NTNU – Trondheim
Norwegian University of
Science and Technology

Cyber Insurance & Insurable Topologies

Håvard Råmundal Halse
Jonas Hoemsnes

Submission date: June 2013
Supervisor: Gergely Biczók, Postdoc
Responsible professor: Jan A. Audestad, Professor II

Norwegian University of Science and Technology
Department of Telematics

Title: Cyber Insurance & Insurable Topologies

Students: Håvard Råmundal Halse & Jonas Hoemsnes

Problem description:

Security breaches are increasingly prevalent in the Internet age causing huge financial losses for companies and their users. Cyber-insurance is a powerful economic concept that can help companies in the fight against such malicious behavior. Earlier research suggests that cyber- insurance has failed to reach its promising potential, although the concept of cyber-insurance has been around since the 1980s. The researchers claims that a functional model for cyber-insurance has to handle its unique problems regarding interdependent security, correlated-risk and asymmetrical-information. These challenges can be described and analyzed by network graphs, and positively some graphs will yield overall higher security (insurable topologies) than other graphs. In order to cater for cyber-insurance, it is essential to understand how to create new or transform existing networks to insurable topologies.

The students will:

- conduct a background study and a market survey to validate the current state of cyber- insurance
- study and characterize graphs describing insurable topologies
- build a model of network formation which gives rise to such insurable topologies
- apply the model to investigate a realistic ecosystem, e.g., cloud computing

Supervisor: Gergely Biczók, Postdoc

Responsible professor: Jan A. Audestad, Professor II

Abstract

Cyber-insurance is a powerful economic concept that can help companies in the fight against cyber-attacks. From the early 90s, most researchers claimed that cyber-insurance had a positive future; it would become a huge economical tool for handling residual cyber-risks.

The market study of the thesis revealed that both the European and US cyber-insurance market is have failed to grasp its promising potential. The US-market has matured more compared to the European-market, but both still have failed compared to the potential market, too fully grasp this potential they need some innovative approaches to handle the unique problems of cyber-insurance.

The thesis proposes several cyber-insurance network formation models, and uses game theory and a simulation tool, Netlogo, to analyze these models. In every model, there are introduced new properties that relate the model to the real world and real insurance products. The results show that insurers can use the insurance-premium as a tool for determining the resulting formation of the network. If the premium is set to the right level, certain structures will evolve, in recent literature these formations have shown to possess properties who make them particularly good for cyber-insurance network. Such as minimizing the average cost, enabling the insurer to calculate the overall risk and possibly increased overall security and utility.

We believe our findings will help the cyber-insurance market evolve, by giving the insurers a proper tool to better analyze and control their cyber-insurance network, in this way helping the cyber-insurance market reach its promised potential.

Further work should try mapping our models and simulations to real world networks in a more convincing way. This could be achieved by finding and introducing better suited risk functions, and by letting nodes choose their neighbors by preference, not randomly.

Preface

This study serves as a master thesis in the 10^{th} semester of our Master of Science degree in Communication Technology at the Norwegian University of Science and Technology.

We would like to thank everyone that have contributed and supported our work throughout this semester. A special thanks is given to our supervisor Gergely Biczók, Postdoc at the Department of Telematics (ITEM), for valuable feedback, ideas and guidance during the project period.

Håvard Råmundal Halse & Jonas Hoemsnes
Trondheim, Norway
June, 2013

Contents

List of Figures

Chapter 1
Introduction to Cyber Insurance

1.1 Motivation

Security breaches are increasingly prevalent in the Internet age, causing huge financial losses for companies and their users. Cyber-insurance is a powerful economic concept that can help companies in the fight against such malicious behavior. Earlier research suggests that cyber-insurance has failed to reach its promising potential, although the concept of cyber-insurance has been around since the 1980s.

The researchers claims that a functional model for cyber-insurance has to handle the problems regarding interdependent security, correlated risk and asymmetrical information. Many researchers have proposed models to solve these problems, but the market still strives to succeed. Another problem with cyber-insurance is to determine the overall risk in the network. If cyber-insurance networks were describable and analyzable by graphs, the calculation of overall risk would be much easier. We ask whether there exist network structures that are superior as cyber-insurance networks compared to other networks. If so, is it possible for insurers to determine the structure of these networks, or even better to create new or transform existing networks into a certain structure?

1.2 Problem definition

In this project, the goal is to analyze the current state of the cyber-insurance market. Study and characterize network structures suited to be used as a cyber-insurance network. A desired structure will possess some characteristics that would be beneficial for a cyber-insurance network. Additionally, we will build a model, which can relate to different real world scenarios, using network formation to force the creation of these structures.

1.3 Reader's guide

Chapter 1 introduces the concept of cyber-insurance and presents a survey of the current cyber-insurance market.
Chapter 2 discusses and summarizes related work.
Chapter 3 shows how graphs can describe real-world networks, such as airline routes and stock markets. The chapter also presents and discusses the properties of graphs well suited for cyber-insurance networks. These graphs are the foundation for the models created in chapter 5.
Chapter 4 presents the basic concepts of graphs and game theory. It also presents the simulation tool, Netlogo, used to simulate the different models of Chapter 5.
Chapter 5 presents different endogenous network formation models, where the nodes are agents, with or without insurance, seeking to establish links with eachother.
Chapter 6 discusses and summarizes the findings in the thesis, with focus on chapter 5.

1.4 Introduction

Security breaches are increasingly prevalent in the Internet age causing huge financial losses for companies and their users. When facing security breaches and risk, there are typically four ways to act [BL08b]:

1. Avoid the risk
2. Retain the risk
3. Self protect and mitigate the risk
4. Transfer the risk

The ICT industry have so far tried to prevent risks with a mixture of options two and three. This has lead to many different techniques and software trying to detect threats and anomalies, to protect the users and infrastructure. Firewalls, intrusion-detection and prevention systems, are some of the solutions. These will reduce the risk, but do not eliminate the risk completely. Although they are all good and needed actions, it is impossible to achieve perfect cyber-security, due to many reasons: Threats are continuously evolving, there will always be accidents and security flaws, attackers have different intentions, network externalities and free-riding in security networks, the lemons-market in security products, misaligned incentives between users and product vendors, and many more. This is why we need cyber-insurance, as an fourth option, to handle the residual risk [LB09, PH13].

Market potential for cyber-insurance. Just like regular insurance, cyber-insurance is an insurance product used to transfer financial risk, associated with computers and network-related incidents, over to a third party. This third party willingly accepts the risk, in exchange for a fee, called insurance premium. The insurance is focused on computer-related issues, and could provide coverage against property loss and theft, data damage, cyber-extortion, loss of income due to denial of service attacks or computer failures, reduced reputation and customers churning due to leaked user information and so on. Traditional insurance policies rarely cover incidents like these, therefore a specialized insurance product is needed to handle these residual risks. I.e. there is a huge potential market for cyber-insurance [PD12].

As mentioned, the concept of cyber-insurance has been around since the 1980s, and has failed to reach its promising potential. There might be several reasons for this slow development, however, it is believed that the main reason so far, is that no model deals with all the unique problems of cyber-insurance at the same time. In addition to the known difficulties of insurance, such as calculating risk, cyber-insurance differs from traditional insurance because it has to deal with the problem of asymmetric information, correlated risk and interdependent security [GGJ+10].

Traditional Insurance. The basic structure of cyber-insurance relates to traditional insurance, where an insurance contract (policy) binds the insurance company to pay a specified amount to the insurance holder whenever an incident occurs. In return, the insurance holder has to pay a fixed monthly or annual fee (premium) to the insurance company. The contract includes a risk assessment of the company's vulnerability and clearly specifies the entitled amount of coverage for each of the different risks. These assessments are used to calculate the companies' premium [Rob12]. Generally, this means that the security is negatively correlated with the premium costs. In cyber-insurance this means that the better the security, the lower the price on the insurance premium.

Generally, to ensure that their business is economically viable, the insurance company will require that insurable risks possess seven distinct characteristics [MCR80]:

1. Large number of similar exposure units: Insurance is based on the principle of pooling resources, where insurance policies are offered to individual members of a large class. Meaning the more customers, the closer the predicted losses will get to the actual losses.

2. Definite loss: A loss should take place at a known time, in a known place and from a known cause. Incidents such as a fire or car crash, are examples where these terms are easy to verify.

3. Accidental loss: The event that triggers a claim should not be something the insurer has discretion or control over.

4. Large loss: The size of the loss must be meaningful from the perspective of the insured. Insurance premiums need to cover both the expected cost of the loss, and in addition, cover all the expenses regarding issuing and administrating policies, adjusting losses and supplying the capital needed to be able to pay claims.

5. Affordable premium: The premium must be proportional to the security offered, otherwise no one will offer/buy the insurance. In the situation where the likelihood of the insured event is high, and the cost is large, it is unlikely that the insurance company will offer the insurance, or if so the premium would be very high.

6. Calculable loss: Both the probability and the cost of an insurable event, has to at least be possible to estimate.

7. Limited risk of catastrophically large losses: If losses happen all at the same time, the likelihood of the insurance company getting bankrupt is high. Therefore, losses are ideally independent and non-catastrophic.

This model will also apply to the risks covered by cyber-insurance. Unfortunately there are additional obstacles regarding cyber-insurance. The three major problems with cyber-insurance are related to; information asymmetry, interdependent security and correlated risk.

Information asymmetry. Information asymmetry arises when one side of a transaction or decision has more or better information than the other party. There are two different cases of information asymmetry. The first one is called adverse selection, where one party simply has less information regarding the performance of the transaction. A good example is when buying health insurance, if a person with bad health purchases insurance, and the information about her health is not available to the insurer, we have a classical adverse selection scenario, where the insurer probably charges too little. We can observe a similar situation for the cyber-industry, where an insurer has no way of confirming whether your network is "healthy", i.e. not contaminated or infected. The other information asymmetry scenario is called moral hazard. It occurs after the signing of the contract, where one party deliberately takes some action that makes the possibility of loss higher, e.g. choosing not to lock your door, since you have insurance. Or in the computer setting, deliberately visiting hostile web-pages, or not using anti-virus software, firewalls or other self-protection software, although you are required to do so. [Pal12].

The task of measuring the level of security is very hard, and in order to lower the premiums people will have an incentive for hiding information about their security level, hence the problem with asymmetry is highly relevant. Another problem occurs on the customer side of the market. For a customer wanting to improve his/her defense mechanisms, the software security market often becomes a lemon's market[1]. It is difficult for the buyer to distinguish the performance of different software products, and thus the reasonable thing to do, is to buy the cheapest. Therefore, the good security products must cost the same as the bad. If the cost of producing good security software is too high, the problem can even result in abandoning the production of good software, because it would not be profitable.

Correlated risk. Another big concern regarding cyber-insurance, is the correlated risk. Among other things, the problem occurs due to the need for standards. Standardization is an important part of the business of computers and computer networks. Generally it enables computers to communicate, install and use different software. A good example is operative systems for personal computers, today we only have a small set of operative systems available, and these systems are standardized, so they can use the same communication channels. The standards generate a lot of the value in the ICT industry, but they also make many threats possible. All systems that use the same standards, create a large number of similar exposure units, i.e. they share common vulnerabilities, which could be exploited at the same time. As we see, this violates the insurance characteristic of limited risk of catastrophically large losses. Thus create a significant difficulty for the cyber-insurance industry, because when a security breach occurs there is a high probability that it will occur to a large number of people, i.e. catastrophic and extreme events occur with a higher probability than in the regular insurance business. To compensate, the logical thing to do would be to raise the premium cost, this could however violate the characteristics of affordable premiums and large losses. If the security breach is large, it could even potentially cause so much damage, that the insurers will not be able to pay all the customers who suffered, and they could go bankrupt.[BS10]

Interdependent security. Another problem in the ICT industry is interdependent security, meaning that you are not only dependent on your own investment in security, but also on everyone else's. Investment in security generates positive externalities, and as public goods, this encourages free riding. Why should I pay for security when I can just free ride on security invested by others? The problem is

[1]Lemon market, the problem of quality uncertainty, was first introduced in a paper [Ake97] by the economist George Akerlof in 1970,and used the market for used cars as an example.[Wik] The conclusion of the paper is that since the buyers lack information to distinguish a bad car(lemon) from a good one(cherry), the buyer will not pay the price the seller wants for a cherry, and the seller will not sell a cherry for the price of a lemon, and thus the lemons drive the cherries out of the market.

that the reward for a user investing in self-protection depends on the security in the rest of the network. i.e. The expected loss due to a security breach at one agent in the network, is not only dependent on this agent's level of investment in security, but also on the security investment done by adjacent agents, and their adjacent agents and so forth. A good example of this is the amount of spam sent every day, which depends on the number of compromised computers. Meaning if you have invested in security software of some kind, you still receive lots of spam because many other people have not invested [Böh10].

Calculating losses As mentioned, a problem in several areas of insurance is the calculation of risk. In cyber-insurance, the unique obstacles contribute to making this particularly difficult. When facing a security breach there are two potential loss classes: [BMR09, MCR80]

- Primary losses or first-degree losses: direct loss of information or data and operating loss. These arise from disuse, abuse or misuse of information. The cost of these arises from recovering, loss of revenue, PR and information sharing costs, hiring of IT specialists etc.
- Secondary losses are indirectly triggered. These are the loss of reputation, goodwill, consumer confidence, competitive strength, credit rating and customer churning.

The cost of the loss from both these classes can be difficult to determine, although the second one is probably the most difficult, since it is challenging to put a value on e.g. how many potential customers did they lose due to the reputation loss, how many customers churned, and what was their value etc. It could also be difficult to determine when the loss happened, where and what caused it. Another problem when calculating losses and determining insurance premiums, is the unavailablity of large amounts of historic data on cyber-crimes, which are needed in many insurance models to calculate the risk, losses and premiums. This problem arises i.a. because many firms do not reveal details about their experienced security breaches. [HH07]

Cyber-insurance instead of security. Another problem with cyber-insurance is actors seeing it as a solution to the problem of being secure. Instead of investing in security, they now have a way of buying their way out. However, as the paper [BL08a] shows, this problem might be solved with the right pricing options, meaning that the insurance companies can create pricing models which make it economically beneficial to invest in security and cyber-insurance. Cyber-insurance can be used as an incentive for buying security. Such models will also make sense for the insurance company, since better security systems yields lower probability for incidents.

As we see there are many problems regarding cyber-insurance, but the insurance industry has been dealing with many difficult problems in other areas of life. Cyber-insurance faces many challenges, but we can't say that internet risks and damages can not be insured. We just need to find a way of helping the insurers to create a better product, i.e. the challenge is to find a way for the insurers to handle these special characteristics, in order to create a healthy cyber-insurance market. [LB09] To help establishing a healthy cyber-insurance market, one needs to know its current status.

1.5 The cyber-insurance market

The market for cyber-insurance emerged in the late 80's, when security software companies began collaborating with insurance companies to offer insurance policies together with their security products. From a marketing perspective, adding insurance helped highlighting the supposedly high quality of the security software. Nevertheless, this new product was a comprehensive solution, which dealt with both risk reduction and residual risk [BL08b]. Continuing into the beginning of the new millennium, several companies started offering standalone cyber-insurance, which sat the frame for the current insurance product. In Norway, startup-companies, such as Safensure AS were established with the goal to deliver cyber-insurance to the Norwegian and European market [dig]. In addition, already well-established insurance companies, such as Gjensidige Nor, started offering insurance products intended for the web-site market. These insurances were created to insure lost income due to malicious hackers, denial of service and other well know cyber-attacks at that time. In 2001 Gjensidige Nor, in cooperation with the German company Tela Versicherung, offered businesses insurance against financial losses due to hacker attacks and sabotage in a range up to 5 million NOK, given that the companies could provide proof that specified security measures were taken [it].

Despite the fact that cyber-insurance has been around for a couple of decades, the market still struggles to gain a foothold. Safensure AS does not longer exist and Gjenside Nor does not advertise a cyber-insurance product anymore. There seems to be many challenges for both buyers and sellers. Buyers face tremendous confusion about cyber risks and their potential impacts on business. The paper [PpD12] points out that people do not know or understand what kinds of risks the cyber-space involves, and how fatal the losses can be. Even when companies have decided to purchase a cyber-insurance, they are confused with what kind of insurance they should purchase, it is difficult to see what it covers, what is a reasonable price etc. Thus, the market for cyber-insurance tends to become a lemon.s market, where the buyer lacks knowledge, and struggles to see the differences between the different insurance contracts.

The UK and US markets. We wanted to reveal the current status of the cyber-insurance market. We limited our survey to the UK and US markets, in addition to the Norwegian market. The first impression reveals that there are several different results and opinions regarding the health of the global cyber-insurance market. The paper [Ins11] studied a sample of 50 organizations in various industry sectors, located in the United States. They showed that on average every company suffered more than one successful attack every week, and argued that successful cyber-attacks could have serious financial consequences. They found that the median cost of cyber-crime in the U.S is $5.9 million per year, ranging from $1.5 million to $36.5 million per company, which is a 56 % increase from 2010.

Another paper [Ris12] collected statistics about cyber-attacks in the UK, and claims that the costs are expected to be £27 billion a year, which makes cyber-crime one of UK's biggest emerging threats. In addition, the paper pointed out that the victims are not only large companies like Google and PlayStation, but also small businesses. Despite these numbers, only 35 % of the companies in the survey had purchased cyber-insurance. This is surprisingly low, since they found no shortage of providers. It was revealed that in the UK there are nine insurers who specializes in cyber-insurance, and in the US around 30-40 insurers.

A UK firm, called CFC underwriting, who offers insurance to small and medium sized businesses, published an article [New] claiming promising numbers for the US cyber-insurance market. On US soil, 20-50% of the businesses purchased either standalone cyber-insurance or benefits from coverage provided in their existing insurance. However, despite recent years' focus on the increasing cyber-crime activity and the catastrophic consequences of having weak security, only 1% of European businesses are enrolled in an insurance program covering cyber-threats. A more optimistic survey, [Pra], pointed out that more and more insurance companies offer cyber-insurance. Yet, of the 13000 companies, only 46 % reported that they were insured against the economic consequences of cyber-attacks.

The media coverage on corporate threats such as Stuxnet and the attacks on Playstation, which lead to a compromise of 77 million user accounts including credit card numbers [Chu], shows that the cyber-threats are growing, and one would assume that we are in need of cyber-insurance. However, even though the number varies, the surveys show that a large share of companies have chosen not to protect themselves against the residual risk of cyber-attacks, by buying cyber-insurance.

The Norwegian market. Our survey of the Norwegian insurance market revealed that specialized cyber-insurance companies, such as Safensure AS, do not exist anymore. Only one out of the five biggest insurance companies[2] offers something

[2]Gjensidige, If Skadeforsikring, DNB, TRYG, Storebrand

similar to a cyber-insurance. Gjensidige Nor offers what they call operation-loss-insurance, which covers expenses due to reconstruction of files and reinstalling software and denial of service attacks. In addition, it is also possible to insure against hacking and sabotage [Nor]. From email correspondence with Gjensidige Nor it was clear that they needed lots of information regarding the company to be able to calculate the insurance premium. They required extensive information about the economic health of the company, and a model of what kind of software and hardware where used with estimated values on each component. Unfortunately, we were not able to obtain the cost of such an insurance.

Market outlook. The survey from [New] claimed that the US cyber-insurance market was much more mature than the European market. A possible reason is the breach notification laws. In the US, 46 states have mandatory breach notification laws, combined with significant penalties for companies failing to protect sensitive data. This means that the US government is creating incentives for firms to buy cyber-insurance. In Europe, only Germany and Austria have similar laws, forcing companies to notify affected customers of data leakage. A recent proposal of the EU wants to introduce the notification law in Europe, and also include penalties for serious data breaches, which could be set as high as 2 % of a company's global revenue [New]. It is proposed that the law should take effect in 2014, although this is highly unlikely, considering the complexity of the effects of this law. Undoubtedly such a law would be a healthy injection to the cyber-insurance market. However, a market based on fear of the consequences of not being insured is not desirable. The ultimate goal for cyber-insurance is to correlate the purchase of cyber-insurance with companies growing desire to invest in more security, and hence lower the risk of being a victim of cyber-crimes. The article claims that the way to meet this goal, is to focus on the serious brand damage a company will experience and not just on the financial loss.

In summary, the cyber-insurance market seems to have a huge potential, but needs some new thinking to fully take advantage of it. We will take another approach and focus on finding network structures that will help the insurers offer fair contracts, which is beneficial for both the customers and the suppliers. Hopefully, this can help in the process of establishing a healthy cyber-insurance market.

Chapter 2
Related work

2.1 Cyber-insurance

While several authors have expressed doubts about the future of cyber-insurance, the authors of [BS10] still have faith in the prevalence of cyber-insurance. The paper describes the three main problems of cyber-insurance as mentioned in chapter 1 of this thesis; information asymmetry, correlated risk and interdependent agents. They argue that a model for cyber-insurance has to overcome each of these obstacles. Instead of presenting a solution, they propose a framework to classify models of cyber-insurance. The framework breaks the modeling down to five key components:

- network environment (nodes controlled by agents, who extract utility. Risk arises here.)
- demand side (agents)
- supply side (insurers)
- information structure, distribution of knowledge among the players.
- organizational environment. Public and private entities whose actions affect network security and agent's security decisions.

The goal is that this unifying framework will help navigating the literature and stimulate research that results in a more formal basis for policy recommendations involving cyber-risk reallocation. [BS10] encourage answering questions such as; under what conditions will a cyber-insurance market thrive? What is the effect of an insurance market, -will the Internet be more secure? Does cyber-insurance contribute to social welfare? Böhme, et.al. also analyze several other papers on cyber-insurance, and show how all of them are touching the problems and key components showed above, but no paper handles all of them. The paper studies other existing models,

and reveals a discrepancy between informal arguments in favor of cyber-insurance and analytic results questioning the viability of a cyber-insurance market.

The paper [MYK06] summarizes the evolution of cyber-insurance, from the early primitive hacker insurance policies, to the modern and comprehensive cyber-insurances. The insurances have evolved, since the insurers have better understanding of the risks and the needs of the businesses. They show how insurers are addressing the adverse selection and moral hazard problem, by classifying the risk level of the insured. They do so by requesting lots of background information about the customers. The paper presents a methodology for calculating the social welfare loss due to adverse selection. The paper is optimistic about the future of cyber-insurance, and concludes that cyber-insurance is making the Internet a safer environment, because insurers are giving users economic incentives to self-protect.

The paper [PGP11] from Pal, et.al. presents a cyber-insurance model which handles both risks due to security (e.g virus) and non-security related features such as power outage and hardware failure. Their model, Aegis, is a simple model in which the user accepts a fraction of loss recovery and the rest is transferred to the insurance company. Pal, et.al. show that when it is mandatory to purchase insurance, risk averse agents would prefer Aegis contracts over traditional cyber-insurance products. The model also gives users an incentive to take greater responsibility in securing their own systems. Hence this answers one of the questions from [BS10]: The overall security of the Internet will increase if Aegis is offered to the market. An interesting result from their analysis is the fact that a decrease/increase in the insurance premium may not always lead to increase/decrease in demand. From the insurer's point of view, this feature means that it might be possible to increase margins without losing market share. Hence, it will be easier to create a market for cyber-insurance.

[PH12] adopts a topological perspective in proposing a mechanism that accounts for the positive externalities (due to purchase of security mechanisms) and network location of users. In addition the authors provide an appropriate way of proportionally imposing fines/rebates on user premiums. This feature relates to our model, where a central node in the network receives a bulk insurance discount, in order to facilitate the creation of star topologies.

[PH13] presents the importance of discriminating network users in insurance contracts. This is done to prevent adverse selection, partly internalizing the negative externalities of interdependent security, achieving maximum social welfare, helping a risk-averse insurer to distribute costs of holding safety capital among its clients, and insurers sustaining a fixed amount of profit per contract. The paper proposes a mechanism to pertinently contract discriminate insured users when having complete network information. This is important since almost every node in the network is

different from other nodes. Hence we need a way of distinguishing good nodes from bad ones by the means of premium price.

High correlation in failures in information systems is a huge concern to the cyber-insurance market. The paper [BK06] introduces a new classification of correlation properties. The authors divide it into two levels, the first level is the correlation within firms, and the second level addresses the global correlations. Further, they create an economic model for risk arrival at these two levels, and use it in simulations to find where a market for cyber-insurance can exist or not. Böhme, et.al. results show that cyber-insurance is best suited for risk-classes where internal correlation is high, and the global correlation is low. This could be one of the reasons for the failure of cyber-insurance so far, the insurers have focused on the wrong markets, i.e. the markets with high global correlation and low internal correlation. One problem with [BK06] is that they do not cover interdependent security neither discuss the problem of information asymmetry.

The papers [BL08a, BL08b] present how risk management on the internet have only introduced methods to reduce the risks, such as firewalls, intrusion detection systems, anti virus etc. But none of these have managed to remove the risk completely. As mentioned in chapter 1, there are four possible ways of removing risk: avoid it, retain it, self-protect and mitigate it or transfer the risk. Most entities on the internet have chosen a mix of of retaining and mitigate by self-protecting. These solutions do not eliminate risk completely, and threats evolve over time. Thus, the only option for completely removing the risk, is to transfer it to a party who willingly accepts it, in exchange for a fee. The key result of these papers is that they show economic reasons for users to not invest in self-protection, and that cyber-insurance will act as an incentive for users to acquire self-protection, i.e. the level of security in the internet will increase with cyber-insurance. The reason for this positive spiral is that investment in insurance will result in overall higher payoff, and since the premiums discriminate users based on the investment in self protection, it will act as a strong incentive to acquire self-protection.

In contrast to the papers [BL08a, BL08b], the paper [SSFW10] claims that in a competitive cyber-insurance market the users' utility will improve, but the network security will worsen relative to a market without cyber-insurance. By competitive insurers Shetty et.al, mean that there are several insurance contracts to choose from, and user choose the one that increases their utility the most. [SSFW10] create and explains two models, in the first one the insurers suffer from information asymmetry regarding the users' investment in security. For most of the parameters used, there will not be offered any insurance in an equilibrium, due to the moral hazard problem. In the second model, the insurer is able to observe the users' security investment, and can thus contract discriminate bad and good users, i.e. no moral hazard is

present. In this model the insurance increases the users' utility, but it does not result in overall better security.

The paper [RKK08] also claims that so far in security management of the internet, there is no solution to the residual risk. To solve this the authors come up with an insurance policy, which they claim can survive in a competitive market. However, this solution does not cover the problem of correlated risk. The actors in the model from [RKK08] are the ISP, who is also the insurer, and the users, the users are of type high- or low-risk users. The question asked is whether a policy that brings profit to the ISPs while protecting the users from risk, exists. Radosavac, et.al find an optimal insurance policy that can be offered to both low and high risk users. Additionally, they also conclude that even when there is an insurance for residual risk, it cannot be guaranteed that a profitable business model exists.

The paper [DS06] describes an interesting network formation game. Although the paper tries to observe susceptibility to sybil attacks in peer-to-peer networks, the approach used for network formation can be related to our thesis. The game's characteristics is as follows: Nodes are either friends or strangers, and the goal of the nodes is to selfishly try to fulfill own communication needs. The nodes' needs is to communicate with as many as possible of their friends. This can be achieved either by direct or indirect connections. Every node has a link budget, i.e. a maximum number of links it can establish, and a set of friends it wants to connect to. [DS06] proposes two random games where nodes might have to take the risk of connecting to non-insured nodes.

1. Random model: Every node in the network initiates a set of friendships with other nodes, denoted F. All nodes have the same link budget $L < F$.

2. Unbalanced Random Mode. The same friendship graph as in the random model is created. However, one of the nodes has a significantly larger link budget ($L_0 > 2F$)

The first model does not result in any equilibrium, except the one where friends only connect to other friends. The other model shows some new insights, when the link budget is comparable to their number of friends, most nodes still choose to only connect to friends. However, when the link budget is set to only one link, except for the rich node, then the resulting equilibrium is a star topology.

2.1.1 Summary

There are many different papers that have described the problems of cyber-insurance, and proposed different models and solutions. However, as we revealed, the cyber-

insurance market is yet far from established and still has lots of potential. Each of the presented models has a slightly different angle towards improving the cyber-insurance market. Although promising result presented in the papers, few improvements have appeared in the market, and it seems that another approach is needed. This is what we intend to do in this thesis. In brief, we will now investigate whether there are any advantageous structures for cyber-insurance and in chapter 5 see if it is possible for networks to evolve endogenously into these structures.

Chapter 3
Graphs and Network Formation

In nature and society, many scenarios can be described using graphs. Infrastructure, such as railroads, water pipelines and electricity grid, societal relationships and disease epidemics, can all be visualized using graphs. Cyber-insurance is no exception, and can also be structured as a graph. This is of interest because, when one can describe a phenomenon with graphs, it is easier to analyze and possibly find some characteristics, hence the graph can be used as an analytic tool [Aud].

Several studies have been done on the characteristics of different graphs, such as E-R graphs and A-B graphs (scale-free graphs), these are thoroughly described in the methodology chapter of this thesis. In addition, one has found special characteristics of star-shaped graphs and cliques. This chapter will highlight which characteristics that are desirable in the cyber-insurance market, and which structures that possess these characteristics. These findings will serve as the foundation of our models, where we try to force these graph structures to emerge.

3.1 Real-world graph structures

As a starting point, let's have a look at a couple of real-world examples of how complex systems with huge amount of data could be structured as graphs. We will see how complex structures become rather intuitive when presented as graphs. By looking at the graph structure, one can determine what type of graph that appears, and hence certain characteristics will apply.

Stock markets. The research paper [Gar07] analyzes the correlation between different stocks in the Greek stock market in year 1997. The authors compared the daily closing price of stock i at day t, and compared the similarity of a pair of stocks i and j by using the correlation coefficient. The idea is that the correlation coefficient between a pair of stocks can be expressed using different distances in a graph structure. A short distance means high correlation and a long distance means

low correlation between the stocks. Normally, this network would be shown as a fully connected graph, which will consist of $\frac{n(n-1)}{2}$ edges, and would be difficult to analyze. However, the new approach presents a clear and understandable graph, consisting of $(n-1)$ edges showing the correlations between the stocks.

The resulting graph can be seen in Figure 3.1, and shows a network consisting of several clusters linked together. Instead of having to analyze a complex system with huge amounts of data, the stock market can be analyzed by its topological properties, such as the high clustering coefficient, i.e a scale free topology, which will among other things point out which stocks have the most influence on others.

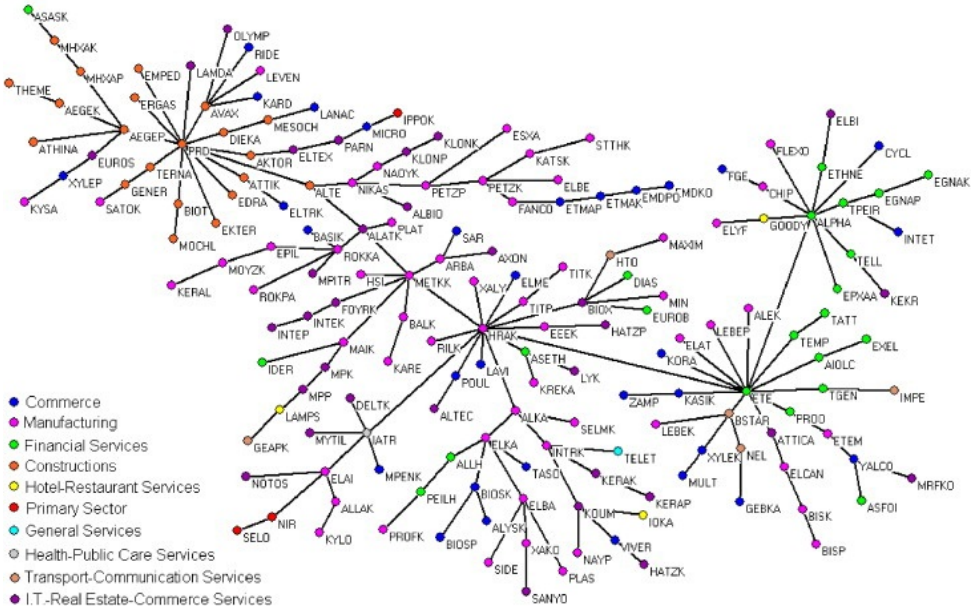


Figure 3.1: Network obtained by comparing two stocks' correlation coefficient in the Greek stock market (Athens Stock Exchange, ASE) in year 1997. The different colors represent the different sectors of economic activity [Gar07].

Airline routes. Another real-world network which shows the same characteristics as scale-free graphs is the map of airline routes. Figure 3.2 shows the US route map of the American airline company SkyWest. The characteristic clustering emerges in the figure, where a majority of the flights departing from either Denver, Chicago or San Francisco. Not surprisingly, these airports are all in the top 7 busiest airports in the US [Faa], and serve as hubs for many of SkyWest's flights. In the

airline industry some airports are called hubs, because that's what they are, - a connection point for major parts of the network of flights. The network of flights, as depicted in Figure 3.2, follows the characteristics of A-B graphs. Hence, as we can confirm from looking at the graph, the network are vulnerable against direct attacks, meaning that shutting down a low degree airport wont create much trouble. However, if one of the hubs is forced to close, it will provoke huge delays throughout the whole network, because a majority of the destinations is interconnected via the hubs.

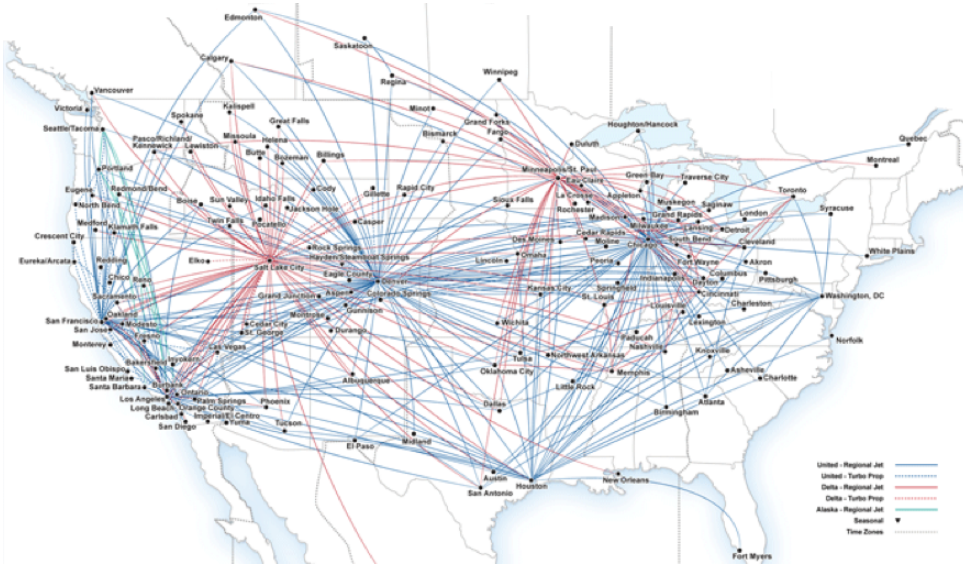


Figure 3.2: SkyWest Airline's combined route map [CfAPA].

Here, both examples can be characterized as scale-free networks, and the work done by Albert and Barabási shows that a large part of natural systems is in fact scale-free graphs [Aud]. Since we are able to determine the graph's type, which in this case is a scale-free graph, we now know that the graph is vulnerable to attacks directed towards the hubs, i.e. the hubs need to be secured. For example, if a delay occurs at an airline hub, these delays will probably cascade throughout the network. This shows the strength of being able to structure systems as graphs. When certain structures appear, one can assume that the network will behave according to a set of rules. This is why we wish to determine whether there are any structures that possess preferred characteristics for cyber-insurance, and then find a proper way to force these formations to evolve.

3.2 Network Structures

Just like stock markets and airline routes, the cyber-insurance market can be described using graphs. The structure that will evolve is dependent on all the nodes and how they connect with each other. The insurer can determine the cost of establishing a link, and thus determine which nodes will connect to eachother. This is what we will try to achieve in our models. However, first we need to shed light on what kind of graph structures that would be desirable to force upon the cyber-insurance market.

To find the proper structure, many different scenarios should be covered. In a network an agent's actions are influenced by its neighborhood structure, i.e. the network connections will affect each individual agent's payoff, meaning that agents are dependent on eachother, and the probability of cascading failures are highly relevant. -If one or more fails, e.g. bankruptcy, failure to deliver at the expected time, system shut down, higher cost etc, then the whole network will be affected. In this case there are several types of networks to consider, every social and economic interaction where an agent's well-being is dependent on externalities as well as on his own actions, is a network worth considering.

We found several interesting papers from evolutionary studies and disease epidemics, which described characteristics in different graph structures. The ones we found appropriate, were those which described the benefits of star- and clique-shaped graphs. These graphs showed characteristics that could be used to make it feasible for both the insurer to offer - and the customer to acquire insurance.

The paper [LHN05] is about evolutionary dynamics and how some structures can amplify or sustain evolution and drift[1]. One aspect of cyber-insurance is risk, and knowledge of how, for example, viruses spread in a network and how to use graph structures to prevent both hackers from entering and virus from spreading, is important. Evolutionary dynamics, and the research of how mutant genes spread throughout a population, as described in the paper, is analogous to this issue. If we can determine some structures where certain nodes are advantageous/disadvantageous, then these structures will have imporant properties, such as sustaining viruses from spreading, or amplify the incentive for obtaining cyber-insurance.

The paper [LHN05], shows that mutants inserted into a circulation graph, will have a fixation probability equal to

$$p_1 = \frac{(1 - \frac{1}{r})}{(1 - \frac{1}{r^N})} \tag{3.1}$$

[1]Drift is the opposite of selective evolution, it is when the network/structure evolve and change at random

Where r represents the relative fitness of the mutant i.e the agents security level, if it is advantageous it will have a certain chance of fixation, and disadvantageous mutants will have a chance of extinction. A circulation graph is a graph that satisfies these two properties:

1. The sum of all edges leaving a vertex is equal for all vertices
2. The sum of all edges entering a vertex is equal for all vertices

A clique is a good example of a circulation graph, and the probability of fixation is as in Eq. (3.1). The fixation probability determines how probable it is that the whole network will eventually be "infected" by the mutant. Which means that it determines the rate of evolution, which relies on both the size of the network and the evolution speed. If the relative fitness of the nodes is high, then the probability of fixation will be low. A probability equal to one means that every node in the network will eventually be affected by the mutant.

An essential part of cyber-insurance is as mentioned earlier, for the insurer to be able to calculate the overall risk of the instance to be insured. Since the probability of fixation can be calculated in circulation graphs, if the insurer knows that the instance is part of a circulation graph, it is possible for the insurer to calculate the probability of fixation in that network. If we can find graphs with an fixation probability that exceeds Eq.(3.1) it is even better, because then the insurer is not only able to calculate the overall probability of fixation, but also to show that the probability of fixation is higher than the one for circulation graphs. [LHN05] shows that such graphs exist, and one example is the star topology, (see Figure 3.3). In this topology the fixation probability is as shown in Eq.(3.2), or more generally Eq.(3.3).

$$p_2 = \frac{(1 - \frac{1}{r^2})}{(1 - \frac{1}{r^{2N}})} \tag{3.2}$$

.

$$p_k = \frac{(1 - \frac{1}{r^k})}{(1 - \frac{1}{r^{kN}})} \tag{3.3}$$

When comparing Eq.(3.1) and Eq.(3.2), we see that the selective difference is amplified from r to r^2, i.e. a star acts as an evolutionary amplifier, favoring advantageous mutants and inhibiting disadvantageous mutants.

There are other graphs where the fixation probability is equal to 3.3, examples are super-stars, such as funnels and metafunnels. These are just more complex star networks. This paper shows that as N gets large, the super-stars will have a

fixation probability, for an advantageous mutant, that converges to 1, and for a disadvantageous mutant converges to 0. As exemplified earlier in this chapter, we know that there are many topologies in our society that are so called scale-free graphs. These graphs have most of their connectivity clustered in a few vertices, which are very similar to a network interconnected by multiple stars, these networks can also be considered as potent selection amplifiers.

The paper [Blu11] present interesting results regarding network formation games. The authors set up a game where the nodes benefit from direct links, but these links also expose them to risk. Each node gains a payoff of a per link it establishes, but it can establish a maximum of δ links. A failure occurs at a node with probability q, and propagates on a link with probability p. If a node fails, it will receive a negative payoff of b, no matter how many links it has established. The characteristics of this game is transferable to how we expect nodes in a cyber-insurance network to interact with eachother. Therefore, the results of the overall payoff change according to different collection of participants.

The results from the model presented by Blumen et.al. shows a situation where clustered graphs achieve a higher payoff when connected to trusted nodes, compared to when connecting with random nodes. Unlike in anonymous graphs, where nodes connect to each other at random, nodes in these graphs share some information with their neighbors, which is used when deciding whether to form a link or not. To further explain these results, they show that there exists a critical point, called *phase transition*, which occurs when nodes have a node degree of $\frac{1}{p}$. At this point a node gets a payoff of $\frac{a}{p}$, and to further increase the payoff the node needs to go

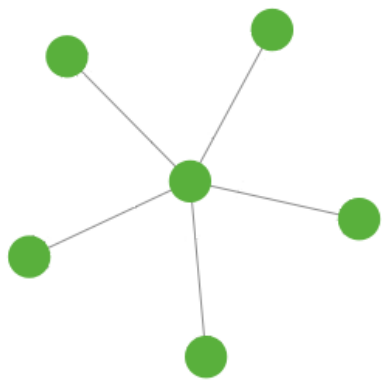
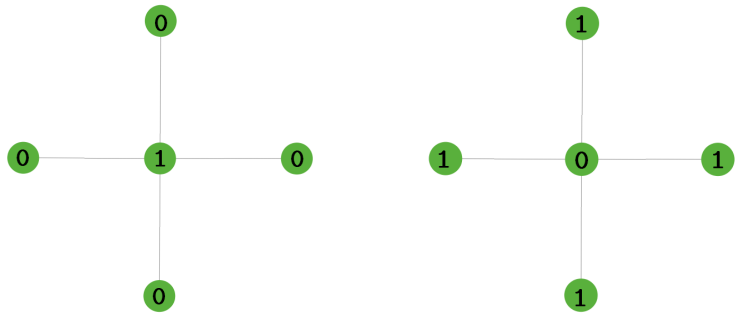


Figure 3.3: A star topology.

into a region with significantly higher failure probability. Because once each node establishes more than $\frac{1}{p}$ links, the contagious edges will with high probability form a large cluster, which results in a rise in probability of node failure, and reduces the overall welfare. From this the paper states that when the minimum welfare exceeds $(1 + f(\delta) * \frac{a}{p})$ we have reached *super-critical payoff*. Otherwise it is called *sub-critical payoff*. Further Easley et.al, show that the only possible way of ending up with super critical payoff, is by forming clustered networks consisting of cliques with slightly more than $\frac{1}{p}$ nodes. However, if the nodes form an anonymous market, by random linking, they can only get sub-critical payoff. In other words, if the nodes can choose who they connect with, and by doing so, create trusted clustered markets, they can achieve a higher payoff by exceeding the critical node degree point.

The paper [GGJ+10] shows how network games evolve when the payoffs are determined not only by your own decisions, but also by your neighbors. This can be used to analyze the star network further. The authors analyze a game on public goods, which is simple but highly relevant for our work. A good example of a public goods is a security product. A security product suffers from strategic substitutes, i.e. if your neighbor acquires a security product, you have less incentive to also acquire the security product. This is because when he acquires it, he gets more secure, and so do you, due to the positive externalities of the product.

The game is set up like this: We have an action space: $X = \{0, 1\}$, where 1 can be considered as acquiring information, taking vaccine, buying security software etc. 0 is not doing so. Each node i has a set of neighbors: N_i, and a payoff function $y_i = x_i + \bar{x}N_i$. The gross payoff to player i is 1 if $y_i >= 1$ and 0 otherwise. But each player also suffers from a cost of $0 < c < 1$ if he chooses action 1. Looking at



(a) Socially Optimal equilibrium, center node choose action 1

(b) Non Socially Optimal equilibrium, leaf nodes choose action 1

Figure 3.4: Figure 3.4a shows the socially optimal equilibrium, and Figure 3.4b shows the non optimal equilibrium.

Figure 3.4, we easily see that there are two equilibriums. One where the center node chooses action 1 and the rest of the nodes choose action 0, and a second equilibrium where all the leaf nodes choose 1 and the center chooses 0. The overall payoff in these two differs from eachother, the latter is not socially optimal because it suffers from a cost equal to: $\#leafnodes * c$, while the other equilibrium only has a total cost of c. It would have been beneficial if we were able to force the game to always end up in the socially optimal equilibrium.

From an insurer's point of view. If an insurance company could identify these star structures, and force them to end up in the socially optimal equilibrium, i.e. minimize the overall cost of link establishment, it would have been very beneficial for both the insurer and the customers. First of all, if the insurer could identify these structures, he could calculate the overall probability of fixation by a contagions node(virus, worm, trojan or other failures). If one could ensure that the center node is protected, one could also calculate the probability of the contagions node being extinguished from the network, and possibly being able to ensure that the network is secure, at least with high probability. One possibility of achieving this could be by offering very cheap insurance to the leaf nodes, and giving the center node an incentive to acquire security products by informing the center node about the probability of failure unless he acquires security, and offer him a decent rebate if he acquires the security product, and a very expensive insurance if not. In this way the insurer could force a rational center node into getting both insurance and a security product, and thus increase the security in the whole network.

This is a simple scenario, analyzing an exogenous network formation[2], but it shows how an insurer can force a star network to end up in the social optimal cost equilibrium. Leading to overall higher security for in the network. We also showed how the insurer could calculate the probabilities of fixation in circulation, star, funnel, meta-funnel and super-star graphs. Can the insurer force cyber-insurance networks to evolve into any of these structures, and at the same time separate the nodes into trusted and untrusted environments? If so, this could contribute significantly to solving the problems of cyber-insurance. The problems of information asymmetry and interdependent risk is reduced. Because, if the insurer knows the network structure, he can calculate the probabilities of failure and catastrophic events. If the network is a star and the insurer can ensure that the center node is secure, the interdependent risk problem is limited to the security of the center node.

[2]Exogenous: The network formation is given. Endogenous: The structure originates from within the network, i.e. the opposite of exogenous

3.3 Research Question

Until now, our thesis has introduced cyber-insurance, presented related work on the issues regarding cyber-insurance and this chapter has presented the properties of different graph structures and briefly introduced the idea of network formation. Generally, the papers in the related work section have presented different models for solving the problems with cyber-insurance. Nevertheless, as we have seen, the cyber-insurance market still fails to evolve, despite all the solutions presented in the different papers. This is why we have chosen to take a different approach. In this chapter, we have shown some structures, especially the star and clique, which could generate benefit for both the insurer and customers in a cyber-insurance market. We will combine the knowledge of these structures and network formation games to investigate networks consisting of nodes, insured or not, wanting to increase their payoff by establishing links with eachother. Is it possible for the insurer to force these networks to evolve endogenously into these structures? We will focus on how the insurer can determine the resulting formation by adjusting the parameter he can control, i.e. the insurance cost. We know that if the insurance premium is too high, no one will buy it. On the other hand, if it is too low, everyone would benefit from having insurance, and insured nodes will make risky decisions, such as connecting to risky nodes. We will try to determine whether it is possible to find the intersections, where the desired structures will evolve, and both the insurer and their customers will benefit from this.

Chapter 4
Methodology

4.1 Graphs

As mentioned in the previous chapter, graphs are good analytical tools when studying complex systems. Since we will use graphs extensively throughout this thesis, it is important to establish an understanding of basic graph properties. Figure 4.1 depicts the basics of an unweighted graph, where the edges are not assigned any value. Weighted edges can be useful to e.g. reflect capacity constraints such as a link's maximum bandwidth, or the length of a road (edge), but will not be used in this thesis. Other common definitions used when describing graphs are listed below [Aud]:

- Edge degree: Number of edges connected with a node.
- Hub: Node with high edge degree.
- Cycle: A chain originating and terminating at the same node.
- Cluster: Subgraph of highly connected nodes.
- Cluster coefficient: Probability for two nodes to be adjacent to a third node.
- Clique: Subgraph where all nodes are adjacent (cluster coefficient = 1).
- Small world graph: Graph with small diameter and large cluster coefficient (e.g. the Internet and A-B graphs, described in section 4.2).

4.2 Random Graphs

Cyber-insurance covers many fields, from financial transactions and software development to computer networks. Many of these fields share a common characteristic, they can all be described as a graph, and often as a random graph. Therefore, the study

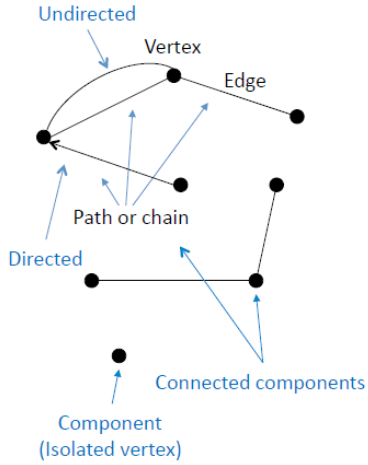


Figure 4.1: General graph [Aud].

of random graphs is of special concern. The research on random graphs is fairly new compared to other mathematical discoveries. The first extensive results where found by Erdös and Rényi in 1959, hence the resulting structures were called E-R graphs. Later and probably with more accurate results were the work of Albert-Barabási in 1999 [Aud], leading to the characterization of A-B graphs.

Erdös-Rényi Graphs. E-R graphs are networks created between a fixed number of n-nodes, where each node connects to another of the $n-1$ nodes with probability p. The resulting graph will on average contain $\frac{n(n-1)p}{2} \approx \frac{n^2p}{2}$ edges [Bol85]. By analysing the graph, the authors found some interesting properties:

- If $p < n^{-2}$ then there is no edges in the graph.
- If $p = c/n$ where c is a constant between $1 < c < log\, n$, the graph will provoke a single large component to grow within the graph.
- If $p > (ln\, n)/n$ then the graph is completely connected.
- If $p = 1/n$ triangles start forming in the graph.

A fully connected E-R graph will have a short diameter similar to the Internet, and thus could be a very good description of structures similar to the Internet. However, the edge degree follows a Poisson distribution, which means that the edge degrees are peaking around the average value [Aud]. E-R graphs do not capture the immense clustering coefficient which is present in networks such as the Internet. In

other words, E-R graphs are not small world graphs, and a different graph structure is needed to model computer networks. An interesting fact about these graphs is their vulnerability. These graphs are very vulnerable against random attacks, such as nature disasters, but robust against directed attacks. Due to the fact that if you remove all edges from one node, little damage is done, since the network is not dependent on only a few nodes. Every node has approximately the same node degree, and it is the sum of all the nodes' connections that creates the network.

Albert-Barabási Graphs. The structure which is believed to be most accurate for modelling computer networks are A-B graphs. A-B graphs are different from E-R graphs since they are scale-free, meaning that the vertices do not have a constant value throughout the entire graph. Albert and Barabási found that the edge degree of each vertex follows a power law distribution; meaning that the probability that the edge degree is g is proportional to $g^{-\gamma}$ where γ is usually a number between 2 and 3. This distribution is called a thick-tail distribution, because there is a significant probability that a node may have a very high degree [Aud]. These graphs are in contrast to E-R-graphs, very vulnerable to directed attacks, because if you take out a hub, mayor parts of the network will be affected. But the graph is very robust against random attacks, which is why most of the networks we observe in nature can be depicted as A-B-graphs. A-B graphs can grow and become scale-free if every new node is connected to one or more already existing node with a probability proportional to the edge degree of that node. The paper present an algorithm that creates A-B graphs and Figure 4.2 shows a graph that evolves from this algorithm:

- A new single vertex is added to the graph.
- This vertex is connected to exactly two other vertices in the graph.
- The probability that the new vertex connects to another vertex is dependent on the edge degree of the other vertex, higher edge degree meaning higher probability
- There is only one edge between two vertices.

In addition to the high clustering coefficient Albert-Barabási showed that A-B-graphs have a fairly small diameter, which can be seen in Figure 4.2. The World Wide Web, neural networks, scientific referencing, co-authorship and many other types of networks are very similar to A-B graphs [Aud].

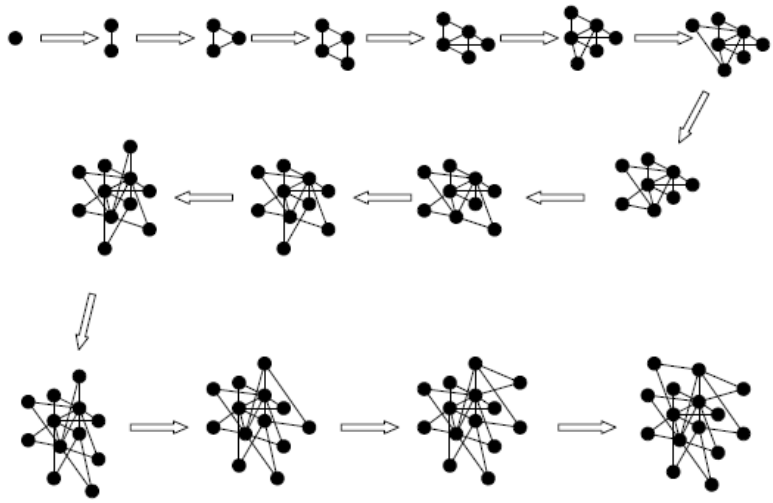


Figure 4.2: Forming an A-B graph in 15 generations [Aud].

4.3 Game Theory

Here we will present some of the game theory concepts used in our models, for more thorough explanation of game theory, see: [NRTV07, Wat08].

One shot game. This type of game assumes that players act at the same time instantly, therefore there is no causality. A game in strategic (normal) form can be described by three elements:

- the set of players $i \in I$, which we take to be the finite set $1, 2, ..., I$.
- the pure-strategy space $s_i \in S_i$ for each player i, where s_i is a possible action of player i.
- and payoff functions U, which give the players utility functions for each profile $s = (s_1, s_2, ..., s_I)$ of strategies.

A general solution concept for games of economic interest is the Nash Equilibrium solution. A Nash Equilibrium is a profile of strategies such that each players strategy is a best response to the other player's strategies.

Nash Equilibrium. A pure strategy profile s^* is a Nash equilibrium if, for all players i

$$U_i(s_i^*, s_{-i}^*) \geq U_i(s_i, s_{-i}^*) \in S_i \tag{4.1}$$

Subgame-perfect equilibrium. A strategy profile s is a subgame perfect equilibrium if it represents a Nash Equilibrium of every subgame of the original game.

Socially optimal. A socially optimal outcome is the set of choices that maximizes the sum of all players' payoffs.

Price of Anarchy. The price of Anarchy (PoA) of a network game, measures the efficiency of the network, by comparing the equilibrium outcome with the socially optimal outcome. The reason for this possible inefficiency is that agents act selfishly and do not necessarily consider other agents' payoff when choosing an action. In our thesis, the price of anarchy will be a number between 0 and 1, where 1 is the socially optimal outcome.

Stackelberg. Also known as a leader-follower game, it introduces multiple stages. The leader commits himself first, chooses his strategy, then the followers respond sequentially. The Stackelberg model can be solved to find the subgame perfect Nash Equilibrium, i.e. the strategy profile that serves each player best, given the strategies of the other players and that entails every player playing in a Nash Equilibrium in every subgame.

Bayesian game. In Bayesian games, information about the other players' characteristics is incomplete. In these types of games, there is one player(the agent) who knows both types, and another player (the principal) who does not know the type of the other player. There are two types of equilibriums in this game: A pooling equilibrium, is an equilibrium where both types of the agent choose the same action, i.e. the principal is not able to distinguish the two types. A separating equilibrium is an equilibrium where the agents of different types choose different actions, and thus the principal is able to determine the agent's type by observing his actions.

Pairwise stability. A graph is pairwise stable if:

1. *No node wishes to delete a link he is involved in.*
2. *If there exists a node which wants to add a link, then the node at the other end of the link does not want to establish this link.*

Pairwise stable networks are robust to one-link deviations, where link severance is unilateral, while link creation is bilateral and under mutual consent of the two involved players [CAİ09].

4.4 Netlogo

In addition to analyzing the different models with game theory, we created a simulator for the models, in a program called Netlogo. Netlogo is a programmable modeling environment for simulating natural and social phenomena. It is well suited for modeling complex systems developing over time [Wil]. Netlogo is well suited to model our complex network formation games, and at the same time Netlogo provided us with a good graphical user interface that enabled us to see the result of the games, and also to easily adjust the different parameters. It was especially of use when facing models that were difficult to analyze, because it gave us a good graphical result, showing how the network evolved, and the final resulting network. In Figure 4.3 we see the user interface, which is used to set up the parameters, start the modeling, and showing the resulting network formation. Figure 4.4 shows how the coding interface looked like. For detailed overview of the code used in our different models, see appendix.

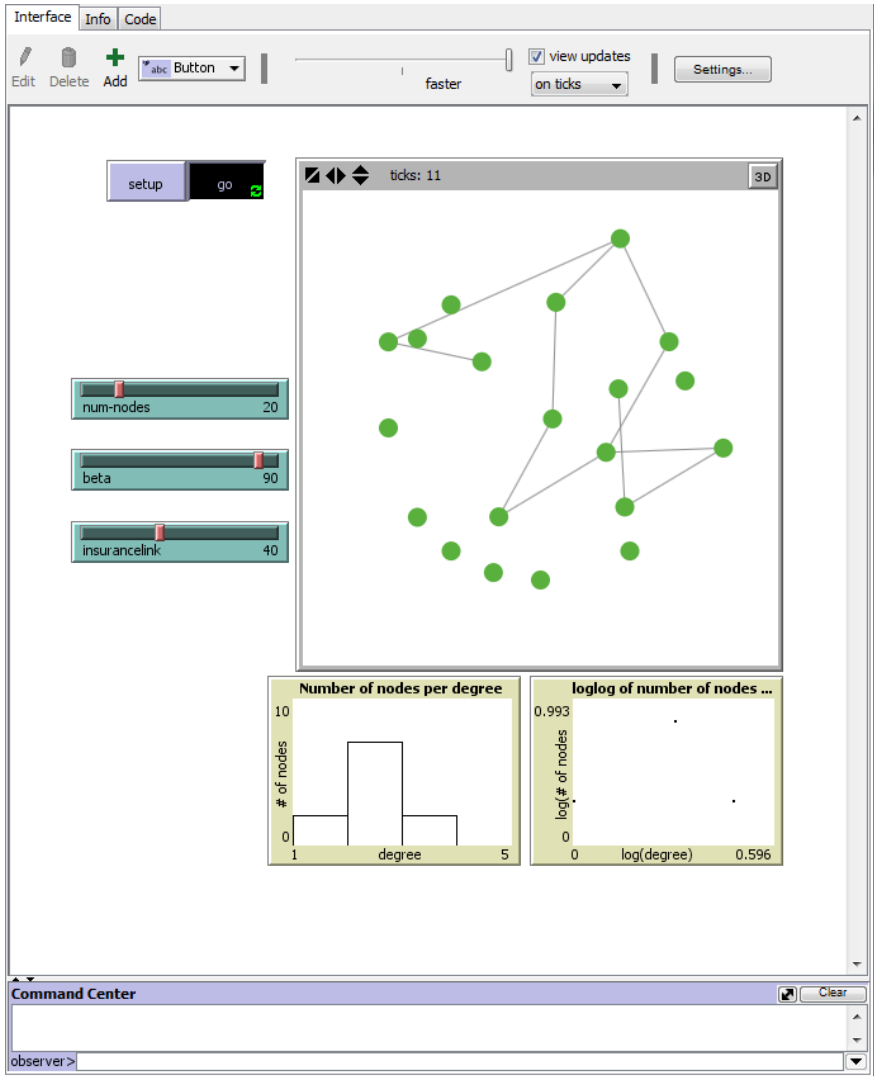


Figure 4.3: The figure shows a screen capture of Netlogo, while we are running one of our simulations.

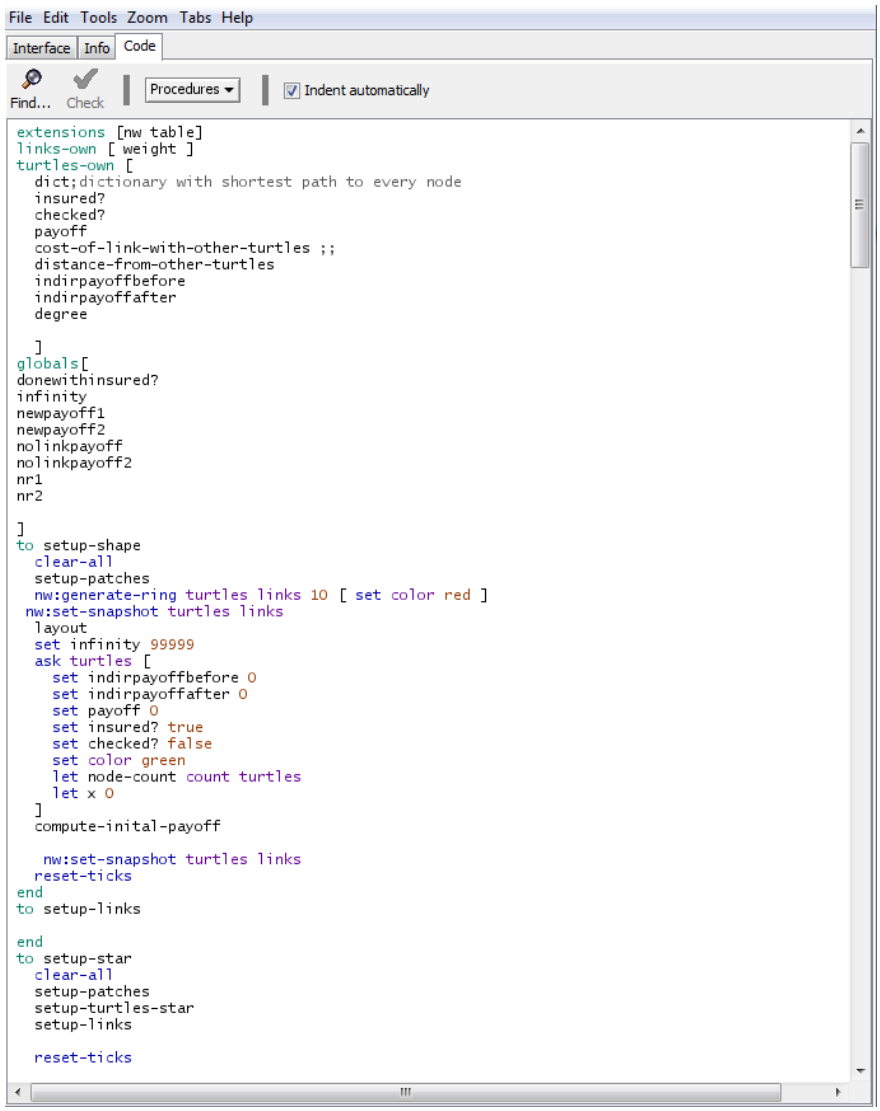


Figure 4.4: The figure shows how the code interface in netlogo looks like.

Chapter 5
Modeling Cyber-Insurance

There are many examples of nodes that need to establish connections with eachother. For example, when a firm is outsourcing tasks, cooperating or depending on other firms in some way. Such scenarios could be modeled, as networks where the nodes represents the firms and the links between them are their dependencies. However, the link between nodes involves some risks, such as: will the company deliver at the reported time, to the reported costs, what happens if it fails to deliver, what if the company goes bankrupt etc. To handle these risks, we need cyber-insurance.

When deciding whether or not to establish a link to a node, the payoff has to be higher in the balance between the expected earnings and the risk of the other party failing to complete the transaction. From the insurer's point of view, a problem with cyber-insurance is to define and calculate risk, because the network structure is undefined. If an insurer were able to predict the network structure, the calculations of overall risk would be realizable. The situation could be even better if the insurer could force a chosen robust network structure to evolve, hence, ensuring a higher total payoff for the network. Examples of such structures are scale-free networks, stars and cliques, as described in the graph theory chapter. In summary, scale-free networks have proven to be very robust against random attacks. Star topologies, or star-like topologies, have a fixation probability that exceeds the fixation probability of circulation graphs. Star structures also have a desirable property of minimizing the average path length, i.e. minimizing the cost spent on establishing links. Finally, the clique has a nice property of being able to achieve super-critical payoff, as showed in [Blu11]. In our thesis we want to focus on the clique and star/star-like structures for the following reasons: Both have been identified to have calculable fixation probability and the possibility of amplifying or suppressing selection and drift. This is favorable, because if the insurer is able to ensure that the nodes have a certain security level, particularly the center node, one can prevent viruses from spreading to other nodes.

Model overview. In our models a node is an agent who is either insured or not, and the nodes' goal is to maximize its payoff, by establishing links with other

nodes. Our goal for the models is to find out if and how an insurer can force these cyber-insurance networks to evolve into the desirable structures found in chapter 3.

In this chapter we will start out with a simple model(model 1: Initial Model) and stepwise add new features to make the models more realistic and applicable to real-world scenarios.

An overview of the models can be seen in Figure 5.1. The first model, albeit unrealistic, shows how the network formation ends up if nodes were able to determine whether other nodes are insured or not, and insured nodes only choose to connect to other insured nodes. Model 2 introduces parameters reflecting the actual cyber-insurance market. This makes the modeling process much more realistic, and we try to find the conditions where the desired network structures evolve. Model 2b is a tiny digression, where we analyze the consequences of a node having incomplete information about the other nodes' type. In model 3 we apply a bonus when a node reaches a desired number of connections. This illustrates a scenario where a node is dependent on other nodes' expertise in order to complete a task. Model 4 analyzes the outcome of adding a bulk-insurance discount, which is a normal phenomenon in the insurance industry, e.g. for each new insurance you purchase, you will receive a discount. In the final model (model 5: Network externalities and discount), we apply model 4 to an already existing model: "the symmetric connection game", in order to analyze the impact of network externalities.

For models 2, 3 and 5, we created a simulator to confirm the results of our calculations. The interested reader can find the source code in Appendix B.

5.1 Model 1: Initial Model

As a starting point, the model is highly simplified in order to show the concept of how cyber-insurance can be used to separate insured and non-insured nodes into two cliques. We assume that every node has complete network information, i.e. it knows how many nodes that exist, and whether or not they are insured. The link establishment process is bidirectional, meaning that both nodes must agree to establish the connection.

For the first model, we assume a set of n nodes that are randomly chosen to be insured or not, as depicted in Figure 5.2a. They all get their own fixed income, and by connecting to other nodes, they can increase their payoff. Non-insured nodes will have a risk of failure, which we model as an expected cost of failure. Therefore, if an insured node chooses to connect to non-insured nodes it will also suffer this expected cost of failure. To simplify the decision process, the model follows a rule that only allows insured nodes to connect to other insured nodes and non-insured nodes can

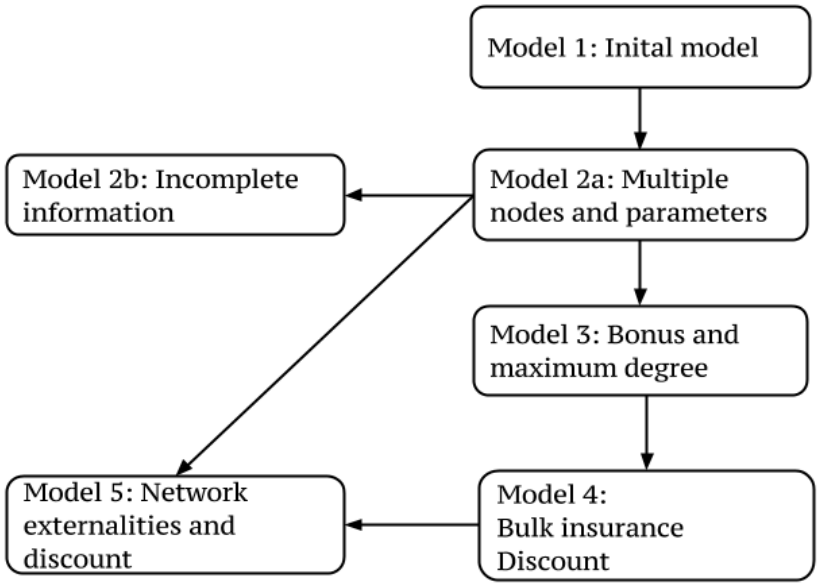
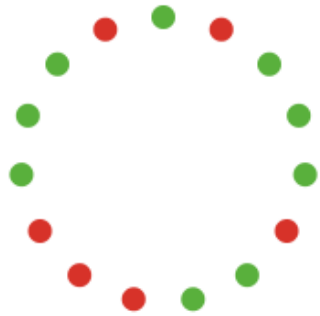


Figure 5.1: The figures show an overview of the different models we have created, and how they relate to each other. For every step, some new features are added to the model.

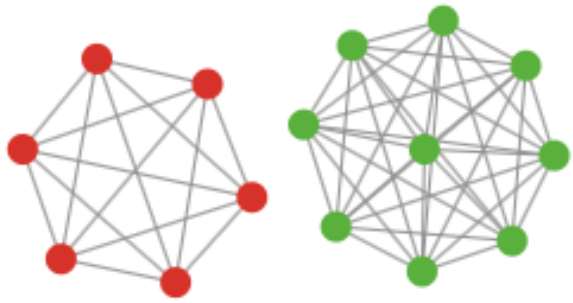
only connect with eachother. The resulting graph will be two fully connected cliques, one consisting of insured nodes and the other of non-insured nodes, as shown in Figure 5.2b.

This dichotomy represents a trusted environment for the insured nodes, because the insured nodes know that each node in the clique is also insured against risks. These nodes will benefit from each connection without having to worry about contagious risks from the connected nodes. A node in the non-insured clique will also experience a change in payoff from the links it has established, as each of the links has a probability of failure. Hence this environment is not trusted, and a link establishment will always involve some risk.

The first model, although very simple, shows a topology where insured agents benefit from being insured, and they are candidates to achieve super-critical payoffs as described in [Blu11].



(a) 15 Nodes randomly chosen to be either insured (green) or non-insured (red).



(b) Two clustered networks. One consisting of insured nodes, the other consisting of non-insured nodes.

Figure 5.2: Shows how nodes connect to eachother according to the model described in section 5.1.

5.2 Model 2: Including Parameters

The first model is highly simplified and suffers from many limitations, among other things, it is too simple to reflect the dynamics of a real world scenario, where nodes will have different variables with different values. To improve the model, we have to introduce parameters that can be adjusted and reflect real world scenarios. It is fair to assume that the insured nodes must pay an insurance premium, and this premium should be dependent on the number of links the node establishes. When two insured nodes establish a link between eachother, they both have to pay a premium, this is to make the game more fair, and more realistic. For example, if the two nodes had different insurance companies, then both companies would charge them for insuring the link. When a node, insured or not, establish link to a non-insured node, risk is present. This risk will be represented as an expected risk cost. Logically, if the

changes in payoff when establishing a link is negative, then no node would want to establish a link. Thus, nodes will also receive a positive change in payoff when establishing links to others. Among other things, this reflect the scenario of receiving benefits from being in a cooperation.

Characteristics of the model. The process of establishing links is a bidirectional decision. The different parameters are denoted as follows: The insurance premium is I_l, the expected risk cost is represented by r. β represents the benefit of establishing a link. Table 5.1 presents an overview of the parameters.

β - income from establishing a direct link
I_l - increased insurance cost per link the node establishes
r - expected risk cost

Table 5.1: Table showing the parameters used in models 1-4

5.2.1 Two nodes scenario

As a starting point for this model which includes parameters, we consider the simplest scenario involving only two nodes. In this game the strategy space of both players consists of four different strategies. A node chooses to be insured or not, and also chooses whether to establish a link to the other node or not. I.e. the different strategies are: Be insured and establish link noted as: IL, be insured and not establish link: $I\overline{L}$. Not insured and establish link: $\overline{I}L$, and not insured and not establish link: $\overline{IL}$. It should be noted that since the decision to establish a connection is bidirectional, both have to choose a strategy where they want to establish a link, for the link establishment to be successful. Hence we end up with the game as shown in Figure 5.3.

As long as both I_l and r are less than β, the only Nash equilibrium in Figure 5.3 is when both nodes choose $\overline{I}L$. If we first look at node A, we see that when node B chooses IL, the best response is $\overline{I}L$, because $\beta > \beta - I_l$. And since the game is symmetric, the same holds for node B. When one of the nodes chooses $\overline{I}L$, the best response will be $\overline{I}L$, because $\beta - r > \beta - I_l - r$, and thus the only Nash equilibrium is when both nodes play $\overline{I}L$.

This means that two nodes will end up in a classic prisoner's dilemma [1] , where

[1]The Prisoner's dilemma was originally framed by Merrill Flood and Melvin Dresher in 1950. The dilemma expresses a situation where two players each has two options whose output depends on the simultaneous choice made by the other. The original dilemma concerns two prisoners who separately decide whether or not to confess to a crime [Dic]. It is a paradox in decision analysis which shows why two individuals might not cooperate, even if it is in their best interest to do so.

Node A \ Node B	IL	$I\overline{L}$	$\overline{I}L$	$\overline{IL}$
IL	$\beta - Il$ $\beta - Il$	0 0	$\beta - Il - r$ β	0 0
$I\overline{L}$	0 0	0 0	0 0	0 0
$\overline{I}L$	β $\beta - Il - r$	0 0	$\beta - r$ $\beta - r$	0 0
$\overline{IL}$	0 0	0 0	0 0	0 0

Figure 5.3: Normal form game, showing the different strategies and the payoffs for the different outcomes. First the payoff of a is written, then the payoff of B.

the best response is actually worse than the socially optimal. In this case it is trivial to see that the socially optimal scenario for both nodes is to choose IL, as long as $I_l < r$. However, the nodes will choose not to buy insurance, since they could risk ending up in a situation where only one of them pays the cost of insurance, and it is better to be the one who does not pay.

Introducing time. One possibility for solving this problem where the two nodes end up choosing not to acquire insurance, is to introduce a leader/follower game. In this game the players do not act at the same time, but in a given order, and they can observe the other players' action. If we consider a game with only two players, player one first selects strategy, insure or not. Then after observing this action player two chooses if he would like to insure or not. Then they choose if they would establish a link or not, in the same order. In this type of game, the leader will benefit from a first mover advantage, because he can now force the game in the direction he prefers. This game can be solved by using backward induction on the extensive form, shown in Figure 5.4, and we find all the different subgame equilibria. We assume that Eq. (5.1) holds. As we can see from Figure 5.4, when we have worked our way back to player 1's first decision, insure or not, we have the following subgame equilibria:$(L_1, \overline{L_1^I}, \overline{L_1^{II}}, L_1^{III}), (I_2, \overline{I_2^I}, L_2^I, \overline{L_2^{II}}, L_2^{III})$. This means that player 1 knows what will happen if he chooses to acquire insurance or not, if he acquires, he will get payoff: $\beta - I_l$, if he does not acquire, he gets: $\beta - r$. This means that as long as $I_l < r$, he will chose to acquire insurance, and by doing so forces the game to end up in an equilibrium where both players acquire insurance and would like to establish a link.

$$I_l < \beta \text{ and } I_l > \beta - r \text{ and } r < \beta \tag{5.1}$$

From this, we see that if the insurance price were set to the right amount, the first player would choose to purchase insurance and will force the outcome of the game to be the socially optimal outcome. The problem with this way of solving the game is that it is very hard to solve for multiple nodes, because the extensive form game becomes extremely complicated. Additionally we have required that the nodes act in a given order, which is not a realistic scenario. Hence, we chose not to use a leader follower game in the other models.

5.2.2 Model 2a: Multiple nodes

Assumptions

To make the modeling possible, we will from now on assume that the type of the nodes is given in advance, i.e. they are chosen to be insured or non-insured with a probability. The reason for this is that, we are focusing on endogenous network

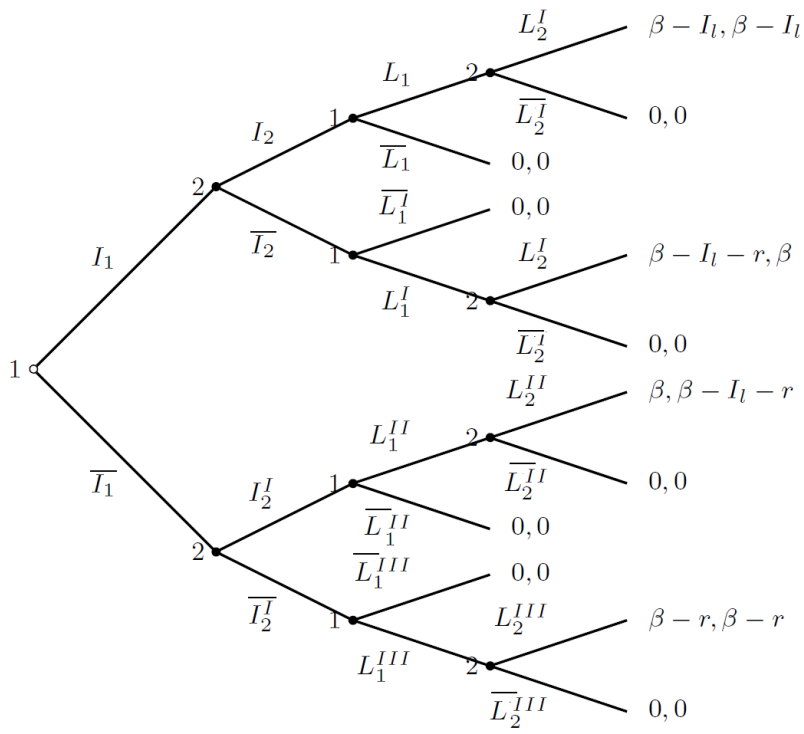


Figure 5.4: Leader follower game, first player 1 chooses to insure or not, then player 2, and then they choose to establish link or not in the same order.

formations, and including this decision process would only drastically increase the complexity of the models.

The model is now made more realistic by introducing multiple nodes. The objective of this model is to find characteristic network formations that will evolve endogenously when the parameters are within certain conditions. For this model we assume that every node has complete information of the network, i.e. every node knows the type of the other players. This is a very strong assumption, but in financial transactions and in cooperative software development networks, it is reasonable to assume that the parties can acquire this type of information prior to establishing a financial contract with each other. As we will show in Section 5.3, when we introduce incomplete information between the nodes, it will not be possible to separate the two types, since they will have to act on beliefs.

Analysis

As mentioned our goal is to find how and when certain network formations evolve. We know that if a node can increase its payoff by establishing a link, it will do so. Thus we can start analyzing the four possible link establishment scenarios, insured to insured, insured to non-insured, non-insured to insured, and non-insured to non-insured. Let U_i denote the payoff of a node with degree i, and let U_{i+1} be the payoff a node will receive if it establishes a new link.

Insured to insured. When two insured nodes are considering establishing a link, they will do so, if and only if both receive a higher payoff. In this scenario the payoff function of adding a link is as shown in Eq.(5.2).

$$U_{i+1} = \begin{cases} \beta - I_l, & \text{if } i = 0 \\ U_i + \beta - I_l, & \text{if } i > 0 \end{cases} \tag{5.2}$$

For a link to be established between two insured nodes, Eq.(5.3) has to hold.

$$I_l < \beta \tag{5.3}$$

Non-insured to insured. The payoff a non-insured node receives by connecting to an insured one is as described in Eq.(5.4). As we see this will always be a positive change in payoff, and thus a non-insured node will always choose to connect to an insured node.

$$U_{i+1} = \begin{cases} \beta, & \text{if } i = 0 \\ U_i + \beta, & \text{if } i > 0 \end{cases} \tag{5.4}$$

Insured to non-insured. The payoff an insured node receives in this scenario is as follows:

$$U_{i+1} = \begin{cases} \beta - I_l - r, & \text{if } i = 0 \\ U_i + \beta - I_l - r, & \text{if } i > 0 \end{cases} \tag{5.5}$$

For this to happen Eq.(5.6) has to hold, since a non-insured node always wants to connect to an insured one, this is the only condition that is needed for this scenario to happen.

$$I_l + r < \beta \tag{5.6}$$

Non-insured to Non-insured. The payoff a non-insured nodes receives when connecting to another non-insured node is as follows:

$$U_{i+1} = \begin{cases} \beta - r, & \text{if } i = 0 \\ U_i + \beta - r, & \text{if } i > 0 \end{cases} \tag{5.7}$$

For this link-establishment scenario to happen Eq.(5.8) has to hold.

$$\beta > r \tag{5.8}$$

Forming a trusted clique. We want to find the conditions for when different network structures will evolve. One desired structure would be to form a trusted clique of only insured nodes, in order to reduce risk and for the nodes to receive super-critical payoff. For this to happen, all insured nodes must connect to each other, i.e. Eq.(5.3) has to hold. Additionally, we need to ensure that insured nodes do not establish links to non-insured nodes. i.e. the following has to hold:

$$I_l + r > \beta \tag{5.9}$$

This gives us the limitation shown in Eq.(5.10) on the insurance link cost.

$$\beta - r < I_l < \beta \tag{5.10}$$

This condition means that if the link insurance cost is between the two boundaries, all the insured nodes will connect with each other, and no other nodes. If the link insurance cost is greater than β, then no insured node will establish any links. And if it is below $\beta - r$, then the insured nodes will also connect to the non-insured ones. It should also be noticed that as long as $r < \beta$, then the non-insured nodes will connect to eachother.

5.2.3 Result and findings

From the analysis we found different conditions for the link establishment process. If Eq.(5.10) is fulfilled, then the network will end up with one clique of only insured nodes. The non-insured nodes will end up in another clique if the risk of connecting to another non-insured node is less than the benefit of establishing the link ($r < \beta$). If the link insurance cost and risk of connecting to non-insured nodes is less than the benefit ($I_l + r < \beta$), then insured nodes will also connect to non-insured nodes. Hence the network will end up in one giant clique.

These findings are independent of the number of players, because we only consider one link at a time, and the change in payoffs is linear and independent of the nodes degree.

Stability and efficiency. When measuring stability in this model, it is easily seen that since the change in payoff when adding links is linear, and non-dependent on the nodes degree, the resulting network will be pairwise stable. It also follows from the definition of a Nash equilibrium, that the resulting network is an equilibrium, since every player has best responded to the other players' best responses, and no

node can increase its payoff by single-handedly changing a strategy. To calculate the efficiency we need to sum up the overall payoff, and compare it to the maximum possible payoff, i.e. we want to find the price of anarchy. The total payoff can be calculated as in Eq.(5.11), where $\sum I \times I$ represents the sum of payoffs achieved from links between insured nodes. $\sum I \times \overline{I}$ is the sum of payoffs achieved from links between non-insured and insured, and $\sum \overline{I} \times \overline{I}$ is the sum of payoffs achieved from links between non-insured and non-insured nodes.

$$U_{total} = \sum I \times I + \sum \overline{I} \times \overline{I} + \sum I \times \overline{I} \tag{5.11}$$

When the parameters are inserted in Eq.(5.11), we get Eq.(5.12), where N_I and $N_{\overline{I}}$, represent the number of insured and non-insured nodes in the network.

$$U_{total} = N_I(N_I - 1)(\beta - I_l) + N_{\overline{I}}(N_{\overline{I}} - 1)(\beta - r) + N_I N_{\overline{I}}(2\beta - r - I_l) \tag{5.12}$$

If we calculate the overall payoff for a network with one clique of insured and another with non-insured, i.e. Eq. (5.10) and $r < \beta$ hold. The total payoff is as shown in Eq.(5.13).

$$U_{total} = N_I(N_I - 1)(\beta - I_l) + N_{\overline{I}}(N_{\overline{I}} - 1)(\beta - r) \tag{5.13}$$

However, this is not the socially best outcome, because there are no links between insured and non-insured nodes, which would have contributed with $2\beta - r - I_l$ for every link, and since $2\beta > r + I_l$ will be true, as long as both the insurance cost and the expected risk cost is less than β. Thus, the socially best outcome would have been one clique, with both insured and non-insured nodes. The formula for calculating the price of anarchy is shown in Eq.(5.14).

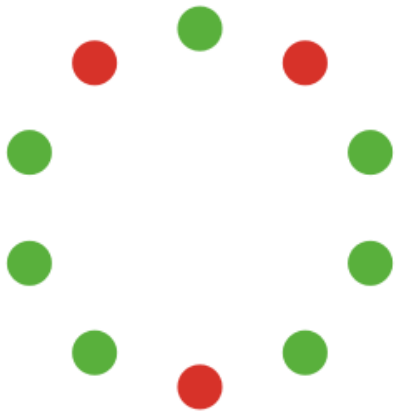
$$PoS = \frac{N_I(N_I - 1)(\beta - I_l) + N_{\overline{I}}(N_{\overline{I}} - 1)(\beta - r)}{N_I(N_I - 1)(\beta - I_l) + N_{\overline{I}}(N_{\overline{I}} - 1)(\beta - r) + N_I N_{\overline{I}}(2\beta - r - I_l)} \tag{5.14}$$

An interesting thing to notice is that the only scenario where the insurer is able to separate the two types of nodes, and at the same time ensuring an efficient and stable outcome, is when there are only links between insured nodes, or between non-insured nodes, or no links at all. This can only happen when $2\beta < I_l + r$, and $I_l > \beta + \beta - r$ or $r > \beta + \beta - I_l$ or if both I_l and r is larger than β.

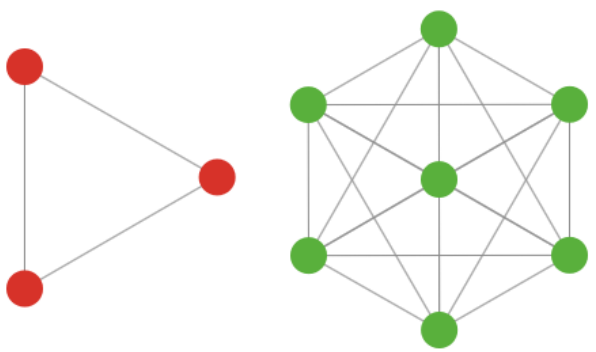
Simulation of the results

To verify that our calculations of the network formation were consistent with the assumptions, we performed different simulations using NetLogo. In the simulator a node is insured with a probability p. The network formation is performed by selecting two random nodes, not neighbors, then both nodes check whether they would prefer to establish a link or not. The rules are as described earlier; when a node is considering establishing a link it chooses to do so if the payoff received is larger than the payoff it already has achieved, and the decision is bilateral. This selection

is repeated until the network is fully connected or no more nodes are willing to establish new links. By selecting nodes at random and checking if both of them would like to connect to each other, we relax the assumption of full network information, because now nodes only get to know if another node is insured or not, when they ask eachother.



(a) Ten nodes randomly, with probability 0.5, chosen to be either insured (green) or non-insured (red).



(b) Two clustered fully connected networks. One consisting of insured nodes the other consisting of non-insured nodes. The link establishment is done by following the rules described above

Figure 5.5: The figure shows the resulting network from a simulation with parameters: $\beta = 0.9$, $I_l = r = 0.5$.

In Figure 5.5 we see the result of a simulation with the parameters: $\beta = 0.9$,

$I_l = r = 0.5$. With these parameters Eq.(5.10) holds, and $r < \beta$. Thus the network formation game ends up in two cliques, one with insured nodes and another with non-insured nodes. The result is shown in Figure 5.5b, and confirms our calculations. The price of anarchy in this scenario is: $PoA = \frac{8}{15}$. In this figure only $n = 10$ nodes are included, this is done to make the figure readable and easy to understand. Similar results were obtained when performing the simulation with larger values of n, but the resulting printouts included to many nodes and links to be readable.

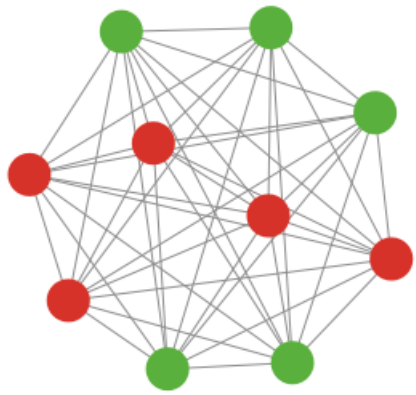
In the next simulations, the parameters were chosen to violate Eq.(5.10). The result can be seen in Figure 5.6. In figure 5.6a we see that when $I_l < \beta - r$, the result is one clique of both insured and non-insured nodes, and the price of anarchy is 1, i.e. this is the socially optimal outcome. In figure 5.6b the insurance cost is $I_l > \beta$, and as we see only non-insured nodes connect to eachother, because the insurance cost per link is higher than the benefit given from connecting to a new node, i.e. the insured ones choose not to establish any connections. The price of anarchy in this scenario is: $PoA = \frac{32}{35}$, and is thus close to the socially optimal outcome.

5.3 Model 2b: Model with incomplete information

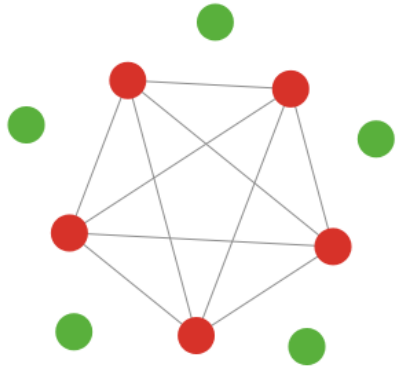
Although we previously mentioned that we chose to assume complete information about other nodes' type. We wanted to get an impression of the complexity when modeling a scenario where some nodes' lack information about the other nodes type. The way we model this is by letting nature select whether a player is insured or not, a node is insured with probability p, and not insured with probability $1 - p$. All nodes know their own type, but in the link establishment process only one node knows the type of the other. The other node only knows the probability of the other node being insured or not. We want to see if it is possible for the nodes with incomplete information to distinguish an insured node from a non-insured one. For this model we will only present the results from the analysis, because the mathematics and analysis is too complex and non-intuitive to include here. The actual analysis and the mathematics of this game can be seen in appendix A.1.

Result and findings. When one player lacks knowledge about the other player, we were only able to find two scenarios where the player with less information could separate the two types of the other node. The first scenario occur when player 2 is insured and $\beta < I_l$, then it is only the non-insured node who wishes to establish a link, and in this way player 2 is able to separate the two types of player 1. However, since the benefit is less than the cost($\beta < I_l$), his best response is to not establish any link.

The other scenario where the node with incomplete information is able to separate the two types of player 1, is when he is not insured and the following is true:



(a) Ten nodes insured with probability 0.5, the parameters where: $\beta = 0.9$, $I_l = 0.3$ and $r = 0.5$, i.e. the link insurance cost, I_l, is violating the condition in Eq.(5.10), and the resulting network is one clique of both insured and non-insured nodes.



(b) Ten nodes insured with probability 0.5, with parameters as before, except for the link insurance cost: $I_l = 0.95$. This resulted in a clique of only non-insured nodes.

Figure 5.6: The figure shows the two possible scenarios that violate Eq.(5.10), 5.6a shows the result when $I_l < \beta - r$ and 5.6b shows the result when $I_l > \beta$.

$r < \beta < I_l + r$. In this scenario it is only the non-insured node who would want to establish a link. Thus, in this scenario the game will end up with a link between two non-insured nodes.

We were also able to find some pooling equilibriums. If the node with incomplete information is insured, a link will be established if the following is true: $\beta + rp - r > I_l$. However, if $I_l < \beta$ but $I_l > \beta + rp - r$, then the pooling equilibrium will be: node 1

wants to establish link, but node 2 rejects. There is also a pooling equilibrium where both nodes want to establish a link, this occurs when node 2 is not insured and the following is true: $\beta - I_l > r$. If the benefit from establishing a link is less than the expected risk cost($\beta < r$), there will be a pooling equilibrium where both players choose not to establish links.

What this shows us is that when one player has incomplete information, it is no longer possible for the insurer to force a network to evolve into a state where a clique of only insured nodes exists. Furthermore, the incentive for establishing links decreases, since the player with less information must act on beliefs. We chose not to simulate this model, since the result would only be a clique of non-insured nodes or a giant clique consisting of both insured and non-insured nodes. In all the other models complete network information is assumed.

5.4 Model 3: Including maximum node degree and bonus

In real world networks, such as in the manufacturing industry, software development firms and many other types of business, in some scenarios a product can usually not be completed without outsourcing some of the work needed. For the manufacturer, it could be beneficial to buy certain parts from others instead of producing them on his own. A software product might need the combined knowledge from different firms. Thus the firm that outsources tasks is dependent on the other firms, and will not reach its goal before the other firms deliver their contribution. When the product is finished the company gets paid. To model this scenario we introduce a maximum node degree, m, per node, which represents the number of partners needed to complete a task. Additionally a bonus γ represents the payoff a node receives when m links are established. Except from this, the game is unchanged.

5.4.1 Analysis

This model is very similar to the earlier models: for nodes to connect to each other, the change in payoff still has to be positive: $U_{i+1} > U_i$. However, we also need to consider the bonus received when reaching the maximum node degree, m. To model this, we add the possible bonus divided on the number of links required to reach the bonus($\frac{\gamma}{m-i}$) in the decision process every time a node is considering to establish a link. In this way the model will change from the earlier models, because now the nodes have more incentive to connect to other nodes, and for every step closer to the goal, the nodes are more willing to accept risk than before. For example, an insured node is more likely to accept a risky link when it only needs one more link to reach the goal, compared to when it need several more links to reach the goal.

The model now introduces a risk factor, because it is not certain that the nodes will obtain enough links, and if not, they will not receive their bonus, and they are stuck with their already established connections.

To analyze this model, let us take a closer look at the four different scenarios of the game. When establishing a link between two insured nodes, the payoff the nodes will receive is as described in Eq.(5.15).

$$U_{i+1} = \begin{cases} \beta - I_l, & \text{if } i = 0 \\ U_i + \beta - I_l, & \text{if } i > 0 \\ U_i + \beta - I_l + \gamma, & \text{if } i = m - 1 \end{cases} \tag{5.15}$$

As described earlier we need to include the possibility of reaching the goal in the decision, and thus for insured nodes to connect to each other, Eq.(5.16) has to hold.

$$\begin{aligned} U_i + \beta - I_l + \frac{\gamma}{m-i} &> U_i \\ \beta - I_l + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> I_l \end{aligned} \tag{5.16}$$

The payoff an insured node receives when connecting to a non-insured node is as follows:

$$U_{i+1} = \begin{cases} \beta - I_l - r, & \text{if } i = 0 \\ U_i + \beta - I_l - r, & \text{if } i > 0 \\ U_i + \beta - I_l - r + \gamma, & \text{if } i = m - 1 \end{cases} \tag{5.17}$$

To establish a connection from an insured node to a non-insured one, the following has to hold:

$$\begin{aligned} U_i + \beta - I_l - r + \frac{\gamma}{m-i} &> U_i \\ \beta - I_l - r + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} - r &> I_l \end{aligned} \tag{5.18}$$

When a non-insured node connects to another non-insured node, this is the payoff they both will receive:

$$U_{i+1} = \begin{cases} \beta - r, & \text{if } i = 0 \\ U_i + \beta - r, & \text{if } i > 0 \\ U_i + \beta - r + \gamma, & \text{if } i = m - 1 \end{cases} \tag{5.19}$$

To establish the connection, the following equation has to hold:

$$\begin{aligned} U_i + \beta - r + \frac{\gamma}{m-i} &> U_i \\ \beta - r + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> r \end{aligned} \tag{5.20}$$

In the case of a non-insured node wanting to establish a link with an insured node, the payoff is a strictly increasing function, see Eq.(5.21), and thus a non-insured node will always connect to an insured node if possible.

$$U_{i+1} = \begin{cases} \beta, & \text{if } i = 0 \\ U_i + \beta, & \text{if } i > 0 \\ U_i + \beta + \gamma, & \text{if } i = m-1 \end{cases} \tag{5.21}$$

5.4.2 Result and findings

If we want a clique of only insured nodes, we have to ensure that insured nodes connect to each other, and that they do not establish connections to non-insured nodes. We know that an insured node would want to connect to another insured node if Eq.(5.16) is satisfied. In the equation we see that the expected bonus per established link is increasing. Thus, if an insured node of degree zero is willing to connect to another insured node, then every insured node with a degree higher than zero would also like to connect to another insured node. To ensure that insured nodes connect to eachother this equation has to hold:

$$\beta + \frac{\gamma}{m} > I_l \tag{5.22}$$

We also want to ensure that insured nodes never establish links with non-insured nodes, from Eq.5.17 we see that this has to hold:

$$\beta + \frac{\gamma}{m-i} - r < I_l \tag{5.23}$$

This can be simplified, if one can ensure that the most risk willing insured node, i.e. the node with degree $m-1$, does not establish links with non-insured nodes. Then we know that no insured node with degree less than $m-1$ will establish links with non-insured nodes. From this we get equation Eq.(5.24).

$$\begin{aligned} \beta + \frac{\gamma}{m-(m-1)} - r &< I_l \\ \rightarrow \quad \beta + \gamma - r &< I_l \end{aligned} \tag{5.24}$$

To summarize, Eq.(5.22) and Eq.(5.24) give the final limitation on the link insurance cost, Eq.(5.25). If this equation is satisfied, the resulting network will contain a clique of only insured nodes.

$$\beta + \gamma - r < I_l < \beta + \frac{\gamma}{m} \tag{5.25}$$

For this to even be possible, $\beta + \gamma - r < \beta + \frac{\gamma}{m}$, i.e. Eq.(5.26) has to hold. This equation reflects that as the risk to bonus ratio gets smaller, it gets more and more difficult to ensure a clique of only insured nodes. When the risk to bonus ratio is less than $1 - \frac{1}{m}$, a clique will never occur. The equation shows that a node would be more and more willing to take a risk, as the reward of doing so increases.

$$\begin{aligned} \gamma - r &< \tfrac{\gamma}{m} \\ 1 - \frac{r}{\gamma} &< \tfrac{1}{m} \\ \rightarrow \quad 1 - \frac{1}{m} &< \tfrac{r}{\gamma} \end{aligned} \tag{5.26}$$

It is also useful to know when non-insured nodes connect to each other. This happens when Eq.(5.19) is satisfied. This equation is dependent on the node degree, and thus for the first link to be established from a non-insured node, the expected payoff has to be higher than the risk($\beta + \frac{\gamma}{m} > r$). If the risk is too high, then the non-insured node must establish links with insured nodes before it coulb be willing to establish risky links.

With these findings, an insurer can determine the outcome of the network formation game by adjusting the insurance cost parameter. If he wants a clique of only insured nodes Eq.(5.25) has to hold. However, it is easy to relax the condition, so that insured nodes only connect to, $j = 1, 2, 3..m$ non-insured nodes. This is done by changing Eq.(5.24) to $\beta + \frac{\gamma}{m-(m-j)} - r < I_l$, which gives us Eq.(5.27). An interesting result in this model is that due to the risk willingness among the nodes, the lower boundary, to ensure separation of insured and non-insured nodes, of the link insurance cost is higher compared to the one found in model 2.

Consequences of not reaching the required number of edges. When a node establishes a link, it does not know whether it will reach the maximum node degree, unless the current node degree is $m - 1$. Hence the node might end up not reaching the desired goal. This can happen if there is not enough nodes willing to establish links. Consequently, nodes who do not reach their goal could end up with a payoff less than U_0.

$$\beta + \frac{\gamma}{j} - r < I_l \tag{5.27}$$

Efficiency and Stability. In this model, the incentive for establishing links is increased compared to model 2. Thus, to maintain a stable network with two cliques the cost of link establishment has to be increased. This increased incentive may result in a higher price of stability, but if every node has received its bonus, then the price of anarchy is 1. The price of anarchy is dependent on the number of nodes in both cliques, and whether its enough nodes for everyone to reach their maximum degree or not. If there are nodes that have not reached their maximum degree in both cliques, then the resulting network is not necessarily the most efficient, and we could be missing a potential payoff due to the cost constraint.

By introducing the maximum degree m we are limiting the problem of price of anarchy, because as long as m is less than the number of insured and number of non-insured nodes, there will be less links established compared to model 2, and overall fewer possible links between insured and non-insured. However, the bonus the nodes receive will contribute to inefficiency, because when nodes do not reach their maximum degree, the potential payoff that could be generated by allowing insured and non-insured nodes to connect, is greater than in model 2.

Simulation of the results

For the first simulation the parameters are set to the following: $\beta = 0.9, I_l = 0.7, r = 0.5, \gamma = 0.2$ and $m = 4$, in order to satisfy condition Eq.(5.25), and enable all nodes to reach their maximum degree.

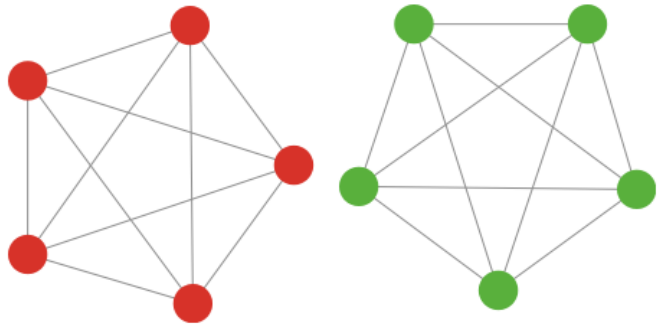


Figure 5.7: Two cliques, one consisting of insured agents the other consists of non-insured. All nodes have reached their goal.

As we see in Figure 5.7, the results were as expected, the cost of insuring a link satisfied the conditions found earlier, and thus the result was two cliques, one consisting of only insured nodes and the other of non-insured nodes. An interesting observation is that β and r is the same as in model 2, but to ensure that only insured nodes connect to eachother, the link insurance cost had to be higher. This is to

compensate for the risk the nodes now are willing to take. The price of anarchy in this scenario is 1, i.e. the socially optimal outcome.

In the second simulation we set the parameter $m = 5$, and kept the other variables unchanged. The resulting network was as expected the same as in the last simulation, but since the nodes did not reach their maximum degree, the price of anarchy is less than one. The price of anarchy can be seen in Eq.(5.28).

$$\begin{aligned} PoA &= \frac{\text{Sum of payoffs}}{\text{Sum of Socially optimal payoffs}} \\ PoA &= \frac{5\times4\times(0.9-0.7)+5\times4\times(0.9-0.5)}{5\times4\times(0.9-0.7)+5\times4\times(0.9-0.5)+5\times(2\times0.9-0.7-0.5+2\times0.2)} \\ PoA &= \frac{12}{17} \end{aligned} \tag{5.28}$$

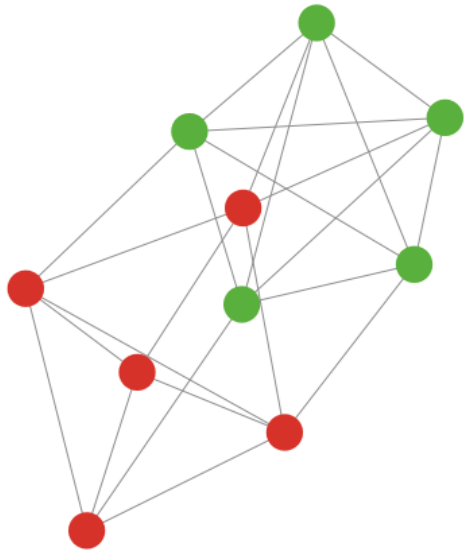
When we changed the link insurance cost, and set it to the same value as in model 2, $I_l = 0.5$, the resulting networks change. Now we found that the insured nodes are willing to establish risky links to reach their maximum degree. Some of the resulting networks can be seen in Figure 5.8. In figure 5.8a the price of anarchy is 0.95, and in figure 5.8b the price of anarchy is 1, i.e. it has reached the socially optimal outcome.

5.5 Model 4: Including bulk insurance discount

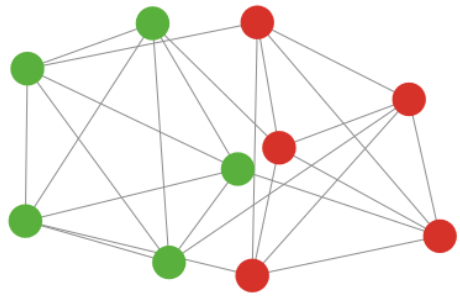
Insurance companies often give a quantum discount when a customer purchases multiple products. From convenience stores, we are used to the slogan "buy one get one for free", and insurers tend to follow the same marketing strategy. It seems to be common for insurance companies to offer discount to their customers if the customers choose to combine some or all of their insurances with them. Several insurance companies in Norway, e.g. Sparebank 1 offers customers up to 25 % discount according to the following rules [Spa].

- 10% discount if the person has signed three different insurances
- 15% discount if the person has signed four different insurances
- 20% discount if the person has signed five or more different insurances
- Plus additional 5% discount if the person is a customer of the bank.

The insurance offered is intended for the individual market and includes among other things: travel insurance, household insurance, car insurance, house insurance, insurance of valuable items and yacht insurance.



(a) One non-insured node has connected to two insured nodes.



(b) Every non-insured node is connected to one insured node, this is the optimal outcome with these parameters.

Figure 5.8: Two possible outcomes when insured nodes are willing to take the risk of connecting to non-insured nodes, to receive their bonus. Figure *a* shows a scenario where one non-insured node has connected to more than one insured node, thus not a socially optimal outcome. Figure *b* shows the optimal outcome.

Inspired by these kinds of discounts on insurance products, we would like to introduce a discount rate dependent on the degree of the node. In a real-world scenario where nodes have an option of acquiring insurance or not, this will make it more attractive for nodes with high degree to acquire insurance, and the discount could act as an incentive for other nodes to also acquire insurance. Therefore, this seems like a reasonable model to include. Since, if you have many links, you will pay

less per link compared to a node having fewer links.

How insurance companies choose to formulate their discount rate might vary. One solution might be to follow a strict 5% discount per new connection, similar to the one from Sparebank 1, or let the discount follow a power law, or a log-function etc. We choose to follow a discount rule which directly reflects the node's degree.

5.5.1 Analysis

The price for adding a new link follows the equation:

$$\frac{I_l}{i+1} \tag{5.29}$$

Here, i is the node's current degree. This means that the more links a node establishes the cheaper the link insurance will be.

Discount model

We start our analysis by applying the discount to model 2. As before we analyze the four different connection scenarios. However, because non-insured nodes do not pay any insurance, it is only the scenario where insured nodes connects to other insured nodes and insured nodes connect to non-insured nodes, that has changed compared to model 2.

When we consider links between insured nodes, we must add the discount to the conditions found in model 2. The condition for establishing links between two insured nodes is shown in Eq.(5.30).

$$\frac{I_l}{i+1} < \beta \tag{5.30}$$

For a link between insured and non-insured nodes to be established, Eq.(5.31) has to hold.

$$\frac{I_l}{i+1} + r < \beta \tag{5.31}$$

Result and findings For an insurer to be able to guarantee that the network ends up in a clique with only insured nodes, we must ensure that the most expensive link establishment, i.e. the first, to another insured node can be achieved. This gives us the same condition as in model 2, i.e. $I_l < \beta$. We also need to ensure that insured nodes do not connect to non-insured nodes, thus we get the final condition in Eq. (5.32), where N_I is the number of insured nodes in the network.

$$(N_I)(\beta - r) < I_l < \beta \tag{5.32}$$

This condition is very strong, because for it to be possible the following has to hold:$\beta - r < \frac{\beta}{N_I}$, and as the number of insured nodes gets higher this gets more and more unlikely. Thus by including bulk discount, the insurer is making it harder for himself to constrain the network formation. The reason for this is that the incentive for establishing links is higher than without discount, and therefore more links will be established.

Stability and efficiency If we compare the total payoff equation in this model, see Eq.(5.33), with the one in model 2 (Eq.(5.12)), we see that the cost for insured nodes has changed, and therefore the payoff achieved from links between insured nodes has increased, and so has the payoff received from potential links between insured and non-insured nodes. As we know, in a scenario where the insurer sets the cost, in a manner that makes the network end up in two cliques, the payoff received from links between insured and non-insured is zero. The potential payoff in a scenario where there are two cliques can be described like this: $(N_I N_{\overline{I}}\beta + N_I(-\sum_{i=N_I}^{N_{\overline{I}}-1} \frac{I_l}{i}))$, and as long as $(N_I N_{\overline{I}}\beta > N_I(-\sum_{i=N_I}^{N_{\overline{I}}-1} \frac{I_l}{i}))$ it would have been socially optimal to have a single clique of both insured and non-insured nodes. When the cost of establishing links decreases and the insurer forces the network formation to end up in two cliques, the price of anarchy will be higher compared to the price of anarchy in model 2. The reason for this is that the incentive for establishing links has increased, and thus for the insurer to be able to constrain the network formation, the cost of establishing links has to be higher.

$$U_{total} = (N_I(N_I-1)\beta - N_I \sum_{i=1}^{N_I-1} \frac{I_l}{i}) + (N_{\overline{I}}(N_{\overline{I}}-1)(\beta - r)) + (N_I N_{\overline{I}}\beta + N_I(-\sum_{i=N_I}^{N_{\overline{I}}-1} \frac{I_l}{i})) \tag{5.33}$$

Discount and Bonus model

We also need to apply the discount to model 3. Still, the only scenarios that has changed in this model is the ones where insured nodes connects to other insured nodes or when insured nodes connects to non-insured nodes.

When insured nodes are considering to establish links with eachother, their payoff functions are as shown in Eq.(5.34).

$$U_{i+1} = \begin{cases} \beta - I_l, & \text{if } i = 0 \\ U_i + \beta - \frac{I_l}{i+1}, & \text{if } i > 0 \\ U_i + \beta - \frac{I_l}{i+1} + \gamma, & \text{if } i = m-1 \end{cases} \tag{5.34}$$

For insured nodes to connect to eachother Eq.(5.35) has to hold.

$$\begin{aligned} U_i + \beta - \frac{I_l}{i+1} + \frac{\gamma}{m-i} &> U_i \\ \beta - \frac{I_l}{i+1} + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> \tfrac{I_l}{i+1} \end{aligned} \tag{5.35}$$

When insured nodes are considering to connect to non-insured nodes, their payoff functions are as shown in Eq. (5.36).

$$U_{i+1} = \begin{cases} \beta - I_l - r, & \text{if } i = 0 \\ U_i + \beta - \frac{I_l}{i+1} - r, & \text{if } i > 0 \\ U_i + \beta - \frac{I_l}{i+1} - r + \gamma, & \text{if } i = m-1 \end{cases} \tag{5.36}$$

For this to happen Eq.(5.37) has to hold.

$$\begin{aligned} U_i + \beta - \frac{I_l}{i+1} + \frac{\gamma}{m-i} - r &> U_i \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> r + \tfrac{I_l}{i+1} \end{aligned} \tag{5.37}$$

5.5.2 Result and findings

We analyze the same scenario as in the other models, namely a clique of only insured nodes. The first step is to guarantee that insured nodes connect to eachother. To ensure that this happens, we need to find the condition for the lowest expected increase in payoff, i.e. at node degree zero. If nodes are willing to establish links at this point, then they will also be willing at all degrees higher than zero. At degree zero there is no discount on the insurance link cost, and thus if Eq.(5.22) from model 3 holds, insured nodes will connect to other insured nodes.

The condition for guaranteeing that insured nodes do not connect to non-insured nodes has changed, we know that if an insured node does not want to establish a link with a non-insured node at degree $m-1$, then neither will any insured node with degree lower than $m-1$ do so. From this we find the condition in Eq.(5.38)

$$\begin{aligned} U_i + \beta - \frac{I_l}{m} + \frac{\gamma}{m-(m-1)} - r &< U_i \\ \beta + \gamma - r &< \tfrac{I_l}{m} \\ \rightarrow \quad m(\beta + \gamma - r) &< I_l \end{aligned}$$

This is a very strong condition, because the only way this can happen is if $\beta+\gamma-r < \frac{1}{m}$. This shows us that when the incentives for establishing links increase, it gets more and more difficult for the insurer to guarantee a clique of only insured nodes. The final condition for ensuring a clique of only insured nodes is shown in Eq.(5.38).

$$m(\beta + \gamma - r) < I_l < \beta + \frac{\gamma}{m} \tag{5.38}$$

The quantum discount results in an overall higher payoff for the insured nodes, since the cost of insuring a new link becomes cheaper. This means that the insured nodes will have a higher incentive to create links, making it harder for the insurer to separate insured and non-insured nodes.

We see that the problem of separating the two node types have increased compared to model 3, meaning that if we have a network where the insurer has managed to separate them, the price of anarchy is also higher compared to a similar scenario in model 3.

5.6 Model 5: Network externalities

In the earlier models, the experienced network effects arose only from a node's neighbours. I.e. when a node established a connection the change in utility were only dependent on fixed variables, and not dependent on the rest of the network. In many real world scenarios it is more realistic that a node will be strongly affected by the indirect connections to other nodes. Social relationships between nodes are good examples of such networks, where each person offer benefits in terms of favors, information etc.

We apply the results from the paper from Jackson and Wolinsky [JW96] and use a network formation game found in [Jac05] to study indirect network effects in our model.

The benefits a player receives in this game are calculated as follows: In addition to the benefit from the direct connection, a node will also benefit from "friends of the friend", and "friends of the friends of the friend" etc. This is achieved by letting the payoff be calculated relative to the distance between the nodes. β now depends on the minimum number of hops to the node, e.g. the benefit of a direct connection is β, the benefit of a friend of a friend is β^2 etc. We want the benefit to decrease with distance, therefore we need the limitation: $0 < \beta < 1$.

Example: Let us consider the network shown in 5.9. Node 1 and node 4 in the network will receive a benefit of $\beta + \beta^2 + \beta^3$ by being connected with nodes 2

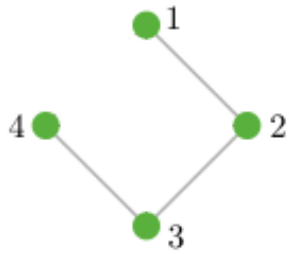


Figure 5.9: Four nodes interconnected with each other.

and 3. $\beta^2 + \beta^3$ represents the indirect benefits from nodes 3 and 4. Nodes 2 and 3 receive a benefit of $\beta + \beta + \beta^2$. For this network to make sense, it is important to also include some cost of having direct connections, or else the rational thing would be to establish a link with everyone. This is done as in earlier models, every node pays a cost for direct connections, but no cost for indirect connections. Thus the total payoff for a node is:

$$\sum_{j \neq i} \beta_{ij}^{d(ij)} - \sum_{j:ij \in g} c_{ij}, \tag{5.39}$$

Where $d(ij)$ represents the shortest path between node i and node j, and c_{ij} represents node i's cost of establishing a link between the two nodes. To simplify the model we choose a symmetric connection process where β and c is set to a fixed global value.

In the paper [JW96], the authors analyze the networks with two different approaches, one with focus on efficiency and the other on stability. The optimal network is of course both efficient and stable, but as we shall see there are some conflicts between efficiency and stability. Matthew, et.al. showed that an efficient network is:

1. *a complete graph* g^N *if* $c < \beta - \beta^2$,
2. *a star encompassing every node if* $\beta - \beta^2 < c < \beta + \frac{(N-2)}{2}\beta^2$,
3. *an empty network (no links) if* $\beta + \frac{(N-2)}{2}\beta^2 < c$.

The most efficient structure is a star structure which encompasses every node. A star structure has the characteristics of minimizing the average path length and uses the minimum number of links $(N-1)$ required for including every node. This structure provides the highest overall payoff for the network, but this network is not necessarily stable.

When analyzing the stability of the network, by using the definition of pairwise stability, Jackson and Wolinsky found four different stability conditions:

1. *a pairwise stable network consists of at most one (non-empty) component,*
2. *if $c < \beta - \beta^2$, the unique pairwise stable network will be a complete graph g^N,*
3. *if $\beta - \beta^2 < c < \beta$, a star encompassing every node will be pairwise stable, although not necessarily the unique pairwise stable graph,*
4. *if $\beta < c$, any pairwise stable network which is nonempty is such that each player has at least two links and is thus inefficient.*

We see that stability condition 2 is the same as efficiency condition 1, and therefore if this condition is fulfilled, the network is both stable and efficient. Condition 3 shows us why the efficient star network is not necessarily stable. If $\beta \leq c < \beta + \frac{(N-2)}{2}\beta^2$ then the efficient network will be a star, but it is not stable.

It should be noticed that it is more beneficial for a node to operate as a leaf node compared to being a center node, due to the cost of direct connections. In a star structure, a leaf node will only have to pay the cost of the link to the center node, and will benefit indirectly for each node connected to the center node. The center node will benefit from each new connection, but, the payoff will only be $\beta - c$ for each connection.

5.6.1 Insurance and connection game

The findings about efficiency and stability are very useful for our model, because if one has knowledge of the different variables it is possible to determine how the network will evolve. Additionally, if you are able to control the variables, you can actually determine the resulting network structure. From the referenced papers, we know that different boundaries on the link cost exists, and how the resulting stable and efficient network will be. Our earlier models show that the cost of establishing a link is the insurance cost and/or the risk cost. From this we can show that if $\beta - \beta^2 < I_l < \beta$ and $r > \beta$, a star with only insured nodes, and no connections between non-insured nodes, is both a stable and an efficient network. If $\beta - \beta^2 < I_l + r < \beta$ and $\beta - \beta^2 < I_l$ and $\beta - \beta^2 < r$, the stable and efficient network is a star consisting of both insured and non-insured nodes. If $I_l < \beta - \beta^2$ all insured nodes will connect to every other insured node, and if $r < \beta - \beta^2$ all non-insured nodes will connect to every other non-insured node. In addition if $r + I_l < \beta - \beta^2$ the resulting network will be a clique of both insured and non-insured nodes. The insurer can thus determine the formation of the network by adjusting the cost parameters.

5.6.2 Homogeneous symmetric connection game

From this point on, the game we will consider is a homogeneous network setting, where every node is considered to be insured. This is done because it will simplify an otherwise very complex model. We are analyzing the resulting network structure, which is easier when only considering one homogeneous cost for every node. Let us look at an example, where the parameters are set to: $\beta = 0.9, I_l = 0.5$. The resulting network from a simulation is shown in Figure 5.10.

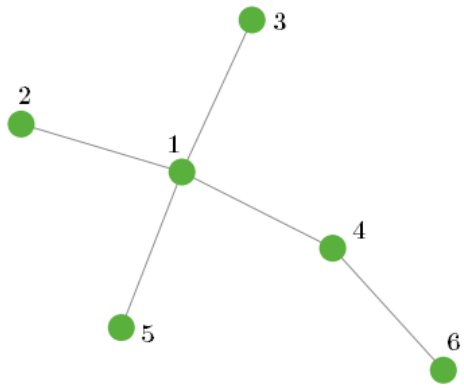


Figure 5.10: The resulting network after a simulation with the parameters $\beta = 0.9, I_l = 0.5$.

As we see this is not an efficient star, but the network is stable. The efficient network would be to delete the link $4, 6$ and adding the link $1, 6$. But since we only consider one link at a time this can not be done. To show this let U_i denote the payoff of node i, the payoffs of the nodes are as described in Eq.(5.40).

$$\begin{aligned} U_1 &= 4\beta + \beta^2 - 4c \\ U_2 = U_3 = U_5 &= \beta + 3\beta^2 + \beta^3 - c \\ U_4 &= 2\beta + 3\beta^2 - 2c \\ U_6 &= \beta + \beta^2 + 3\beta^3 - c \end{aligned} \tag{5.40}$$

Node 6 would benefit from adding the link $1, 6$, but node 1 is not willing to do so, because then it must pay an extra cost, and since $\beta^2 > \beta - c$, the network is stable, but not efficient.

From this we see that, even when the most efficient and stable network is a star, we can not guarantee that the network formation game will end up in a star. This is because we only consider one link at a time, and not the whole network.

A star is not possible with high *n*. In the paper [Jac05] the authors came up with the following proposition: Consider the symmetric connections model in the case where $\beta - \beta^2 < c < \beta$. As the number of nodes grows, the probability that a stable state (under the process where each link has an equal probability of being identified) is reached with the efficient network structure of a star goes to zero. But if a network reaches the efficient star structure, it is also pairwise stable, and will remain a star. We confirmed this when running multiple simulations. When we used few nodes the resulting network often became a star, but as the number of nodes increased the network rarely became a star.

However, the structure of the networks that evolve is very similar to a scale-free network. There are many nodes with low node degree, and few with a high node degree. One example of this is shown in Figure 5.11. There are only ten nodes, but the network has the properties of a scale-free network. Two nodes have a degree of 4, and the rest have a degree of one or two.

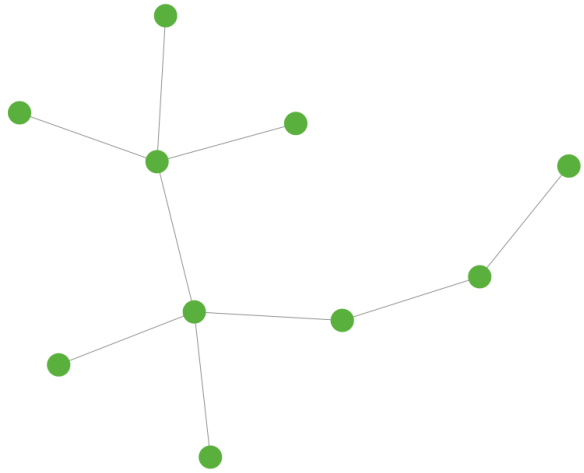


Figure 5.11: The resulting network after a simulation with the parameters described earlier and 10 nodes.

Bulk insurance. As noted before it is not preferable to be the center node, due to the cost of all the direct links. In a model with bulk insurance discount, the extra cost for the center node would decrease significantly. This could be used to increase the probability of reaching a star formation.

Using the discount formula from the previous model, we end up with Eq.(5.41)

to achieve an efficient and stable star topology. i represents the node degree.

$$\beta - \beta^2 < \frac{i_l}{i+1} < \beta \tag{5.41}$$

An interesting property of the discount model is that the conditions for efficient and stable networks will change. Because when the node degree increases, the insurance cost might reach the critical degree g, and the best strategy for a node with degree g or higher, is to connect to every node, as shown in Eq.(5.42). The critical degree occurs when a node's optimal strategy changes from relaying on indirect connections to connecting to every node.

$$\frac{I_l}{g} < \beta - \beta^2 \tag{5.42}$$

This is possible when $g < n$, where n represents the number of nodes in the network. The stability condition has changed for a node with a critical degree. The stable and efficient condition for this node is, as shown earlier, to have a direct connection to every other node. Thus if we have a star topology, both the leaf nodes and the center node are stable, and the center node has been compensated for its role in the network.

Since the networks formed are similar to scale-free networks, we can calculate the probability of a node having degree g, see Eq.(5.43). γ is the power law parameter, as described in Chapter 4.

$$P(g) = g^{-\gamma} \tag{5.43}$$

When a node i reaches the critical degree g its optimal strategy is to connect to every node, since the payoff generated from direct connections is larger than any indirect connection. In general, nodes prefer to connect to nodes with high connectivity[2], and will thus prefer to connect to this node compared to nodes with a degree lower than g. In this way, nodes will connect to the node who has a degree greater than or equal to g, and remove the links to their low-degree nodes which they can instead reach through the node with high connectivity.

Let us consider a case with n nodes, and two of these nodes, i and j, have an equal degree larger than g. The rest of the nodes have a degree of one or zero. If there exists a node with degree zero, it would prefer to be connected to i or j, and so will i and j, so this will eventually happen. If a node connected to i is considering connecting to j, or vice versa, it will do so because j can offer a higher connectivity than i. Now j has a higher degree than i, and thus every node would prefer to connect to j over i. This will eventually result in a star formation, with j as the center node. From this we get the conjectures:

[2]A node with high degree implies a node with high connectivity.

Conjecture 1. If the critical degree ratio is low, i.e. the ratio between critical degree and number of nodes in the network, the resulting network will with high probability be a clique.

Conjecture 2. If the critical degree ratio is at a medium level, the resulting network will with high probability be a star.

Conjecture 3. If the critical degree ratio is high, the resulting network will with high probability be a star-like/scale-free structure.

A numerical example of the boundaries between the different structures, we found from our simulation (described in the next section) with 20 nodes is the following: As seen in Figure 5.12 a critical degree of 1-5 applies to conjecture 1, 6-12 applies to conjecture 2 and 13-20 applies to conjecture 3.

Results and findings

To prove the conjectures above, we created a simulator. The rules of the simulator are as following: Every round of the game,two random nodes, not neighbors, are selected, and asked if they would want to establish a link. The link establishment is a symmetric decision, i.e. the link is established if it result in an increased payoff for both nodes. If the link is added, we check if either of the nodes would prefer to delete some of their already existing links, this decision is asymmetric. A link will be deleted if the node will achieve a higher payoff without it. Then we ask the rest of the nodes if they would like to delete any links. This procedure is repeated as long as it is possible to add new links. The payoff function of each node is as described earlier (see Eq.(5.39)), except that the cost is now dependent on the degree of the node. For the simulations to be realizable, we had to set the number of nodes to 20, or else the computational time would be to high. For every critical degree, from three to nineteen, we ran 50 simulations, and noted the resulting network formation. We chose to start from critical degree equal three, since any number below would result in a clique, because it would be more beneficial to be directly connected to every node.

We know that if Eq.(5.41) is satisfied for all i, then the efficient and stable state is a star. But a more interesting scenario occurs when we have a graph where one or more of the nodes reaches the critical degree. -Will the final structure be scale-free, a star or simply just unstructured? The results from the simulation can be seen in Figure 5.12, 5.13 and 5.15. As we see from Figure 5.12, the probability of the resulting network being a star suddenly increases from zero to 42% at critical degree five to six, and then jumps from 42 to 70-, 86-,96-, 98% at critical degree six to nine. These results confirm our conjectures, and show that the discount can drastically increase the probability of the network ending up in a star.

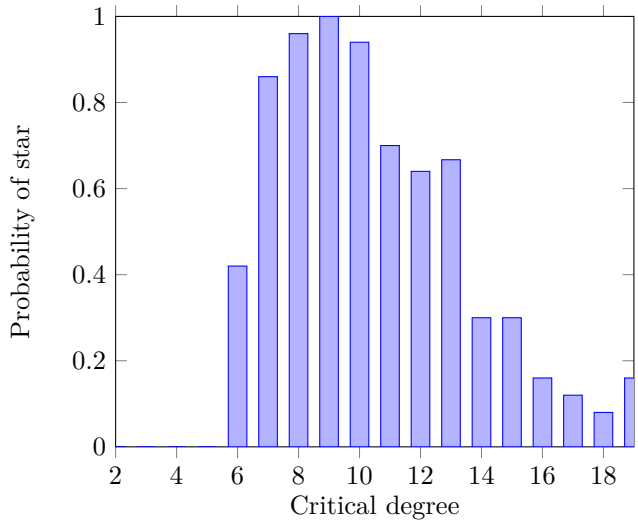


Figure 5.12: Shows the probability of the network ending up in a star, given different critical degrees.

From Figure 5.13 we can observe that the opposite is happening when the critical degree is increased; the probability of the resulting network being a clique drastically decreases. As we can see with a critical degree of seven or higher, it is very unlikely that we end up with a clique. These findings support our conjectures.

An interesting comparison can be made between the emergence of a star versus a clique. Figure 5.14 shows a plot of the network resulting in a star and another plot of the probability for the resulting network to become a clique. As we can see, from a critical degree of five to seven, the resulting network structure, changes from almost certainly ending up in a clique, to almost certainly ending up in a star structure. The reason is as mentioned before that when the critical degree is low, the likelihood of many nodes reaching the critical degree is high. And none of these would like to delete any links. Hence we end up with a clique. The reason why we end up with star structures is because it is less likely that many nodes end up reaching the critical degree, hence most of the nodes still prefer to rely on indirect links, but the ones that reach the critical degree prefer to connect to everyone. Since the nodes with critical degree, have high connectivity, nodes will prefer to be connected with these, compared with other nodes. Nodes prefer to be connected to the ones with critical degree, the nodes with critical degree would like to connect to everyone, and thus the structure evolves into a star, with the critical degree node in the center.

In Figure 5.12 when the critical degree gets closer to the number of nodes in the network, the probability of the network evolving into a star decreases. However, in

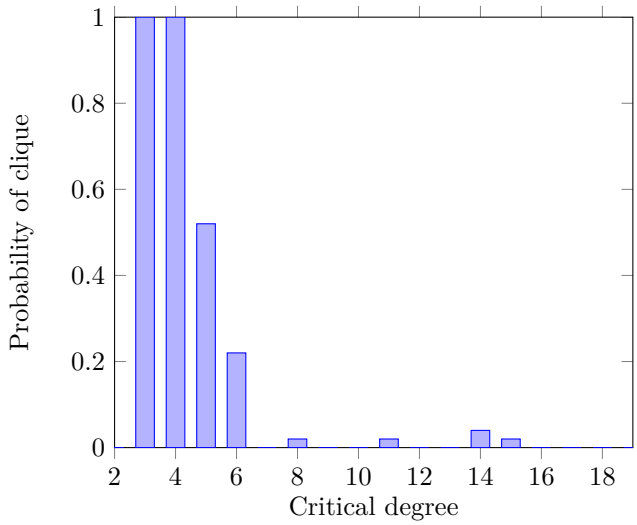


Figure 5.13: Shows the probability of the network ending up in a clique, given different critical degrees.

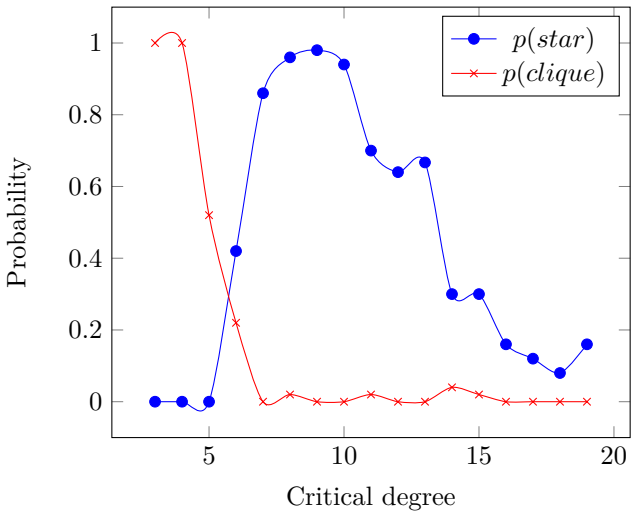


Figure 5.14: Shows the comparison between the probability of the network ending up in a star (blue) or clique (red), given different critical degrees.

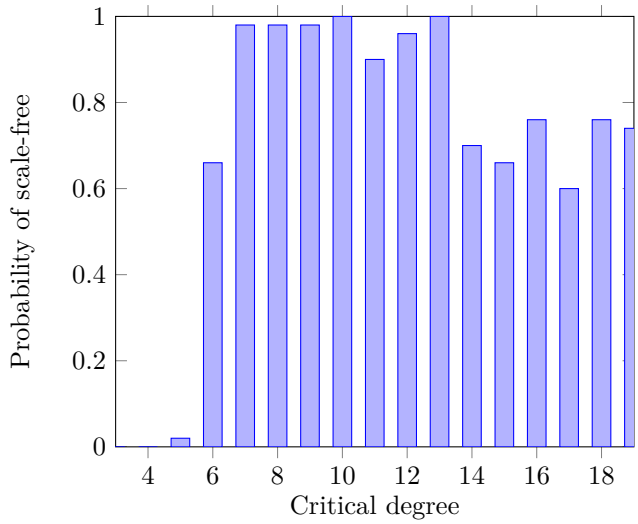


Figure 5.15: Shows the probability of the network ending up in a scale-free structure, given different critical degrees.

Figure 5.15, we have plotted the probability of the network evolving into a network where only a few(2-4) nodes end up with a high degree, but not necessarily a critical degree. As we see, this occurs with high probability from critical degree six and up. These networks are so called scale-free networks (A-B graphs, described in the methodology chapter), because there are a few hubs, that account for most of the connectivity in the network. The reason why we end up with a scale-free network is because nodes prefer to be connected with nodes with high connectivity, and thus will delete links to nodes with low connectivity. This is very similar to the simple model that creates scale-free networks, where the probability of connecting to a node is proportional to the degree of the node.

Price of Anarchy. Another interesting thing is the average price of anarchy as function of the critical degree. The price of anarchy was calculated by taking the average total payoffs and dividing on the optimal payoff. The result can be seen in Figure 5.16.

We see that the price of anarchy for the first critical degrees is 1, and then decreases until degree six, and at seven it increases again. This is because at degree one to five, the socially optimal structure is a clique. At degree six, a clique and a star, are almost equally good, and at degree seven and up, a star structure is the socially optimal outcome. In other words, when the cost is low, a clique is the optimal structure, and when the cost is high a star is the optimal structure.

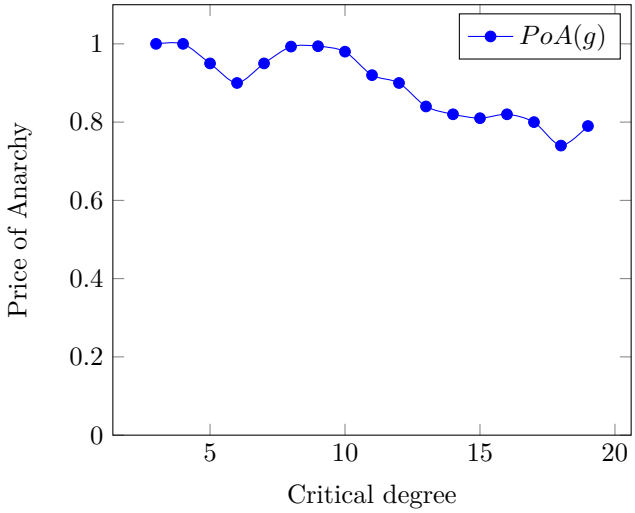
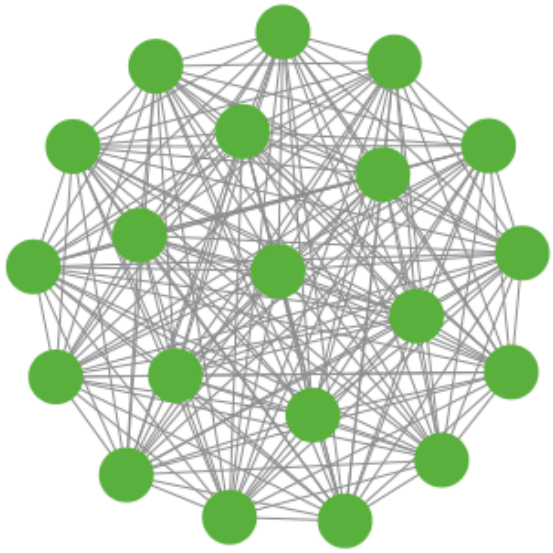


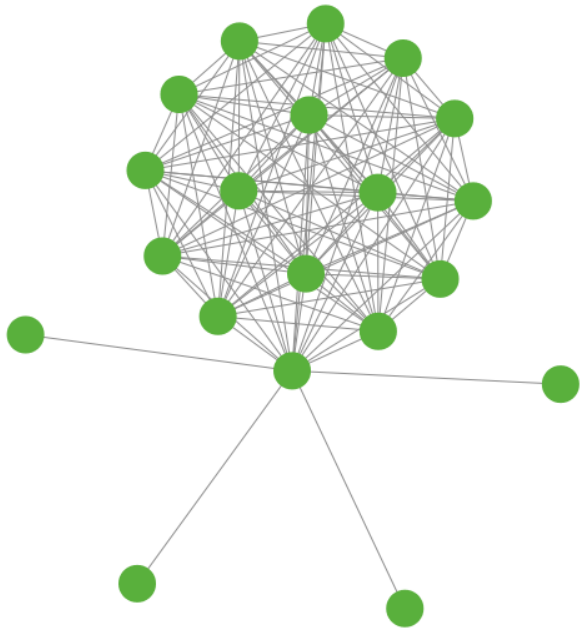
Figure 5.16: Shows the price of anarchy as a function of critical degree

This further improves our findings, because we have now shown how an insurer can determine the resulting network formation by changing the cost. In addition, the formation that evolves has a price of anarchy close to 1.

Example structures from the simulation. In Figure 5.17 we see two of the many possible outcomes when the critical degree is achieved at a low node degree. As we see, most of the nodes have reached the critical degree, and thus connected to every other node. In Figure 5.18 we see one example of a scalefree network, and the standard star network, both with twenty nodes, and results from the simulations when the critical degree was set to a value above six.

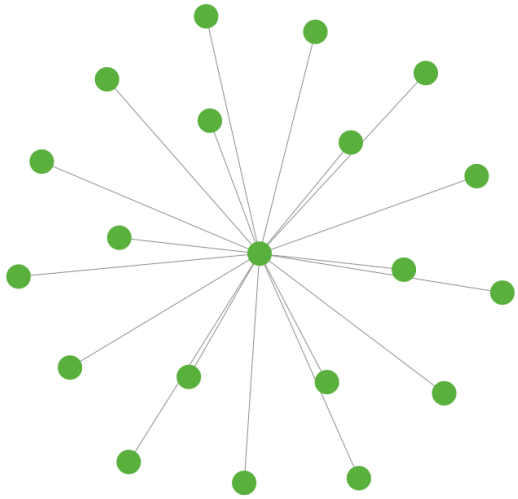


(a) A clique consisting of twenty nodes.

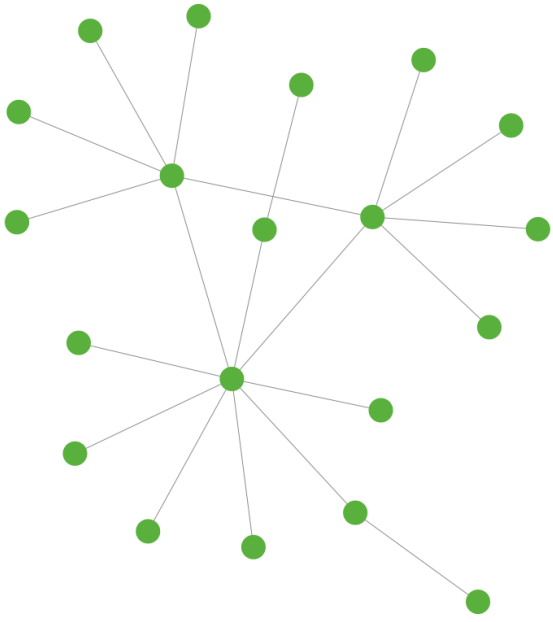


(b) A network with high average node degree, but not a clique.

Figure 5.17: Two different outcomes of the simulations where the critical degree is low



(a) A star consisting of twenty nodes



(b) A scalefree network with twenty nodes, where three nodes account for most of the connectivity.

Figure 5.18: Two different outcomes from running simulations with a high critical degree.

Chapter 6
Summary

6.1 Discussion

From our background study, it was revealed that the current market for cyber-insurance is far from healthy, and many have failed in attempts to establish a cyber-insurance market, also here in Norway. As described in the introduction, there are certain obstacles that are unique for cyber-insurance, and arguably these are the reasons why cyber-insurance has not emerged as expected. However, we believe that there is a need for cyber-insurance, and that our new approach of analyzing the cyber-insurance market through graphs and network formation games could help establishing and improving the market.

We studied a variety of different network formation games, in order to find out if there were any superior network topologies that would fit as a cyber-insurance network, were ideally both the insurer and customers get a higher payoff from purchasing cyber-insurance. We found that star and clique networks had appropriate characteristics, not only do they have calculable fixation probability, but they could also generate better security and overall higher payoff for the nodes. With these networks in mind, we wanted to find a way of forcing networks to evolve into these structures. We found that insurers could adjust the insurance premium in order to control the formation of networks. If the price is set to the right level, networks with calculable risk will evolve, and if the insurer is able to separate the nodes into two different networks, one consisting of trusted, insured nodes, the other of non-insured nodes, the trusted nodes can even further increase their payoff, compared to a non-trusted network. The insurer now possesses a tool for setting the insurance premium properly, possible resulting in better products for both the customer and the insurer.

We created several different models, where the first model showed a very simple and naïve way for the insurer to separate insured and non-insured nodes into two cliques. To make the model more applicable to real-world scenarios, we created

several models, and for each model we added some new features. To get an overview of the models we created, we refer to Figure 5.1.

In model 2 we made model 1 realizable, by including the parameters expected cost of risk, insurance cost and the benefit per link. Then we analyzed the parameters and found out when and how different network structures would evolve. By adjusting the insurance cost to the right level, the insurer can make the network formation game end up in a giant clique of both insured and non-insured nodes, or a clique of only insured and another of only non-insured. The condition for separating insured from non-insured nodes are: $\beta - r < I_l < \beta$, additionally if $\beta > r$, the non-insured nodes will also form a clique, and the resulting network will be two cliques. The solution is also stable, since the resulting network consists of one or two cliques, it is not possible to add any more links. Because the change in payoff is linear and non-dependent on the rest of the network, when a link is added, there is no reason to remove it later. This holds for models 1,2,3 and 4. We also showed that when the insurer sets the cost such that the network ends up in two cliques, it is not the socially optimal, because the network will suffer from the lost benefits of connections between insured and non-insured nodes, i.e. it has a price of anarchy less than 1.

In model 2b, we showed that to be able to separate the networks into two cliques, the nodes must know the other nodes' types. Otherwise, the nodes will have an incentive to pretend to be an insured node, which will result in an untrusted network. We think it is reasonable to assume that nodes in a real world-scenario know whether their transactional partner has insurance or not, therefore we chose not to include this uncertainty in the other models.

In model 3 we applied the model to certain real-world scenarios, such as software development firms/chains, or other networks where the final product is dependent on the collaboration of multiple participants. This was done by including a bonus, which is first received when a node reaches the desired number of links (called max-degree). This made the separation process of insured and non-insured nodes more difficult for the insurer. Due to the possibility of achieving the bonus, a node will have more incentive to establish links, and is thus more accepting towards establishing links with risky nodes. The conditions for separating insured and non-insured nodes in this scenario are: $\beta + \gamma - r < I_l < \beta + \frac{\gamma}{m}$. For the separation of insured and non-insured nodes to be possible, the following has to hold: $1 - \frac{1}{m} < \frac{r}{m}$. As we see, as γ and/or m increases, this gets more and more difficult to achieve.

In Model 4 we tried to implement a common feature used by insurance companies, bulk discount, in order to see how this affected the network formation. The cost of insuring a link is now dependent on the node's degree. We implemented this feature on both model 2 and 3, which resulted in even higher incentive for insured nodes

to establish links with non-insured nodes. The reason is intuitive, since the cost of doing so decreases as the node degree increases. When we applied the discount on model 2, the conditions for ensuring separation of insured and non-insured nodes were: $N_I(\beta - r) < I_l < \beta$, where N_I represents the number of insured nodes in the network. This condition is very strong, because for the separation to be possible the following has to hold: $N_I(\beta - r) < \beta$. As we see, it is now more difficult for the insurer to separate insured and non-insured nodes, compared to model 2, because now the lower boundary on the insurance cost is multiplied with the number of insured nodes in the network ($N_I \times (\beta - r)$).

When applying the discount to model 3, the condition to ensure separation becomes: $m(\beta + \gamma - r) < I_l < \beta + \frac{\gamma}{m}$, and as in the other models, this further complicates the separation process for the insurer.

We also showed that the price of anarchy is even higher when applying discount to model 2. This is because the costs are decreasing, and thus when we have two separate cliques, the potential lost payoff between them will increase. We were not able to calculate the price of anarchy in model 3, because the calculation of the optimal solution is too complex when the bonus and max degree are introduced. However, we can see that since the incentive for establishing links has increased, and thus the insurer has to set a higher price to compensate for this, the price of anarchy will be less than 1, i.e. the more incentive for link establishment you have, the harder it gets to ensure separation of the nodes.

In our last model we applied our model 4 (discount) to an already existing model, "the symmetric connection game". In this old game it has been shown that there are three different efficient and stable networks, clique, star and an empty network, that arise under certain cost conditions. If $I_l < \beta - \beta^2$, the efficient and stable network is a clique. If $\beta - \beta^2 < I_l < \beta$ a star is both stable and efficient. If $I_l > \beta + \frac{N-2}{2}\beta^2$ an empty network is both stable and efficient. In general, a clique is the most efficient if the cost of establishing links is less than the benefit gained from indirect connections. A star is the most efficient if the cost is higher than the benefit from indirect connections, but less than the benefit of direct connections. Unfortunately, it is proved that as the number of nodes in the networks increases, the probability of the network ending up in star goes to zero. However, when we applied our insurance discount to this model, we found conjectures saying that, setting the cost to the right level, one can with high probability ensure that either a clique, a star or a scale-free structure will evolve. This changes the connection game drastically, because now the insurer is able to force the network into three possible network formations, where the star has a fixation probability that exceeds the cliques. The insurer can use these findings to ensure that one of the beneficial structures, star or clique evolves. If the insurer is able to force a star to evolve, this can be used to drastically increase the

overall security, and at the same time minimize the overall link cost.

Limitations and future work One limitation to our work, and a suggestion for future work, is a requirement for mapping our models and simulations to real world networks in a more convincing way. Real-world networks are not random, nodes may prefer to talk to nodes with high degree or low degree, i.e. the payoff function has to be changed.

Another limitation and suggestion for future work is that we have assumed additive risk. It is reasonable to assume that the probability of failure increases if a node accepts more and more links to non-trusted nodes. However, we were not able to determine whether the risk parameter increases according to an additive, exponential, logarithmic distribution or something completely different. By introducing a complex risk function, we would only have distorted the goal of the models. The decision to use additive risk was taken due to the simplicity of the function and the fact that we do not know for sure what the distribution actually looks like.

Another interesting thing to research, is the game of choosing insurance or not. In future work this could be applied to our models, but this could also possibly be too complex, and only disrupt the models.

6.2 Conclusion

The current market for cyber-insurance is far from healthy, and many have failed to establish a cyber-insurance market. However, we believe that there is a need for cyber-insurance, and that our new approach of analyzing the cyber-insurance market through graphs and network formation games could help improving and establishing a better market.

We surveyed literature on networks and risk, and found recent literature that showed how graphs like cliques, star, super-star, funnel and meta-funnel all have a calculable fixation probability, and that stars and funnels fixation probability exceeds the one of a clique. With these structures in mind, we created and analyzed different network formation games, and tried to find link-cost constraints, which enabled these structures to evolve.

In models one to four, we found cost constraints to separate insured and non-insured nodes into two cliques. For each model, we added some new features that made the model more applicable to real world scenarios, and for every feature added, it became more difficult for the insurer to separate the two types of nodes. This is due to the increased incentive for establishing links, and thus the nodes became more and more accepting towards risk.

In the last model, we introduced the concept of bulk insurance into an already existing network formation game, "the symmetric connection game", and showed that this enabled the insurer to determine, with high probability, when and how, cliques, stars or scale-free network would evolve. We showed that at a point, called critical degree, a node's optimal strategy would change from relying on indirect connections, to suddenly wanting to connect to everyone. If the critical degree is set to the right level, one can ensure that the different structures evolve. If the critical degree is set to a low degree, a clique will most certainly evolve, at a medium level, a star will evolve, and at a high level, a scale-free network will evolve. We proved this by performing multiple simulations, 50 simulations for every critical degree. What makes this a very interesting finding, is that in the connection game, earlier research has proven that as the number of nodes increases, the probability of the network reaching a star goes towards zero. However, by introducing a discount, that will subsidize the center node, one can drastically increase the probability of the network ending up in a star.

In summary, we have shown how insurers can determine the resulting networks, by adjusting the insurance cost, for several network formation games. We have also showed how insurers can be assisted in calculating the overall probability of fixation. We found these conditions for several models, with different properties that relate them to the real world and other insurance products. We believe our findings can help the cyber-insurance market evolve into a very interesting and useful market.

References

[Ake97] George A Akerlof. The market for" lemons": Quality uncertainty and the market mechanism. *Readings in Microeconomic Theory*, page 285, 1997.

[And10] R.J. Anderson. *Security Engineering: A guide to building dependable distributed systems*. Wiley, 2010.

[Aud] Jan A. Audestand. Some aspects concerning the vulnearbility of the computerized society. http://www.item.ntnu.no/_media/academics/courses/ttm6/vulnerability.pdf. Accessed: 20/02/2013.

[BK06] Rainer Böhme and Gaurav Kataria. Models and measures for correlation in cyber-insurance. In *Fifth Workshop on the Economics of Information Security*, 2006.

[BL08a] Jean Bolot and Marc Lelarge. Cyber insurance as an incentive for internet security. *Managing information risk and the economics of security*, pages 269–290, 2008.

[BL08b] Jean C Bolot and Marc Lelarge. A new perspective on internet security using insurance. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages 1948–1956. IEEE, 2008.

[Blu11] Easley D. Kleinber J. Kleinberg R. anad Tardon E. Blumen, L. Network formation in the presence of contagious risk. 2011.

[BMR09] T. Bandyopadhyay, V.S. Mookerjee, and R.C. Rao. Why it managers don't go for cyber-insurance products. *Communications of the ACM*, 52(11):68–73, 2009.

[Böh10] Rainer Böhme. Towards insurable network architectures. *Information Technology*, 2010, 2010.

[Bol85] B. Bollobás. Random graphs. *Academic Press*, 1985.

[Bro] RTM Insurance Brokers. Rtm's hackersforsikring. http://www.hackerforsikring.dk/index.html. Accessed: 13/02/2013.

[BS10] R. Böhme and G. Schwartz. Modeling cyber-insurance: Towards a unifying framework. *Proceedings of GameSec*, 2010, 2010.

[CAİ09] Antoni Calvó-Armengol and Rahmi İlkılıç. Pairwise-stability and nash equilibria in network formation. *International Journal of Game Theory*, 38(1):51–79, 2009.

[CfAPA] CAPA Centre for Asia Pacific Aviation. Skywest airlines. http://centreforaviation.com/profiles/airlines/skywest-airlines-oo. Accessed: 08/04/2013.

[Chu] Emily Chung. Playstation data breach deemed in 'top 5 ever'. http://www.cbc.ca/news/business/story/2011/04/27/technology-playstation-data-breach.html. Accessed: 2/05/2013.

[CoA] Travelers Casualty and Surety Company of America. Cyberrisk. https://www.travelers.com/business-insurance/management-professional-liability/Cyber-Risk.aspx. Accessed: 31/01/2013.

[Dic] Oxford Dictionaries. Prisoner's dilemma. http://oxforddictionaries.com/definition/english/prisoner's%2Bdilemma. Accessed: 25/04/2013.

[dig] digi.no. Vil forsikre alt og alle på nett. http://www.digi.no/39107/vil-forsikre-alt-og-alle-paa-nett. Accessed: 18/02/2013.

[DS06] George Danezis and Stefan Schiffner. On network formation,(sybil attacks and reputation systems). In *DIMACS Workshop on Information Security Economics*, pages 18–19, 2006.

[EK12] D. Easley and J. Kleinberg. Networks, crowds, and markets: Reasoning about a highly connected world, 2012.

[Faa] faa Federal aviation administration. Calendar year 2011 primary airports. http://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/passenger/media/cy11_primary_enplanements.pdf. Accessed: 08/04/2013.

[Gar07] Argyrakis P. Garas, A. Correlation study of the athens stock exchange. 2007.

[GGJ+10] A. Galeotti, S. Goyal, M.O. Jackson, F. Vega-Redondo, and L. Yariv. Network games. *The review of economic studies*, 77(1):218–244, 2010.

[HH07] Hemantha S Herath and Tejaswini C Herath. Cyber-insurance: Copula pricing framework and implications for risk management. In *Workshop on the Economics of Information Security (WEIS), Carnegie Mellon University, Pittsburgh, PA*, 2007.

[Ins11] Ponemon Institute. Second annual cost of cyber crime study, benchmark study of u.s: Companies. Technical report, Ponemon Institute, Aug 2011.

[it] Dagens it. Forsikring mot hackere. http://www.dagensit.no/arkiv/article1345297.ece. Accessed: 14/02/2013.

[Jac05] M.O. Jackson. A survey of network formation models: Stability and efficiency. *Group Formation in Economics: Networks, Clubs and Coalitions, ed. G. Demange and M. Wooders*, pages 11–57, 2005.

[JW96] Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.

[KH03] Howard Kunreuther and Geoffrey Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2-3):231–249, 2003.

[LB09] Mark Lelarge and Jean Bolot. Economic incentives to increase security in the internet: The case for insurance. In *INFOCOM 2009, IEEE*, pages 1494–1502. IEEE, 2009.

[LHN05] Erez Lieberman, Christoph Hauert, and Martin A Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 2005.

[MCR80] R.I. Mehr, E. Cammack, and T. Rose. *Principles of insurance*. RD Irwin, 1980.

[MYK06] Ruperto P Majuca, William Yurcik, and Jay P Kesan. The evolution of cyberinsurance. *arXiv preprint cs/0601020*, 2006.

[New] Graeme Newman. Cyber liability in europe: What insurers should knowl. http://www.cfcunderwriting.com/media/news-articles/european-cyber.aspx. Accessed: 14/02/2013.

[Nor] Gjensidige Nor. Medlemsfordeler hos gjensidige 2012 - nal. http://www.arkitektur.no/gjensidige?iid=372345&pid=NAL-Article-Files.Native-InnerFile-File. Accessed: 14/02/2013.

[NRTV07] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge University Press, 2007.

[Pal12] Ranjan Pal. Cyber-insurance for cyber-security a solution to the information asymmetry problem. May 2012.

[PD12] National Protection and Programs Directorate. Cybersecurity insurance workshop readout report. *U.S. Department of Homeland Security*, 2012.

[PGP11] Ranjan Pal, Leana Golubchik, and Konstantinos Psounis. Aegis a novel cyber-insurance model. In *Decision and Game Theory for Security*, pages 131–150. Springer, 2011.

[PH12] Ranjan Pal and Pan Hui. Cyberinsurance for cybersecurity a topological take on modulating insurance premiums. *ACM SIGMETRICS Performance Evaluation Review*, 40(3):86–88, 2012.

[PH13] Ranjan Pal and Pan Hui. On differentiating cyber-insurance contracts a topological perspective. *Internet Management Conference*, 2013.

[PpD12] National Protection and U.S. Department of Homeland Security programs Directorate. Cybersecurity insurance workshop readout report, Nov 2012.

[Pra] Mary K. Pratt. Cyber insurance offers it peace of mind – or maybe not. http://www.computerworld.com/s/article/9223366/Cyber_insurance_offers_IT_peace_of_mind_or_maybe_not?taxonomyId=17&pageNumber=1. Accessed: 31/01/2013.

[Ris12] Stratic Risk. Evolving cyber cover. http://www.strategic-risk.eu/Journals/2012/02/22/i/j/w/RiskFinancing_Mar12.pdf, March 2012. Accessed: 31/01/2013.

[RKK08] Svetlana Radosavac, James Kempf, and Ulaş C Kozat. Using insurance to increase internet security. In *Proceedings of the 3rd international workshop on Economics of networked systems*, pages 43–48. ACM, 2008.

[Rob12] N. Robinson. Incentives and barriers of the cyber insurance market in europe. 2012.

[Spa] Sparebank1. Spar inntil 25 https://www2.sparebank1.no/sr-bank/forsikring/skadeforsikring/fa-rabatt-pa-forsikringer/. Accessed: 09/04/2013.

[SSFW10] Nikhil Shetty, Galina Schwartz, Mark Felegyhazi, and Jean Walrand. Competitive cyber-insurance and internet security. In *Economics of Information Security and Privacy*, pages 229–247. Springer, 2010.

[Wat08] Joel Watson. *Strategy: An introduction to game theory*. WW Norton, 2008.

[Wat11] Tower Watson. Despinte increasing cyber threats, most companies are not buying network liability policies. http://www.towerswatson.com/press/4482, May 2011. Accessed: 31/01/2013.

[Wik] Wikipedia. The market for lemons. http://en.wikipedia.org/wiki/The_Market_for_Lemons. Accessed: 13/02/2013.

[Wil] Uri Wilensky. Netlogo, programmable modeling enviroment. Accessed: 15/02/2013.

Appendix A
Analysis

A.1 Analysis of model-2b: Incomplete information

In this section we present the analysis and mathematics for model 2b: Incomplete information.

When facing a game with incomplete information, there exists two types of equilibriums, one where node 2 is able to seperate node 1's type, called separating equilibrium. The other is where he is not able to separate them, called pooling equilibrium. We have two types of node, type 1 ($t1$): insured and type 2 ($t2$): not insured.

Node 2 is insured. There are two different games to model, one where node 2 is insured, and the other where he is not insured. We start with the one where he is insured. Node 1's type is chosen randomly by nature, with probability p of being type 1 and $1-p$ of being type 2.

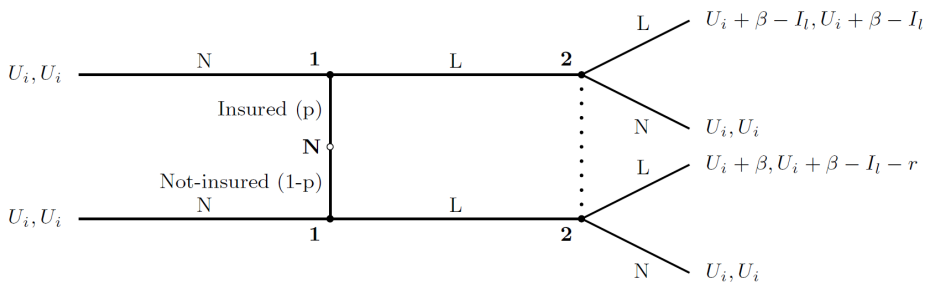


Figure A.1: Signalling game with two nodes, node 1's type choosen by nature, node 2 is insured. Node 1 have complete information, node 2 suffer from incomplete information, and act on best response functions based on beliefs.

In the extensive-form shown in Figure A.1, we see that $t2's$ strategy L dominates N, and thus $t2$ will never play N.

Separating equilibrium. Since node 1 will never play N as type 2, there are only one possible separating equlibrium, type 1 plays L and type 2 plays N. Hence node 2's beliefs are as in Eq.(A.1).

$$\sigma_1(t_i) = \begin{cases} N, & \text{if } t1 \\ L, & \text{if } t2 \end{cases} \tag{A.1}$$

Let $\mu_1(t_i|N)$, denote the probability that node 1 is of type t_i. By using bayes rule we get this equation:

$$\mu_1(t_1|N) = \frac{P(N|t_1)P(t_1)}{P(N)} = \frac{P(N|t_1)P(t_1)}{P(N|t_1)P(t_1) + P(N|t_2)P(t_2)} \tag{A.2}$$

With node 2's belief, we get that $\mu_1(t_1|N) = 1$ and $\mu_1(t_2|L) = 1$. We can now calculate node 2's expected utility from playing L and N:

$$\begin{aligned} EU_2(L,L) &= \mu_1(t_1|L)U_2(L,L;t_1) + \mu_1(t_2|L)U_2(L,L;t_2) \\ &\rightarrow \qquad EU_2(L,L) = U_i + \beta - I_l - r \qquad \text{(A.3)} \\ EU_2(N,L) &= \mu_1(t_1|L)U_2(N,L;t_1) + \mu_1(t_2|L)U_2(N,L;t_2) \\ &\rightarrow \qquad EU_2(N,L) = U_i \qquad \text{(A.4)} \end{aligned}$$

From these two equations we see that the best response of node $2(BR_2)$ when he observes the other node choosing action L is:

$$BR_2(L) = \begin{cases} L, & \text{if } \beta - r \geq I_l \\ N, & \text{if } \beta - r < I_l \end{cases} \tag{A.5}$$

Node 2's expected utility when type 1 chooses N, is easily seen to be U_i. To confirm if this is a separating equilibrium we must see if node 1 has any incentive to deviate from the strategies in node 2's belief. Type 2 will never deviate, so lets investigate type 1. In order to get node 1 to be willing to play N when he knows node 2's best response function, the following must hold: $\beta < I_l$. If this is true, then node 2's best response is to play N. I.e. the only separating equilibrium is the following:

$$\beta < I_l \tag{A.6}$$

$$\sigma_1 = \begin{cases} N, & \text{if } t1 \\ L, & \text{if } t2 \end{cases} \tag{A.7}$$

$$BR_2(\sigma_1) = N \tag{A.8}$$

This means that in a separating equilibrium, the game will end up with no link establishment.

Pooling equilibrium. In a pooling equilibrium node 2 will not be able to distinguish the two types, and since $t1$'s strategy L dominates N, i.e. there is only one possible equilibrium, the one where both types of node 1 plays L.

$$\sigma_1(t_i) = \begin{cases} L, & \text{if } t1 \\ L, & \text{if } t2 \end{cases} \tag{A.9}$$

By using bayes rule we get that $\mu(t_1|L) = p$ and $\mu(t_2|L) = 1 - p$. Node 2's expected utility is then:

$$\begin{aligned} EU_2(L, L) &= p(U_i + \beta - I_l) + (1 - p)(U_i + \beta - I_l - r) \\ \rightarrow \quad EU_2(L, L) &= U_i + \beta - I_l - r + pr \qquad (A.10) \\ EU_2(N, L) &= U_i \qquad (A.11) \end{aligned}$$

From this we get node2's best response:

$$BR_2(L) = \begin{cases} L, & \text{if } \beta + rp - r \geq I_l \\ N, & \text{if } \beta + rp - r < I_l \end{cases} \tag{A.12}$$

By using this best response function, node 1 sees that as long as $\beta > I_l$ he will never deviate from node 2's beliefs. Hence, it is a pooling equilibrium where both nodes choose L, as long as $\beta > I_l$ and $\beta + rp - r > I_l$. We also know that: $rp - r \leq 0$ is allways true, and thus there also exists a pooling equilibrium where node 1, plays L, and node 2, plays N. This equilibrium will occur when $\beta > I_l$ and $\beta + rp - r < I_l$.

Node 2 not insured. Here we will analyze the game when node 2 is not insured. The rules of the game are as before, the only thing that has changed is the type of node 2, and thus the payoffs are different and we need to see if there exists separating and pooling equilibrium in this game as well.

Separating equilibrium. In this game there is no dominant strategy for node 1, thus we have to check for the two possible separating equilibriums. We start with the separating equilibrium with the beliefs shown in Eq.(A.13).

$$\sigma_1(t_i) = \begin{cases} L, & \text{if } t1 \\ N, & \text{if } t2 \end{cases} \tag{A.13}$$

With the beliefs in Eq.(A.13), this is node 2's expected payoffs:

$$EU_2(L, L) = (U_i + \beta) \tag{A.14}$$

$$EU_2(N, L) = (U_i) \tag{A.15}$$

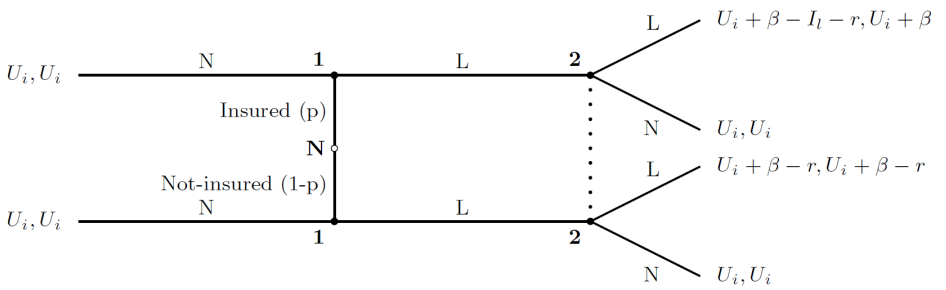


Figure A.2: Signalling game with two nodes, node 1's type choosen by nature, node2 is not insured. Node 1 have complete information, node 2 suffer from incomplete information, and act on best response functions based on beliefs.

From this we see that his best response when node 1's action is L, is to allways play L:

$$BR_2(L) = L \tag{A.16}$$

To see if this is an equilibrium, we have to see if node 1 has any incentive to deviate. We need to check for the two types of node 1: If $\beta > r$ then type 2 would deviate, because he could achieve a higher payoff by playing L, given the beliefs of node 2 in Eq.(A.13). Hence we know that for this to be an equilibrium, the following has to hold

$$\beta < r \tag{A.17}$$

When analyzing from node 1 type 1's perspective, for him to play L, this has to hold: $U_i + \beta - I_l - r > U_i$. The only way this can hold is if $\beta > I_l + r$. We see that Eq.(A.17) is violating this condition, and thus we have no separating equilibrium with the beliefs in Eq.(A.13).

Now lets look at the other possible separating equilibrium, see Eq.(A.18).

$$\sigma_1(t_i) = \begin{cases} N, & \text{if } t1 \\ L, & \text{if } t2 \end{cases} \tag{A.18}$$

Node 2's expected payoffs are as follows:

$$EU_2(L, L) = U_i + \beta - r \tag{A.19}$$

$$EU_2(N, L) = U_i \tag{A.20}$$

From this we get the best response function:

$$BR_2(L) = \begin{cases} L, & \text{if } \beta \geq r \\ N, & \text{if } \beta < r \end{cases} \tag{A.21}$$

For this to be a separating equilibrium, we need to see if node 1 would deviate from node 2's beliefs. Type $t1$ will not deviate as long as $\beta < I_l + r$. Type $t2$ will not deviate if $\beta \geq r$, if this condition is true, we see that node 2 will play L. I.e. the only separating equilibrium that exists is when node 2 plays L, node 1 of type $t1$ plays N and node 1 of type $t2$ plays L. For this to happen we get this condition on β.

$$I_l + r > \beta > r \tag{A.22}$$

Pooling equilibrium. Two possible, one where both types of node 1 plays L, and one where both types plays N. Lets first analyze the one where both types of node 1 plays L.

$$\sigma_1(t_i) = \begin{cases} L, & \text{if } t1 \\ L, & \text{if } t2 \end{cases} \tag{A.23}$$

With the beliefs shown above, node 2's expected payoffs are:

$$EU_2(L) = p(U_i + \beta) + (1-p)(U_i + \beta - r)$$

$$EU_2(L) = U_i + \beta - r + pr \tag{A.24}$$

$$EU_2(N) = U_i \tag{A.25}$$

From this we get the best response function :

$$BR_2(L) = \begin{cases} L, & \text{if } \beta \geq r - pr \\ N, & \text{if } \beta < r - pr \end{cases} \tag{A.26}$$

Will node 1 deviate knowing this? Type $t1$ will not deviate as long as: $\beta - I_l \geq r$, and type $t2$ will not deviate as long as $\beta > r$. From this we get the final condition, if $\beta - I_l \geq r$ then there exists a pooling equilibrium where both types of node 1 plays L and node 2 also play L. From this we see that the other pooling equilibrium where both types of node 1, plays N, will only occur when $\beta < r$ and $\beta < I_l + r$.

Appendix B
Simulation models

B.1 Model 2: Including parameters

Here is the Netlogo source code, used to create the simulator for model 2.

```
turtles-own[
  insured?
  checked?
  numberofedges
  payoff
]
globals[
  numberofinsued
  numberofnotinsured
  donewithinsured?
  donewithnotinsured?
  ]
to setup
  clear-all
  setup-turtles
  reset-ticks
  set numberofinsued 0
  set donewithinsured? false
  set donewithnotinsured? false
  setup-patches
end

to setup-turtles
  set-default-shape turtles "circle"

  crt num-nodes
  layout-circle turtles max-pxcor - 20
  ask turtles [
    set payoff 0
    set insured? false
    set checked? false
    set color red

     if (random-float 100.0 <(prob-
insured))[
       set color green
       set insured? true
       set numberofinsued (numberofinsued
+ 1)
       ]
  ]
  ;ask turtles [ set label who set label-
color black]
end

to setup-patches
ask patches [
  set pcolor white
  ]
end

to go
   if not donewithinsured? [
    add-edge
  ]
  tick
end

to add-edge
 let node1 one-of turtles with[not
checked?]
 if node1 = nobody
 [
   display
   user-message "insured clique finished"
   stop
   ]
  ask node1[
    let node2 one-of turtles with [not link-
neighbor? node1 and (self != node1) and
not checked?]

    ifelse node2 = nobody
    [
      set checked? true
      add-edge
     ]
    [
     let nolinkpayoff payoff
     ifelse insured?
     [
       ;node1 is insured
       ask node2
       [
         let nolinkpayoff2 payoff
         ifelse insured?
         [
```

```
        ;node2 and node1 insured
        let newpayoff1 (nolinkpayoff +
(beta / 100) - (insurancelink / 100 ))
        let newpayoff2 (nolinkpayoff2 +
(beta / 100) - (insurancelink / 100 ))
        if newpayoff1 > nolinkpayoff and
newpayoff2 > nolinkpayoff2
        [
          ;add link
          create-link-with node1
          set payoff newpayoff2
          ask node1[
            set payoff newpayoff1

          ]
        ]
        ;done with adding link
        ]
       [;begin else
        ;node2 not insured
        let newpayoff1 (nolinkpayoff +
(beta / 100) - (risk / 100) - (insurancelink
/ 100 ))
        let newpayoff2 (nolinkpayoff2 +
(beta / 100))
        if newpayoff1 > nolinkpayoff and
newpayoff2 > nolinkpayoff2
        [
          ;add link
          create-link-with node1
          set payoff newpayoff2
          ask node1[
            set payoff newpayoff1
          ]
        ]
        ;done with adding link

        ];end else
     ];done with node2
     ]
    [
     ;node1 not insured
     ask node2
     [
        let nolinkpayoff2 payoff
        ifelse insured?
        [
         ;node2 insured and node1 not
insured
        let newpayoff1 (nolinkpayoff +
(beta / 100))
        let newpayoff2 (nolinkpayoff2 +
(beta / 100) - (risk / 100) - (insurancelink
/ 100 ))
        if newpayoff1 > nolinkpayoff and
newpayoff2 > nolinkpayoff2
        [
          ;add link
          create-link-with node1
          set payoff newpayoff2
          ask node1[
            set payoff newpayoff1

          ]
        ]
        ;done with adding link
        ]
       [;begin else
        ;node2 and node1 not insured
        let newpayoff1 (nolinkpayoff +
(beta / 100) - (risk / 100))
        let newpayoff2 (nolinkpayoff2 +
(beta / 100) - (risk / 100))
        if newpayoff1 > nolinkpayoff and
newpayoff2 > nolinkpayoff2
        [
          ;add link
          create-link-with node1
          set payoff newpayoff2
          ask node1[
            set payoff newpayoff1
          ]
        ]
        ;done with adding link

        ];end else
     ];done with node2
```

```
      ]                                                    ]
    ;set color green                                     end
    ;add-edge
    ]
  ]
  layout
end

to add-edge-not-insured
 let node1 one-of turtles with[not
insured? and not checked?]
 if node1 = nobody
 [
  ;display
  ;user-message "non-insured clique
finished"
  stop
  ]
 ask node1[
   let node2 one-of turtles with [not
insured? and not link-neighbor? node1
and (self != node1) and not checked?]
   ifelse node2 = nobody
  [
    display
    set donewithnotinsured? true
    set checked? true
    add-edge-not-insured
    ]
 [
  create-link-with node2
  add-edge-not-insured
   ]

  ]
  layout

end

to layout
 repeat 10 [
   layout-spring (turtles with [any? link-
neighbors]) links 0.4 6 1
   display  ;; so we get smooth animation
```

B.2 Model 3: Including maximum node degree and bonus

Here is the Netlogo source code, used to create the simulator for model 3.

```
turtles-own[
 insured?
 checked?
 payoff
 m
 explored?
 ;numberofinsured
 ;numberofnotinsured
 max?
]
globals[
 numberofcliques
 component-size      ;; current running size of
component being explored
 giant-component-size  ;; size of largest connected
component
 components
 donewithinsured?
 donecounting?
 done?
 ]
to setup
 clear-all
 setup-turtles
 reset-ticks
 set-max-degree
 set donewithinsured? false

 setup-patches
end
to set-max-degree
 ask turtles[
  ifelse random-max-degree?
  [
    set m ((random 5) + 1)
    ]
  [
    set m max-degree
    ]
 ]
end
to setup-turtles
 set-default-shape turtles "circle"
 set numberofcliques 0
 set done? false
 set donecounting? false
 crt num-nodes
 layout-circle turtles max-pxcor - 20
 ask turtles [
  set payoff 0

  set max? false
  ;set numberofinsured 0
  ;set numberofnotinsured 0
  set insured? false
  set checked? false
  set color red

   if (random-float 100.0 <(prob-insured))[
    set color green
    set insured? true
    ]
 ]
 ;ask turtles [ set label who set label-color black]

end
to show-label
 ask turtles[

 ifelse show-payoff?
   [ set label payoff]
   [ set label ""]
 ]
end

to setup-patches
ask patches [
 set pcolor white
 ]
end

to go
 show-label
 if done? and not donecounting?
 [
  find-all-components
 ]
 if not donewithinsured? [
  add-edge
 ]
 ;if not donewithnotinsured?[
 ;  add-edge-not-insured
 ; ]
 tick
end

to add-edge
 let node1 one-of turtles with[not checked? and
not max?]
 if node1 = nobody
 [
  set done? true
  display
  user-message "insured clique finished"
  stop
  ]
```

```
 ask node1[
  if (m - (count(link-neighbors))) <= 0
  [
   set max? true
   add-edge
  ]
  let node2 one-of turtles with [not link-neighbor?
node1 and (self != node1) and not checked? and
not max?]

  ifelse node2 = nobody
  [
   set checked? true
   add-edge
  ]
  [
   let nolinkpayoff payoff
   let n1m m
   let n1numberofinsured (count(link-neighbors
with[insured?]))
   let n1numberofnotinsured (count(link-neighbors
with[not insured?]))
   ifelse insured?
   [
    ;node1 is insured
    ask node2
    [
     if (m - (count(link-neighbors))) <= 0
     [
      set max? true
      add-edge
     ]
     let nolinkpayoff2 payoff
     let n2m m
     let n2numberofinsured (count(link-neighbors
with[insured?]))
     let n2numberofnotinsured (count(link-
neighbors with[not insured?]))
     ifelse insured?
     [
      ;node2 and node1 insured
      let g1 ((gamma / 100 ) / (n1m -
n1numberofinsured - n1numberofnotinsured ) )
      let g2 ((gamma / 100 ) / (n2m -
n2numberofinsured  - n2numberofnotinsured ))
      let newpayoff1 (nolinkpayoff + (beta / 100 ) -
(insurancelink / 100 ) + g1)
      let newpayoff2 (nolinkpayoff2 + (beta / 100 )
- (insurancelink / 100 ) + g2)
      if newpayoff1 > nolinkpayoff and newpayoff2
> nolinkpayoff2
      [
       ;add link
       create-link-with node1
       set payoff (newpayoff2 - g2)
       ;set numberofinsured (numberofinsured +
1)
       if (m - (count(link-neighbors))) <= 0
       [;set max true
        set payoff (payoff + (gamma / 100 ))
        set max? true
        ]
       ask node1[
        set payoff (newpayoff1 - g1)
        ;set numberofinsured (numberofinsured +
1)
        if (m - (count(link-neighbors))) <= 0
       [;set max true
        set payoff (payoff + (gamma / 100 ))
        set max? true
        ]

       ]
      ]
      ;done with adding link
      ]
     [;begin else
      ;node2 not insured
      let g1 ((gamma / 100 ) / (n1m -
n1numberofinsured - n1numberofnotinsured ) )
      let g2 ((gamma / 100 ) / (n2m -
n2numberofinsured  - n2numberofnotinsured ))
      let newpayoff1 (nolinkpayoff + (beta / 100 ) -
(risk / 100 ) - (insurancelink / 100 ) + g1)
      let newpayoff2 (nolinkpayoff2 + (beta / 100 )
+ g2)
      if newpayoff1 > nolinkpayoff and newpayoff2
> nolinkpayoff2
      [
       ;add link
       create-link-with node1
       set payoff (newpayoff2 - g2)
       ;set numberofinsured (numberofinsured +
1)
       if (m - (count(link-neighbors))) <= 0
       [;set max true
        set payoff (payoff + (gamma / 100 ))
        set max? true
        ]
       ask node1[
        set payoff (newpayoff1 - g1)
        ;set numberofnotinsured
(numberofnotinsured + 1)
        if (m - (count(link-neighbors))) <= 0
       [;set max true
        set payoff (payoff + (gamma / 100 ))
```

```
          set max? true
          ]
         ]
        ]
        ;done with adding link

        ];end else
     ];done with node2
     ]
    [
     ;node1 not insured
     ask node2
     [
      if (m - (count(link-neighbors))) <= 0
      [
       set max? true
       add-edge
      ]
      let nolinkpayoff2 payoff
      let n2m m
      let n2numberofinsured (count(link-neighbors
with[insured?]))
      let n2numberofnotinsured (count(link-
neighbors with[not insured?]))
      ifelse insured?
      [
       ;node2 insured and node1 not insured
       let g1 ((gamma / 100 ) / (n1m -
n1numberofinsured - n1numberofnotinsured ) )
       let g2 ((gamma / 100 ) / (n2m -
n2numberofinsured  - n2numberofnotinsured ))
       let newpayoff1 (nolinkpayoff + (beta / 100 ) +
g1)
       let newpayoff2 (nolinkpayoff2 + (beta / 100 )
- (insurancelink / 100 ) - (risk / 100 ) + g2)
       if newpayoff1 > nolinkpayoff and newpayoff2
> nolinkpayoff2
       [
        ;add link
        create-link-with node1
        set payoff (newpayoff2 - g2)
        ;set numberofnotinsured
(numberofnotinsured + 1)
        if (m - (count(link-neighbors))) <= 0
        [;set max true
         set payoff (payoff + (gamma / 100 ))
         set max? true
         ]
        ask node1[
         set payoff (newpayoff1 - g1 )
         ;set numberofinsured (numberofinsured +
1)
         if (m - (count(link-neighbors))) <= 0
          [;set max true
          set payoff (payoff + (gamma / 100 ))
          set max? true
          ]

         ]
        ]
        ;done with adding link
        ]
       [;begin else
        ;node2 and node1 not insured
        let g1 ((gamma / 100 ) / (n1m -
n1numberofinsured - n1numberofnotinsured ) )
        let g2 ((gamma / 100 ) / (n2m -
n2numberofinsured  - n2numberofnotinsured ))
        let newpayoff1 (nolinkpayoff + (beta / 100 ) -
(risk / 100 ) + g1)
        let newpayoff2 (nolinkpayoff2 + (beta / 100 )
- (risk / 100 ) + g2)
        if newpayoff1 > nolinkpayoff and newpayoff2
> nolinkpayoff2
        [
         ;add link
         create-link-with node1
         set payoff (newpayoff2 - g2 )
         ;set numberofnotinsured
(numberofnotinsured + 1)
         if (m - (count(link-neighbors))) <= 0
         [;set max true
          set payoff (payoff + (gamma / 100 ))
           set max? true
          ]
         ask node1[
          set payoff (newpayoff1 - g1)
          if (m - (count(link-neighbors))) <= 0
          [;set max true
           set payoff (payoff + (gamma / 100 ))
           set max? true
          ]

          ;set numberofnotinsured
(numberofnotinsured + 1)
         ]
        ]
        ;done with adding link
        ];end else
     ];done with node2
     ]
    ;set color green
    ;add-edge
    ]
   ]
   layout
```

```
end

to add-edge-not-insured
 let node1 one-of turtles with[not insured? and not
checked?]
 if node1 = nobody
 [
  ;display
  ;user-message "non-insured clique finished"
  stop
  ]
 ask node1[
   let node2 one-of turtles with [not insured? and
not link-neighbor? node1 and (self != node1) and
not checked?]

   ifelse node2 = nobody
  [
    display
    set checked? true
    add-edge-not-insured
    ]
 [
  create-link-with node2
  add-edge-not-insured
   ]
  ]
  layout
end
to find-all-components
 set components []
 set giant-component-size 0

 ask turtles [ set explored? false ]
 ;; keep exploring till all turtles get explored
 loop
 [
  ;; pick a turtle that has not yet been explored
  let start one-of turtles with [ not explored? ]
  if start = nobody [
    set donecounting? true
    display
    user-message "Done counting cliques"
    stop ]
  ;; reset the number of turtles found to 0
  ;; this variable is updated each time we explore
an
  ;; unexplored turtle.
  set component-size 0
  ask start [ explore ]
  set numberofcliques numberofcliques + 1
  ;; the explore procedure updates the
component-size variable.
  ;; so check, have we found a new giant
component?
  if component-size > giant-component-size
  [
   set giant-component-size component-size
  ]
  set components lput component-size
components
 ]
end

;; finds all turtles reachable from this turtle
to explore ;; turtle procedure
 if explored? [ stop ]
 set explored? true
 set component-size component-size + 1
 ask link-neighbors [ explore ]
end

to layout
 repeat 10 [
  layout-spring (turtles with [any? link-neighbors])
links 0.4 6 1
  display  ;; so we get smooth animation
 ]
end
```

B.3 Model 5: Network externalities

Here is the Netlogo source code, used to create the simulator for model 5.

```
extensions [nw table]
links-own [ weight ]
turtles-own [
 dict;dictionary with shortest path to every node
 insured?
 checked?
 payoff
 cost-of-link-with-other-turtles ;;
 distance-from-other-turtles
 indirpayoffbefore
 indirpayoffafter
 degree
 ]
globals[
donewithinsured?
infinity
newpayoff1
newpayoff2
nolinkpayoff
nolinkpayoff2
nr1
nr2

]
to setup-shape
 clear-all
 setup-patches
 nw:generate-ring turtles links 10 [ set color red ]
 nw:set-snapshot turtles links
 layout
 set infinity 99999
 ask turtles [
  set indirpayoffbefore 0
  set indirpayoffafter 0
  set payoff 0
  set insured? true
  set checked? false
  set color green
  let node-count count turtles
  let x 0
 ]
 compute-inital-payoff
  nw:set-snapshot turtles links
 reset-ticks
end
to setup-star
 clear-all
 setup-patches
 setup-turtles-star
 reset-ticks
end
to setup
 clear-all
 setup-patches
 setup-turtles
 reset-ticks
end

to setup-turtles-star
 setup-patches
 set-default-shape turtles "circle"
 nw:generate-star turtles links num-nodes
 nw:set-snapshot turtles links
 layout
 set infinity 99999
 ask turtles [
  set indirpayoffbefore 0
  set indirpayoffafter 0
  set payoff 0
  set insured? true
  set checked? false
  set color green
  let node-count count turtles
  let x 0
 ]
 compute-inital-payoff
  nw:set-snapshot turtles links
end
to setup-turtles
 set-default-shape turtles "circle"
 set infinity 99999
 crt num-nodes
 layout-circle turtles max-pxcor - 20
 ask turtles [
  set indirpayoffbefore 0
  set indirpayoffafter 0
  set payoff 0
  set insured? true
  set checked? false
  set color green
  let node-count count turtles
  let x 0
 ]
  nw:set-snapshot turtles links
 ;ask turtles [ set label who set label-color black
end

to compute-inital-payoff
 find-path-lengths
 ask turtles [
  set degree count link-neighbors
  let nr who
  let i 0
  let j 1
  set payoff 0
  foreach distance-from-other-turtles [
```

```
    if( ? < 999 )[
      if(? != 0)[
        set payoff (payoff +( (beta / 100) ^ ? ))
      ]
      if( ? = 1)[
        set payoff (payoff - ((insurancelink / 100 )  /
(j)))
        set j j + 1
    ]
    ]
  ]
 ]
end

to setup-patches
ask patches [
 set pcolor white
 ]
end

to go
   add-edge-simpler
   delete
   layout
   tick
end
to delete
 let i 0
 while [i < count turtles ]
 [
  check-delete i
  set i i + 1
  ]
end
to add-edge-simpler
 set newpayoff1 -1
 set newpayoff2 -1
 set nolinkpayoff 0
 set nolinkpayoff2 0
 compute-inital-payoff
 let node1 one-of turtles
 if( node1 = nobody)[
  display
  user-message "ferdig"
  stop
 ]
 set nr1 0
 set nr2 0
 let link? false
 ask node1[
  set nolinkpayoff payoff
  set nr1 who
   let node2 one-of turtles with [not link-neighbor?
node1 and (self != node1) and not checked?]
   ifelse node2 = nobody
   [
    set checked? true
   ]
   [
    ask node2 [set nr2 who
     set nolinkpayoff2 payoff
     ]
    set link? true
   ]
  ]
  if( link?)[
  create-and-check-path nr1 nr2
  check-delete nr1
  check-delete nr2
  ]
end

to setup-indivudal-map
 let j 0
 let c count turtles
 while [j < c][
 ask turtle j[
 let i 0
 set dict table:make
 while [i <= c - 1][
 if j != i[
  table:put dict i nw:path-to turtle i
  ]
 set i i + 1
 ;end while
 ]
 ;end ask
 ]
 set j j + 1
 ;end while
 ]

end

to check-delete[a]
 let i 0
 let opay -1
 let dist []
 ask turtle a[
 set opay payoff
 set dist distance-from-other-turtles
 ]
  foreach dist
  [
   if( ? = 1)[
```

```
    ;neighbors
    ;i is the turtle nr
    ask link a i[
     die
    ]
    nw:set-snapshot turtles links
    find-path-lengths
    compute-inital-payoff
    ask turtle a[
     if(payoff < opay)
     [
       ;do not delete link
       create-link-with turtle i [ set weight 2.0 ] ]
     nw:set-snapshot turtles links
     ]
     find-path-lengths
     compute-inital-payoff
    ]
     set i i + 1
   ]

end

to create-and-check-path[a b]
  let temp []
  let nextloop? true
  ;create temporary table of pathes from 0 to 2.
  ask turtle a [set temp nw:path-to turtle b]

  let len length temp
  ask turtle a [ create-link-with turtle b [ set weight
2.0 ] ]
  let nlink link a b
  nw:set-snapshot turtles links
  setup-indivudal-map
  find-path-lengths

  let t []
  let i 0
  let oldneighbor -1

  nw:set-snapshot turtles links
  find-path-lengths
  compute-inital-payoff
  if ( ([payoff] of turtle a ) <  nolinkpayoff or
([payoff] of turtle b ) <  nolinkpayoff2 )[
    ;remove new link, and recreate the old.
    ask link a b[
      die
      nw:set-snapshot turtles links
    ]
    if(oldneighbor != -1)[
      ask turtle a [ create-link-with oldneighbor [ set
weight 2.0 ] ]
      nw:set-snapshot turtles links
    ]
  ]
  find-path-lengths
  compute-inital-payoff
  setup-indivudal-map

end

to layout
  repeat 10 [
    layout-spring (turtles with [any? link-neighbors])
links 0.4 6 1
    display  ;; so we get smooth animation
  ]
end

to find-path-lengths
  ;; reset the distance list
  ask turtles
  [
    set distance-from-other-turtles []
  ]

  let i 0
  let j 0
  let k 0
  let node1 one-of turtles
  let node2 one-of turtles
  let node-count count turtles
  ;; initialize the distance lists
  while [i < node-count]
  [
    set j 0
    while [j < node-count]
    [
      set node1 turtle i
      set node2 turtle j
      ;; zero from a node to itself
      ifelse i = j
      [
        ask node1 [
          set distance-from-other-turtles lput 0
distance-from-other-turtles
        ]
      ]
      [
        ;; 1 from a node to it's neighbor
        ifelse [ link-neighbor? node1 ] of node2
        [
          ask node1 [
```

```
          set distance-from-other-turtles lput 1
distance-from-other-turtles
        ]
       ]
       [
        ask node1 [
          set distance-from-other-turtles lput infinity
distance-from-other-turtles
        ]
       ]
     ]
     set j j + 1
   ]
   set i i + 1
  ]
  set i 0
  set j 0
  let dummy 0
  while [k < node-count]
  [
   set i 0
   while [i < node-count]
   [
    set j 0
    while [j < node-count]
    [
     set dummy ( (item k [distance-from-other-
turtles] of turtle i) +
                (item j [distance-from-other-turtles] of
turtle k))

      if dummy < (item j [distance-from-other-
turtles] of turtle i)
      [
       ask turtle i [
         set distance-from-other-turtles replace-item
j distance-from-other-turtles dummy

        ]
       ]
       set j j + 1
     ]
     set i i + 1
   ]
   set k k + 1
  ]

end
```