

Cyber-insurance & endogenous network formation

Håvard Råmundal Halse¹, Jonas Hoemsnes¹ and Gergely Biczók²

¹ Dept. of Telematics
Norwegian University of Science and Technology
{havahals, jonashoe}@stud.ntnu.no

² Dept. of Telematics
Norwegian University of Science and Technology
gbiczok@item.ntnu.no

Abstract. Cyber-insurance is a powerful economic concept that can help companies in the fight against cybercrime. From the early 80s, several researchers claimed that cyber-insurance had a bright future, were it would become a huge economical tool for handling residual cyber-risks. However, both the European and US cyber-insurance market have failed to grasp its promising potential. To fully grasp this potential they need innovative approaches to handle the unique problems of cyber-insurance. This paper find and characterize network structures with properties that make them superior as cyber-insurance. And creates several models for forming these network structures. In every model, new properties that relate the model to the real-world and real-world insurance products are added. The results show that insurers can use the insurance premium as a tool for determining the resulting formation of the network, and if set to the right level, these superior structures will evolve. We believe our findings could help the cyber-insurance market evolve, by giving the insurers a proper tool to better analyze and control formation of cyber-insurance networks.

1 Introduction to cyber-insurance

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 [BO1]:

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, mis-aligned 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 [BO2,RA1].

The paper [BS1] describes the three main problems of cyber-insurance information asymmetry, correlated risk and interdependent agents.

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.

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. 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.

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 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. Bohme et. al argue that a model for cyber-insurance has to overcome each of these obstacles. They analyze several other papers on cyber-insurance, and show how all of them are touching upon these problems, but mostly they focus on only one or two of these, not all three.

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 [BO3]. Continuing into the beginning of the new millennium, several companies started offering standalone cyber-insurance, which sat the frame for the current insurance product. However, as found in the papers [PO1,ST1,GN1] both the US and the european market have not managed to reach its promising potential. Companies are weekly suffering from successful attacks, but most firms still have not acquired cyber-insurance. The cyber-insurance market seem to have a huge potential, but needs some new thinking to fully take advantage of it. We will take a new approach where we

focus on finding network structures that will be beneficial for cyber-insurance, and see if it is possible for insurers to force these structures to evolve.

2 Cyber-insurance 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 each other. 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 each other, 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 [?] is about evolutionary dynamics and how some structures can amplify or sustain evolution and drift³. 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 important properties, such as sustaining viruses from spreading, or amplify the incentive for obtaining cyber-insurance.

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

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

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

³ Drift is the opposite of selective evolution, it is when the network/structure evolve and change at random

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. (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.(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.

[?] shows that such graphs exist, and one example is the star topology. In this topology the fixation probability is as shown in Eq.(2), or more generally Eq.(3).

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

$$p_k = \frac{(1 - \frac{1}{r^k})}{(1 - \frac{1}{r^{kN}})} \quad (3)$$

When comparing Eq.(1) and Eq.(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, 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 [?] 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 each other. 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 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.

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 contagious 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 contagious 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 formation⁴, 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.

2.1 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 each other. 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.

3 Modeling Cyber-insurance

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

⁴ Exogenous: The network formation is given. Endogenous: The structure originates from within the network, i.e. the opposite of exogenous

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.

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.(4).

$$U_{i+1} = \begin{cases} \beta - I_l, & \text{if } i = 0 \\ U_i + \beta - I_l, & \text{if } i > 0 \\ U_i + \beta - I_l + \gamma, & \text{if } i = m - 1 \end{cases} \quad (4)$$

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) has to hold.

$$\begin{aligned} U_i + \beta - I_l + \frac{\gamma}{m-i} &> U_i \\ \beta - I_l + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> I_l \end{aligned} \quad (5)$$

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

$$U_{i+1} = \begin{cases} \beta - I_l - r, & \text{if } i = 0 \\ U_i + \beta - I_l - r, & \text{if } i > 0 \\ U_i + \beta - I_l - r + \gamma, & \text{if } i = m - 1 \end{cases} \quad (6)$$

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

$$\begin{aligned} U_i + \beta - I_l - r + \frac{\gamma}{m-i} &> U_i \\ \beta - I_l - r + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} - r &> I_l \end{aligned} \quad (7)$$

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

$$U_{i+1} = \begin{cases} \beta - r, & \text{if } i = 0 \\ U_i + \beta - r, & \text{if } i > 0 \\ U_i + \beta - r + \gamma, & \text{if } i = m - 1 \end{cases} \quad (8)$$

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

$$\begin{aligned} U_i + \beta - r + \frac{\gamma}{m-i} &> U_i \\ \beta - r + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> r \end{aligned} \quad (9)$$

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.(10), and thus a non-insured node will always connect to an insured node if possible.

$$U_{i+1} = \begin{cases} \beta, & \text{if } i = 0 \\ U_i + \beta, & \text{if } i > 0 \\ U_i + \beta + \gamma, & \text{if } i = m - 1 \end{cases} \quad (10)$$

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) 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 each other this equation has to hold:

$$\beta + \frac{\gamma}{m} > I_l \quad (11)$$

We also want to ensure that insured nodes never establish links with non-insured nodes, from Eq.6 we see that this has to hold:

$$\beta + \frac{\gamma}{m-i} - r < I_l \quad (12)$$

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.(13).

$$\begin{aligned} \beta + \frac{\gamma}{m-(m-1)} - r < I_l \\ \rightarrow \quad \beta + \gamma - r < I_l \end{aligned} \quad (13)$$

To summarize, Eq.(11) and Eq.(13) give the final limitation on the link insurance cost, Eq.(14). If this equation is satisfied, the resulting network will contain a clique of only insured nodes.

$$\beta + \gamma - r < I_l < \beta + \frac{\gamma}{m} \quad (14)$$

For this to even be possible, $\beta + \gamma - r < \beta + \frac{\gamma}{m}$, i.e. Eq.(15) 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.

$$\begin{aligned} \gamma - r &< \frac{\gamma}{m} \\ 1 - \frac{r}{\gamma} &< \frac{1}{m} \\ \rightarrow \quad 1 - \frac{1}{m} &< \frac{r}{\gamma} \end{aligned} \quad (15)$$

It is also useful to know when non-insured nodes connect to each other. This happens when Eq.(8) 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 could 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.(14) 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.(13) to $\beta + \frac{\gamma}{m-(m-j)} - r < I_l$, which gives us Eq.(16). 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 .

$$\beta + \frac{\gamma}{j} - r < I_l \quad (16)$$

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.

4.3 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.(14), and enable all nodes to reach their maximum degree.

As we see in Figure 1, 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

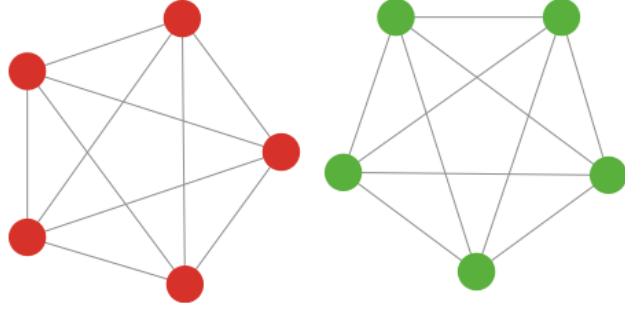


Fig. 1. Two cliques, one consisting of insured agents the other consists of non-insured. All nodes have reached their goal.

that only insured nodes connect to each other, 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.(17).

$$\begin{aligned}
 PoA &= \frac{\text{Sum of payoffs}}{\text{Sum of Socially optimal payoffs}} \\
 PoA &= \frac{5 \times 4 \times (0.9 - 0.7) + 5 \times 4 \times (0.9 - 0.5)}{5 \times 4 \times (0.9 - 0.7) + 5 \times 4 \times (0.9 - 0.5) + 5 \times (2 \times 0.9 - 0.7 - 0.5 + 2 \times 0.2)} \\
 PoA &= \frac{12}{17}
 \end{aligned} \tag{17}$$

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 4.3. In figure 2 the price of anarchy is 0.95, and in figure 3 the price of anarchy is 1, i.e. it has reached the socially optimal outcome.

[Two possible outcomes when insured nodes are willing to take the risk of connecting to non-insured nodes, to receive their bonus.] 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.

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

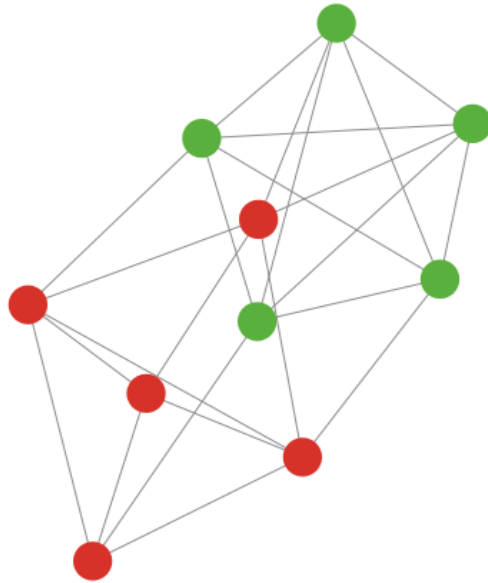


Fig. 2. One non-insured node has connected to two insured nodes.

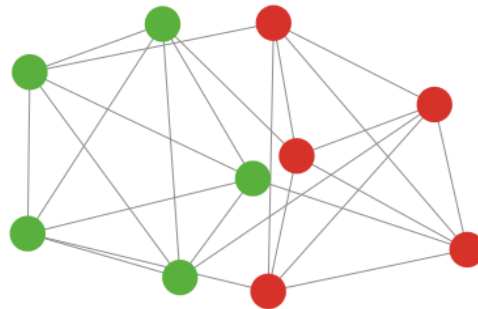


Fig. 3. Every non-insured node is connected to one insured node, this is the optimal outcome with these parameters.

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 [?].

- 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.

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.

5.1 Analysis

The price for adding a new link follows the equation:

$$\frac{I_l}{i+1} \quad (18)$$

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.(19).

$$\frac{I_l}{i+1} < \beta \quad (19)$$

For a link between insured and non-insured nodes to be established, Eq.(20) has to hold.

$$\frac{I_l}{i+1} + r < \beta \quad (20)$$

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. (21), where N_I is the number of insured nodes in the network.

$$(N_I)(\beta - r) < I_l < \beta \quad (21)$$

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.(22), with the one in model 2 (Eq.(??)), 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: $(N_I N_{\bar{I}} \beta + N_I (-\sum_{i=N_I}^{N_{\bar{I}}-1} \frac{I_l}{i}))$, and as long as $(N_I N_{\bar{I}} \beta > N_I (-\sum_{i=N_I}^{N_{\bar{I}}-1} \frac{I_l}{i}))$ 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.

$$U_{total} = (N_I(N_I-1)\beta - N_I \sum_{i=1}^{N_I-1} \frac{I_l}{i}) + (N_{\bar{I}}(N_{\bar{I}}-1)(\beta-r)) + (N_I N_{\bar{I}} \beta + N_I (-\sum_{i=N_I}^{N_{\bar{I}}-1} \frac{I_l}{i})) \quad (22)$$

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.(23).

$$U_{i+1} = \begin{cases} \beta - I_l, & \text{if } i = 0 \\ U_i + \beta - \frac{I_l}{i+1}, & \text{if } i > 0 \\ U_i + \beta - \frac{I_l}{i+1} + \gamma, & \text{if } i = m - 1 \end{cases} \quad (23)$$

For insured nodes to connect to eachother Eq.(24) has to hold.

$$\begin{aligned} U_i + \beta - \frac{I_l}{i+1} + \frac{\gamma}{m-i} &> U_i \\ \beta - \frac{I_l}{i+1} + \frac{\gamma}{m-i} &> 0 \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> \frac{I_l}{i+1} \end{aligned} \quad (24)$$

When insured nodes are considering to connect to non-insured nodes, their payoff functions are as shown in Eq. (25).

$$U_{i+1} = \begin{cases} \beta - I_l - r, & \text{if } i = 0 \\ U_i + \beta - \frac{I_l}{i+1} - r, & \text{if } i > 0 \\ U_i + \beta - \frac{I_l}{i+1} - r + \gamma, & \text{if } i = m - 1 \end{cases} \quad (25)$$

For this to happen Eq.(26) has to hold.

$$\begin{aligned} U_i + \beta - \frac{I_l}{i+1} + \frac{\gamma}{m-i} - r &> U_i \\ \rightarrow \quad \beta + \frac{\gamma}{m-i} &> r + \frac{I_l}{i+1} \end{aligned} \quad (26)$$

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.(11) 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.(27)

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

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.(27).

$$m(\beta + \gamma - r) < I_l < \beta + \frac{\gamma}{m} \quad (27)$$

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.

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 [?] and use a network formation game found in [?] 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: $0 < \beta < 1$.

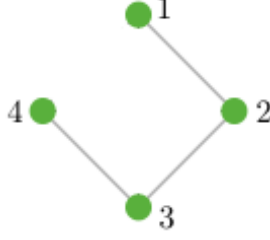


Fig. 4. Four nodes interconnected with each other.

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

$$\sum_{j \neq i} \beta_{ij}^{d(ij)} - \sum_{j:ij \in g} c_{ij}, \quad (28)$$

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 [?], 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.

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.

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.

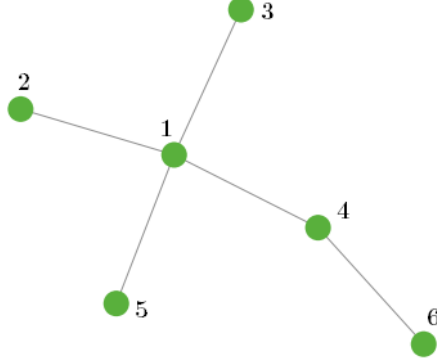


Fig. 5. The resulting network after a simulation with the parameters $\beta = 0.9$, $I_l = 0.5$.

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.(29).

$$\begin{aligned}
 U_1 &= 4\beta + \beta^2 - 4c \\
 U_2 &= U_3 = U_5 = \beta + 3\beta^2 + \beta^3 - c \\
 U_4 &= 2\beta + 3\beta^2 - 2c \\
 U_6 &= \beta + \beta^2 + 3\beta^3 - c
 \end{aligned} \tag{29}$$

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 [?] 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 6. 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.

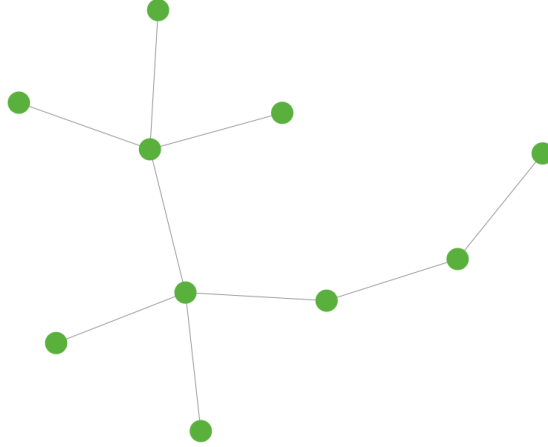


Fig. 6. The resulting network after a simulation with the parameters described earlier and 10 nodes.

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

$$\beta - \beta^2 < \frac{i_l}{i + 1} < \beta \quad (30)$$

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.(31). The critical degree occurs when a node's optimal strategy changes from relaying on indirect connections to connecting to every node.

$$\frac{I_l}{g} < \beta - \beta^2 \quad (31)$$

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.(32). γ is the power law parameter, as described in Chapter 4.

$$P(g) = g^{-\gamma} \quad (32)$$

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 connectivity⁵, 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:

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 7 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

⁵ A node with high degree implies a node with high connectivity.

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.(28)), 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 too 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.(30) 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 7, 8 and 10. As we see from Figure 7, 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.

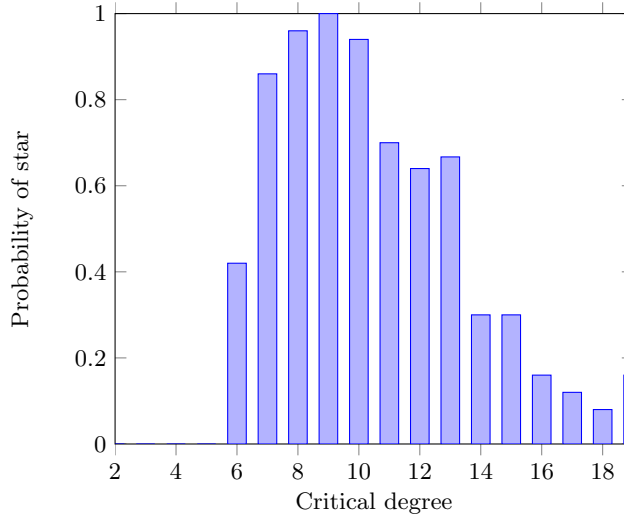


Fig. 7. Shows the probability of the network ending up in a star, given different critical degrees.

From Figure 8 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.

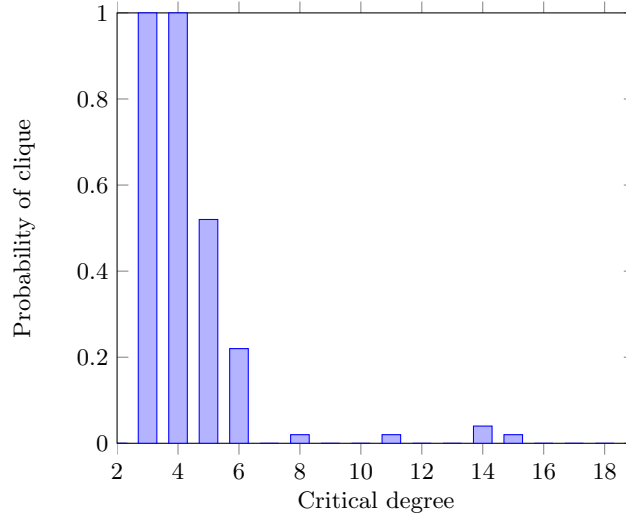


Fig. 8. Shows the probability of the network ending up in a clique, given different critical degrees.

An interesting comparison can be made between the emergence of a star versus a clique. Figure 9 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 7 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 10, 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

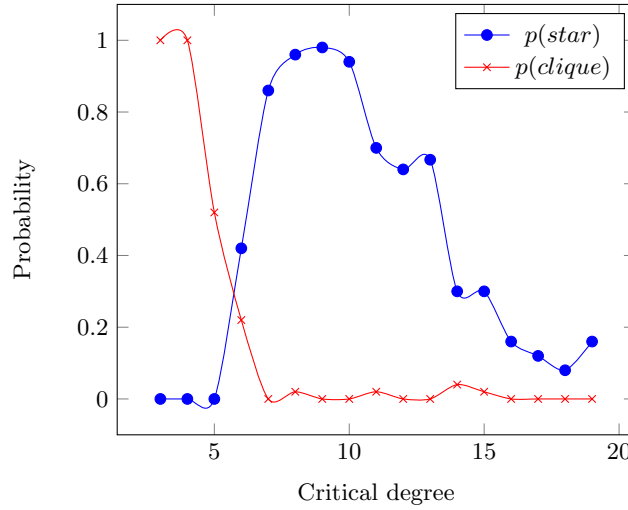


Fig. 9. Shows the comparison between the probability of the network ending up in a star (blue) or clique (red), given different critical degrees.

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 11.

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.

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. Two different outcomes of the simulations where the critical degree is low

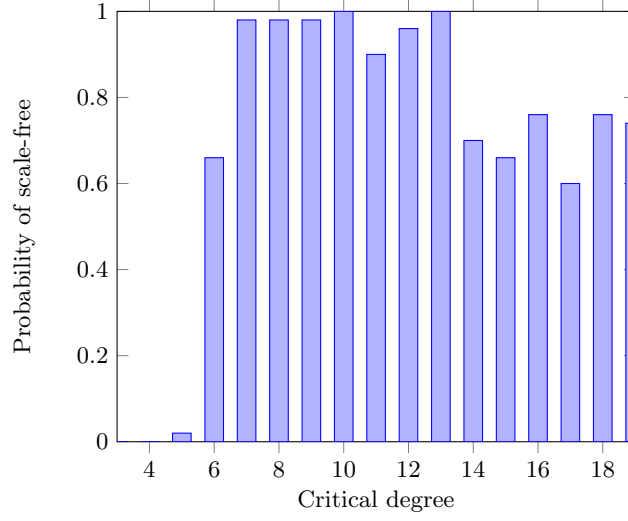


Fig. 10. Shows the probability of the network ending up in a scale-free structure, given different critical-free.

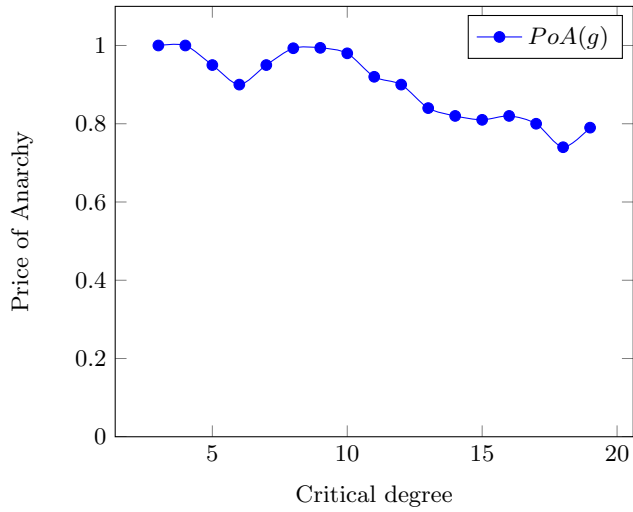


Fig. 11. Shows the price of anarchy as a function of critical degree

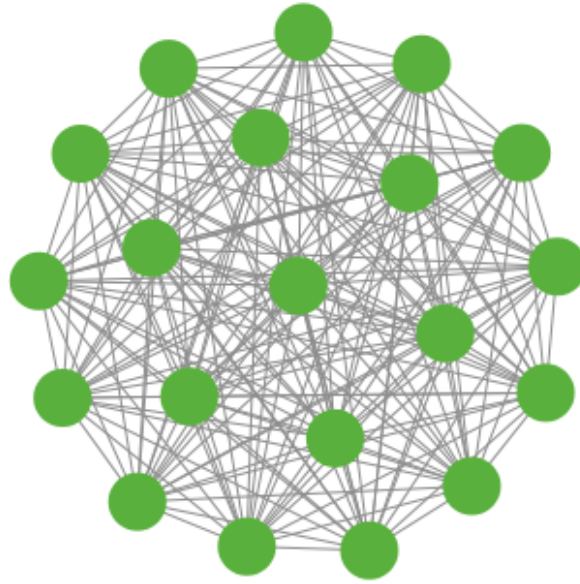


Fig. 12. A clique consisting of twenty nodes.

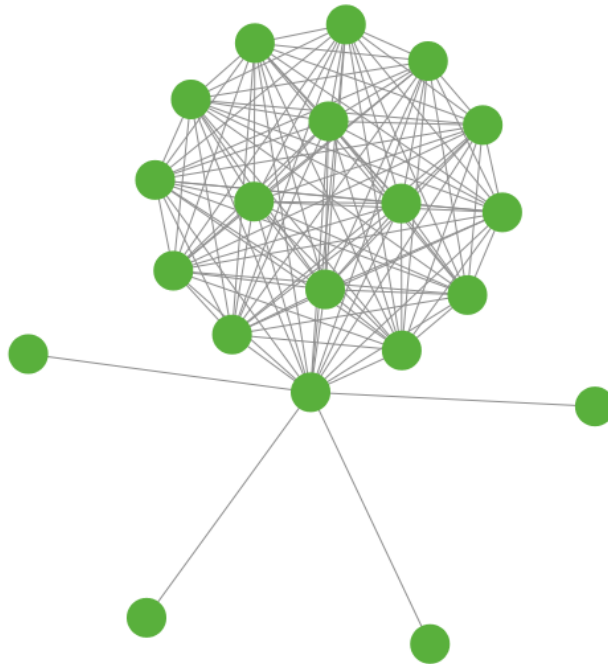


Fig. 13. A network with high average node degree, but not a clique.

In Figure 6.2 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 6.2 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.

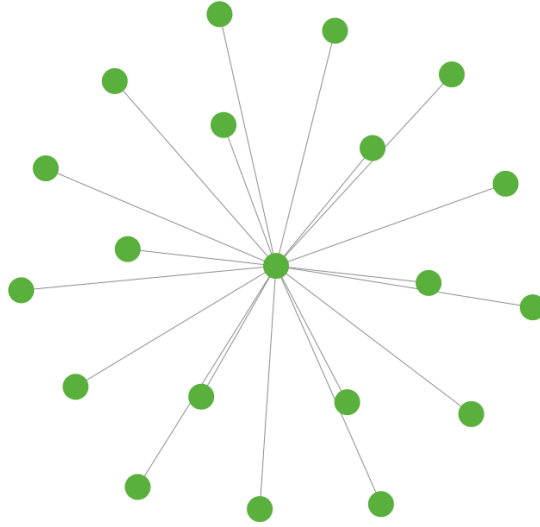


Fig. 14. A star consisting of twenty nodes

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

7 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

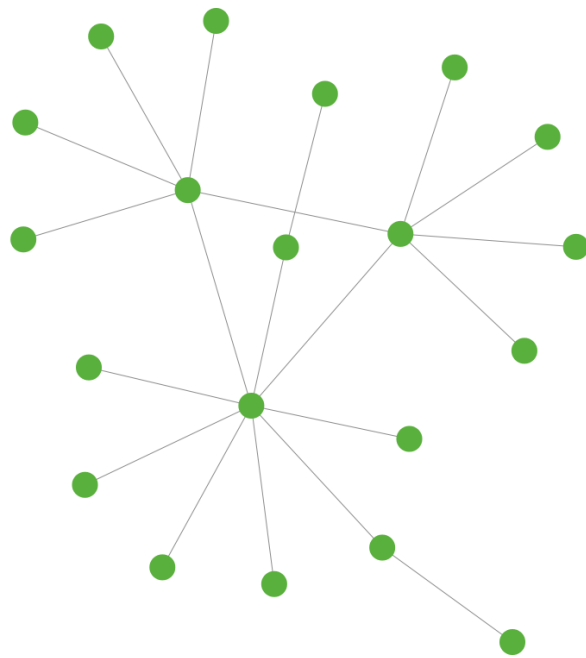


Fig. 15. A scalefree network with twenty nodes, where three nodes account for most of the connectivity.

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, possibly 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 ??.

In model 2a 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 nodes and another of only non-insured nodes. 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 pairwise stable, since 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. And since the resulting network consists of one or two cliques, it is not possible to add any more links. 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. When we included both bonus and discount, the calculation of price of anarchy became too complex. However, we see that the incentive for establishing links has increased, and thus the insurer has to set a higher price to compensate for this, and therefore the potential price of anarchy is even higher 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, by 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 to map 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. In addition, the decision to use additive risk were taken due to the simplicity of the function and the fact that we do not know how a real-world risk distribution actually looks like. By introducing a complex risk function, we would only have distorted the goal of our models. i.e. suggestions for improving our models is to introduce more realistic payoff functions.

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

8 Conclusion

So far, cyber-insurance has failed to reach its promising potential, and many have failed to establish a sustainable cyber-insurance market. We believe that cyber-insurance is an essential part of the internet economy, 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 structure of insurance 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 viable and better market.

References

- [BO1] Bolot, Jean and Lelarge, Marc: Cyber insurance as an incentive for Internet security Managing information risk and the economics of security (2008) 269–290
- [BO2] Lelarge, Mark, and Jean Bolot: Economic incentives to increase security in the internet: The case for insurance. INFOCOM (2009), IEEE.
- [BO3] Bolot, Jean C and Lelarge, Marc A new perspective on internet security using insurance INFOCOM The 27th Conference on Computer Communications. IEEE (2008) 1948–1956
- [RA1] Pal, Ranjan and Hui, Pan On Differentiating Cyber-Insurance Contracts A Topological Perspective Internet Management Conference (2013), IEEE.
- [PO1] Ponemon Institute Second Annual Cost of Cyber Crime Study, Benchmark Study of U.S: Companies Ponemon Institute (2011)
- [ST1] Evolving cyber cover http://www.strategic-risk.eu/Journals/2012/02/22/i/j/w/RiskFinancing_Mar12.pdf Accessed: 31/01/2013
- [GN1] Graeme Newman Cyber liability in Europe: What insurers should knowL CFC Underwriter <http://www.cfcunderwriting.com/media/news-articles/european-cyber.aspx> Accessed: 14/02/2013
- [BS1] Böhme, R. and Schwartz, G. Modeling cyber-insurance: Towards a unifying framework Proceedings of GameSec (2010)
- [CE1] Clarke, F., Ekeland, I.: Nonlinear oscillations and boundary-value problems for Hamiltonian systems. Arch. Rat. Mech. Anal. **78** (1982) 315–333
- [CE2] Clarke, F., Ekeland, I.: Solutions périodiques, du période donnée, des équations hamiltoniennes. Note CRAS Paris **287** (1978) 1013–1015
- [MT1] Michalek, R., Tarantello, G.: Subharmonic solutions with prescribed minimal period for nonautonomous Hamiltonian systems. J. Diff. Eq. **72** (1988) 28–55

- [Ta1] Tarantello, G.: Subharmonic solutions for Hamiltonian systems via a \mathbb{Z}_p pseudoindex theory. *Annali di Matematica Pura* (to appear)
- [Ra1] Rabinowitz, P.: On subharmonic solutions of a Hamiltonian system. *Comm. Pure Appl. Math.* **33** (1980) 609–633