H., Pan, R., and Kaski, K. (2011). Emergence of bursts and communities in evolving weighted networks. *PloS one*, 6(8):e22687.

Johnson, C. and Gilles, R. (2000). Spatial social networks. *Review of Economic Design*, 5(3):273–299.

Johnson, J., Boster, J., and Palinkas, L. (2003). Social roles and the evolution of networks in extreme and isolated environments. *Journal of Mathematical Sociology*, 27(2-3):89–121.

Jones, J. and Handcock, M. (2003). An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1520):1123–1128.

Kacperski, K. and Hołyst, J. A. (2000). Phase transitions as a persistent feature of groups with leaders in models of opinion formation. *Physica A: Statistical Mechanics and its Applications*, 287(3):631–643.

Kaluza, P., Kölzsch, A., Gastner, M., and Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48):1093–1103.

Kaufmann, P., Stagl, S., and Franks, D. (2009). Simulating the diffusion of organic farming practices in two new eu member states. *Ecological Economics*, 68(10):2580–2593.

Keeling, M. (2005). The implications of network structure for epidemic dynamics. *Theoretical Population Biology*, 67(1):1–8.

Keeling, M. and Eames, K. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4):295–307.

Killworth, P. and Bernard, H. (1976). Informant accuracy in social network data. *Human Organization*, 35(3):269–286.

Kimura, D. and Hayakawa, Y. (2008). Coevolutionary networks with homophily and heterophily. *Physical Review E*, 78(1):016103.

Klimt, B. and Yang, Y. (2004). The enron corpus: A new dataset for email classification research. *ECML 2004*, LNAI 3201:217–226.

Kossinets, G., Kleinberg, J., and Watts, D. (2008). The structure of information pathways in a social communication network. In *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '08, pages 435–443. ACM.

Kossinets, G. and Watts, D. (2006). Empirical analysis of an evolving social network. *Science*, 311:88–90.

Kossinets, G. and Watts, D. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2):405–450.

Kozma, B. and Barrat, A. (2008). Consensus formation on adaptive networks. *Physical Review E*, 77(1):016102.

Krackhardt, D. (1994). *The post-bureaucratic organization: New perspectives on organizational change*, chapter Constraints on the interactive organization as an ideal type, pages 211–222. Sage Thousand Oaks, CA.

Krause, U. (2000). A discrete nonlinear and non-autonomous model of consensus formation. *Communications in difference equations*, pages 227–236.

Krebs, V. (2002). Mapping networks of terrorist cells. *Connections*, 24(3):43–52.

Kumpula, J., Onnela, J., Saramäki, J., Kaski, K., and Kertész, J. (2007). Emergence of communities in weighted networks. *Phys. Rev. Lett.*, 99:228701.

Kumpula, J., Onnela, J.-P., Saramäki, J., Kertesz, J., and Kaski, K. (2009). Model of community emergence in weighted social networks. *Computer Physics Communications*, 180(4):517–522.

Laumann, E., Marsden, P., and Prensky, D. (1989). *Research methods in Social Network Analysis*, volume 61, chapter 3: The boundary specification problem in network analysis, pages 61–87. George Mason University Press Fairfax, VA.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabsi, A., Brewer, D., Christakis, N., Contractor, N.and Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., and Van Alstyne, M. (2009). Computational social science. *Science*, 323(5915):721–723.

Leskovec, J., Backstrom, L., Kumar, R., and Tomkins, A. (2008). Microscopic evolution of social networks. In *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 462–470. ACM.

Leskovec, J. and Horvitz, E. (2008). Planetary-scale views on a large instant-messaging network. In *Proceeding of the 17th international conference on World Wide Web*, pages 915–924. ACM.

Leskovec, J., Huttenlocher, D., and Kleinberg, J. (2010). Signed networks in social media. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 1361–1370. ACM.

Lewis, T. (2011). *Network science: Theory and applications*. John Wiley & Sons.

Leykum, L., Kumar, P., Parchman, M., McDaniel, R., Lanham, H., and Agar, M. (2012). Use of an agent-based model to understand clinical systems. *Journal of Artificial Societies and Social Simulation*, 15(3):2.

Liljeros, F., Edling, C., Amaral, L., Stanley, H., and Åberg, Y. (2001). The web of human sexual contacts. *Nature*, 411(6840):907–908.

Lusseau, D., Schneider, K., Boisseau, O., Haase, P., Slooten, E., and Dawson, S. (2003). The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54(4):396–405.

Manna, S. S. and Sen, P. (2002). Modulated scale-free network in euclidean space. *Phys. Rev. E*, 66:066114.

Marsden, P. (1990). Network data and measurement. *Annual review of sociology*, 16:435–463.

Marsili, M., Vega-Redondo, F., and Slanina, F. (2004). The rise and fall of a networked society: A formal model. *Proceedings of the National Academy of Sciences of the United States of America*, 101(6):1439–1442.

McPherson, J., Popielarz, P., and Drobnic, S. (1992). Social networks and organizational dynamics. *American Sociological Review*, 57(2):153–170.

McPherson, M., Smith-Lovin, L., and Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.

Milgram, S. (1967). The small world problem. *Psychology today*, 2(1):60–67.

Miller, J. H. and Page, S. E. (2009). *Complex Adaptive Systems: An Introduction to Computational Models of Social Life: An Introduction to Computational Models of Social Life*. Princeton University Press.

Miritello, G., Moro, E., Lara, R., Martínez-López, R., Belchamber, J., Roberts, S., and Dunbar, R. (2013). Time as a limited resource: Communication strategy in mobile phone networks. *Social Networks*, 35:89–95.

Monge, P. and Contractor, N. (2001). Emergence of communication networks. *The new handbook of organizational communication: Advances in theory, research, and methods*, pages 440–502.

Monge, P., Rothman, L., Eisenberg, E., Miller, K., and Kirste, K. (1985). The dynamics of organizational proximity. *Management Science*, 31(9):1129–1141.

Moreno, Y., Pastor-Satorras, R., and Vespignani, A. (2002). Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 26(4):521–529.

Morgan, D., Neal, M., and Carder, P. (1996). The stability of core and peripheral networks over time. *Social Networks*, 19(1):9–25.

Moss, S. (2008). Policy modelling: Problems and prospects. Technical report, Technical report, CPM Report.

Motter, A. and Lai, Y. (2002). Cascade-based attacks on complex networks. *Physical Review E*, 66(6):065102.

Mukherjee, A., Choudhury, M., Peruani, F., Ganguly, N., and Mitra, B., editors (2013). *Dynamics On and Of Complex Networks, Volume 2*, volume 2. Springer.

Nardini, C., Kozma, B., and Barrat, A. (2008). Who's talking first? consensus or lack thereof in coevolving opinion formation models. *Phys. Rev. Lett.*, 100:158701.

Newman, M. (2001). Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132.

Newman, M. (2002). Assortative mixing in networks. *Physical Review Letters*, 89(20):208701.

Newman, M. (2003a). The structure and function of complex networks. *SIAM review*, 45(2):167–256.

Newman, M. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582.

Newman, M. and Park, J. (2003). Why social networks are different from other types of networks. *Physical Review E*, 68(3):036122.

Newman, M. E. J. (2003b). Properties of highly clustered networks. *Physical Review E*, 68:026121.

Newman, M. E. J. (2010). *Networks: an introduction.* Oxford University Press.

Newman, M. E. J. (2011). Complex systems: a survey. *Am. J. Phys.*, 79:800–810.

Noble, J., Davy, S., and Franks, D. (2004). Effects of the topology of social networks on information transmission. In *From Animals to Animats 8: Proceedings of the Eighth International Conference on Simulation of Adaptive Behavior*, pages 395–404. The MIT Press.

Noble, J., Silverman, E., Bijak, J., Rossiter, S., Evandrou, M., Bullock, S., Vlachantoni, A., and Falkingham, J. (2012). Linked lives: the utility of an agent-based approach to modeling partnership and household formation in the context of social care. In *Simulation Conference (WSC), Proceedings of the 2012 Winter Simulation Conference*, pages 1–12. IEEE.

Onnela, J., Saramäki, J., Hyvönen, J., Szabó, G., Argollo de Menezes, M., Kaski, K., Barabási, A., and Kertész, J. (2007a). Analysis of a large-scale weighted network of one-to-one human communication. *New Journal of Physics*, 9:179.

Onnela, J., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., and Barabási, A. (2007b). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336.

Pacheco, J., Traulsen, A., and Nowak, M. (2006). Coevolution of strategy and structure in complex networks with dynamical linking. *Physical review letters*, 97(25):258103.

Padgett, J. and Ansell, C. (1993). Robust action and the rise of the medici, 1400-1434. *American journal of sociology*, 98(6):1259–1319.

Palla, G., Barabasi, A., and Vicsek, T. (2007). Quantifying social group evolution. *Nature*, 446(7136):664–667.

Panzarasa, P., Opsahl, T., and Carley, K. (2009). Patterns and dynamics of users' behavior and interaction: Network analysis of an online community. *Journal of the American Society for Information Science and Technology*, 60(5):911–932.

Papa, M. (1990). Communication network patterns and employee performance with new technology. *Communication Research*, 17(3):344–368.

Park, K., Lai, Y., Gupte, S., and Kim, J. (2006). Synchronization in complex networks with a modular structure. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(1):015105–015105.

Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63(6):066117.

Pattison, P. and Robins, G. (2004). Building models for social space: neighourhood-based models for social networks and affiliation structures. *Mathématiques et sciences humaines. Mathematics and social sciences*, 168(168):11–29.

Penrose, M. (2003). *Random geometric graphs*, volume 5. Oxford University Press Oxford.

Poncela, J., Gómez-Gardeñes, J., Traulsen, A., and Moreno, Y. (2009). Evolutionary game dynamics in a growing structured population. *New Journal of Physics*, 11(8):083031.

Popielarz, P. (1999). In(voluntary) association: A multilevel analysis of gender segregation in voluntary organizations. *Gender & Society*, 13(2):234–250.

Potterat, J., Muth, S., Rothenberg, R., Zimmerman-Rogers, H., Green, D., Taylor, J., Bonney, M., and White, H. (2002). Sexual network structure as an indicator of epidemic phase. *Sexually transmitted infections*, 78(suppl 1):i152–i158.

Potterat, J., Rothenberg, R., and Muth, S. (1999). Network structural dynamics and infectious disease propagation. *International journal of STD & AIDS*, 10(3):182–185.

Prettejohn, B. and McDonnell, M. (2011). Effect of network topology in opinion formation models. In Guttmann, C., Dignum, F., and Georgeff, M., editors, *Collaborative Agents - Research and Development*, volume 6066. Springer Berlin, Heidelberg.

Raub, W., Buskens, V., and van Assen, M. (2011). Micro-macro links and microfoundations in sociology. *The Journal of Mathematical Sociology*, 35:1–25.

Ravasz, E., Somera, A., Mongru, D., Oltvai, Z., and Barabási, A. (2002). Hierarchical organization of modularity in metabolic networks. *Science*, 297(5586):1551–1555.

Read, J., Eames, K., and Edmunds, W. (2008). Dynamic social networks and the implications for the spread of infectious disease. *Journal of the Royal Society Interface*, 5(26):1001–1007.

Reis, H., Maniaci, M., Caprariello, P., Eastwick, P., and Finkel, E. (2011). Familiarity does indeed promote attraction in live interaction. *Journal of personality and social psychology*, 101(3):557–570.

Roberts, S., Wilson, R., Fedurek, P., and Dunbar, R. (2008). Individual differences and personal social network size and structure. *Personality and individual differences*, 44(4):954–964.

Robins, G. and Pattison, P. (2001). Random graph models for temporal processes in social networks. *Journal of Mathematical Sociology*, 25(1):5–41.

Robins, G., Pattison, P., Kalish, Y., and Lusher, D. (2007a). An introduction to exponential random graph (p^*) models for social networks. *Social networks*, 29(2):173–191.

Robins, G., Snijders, T., Wang, P., Handcock, M., and Pattison, P. (2007b). Recent developments in exponential random graph (¡ i¿ p¡/i¿*) models for social networks. *Social Networks*, 29(2):192–215.

Rosenquist, J., Murabito, J., Fowler, J., and Christakis, N. (2010). The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine*, 152(7):426–433.

Rosvall, M. and Sneppen, K. (2009). Reinforced communication and social navigation generate groups in model networks. *Physical Review E*, 79(2):026111.

Rothenberg, R., Potterat, J., and Woodhouse, D. (1996). Personal risk taking and the spread of disease: beyond core groups. *Journal of Infectious Diseases*, 174(Supplement 2):S144–S149.

Rothenberg, R., Potterat, J., Woodhouse, D., Muth, S., Darrow, W., and Klovdahl, A. (1998). Social network dynamics and hiv transmission. *Aids*, 12(12):1529–1536.

Salathé, M. and Jones, J. (2010). Dynamics and control of diseases in networks with community structure. *PLoS computational biology*, 6(4):e1000736.

Salzinger, L. (1982). The ties that bind: The effect of clustering on dyadic relationships. *Social Networks*, 4(2):117–145.

Sampson, S. (1968). *A novitiate in a period of change: An experimental and case study of social relationships*. PhD thesis, Cornell University, September.

Santos, F., Pacheco, J., and Lenaerts, T. (2006). Cooperation prevails when individuals adjust their social ties. *PLoS Computational Biology*, 2(10):e140.

Scott, J. (1988). Social network analysis. *Sociology*, 22(1):109.

Shaw, L. and Schwartz, I. (2010). Enhanced vaccine control of epidemics in adaptive networks. *Physical Review E*, 81(4):046120.

Shirley, M. and Rushton, S. (2005). The impacts of network topology on disease spread. *Ecological Complexity*, 2(3):287–299.

Sinclair, M., Siemieniuch, C., Haslam, R., Henshaw, M., and Evans, L. (2012). The development of a tool to predict team performance. *Applied ergonomics*, 43(1):176–183.

Skyrms, B. and Pemantle, R. (2000). A dynamic model of social network formation. *PNAS*, 97(16):9340–9346.

Smith-Lovin, L. and McPherson, J. (1993). You are who you know: A network approach to gender. *Theory on gender/feminism on theory*, pages 223–251.

Snijders, T. (1996). Stochastic actor-oriented models for network change. *Journal of mathematical sociology*, 21(1-2):149–172.

Snijders, T. (2002). The statistical evaluation of social network dynamics. *Sociological methodology*, 31(1):361–395.

Snijders, T., Van de Bunt, G., and Steglich, C. (2010). Introduction to stochastic actor-based models for network dynamics. *Social networks*, 32(1):44–60.

Snijders, T. and van Duijn, M. (1997). *Simulating Social Phenomena*, volume Lecture Notes in Economics and Mathematical Systems, chapter Simulation for statistical inference in dynamic network models, pages 493–512. Springer.

Solé, R. and Montoya, M. (2001). Complexity and fragility in ecological networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1480):2039–2045.

Spraragen, M., Landwehr, P., Ranganathan, B., Zyda, M., Carley, K., Chang, Y., and Mahaswaran, R. (2013). Cosmopolis: A massively multiplayer online game for social and behavioral research. *Journal of Artificial Societies and Social Simulation*, 16(1).

Steglich, C., Snijders, T., and West, P. (2006). Applying siena. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 2(1):48–56.

Stokman, F. and Doreian, P. (1996). Concluding remarks. *The Journal of Mathematical Sociology*, 21(1-2):197–199.

Strogatz, S. (2001). Exploring complex networks. *Nature*, 410:268–276.

Sutcliffe, A., Dunbar, R., Binder, J., and Arrow, H. (2012). Relationships and the social brain: Integrating psychological and evolutionary perspectives. *British journal of psychology*, 103(2):149–168.

Szell, M. and Thurner, S. (2010). Measuring social dynamics in a massive multiplayer online game. *Social Networks*, 32(4):313–329.

Takács, K., Janky, B., and Flache, A. (2008). Collective action and network change. *Social Networks*, 30(3):177–189.

Toivonen, R., Kovanen, L., Kivelä, M., Onnela, J., Saramäki, J., and Kaski, K. (2009). A comparative study of social network models: Network evolution models and nodal attribute models. *Social Networks*, 31(4):240–254.

Toivonen, R., Kumpula, J., Saramäki, J., Onnela, J., Kertész, J., and Kaski, K. (2007). The role of edge weights in social networks: modelling structure and dynamics. *Noise and Stochastics in Complex Systems and Finance*, 6601(1):66010B.

Toivonen, R., Onnela, J., Saramäki, J., Hyvönen, J., and Kaski, K. (2006). A model for social networks. *Physica A*, 371:851–860.

Van den Bulte, C. and Moenaert, R. (1998). The effects of r&d team co-location on communication patterns among r&d, marketing, and manufacturing. *Management Science*, 44(11-Part-2):S1–S18.

van Duijn, M., Zeggelink, E., Huisman, M., Stokman, F., and Wasseur, F. (2003). Evolution of sociology freshmen into a friendship network. *Journal of Mathematical Sociology*, 27(2-3):153–191.

Van Segbroeck, S., Santos, F., Lenaerts, T., and Pacheco, J. (2011). Selection pressure transforms the nature of social dilemmas in adaptive networks. *New Journal of Physics*, 13(1):013007.

Van Segbroeck, S., Santos, F., and Pacheco, J. (2010). Adaptive contact networks change effective disease infectiousness and dynamics. *PLoS Computational Biology*, 6(8):e1000895.

Van Segbroeck, S., Santos, F. C., Lenaerts, T., and Pacheco, J. M. (2009). Reacting differently to adverse ties promotes cooperation in social networks. *Physical Review Letters*, 102(5):058105.

Vázquez, A. (2003). Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Phys. Rev. E*, 67:056104.

Vazquez, F., Eguíluz, V., and Miguel, M. (2008). Generic absorbing transition in coevolution dynamics. *Physical review letters*, 100(10):108702.

Vespignani, A. (2010). The fragility of interdependency. *Nature*, 464:984–985.

Volz, E. (2008). Susceptible-infected-recovered epidemics in populations with heterogeneous contact rates. *The European Physical Journal B-Condensed Matter and Complex Systems*, 63(3):381–386.

Volz, E. and Meyers, L. (2007). Susceptible–infected–recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society B: Biological Sciences*, 274(1628):2925–2934.

Wang, D., Shi, X., McFarland, D., and Leskovec, J. (2012). Measurement error in network data: A re-classification. *Social Networks*, 34:396–409.

Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge university press.

Watts, D. (1999a). *Small worlds: the dynamics of networks between order and randomness*. Princeton university press.

Watts, D. J. (1999b). Networks, dynamics, and the smallworld phenomenon. *American Journal of Sociology*, 105(2):493–527.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.

Weeks, M., Clair, S., Borgatti, S., Radda, K., and Schensul, J. (2002). Social networks of drug users in high-risk sites: Finding the connections. *AIDS and Behavior*, 6(2):193–206.

Wellman, B. (1926). The school child's choice of companions. *The Journal of Educational Research*, 14(2):126–132.

Wesolowski, A., Eagle, N., Tatem, A. J., Smith, D. L., Noor, A. M., Snow, R. W., and Buckee, C. O. (2012). Quantifying the impact of human mobility on malaria. *Science*, 338(6104):267–270.

West, D. (2000). *Introduction to graph theory*. Prentice Hall.

White, D. and Harary, F. (2001). The cohesiveness of blocks in social networks: Node connectivity and conditional density. *Sociological Methodology*, 31(1):305–359.

Wong, L., Pattison, P., and Robins, G. (2006). A spatial model for social networks. *Physica A: Statistical Mechanics and its Applications*, 360(1):99–120.

Wu, F., Huberman, B., Adamic, L., and Tyler, J. (2004). Information flow in social groups. *Physica A: Statistical Mechanics and its Applications*, 337(1):327–335.

Wu, J.-j., Gao, Z.-y., and Sun, H.-j. (2006). Cascade and breakdown in scale-free networks with community structure. *Physical Review E*, 74(6):066111.

Wu, X. and Liu, Z. (2008). How community structure influences epidemic spread in social networks. *Physica A: Statistical Mechanics and its Applications*, 387(2-3):623–630.

Xulvi-Brunet, R. and Sokolov, I. (2002). Evolving networks with disadvantaged long-range connections. *Physical Review E*, 66:026118.

Xulvi-Brunet, R. and Sokolov, I. M. (2007). Growing networks under geographical constraints. *Phys. Rev. E*, 75:046117.

Yook, S., Jeong, H., and Barabási, A. (2002). Modeling the internet's large-scale topology. *Proceedings of the National Academy of Sciences*, 99(21):13382–13386.

Yook, S., Jeong, H., Barabási, A., and Tu, Y. (2001). Weighted evolving networks. *Physical Review Letters*, 86(25):5835–5838.

Zachary, W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33(4):452–473.

Zahn, G. (1991). Face-to-face communication in an office setting the effects of position, proximity, and exposure. *Communication Research*, 18(6):737–754.

Zanette, D. and Gil, S. (2006). Opinion spreading and agent segregation on evolving networks. *Physica D: Nonlinear Phenomena*, 224(1):156–165.

Zanette, D. and Risau-Gusmán, S. (2008). Infection spreading in a population with evolving contacts. *Journal of biological physics*, 34(1):135–148.

Zimmermann, M., Eguíluz, V., and San Miguel, M. (2004). Coevolution of dynamical states and interactions in dynamic networks. *Physical Review E*, 69(6):065102.

zu Erbach-Schoenberg, E., Bullock, S., and Brailsford, S. (forthcoming 2014). A model of spatially constrained social network dynamics. *Social Science Computer Review, Special issue on Social interaction*, 32(3).

zu Erbach-Schoenberg, E., McCabe, C., and Bullock, S. (2011). On the interaction of adaptive timescales on networks. In *Proceedings of the 11th European Conference on Artificial Life*, pages 900 – 907. MIT Press.