Scale-Free Networks

The Impact of Fat Tailed Degree Distribution on Diffusion and Communication Processes

The Authors

Oliver Hein
Michael Schwind
Wolfgang König

Dipl.-Inform. Oliver Hein
Dipl.-Wirtsch.-Ing. Michael Schwind
Prof. Dr. Wolfgang König
Johann Wolfgang Goethe-Universität
Institut für Wirtschaftsinformatik
Mertonstr. 17
60325 Frankfurt am Main
{ohein, schwind, koenig}@is-frankfurt.de

Executive Summary

The study of network topologies provides interesting insights into the way the principles on which the construction of connected systems are based influence diffusion dynamics and communication processes in many socio-technical systems.

- Empirical research has shown that there are principles for the construction of social networks and their technical derivatives, like e-mail networks, the Internet, publication co-authoring, or business collaboration.
- Such real world networks attach new members over time and the mode of attachment prefers existing members that are already well connected. This principle is called "preferential attachment" and leads to the emergence of "scale-free" networks.
- Scale-free networks seem to be a better fit for the description of real world networks than the random networks used so far. Their behavior in terms of diffusion and communication processes is fundamentally different from that of random networks.
- To illustrate the value of scale-free networks for applications in information systems research, examples will be given to illustrate their usefulness for real world network modeling. A communication network of security traders will show what impact network topology has on the dynamics of complex socio-technical systems.

Stichworte: Random Networks, Scale-Free Networks, Communication Networks, Socio-Technical Networks, Diffusion

1 Introduction

The widespread presence of networked systems in technical applications as well as in the business world makes network research useful for future applications in information systems (IS) research. Research into networks means not just analyzing the topology of networks, but also examining the dynamics of the processes that take place in them. Research into diffusion processes in networks has been shown to provide especially fruitful insights into several kinds of IS domains such as e.g. information technology (IT) robustness in system failure situations, denial-of-services attacks, and computer virus spreading.

For decades network research was dominated by random networks, also called Erdös and Renyi networks or just ER-networks [ErRe59]. ER-networks are random in the sense that the links between the nodes of a network are created with the help of a random variable. Recently, following the introduction of what were called "scale-free" networks by Barabasi and Albert [BaAl99], the field has received growing attention from scientific research. Scale-free networks seem to match real world applications much better than ER-network models [Bara03]. The term *scale-free* refers to the distribution principle of how many links there are per node.

This article is intended to provide an introduction into the metrics of modern network analysis. The basic network measures like the degree distribution, the clustering coefficient and the average path length will be presented to enable a better understanding of the two network models "ER" and "scale-free". Diffusion processes in networks depend very much on network topology. Both network models will be examined in terms of their different diffusion behavior. Finally, the potential impact for real-life applications will be shown by the introduction of a new application of scale-free networks: The simulation of the implications of network topology for the price building process in a security trader network that relies on a communication network.

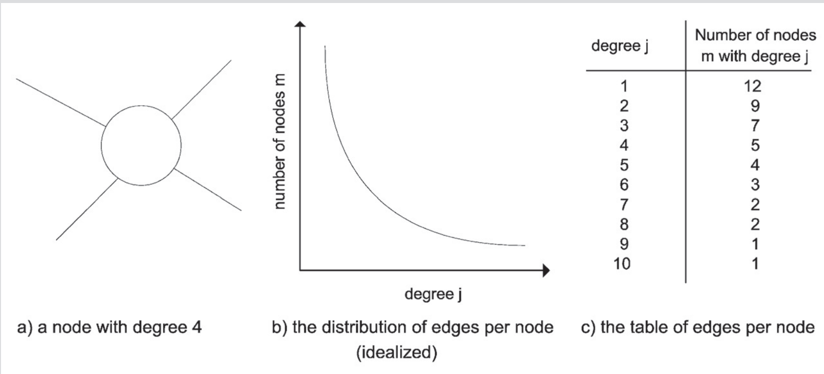

Figure 1 The Basic Concept of a Network Degree Distribution

The article is intended to emphasize the importance of choosing the right network model for the right purpose. The fundamental differences between ER- and scale-free networks will become apparent in the sample application presented.

2 Network Models

This paragraph presents ER- and scale-free networks as well as their main characteristics and the processes by which they are created. ER-networks, because of their diminishing importance for modeling real world networks, are only briefly introduced to show their important differences from scale-free networks. The basics of network analysis are discussed first to enable a better understanding.

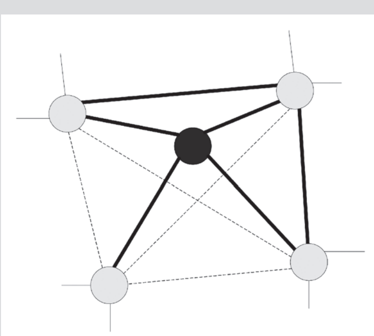

Figure 2 The Bold Node is Connected with Bold Edges to its Four Direct Neighbors

2.1 Basics of Network Analyses

In an undirected network graph, the *degree* of a node is defined as the number of edges it possesses. Figure 1a shows an example of a node with degree 4. If all the nodes of a network that share the same degree are counted and the results are sorted by increasing degree a function like that in Figure 1c is most likely to occur. There are no nodes with degree zero, because these nodes would not be connected to the network. The network in figure 1c exhibits a majority of nodes with degree one to three and rather fewer nodes with a degree greater than seven. Chart 1b is idealized; the degree distribution of networks does not have to be continuous.

If one node of the network in figure 1c were randomly chosen, the probability of obtaining a node with only one to three edges is much higher than that of obtaining a node with a degree higher than seven. Therefore it is possible to define a *probability distribution function* $P(j, v, N)$ that returns the probability that node v has j edges within network N [DoMe03]. The concept of the degree distribution of a network has important consequences for the properties of a network and will be frequently used in the paragraphs that follow. It will be seen that deviation from the normal distribution will lead to new results in terms of diffusion within networks.

C*lustering* within networks is another important factor when analyzing networks. It is interesting to know how well nodes are interconnected within a specified area of the network. A similar task is the question how many of a person's friends also know each other. Figure 2 presents an example of a person (bold node with four bold edges) with four friends. Out of six possible friendship relationships (dashed lines) between the four friends (grey circles), only two friendships exist (bold edges).

Using the ratio between existing and possible relations, a *clustering coefficient* may be computed. If a node has z nearest neighbors, a maximum of $z(z-1)/2$ edges is possible between them. Watts and Strogatz defined the *clustering coefficient* for node v as the ratio of the number i of existing nodes to the possible number of edges between the direct neighbors of node v [WaSt98]:

$$C^v = \frac{2i}{z(z-1)} . \quad (1)$$

As for the degree distribution, the clustering coefficient plays an important role when analyzing networks in terms of important properties like diffusion, which will be discussed in subsequent paragraphs.

The *path length* between two nodes of a network is defined as the number of edges between them. The *minimal path length* is the shortest path between two nodes. The *average path length* is the average of all the minimal path lengths between all pairs of nodes in a network.

2.2 ER-Networks

Since the seminal paper of Erdös and Renyi in 1959, the *random* network theory dominated scientific thinking [Bara03]. Real world networks had been thought to be too complex to understand and therefore held to be random. In the absence of other well-understood network models, random networks were widely used when modeling networks.

The process of creating an ER-network depends on probability p. For a network with n nodes each possible pair of distinct nodes are connected with an edge with probability p.

An ER-network has the property that the majority of nodes have a degree that is close to the average degree of the overall network and that there is not much deviation from the average below and above it. It has been shown that the distribution of links follows a Poisson distribution (figure 3). Knowing the degree distribution, the average path length, and the clustering coefficient of ER-networks, it is feasible to analyze their different behavior compared to scale-free networks.

ER-networks are still used in some types of models. The following chapters will show that ER-networks may not be applicable for every purpose, because they lack some of the properties of other classes of networks.

2.3 Scale-Free Networks

When Albert et al. [AlJB99] started to map the internet in 1999, they did not know that they were about to influence network research in a sustained way. Because of the diverse interests of every internet user and the gigantic number of web pages, the linkages between web pages were thought to be randomly linked as a random network. The results of their study have disagreed with this expectation in a surprising way. Only a few pages have the majority of links, whereas most pages are only very sparsely connected. More than 80% of all pages visited have 4 links or less, only 0.01% of the pages are linked to more than 1,000 other pages (some up to two million).

If the *nodes* with one, two, three, etc. connections are counted, and the numbers are plotted into a chart as in the middle of figure 4, the distribution of edges becomes visible. There are many nodes with only a few links and a few nodes with a large number of links. If the distribution of edges is plotted in a logarithmic chart, as in figure 4 right, the *power-law* nature of the distribution appears in a straight curve with a slope of $-\gamma$ from equation 2. The difference from ER-networks now becomes evident (figure 4 left). The probability distribution function $P(j)$ of the degree j of scale-free networks is described by:

$$P(j) \approx j^{-\gamma} \qquad (2)$$

with $j > 0$ and $\gamma > 0$, with γ called the *scale-free exponent.* The term *scale-free* is a mathematical expression which stands for power-law distributions as in equation 2. The power-law distributions belong to the class of *leptokurtic* or *"fat-tailed"* distributions. They deviate from the Poisson distribution in two ways: they have higher peaks and "fat" tails. Figure 5 shows a power-law distribution in relation to a Poisson distribution. The peak on the left side of the chart is higher and the tail on the right side is thicker, therefore called "fat tail". Compared to the Poisson distribution, the power-law distribution shows the greater probability of nodes with a number of links close to the average of all links (see the peak with 2 links in the example of figure 5), and a greater probability of nodes with many links (4 or more links in the tail in figure 5).

Since the important observation of Albert et al. [AlJB99], scientists have empirically analyzed many real world networks with the scale-free property. The physical structure of the internet (router level, domain level, web links), social networks like e-mail networks, the structure of software modules, and many more examples show surprising scale-free structures (table 1).

γ is the degree exponent and ranges between 1.8 and 2.5. The factor is derived by counting the number of edges per node and plotting the results by increasing degree within a chart with log-log-scale. The slope of the resulting straight line is γ (figure 4 right). What is the reason for the similar γ in table 1 and how do the different networks compare? First of all they are all *evolving* networks, which means that they develop over time. ER-networks in contrast are created with the number of nodes already fixed. Second, they are created by *preferential attachment*, which means that newly introduced nodes connect with a higher probability to existing nodes with many links. If the evolution of networks is the appropriate explanation for the distinct scale-free property, then the growth path of networks deserves particular attention. Starting with a small number of nodes, like the internet in 1990, new nodes are added over time. Older nodes therefore have a greater probability of acquiring new links. Furthermore, already well-connected nodes are much easier to find and will in consequence acquire more new links. Coupled with the expanding nature of many networks this might explain the occurrence of hubs, which hold a much higher number of links than the average node. Figure 6 presents an idea of how scale-free networks may evolve over time for the first 8 steps. At each step, a new node is added with two new edges. The new edges are connected to existing nodes with a probability that corresponds to the number of edges of the existing nodes. The "richest" nodes will receive most of the newly added edges.

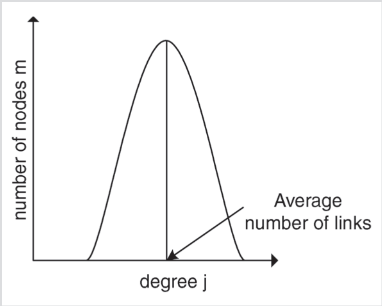

Figure 3 The Poisson Degree Distribution of the Number of Edges of ER-Networks with the Characteristic Bell Shape around the Average Number of Edges per Node [DoMe03]

The generalized creation process of a scale-free-network is defined by Barabasi and Albert [BaBo03]: Start with a small number of nodes m_0. During each step, insert a new node with connections to $m \leq m_0$ existing nodes. The probability Π_v of node v being connected to a new node depends on its degree j_v. The higher its degree, the greater the probability that it will receive new connections ("the rich get richer"):

$$\Pi_v = j_v \Big/ \sum_k j_k \,. \qquad (3)$$

The process of preferential attachment, as defined by Barabasi and Albert, is *linear*. A

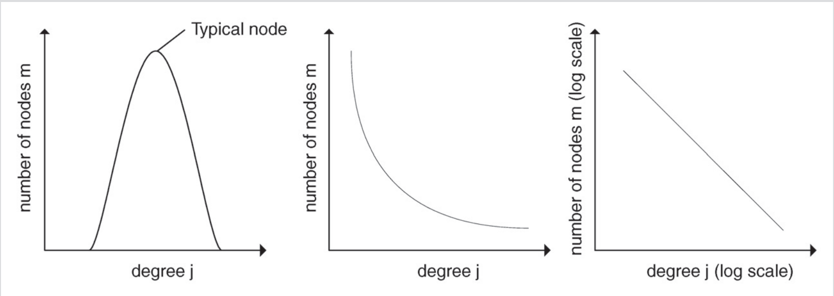

Figure 4 Poisson Distribution of Links for ER-Networks (left) and Scale-free Networks (middle, right in log-log scale with the characteristic straight slope of a power law) [BaBo03]

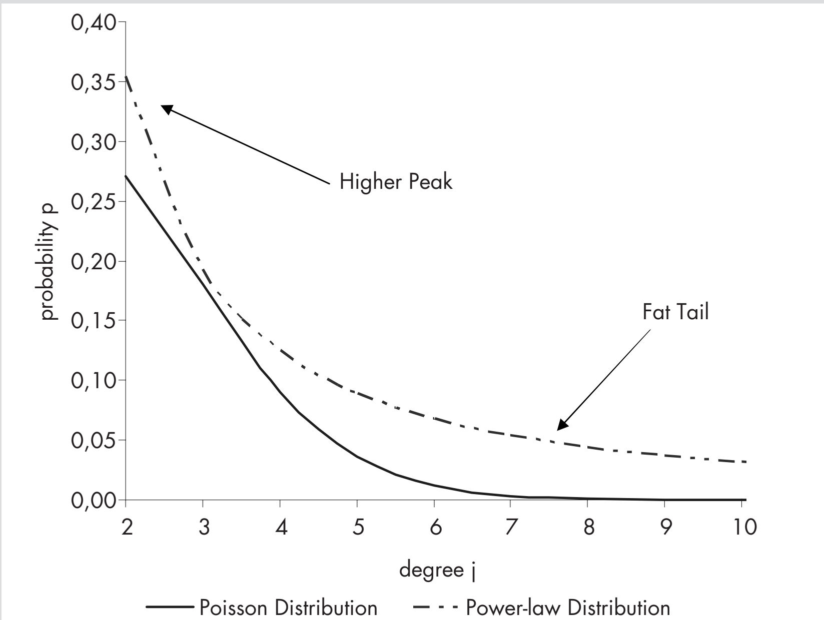

Figure 5 Poisson Distribution and Power-Law Distribution

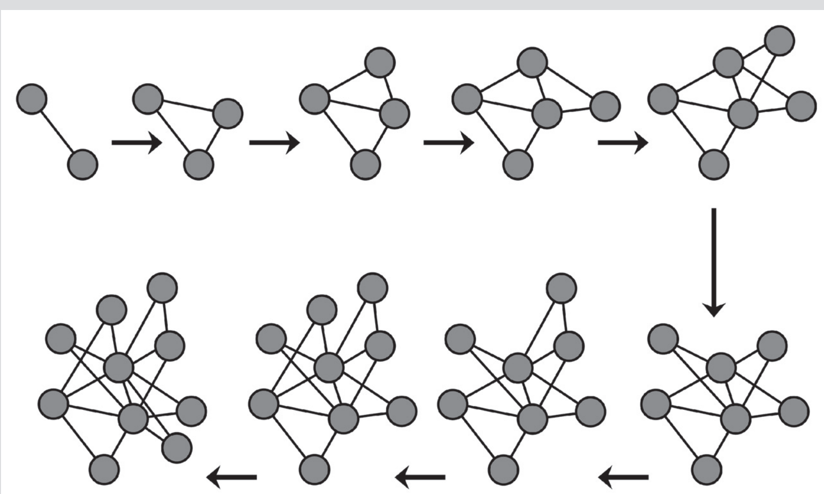

Figure 6 The Emergence of a Scale-Free Network as a Result of the Preferential Attachment of Newly Added Edges and Nodes to Already Existing Nodes [BaBo03]

new node is twice as likely to link to an existing node that has twice as many links as its neighbor. If the probability Π_v is increased even more the network evolves to a star topology with one central hub [BaBo03].

To return to the examples of table 1, it may now be possible to explain why these networks are similar in terms of network measures like the scale-free exponent γ. When scale-free networks are constructed, as Barabasi and Albert suggested, the resulting scale-free exponent γ is, as in the examples in table 1, between 1.8 and 2.5, depending on how many links m a new node will link to existing nodes. The high correlation of empirically examined real world networks with networks generated by the Barabasi-Albert-algorithm indicates that preferential attachment does indeed seem to be a valid explanation of how this network structure evolves.

In the following section, the examples from table 1 will be discussed to get a better understanding of how scale-free networks may arise and be analyzed. It may help to enable the reader to use the right network model for the right application. The scale-free exponent γ makes the examples from table 1 comparable. They all are similar in the sense that the probability distribution of their links is almost the same. However, only almost, because e-mail networks with a γ of 1.81 and software with a γ of 2.5 show the greatest difference among the examples in table 1. How can this difference be explained? The scale-free exponent γ within $P(j) \approx j^{-\gamma}$ determines the probability $P(j)$ of the occurrence of nodes with degree j within the network. With a rising degree j the probability of the existence of nodes with degree j within the network decreases. In other words: the steeper the slope, the more concentrated the analyzed network is. It means that a higher γ leads to a lower probability of nodes with many links. Carried forward to the example of e-mail and software networks, e-mail networks with a γ of 1.81 have a higher probability of nodes with many links, than is the case for software networks with a γ of 2.5. Therefore, it may be concluded that e-mail networks possess a higher quantity of "super-nodes" with many links compared to software networks.

Let us apply these findings to the five networks shown in table 1: The structure of the Internet can be studied on different levels (table 1, line 1–3). The physical level is composed of the servers, switches and routers. If a router is taken as a node and its connections to other routers as links, a complex network arises and can be studied for its network properties. A scale-free exponent γ of 2.1 (table 1, line 2) for an observed router network of 228.298 routers shows the existence of some major router hubs and many small routers within their network periphery. The linkage between web pages – WWW – is another Internet structure that forms a large network (table 1, line 3). For the WWW it is possible to analyze the in-degree, how many links lead to a web page, and the out-degree, how many links lead out of the web page to other web pages. If the web pages are taken as nodes and the links as edges between them another large network is created. As in the example with the routers, the distribution of links between the web pages of the WWW seems to have the scale-free property.

Yet another internet network is the system of Internet domains and sub-domains. They may be seen as another structural level of the Internet (table 1, line 1). Domain

Table 1 A Choice of Some Real World Networks within the IS Domain that Share the Scale-Free Property [WaCh03]

Network	Size n	Clustering coefficient C	Degree exponent γ
Internet, domain level [VaPV02]	32711	0.24	2.1
Internet, router level [VaPV02]	228298	0.03	2.1
WWW [AlJB99]	153127	0.11	$\gamma_{\text{in}} = 2.1$ $\gamma_{\text{out}} = 2.45$
E-mail [EbMB02]	56969	0.03	1.81
Software [VaFS02]	1376	0.06	2.5

and sub-domain names inter-linked according to their hierarchy describe another Internet related scale-free network. The three levels of the Internet mentioned are three distinct networks, each one created for a different purpose. From the physical network of routers and cables, to the organization of parts of the Internet in the form of domains and sub-domains, to the structure of web pages, which are linked according to their content, all three networks share the same scale-free exponent γ of 2.1 (in the case of the WWW $\gamma_{\text{in}} = 2.1$).

An e-mail network (table 1, line 4) is less a technical network and more a social network, because it represents the social contacts of users to other users. If an e-mail user is seen as a node, the links to other nodes are created by tracing the e-mails sent and received between the users. If a user A sends an e-mail to user B, an edge between node A and B is drawn. Even though an e-mail network represents an area of human communication habits, it seems to have a similar structure, seen from the standpoint of the degree distribution, to a technical infrastructure like the Internet. With a scale-free exponent γ of 1.81 it may be assumed that few people use e-mail intensively with contacts to many other people, while most e-mail users only write e-mails to a small group of receivers.

Valverde et al. [VaFS02] analyzed the class diagram of the Java Development Framework (JDK 1.2) (table 1, line 5). In the case of software networks, we can analyze not only software structures that have evolved over time, but also those which were created intentionally for the goal of more efficient use of individual software modules. The modular structure of modern software may be represented by class diagrams. Software modules are taken as the nodes of a directed graph with links according to their interaction. A lower degree coefficient in the scale-free structures of the class diagrams, as found by Valverde et al. [VaFS02], could lead to the assumption that a lower γ indicates an improved software design process. The network analysis may therefore provide a tool for verifying the outcome of the efficient reuse of software modules.

3 Diffusion in Networks

The spread of epidemics within networks is an interesting subject that everyone has occasionally been confronted with in real life. The computer worm Melissa and many other examples of epidemics have something in common. They are infections that disperse through complex networks by spreading from one individual to their direct neighbors. The prediction of epidemics is especially important to prevent damage to life as well as to material goods, and depends on factors like how fast a neighbor becomes infected, how long they remain infectious and how the network the individual is a part of is structured. The diffusion of viruses or information through complex networks may be studied with the help of simulation models, to produce predictions of the course of epidemics [Keel99].

In the following three sections, a basic model for diffusion alongside its application on ER- and scale-free networks will be presented.

3.1 SIS / SIR-Model

The mathematical framework for the study of epidemics goes back to Kermack and McKendrick 1927 [KeMc27]. They invented the SIR-model, which stands for *susceptible-infected-removed*. An individual may be in one of the three states. If a susceptible person meets an infected person, for example the direct neighbor within a network, his or her status changes from susceptible to infected with probability δ. With probability ν a person may change his or her status from infected to removed, which means the person is either immune or has died. The SIS-model, a variant of the SIR-model, changes the individual back to the susceptible status and another infection may recur. How an infection spreads depends on $\lambda = \delta/\nu$ and the structure of the network. The course of an infection can be simulated using an iterative network model. The intended number of time steps stands for the iterations that are used; δ and ν are reviewed per time step for every node within the network. The status of every node is changed over the course of the intended number of time steps. For every time step the number of infected nodes is recorded. An infection path will arise that is characteristic for $\lambda = \delta/\nu$ and the network topology used.

Within ER-networks there is a constant $\lambda_c > 0$. If $\lambda_c > \delta/\nu$ the epidemic dies out quickly. With $\lambda_c < \delta/\nu$ the epidemic spreads over the whole network [AlBa02; Leve02; LlMa01; Lope04]. Pastor-Satorras and Vespignani found that within scale-free networks the constant λ_c is close to zero [MaLo01; PaVe01]. This means that within scale-free networks there is no resistance to infections, and that an infection once acquired quickly develops into an epidemic. The diffusion behavior of epidemics therefore depends on the infection and recovery rate on the one hand and the topology of the network on the other. In the following sections, the two network types will be presented in the light of the SIS-model.

3.2 Diffusion and ER-Networks

The degree distribution of ER-networks peak at an average value $\langle j \rangle$ and decay exponentially for $j \gg \langle j \rangle$ and $j \ll \langle j \rangle$ [MoPV02]. It has been shown that under these assumptions for the epidemic threshold of ER-networks there is a critical non-zero value λ_c:

$$\lambda_c = 1/\langle j \rangle \,. \tag{4}$$

If λ is below λ_c the infection dies, if λ is above λ_c the infection spreads and becomes persistent [PaVe02]. If the infection and recovery rates of the network members are known, the number of nodes m that may be infected within ER-networks is shown in figure 7. Until λ_c is reached the infection is not able to remain in the network, it dies out more quickly than it is able to infect new members of the network. If λ is greater than λ_c the infection can remain in the network and infects broader regions of the system. The higher the λ chosen, the more regions of the network may be reached during the course of an infection.

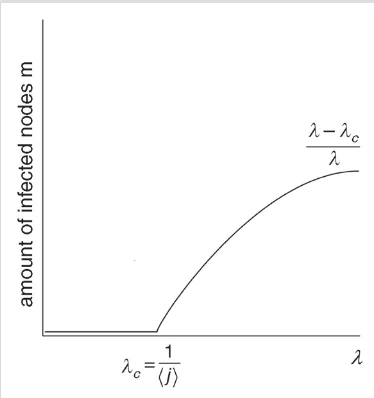

Figure 7 Spreading of an Infection within ER-Networks Depending on λ

3.3 Diffusion and Scale-Free Networks

For scale-free networks it has been shown that there is no critical value $\lambda_c > 0$ below which an infection will not spread through the system [DoMe03; MaLo01]. Figure 8 shows the long-run prevalence of an infection in a scale-free network. The difference from ER-networks in figure 8 is evident. There is no epidemic threshold, and the infection starts to conquer distant parts of the system from the outset.

If the scale-free exponent γ in the degree distribution $P(j) = j^{-\gamma}$ of the scale-free system is less or equal to 3, as is the case for many real life networks (table 1), there is no epidemic threshold, which means that the network is extremely sensitive to the spread of infections [DoMe03]. That finding explains why the WWW is so vulnerable to computer viruses.

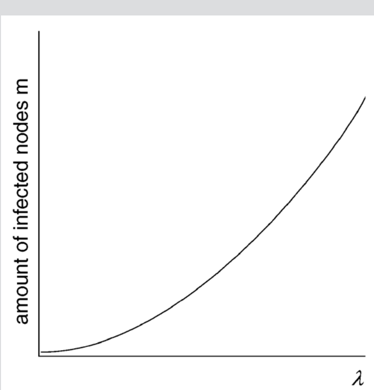

Figure 8 Spreading of an Infection within Scale-Free Networks Depending on λ.

4 A Scale-Free Communication Network of Security Traders

Network theory has made significant progress during the last seven years, as a result of the publications of Watts, Strogatz [WaSt98] and Barabasi, Albert [BaAl99]. The most important progress is that network theory was taken out of the ivory towers of purely mathematical interest and brought into everyday life by fitting theoretical expectations to empirical findings. What is surprising is how powerful the analysis of real-world networks using to the scale-free paradigm is, in terms of explanation and prediction. It seems that they are able to capture an important part of natural mechanisms that lead to higher efficiency and robustness. Dorogovtsev, Mendes [DoMe03] is an appropriate source for a broader review of real-world applications.

In the following an artificial stock market in which the traders communicate by a scale-free communications network is presented. It will be shown that the topology of the communication network influences the price building process.

What is the impact of general human interaction within specific network topologies with regard to the outcome of joint goals? A broad empirical basis is necessary to secure sufficient evidence. A potential testing ground for a network theory of human interaction could be the international financial markets. They deliver a large amount of second-by-second data and may be viewed as the aggregation of human interaction through a price building process. Simulated model data could be tested against time series taken from real markets and could potentially lead to a better understanding of joint interaction within a network environment.

The decision making process of individual security investors is a still undiscovered field. The assumption of perfectly rational investors seems not to hold any more [Shle00]. The influence of factors other than the fundamental value or the technical analysis of a security seems to guide the behavior of investors [Mand04; Shil01]. The herding effect of groups of investors during some market situations, like booms and crashes, seems to play an especially important role [Bane93]. Herding is the result of the exchange of information between two or more investors and the willingness to deviate from rational valuation rules. To what extent does herding influence the decision making process of an individual security trader and what are the consequences for price building at a security exchange?

We want to answer this question, by means of a security market simulation (micro-simulation) of individual security traders connected by a well-defined communication network. In this way, the effects of herding may be measured. The quality of the communication network plays a central role when modeling the real-life exchange of information between investors. We choose a scale-free network as the communication model between agents. Simulation models always need validation, without which the results would be useless. When speaking of stock market simulations, the properties of the price development generated by the simulation model are decisive. Real stock markets have important statistical properties, called *universal stylized facts*. If the daily percentage changes in stocks or currencies are reviewed, the distribution of the changes shows many small and some very large fluctuations (also called leptokurtic or "fat-tailed" distribution of changes). This statistical property can only be achieved by complex systems and is hard to generate. Another stylized fact is the correlation between the volatility of prices and the volume of shares traded. It means that volume rises with more widely fluctuating stock prices. The clustering of volatility, as large fluctuations tend to cluster together, is also experienced when examining real markets [Cont01]. A valid stock market simulation must be able to reproduce all the above-mentioned and other stylized facts of capital markets. This is a difficult task, but is at the moment the only possible way to validate stock market simulations with empirical findings.

The latest models of capital market simulations [AHLP93; CoBo00; LeLS00; Lux00] are already able to reproduce the universal stylized facts of financial markets. They differ in two key points from the model presented. First, they do not use an auction method that is in use on real markets, but mostly a function that maps supply and demand to a price. The authors chose a realistic auction method that is in use at the Frankfurt Stock Exchange, namely the "Kassakurs"-Method [ScPr95]. Second, herding seems to be a major determinant for bubbles and crashes on the stock exchange. This already well studied phenomenon [Bane92] cannot be found in any capital market simulation models known to the authors. Therefore, a communication infrastructure was introduced to enable agents to share information. The resulting diffusion of information through a network with a well-defined topology is analyzed with regard to the properties of the simulated prices on the artificial stock exchange.

Multi-agent models with inter-agent communication are able to model the complex interactions between the market participants within the controlled environment of the computer lab. The structure of the communication network has proved to be a driving force for the price building process within security markets [HeSc05]. Information diffuses like an epidemic through a communication network that has a scale-free structure. Agents adapt their beliefs as a result of the information they get from their direct neighbors within the communication network. If the majority of the neighbors decide for a specific trading strategy, the agent adapts its strategy. If the simulation results are reviewed, it becomes evident that the scale-free structure of the communication network is the cause of extreme fluctuations within the price building process [HeSS05]. The power of opinion-making agents, the hubs of the scale-free network, is the cause of rectified trading strategies within the agent community. Their influence leads to homogenous buy and sell behavior on the part of a majority of the agents and in consequence to an imbalance between buy and sell orders. The majority of agents buys and sells at the same time and causes prices to fall and rise with higher fluctuations. Only when extreme levels are reached does the price revert, and the influence of the direct neighbors weigh less than actual losses an agent experiences. In consequence, the agent returns to its original trading strategy. Figure 9 shows a typical simulation run with 200 agents who communicate through a scale-free network. The price

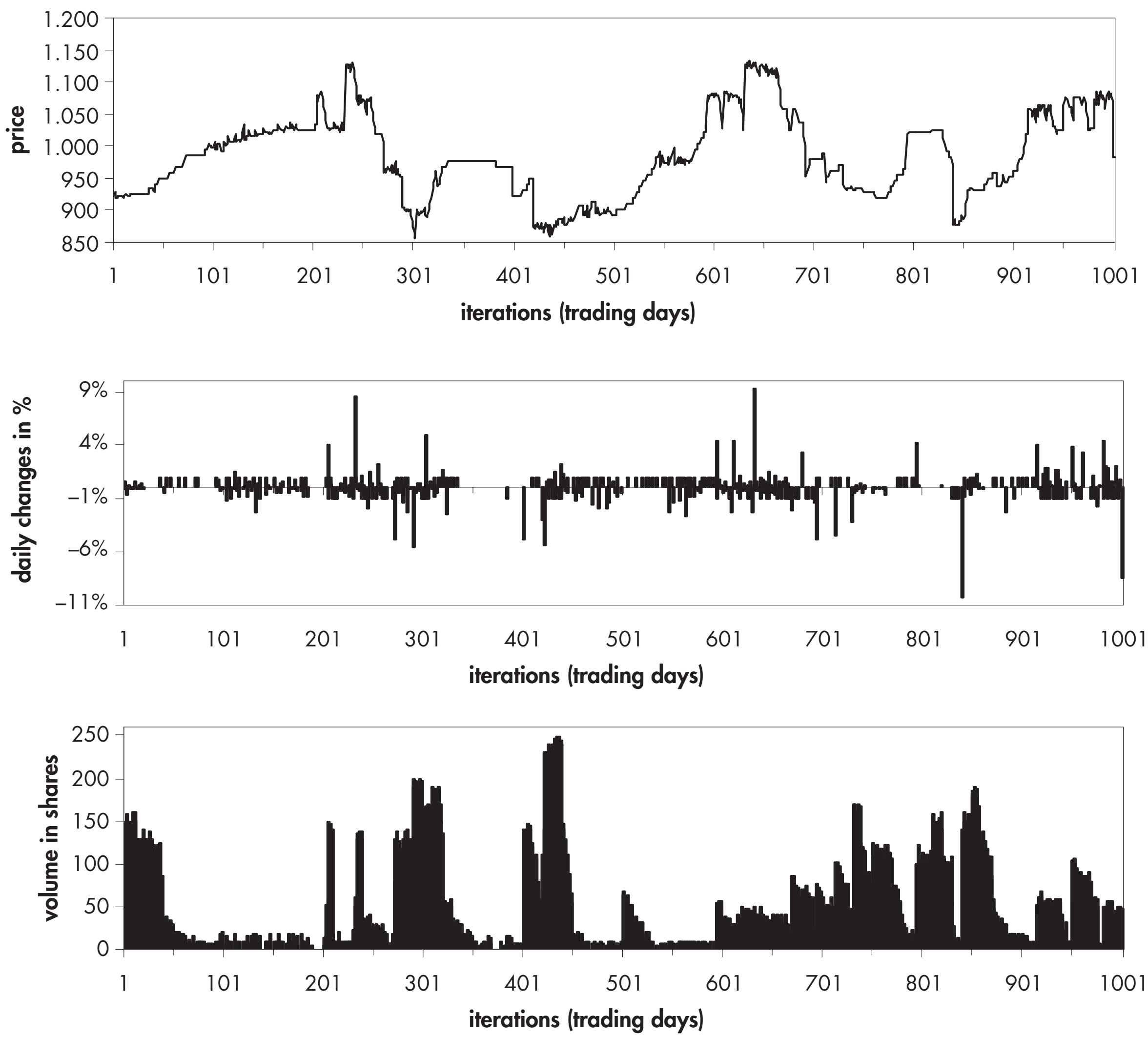

Figure 9 Daily Prices, Daily Price Changes and the Daily Volume of Shares Traded of a Simulation of an Artificial Stock Market with a Scale-Free Communication Network [HeSS05]

Table 2 Statistical Moments of Simulation Runs with ER- and Scale-Free Communication Networks [HeSS05]

	"ER"	"scale-free"	difference in %
Max. daily price change	7.06%	9.30%	+31.72%
Min. daily price change	−5.15%	−10.28%	+99.61%
Standard deviation of daily price changes	0.87%	0.99%	+13.79%
Kurtosis	17.90	29.71	+65.97%
Hill-estimator	1.52	1.48	− 2.63%

chart (top) exhibits the ups and downs of the price of an artificial stock, while the chart of the daily changes in percent (middle) indicates some clustering of daily changes and the volume chart (below) shows some volume-volatility clustering when compared to the chart in the middle. Even visually, some typical phenomena of capital markets can be recognized. The statistical numbers are reviewed later in table 2.

To make certain that the scale-free network causes the price fluctuations, another communication network is needed as a benchmark. In parallel to the scale-free network, an ER-network was used for the simulation runs. The results show that the structure of the ER-network influences the prices on the artificial market in another way. There are no opinion leaders in an ER-network. Every trader has the same amount of influence; the diffusion of opinions through the network is slower and sometimes an opinion diminishes before a majority may be able to adopt a different trading strategy. The result is less volatility and less extreme daily moves. Table 2 presents the difference between an ER- and a scale-free communication network in terms of market price volatility. Some statistical measures are needed for the analysis of the different properties. A kurtosis > 3 signals by definition a deviation from the normal distribution, which indicates a "fat-tailed" distribution which means a higher probability for large price changes, also called *returns*, than in the normal distribution. The Hill estimator examines the tails; lower values stand for heavier tails in the return distribution and therefore a more extreme distribution of returns between small returns and large returns. A Hill estimator value around 1 and 2 indicates explosive volatility after longer periods of quietness, which is commonly experienced on real financial markets. The results are comparable with the behavior of real stock markets [HeSc05]. They indicate that the simulation runs with a scale-free communication network yields higher volatility in daily price changes than with an ER-network. In every volatility measure the results are increased. The simulation results seem to support the intuitive expectation that the diffusion process on scale-free networks leads to a quicker distribution of profitable market opinions. The buying and selling within the model is synchronized and larger groups of agents buy and sell at the same time. The consequence is that there are not enough buy or sell orders that can be matched, so that the price needs to change in greater quantities which leads to higher volatility.

Micro-simulations of artificial stock markets with inter-agent communication networks seem to hint in the direction of the herding effect having a major influence on the price building process. The effect of herding is made visible and reproducible using scale-free networks in a multi-agent simulation.

5 Conclusion

Network topology does matter. As the applications of a scale-free network within a communication structure of security traders presented above show, the topology of the network plays a crucial role for the dynamics of an artificial stock market. The network topology brings a new dimension into the simulation, as not only local decisions, but also the spatial organization of the underlying communication system influences the development of asset returns in the network of traders investigated. The price of an asset in the system depends not only on system time, but also on the relative geographical position of the individual

Abstract

Scale-Free Networks – The Impact of Fat Tailed Degree Distribution on Diffusion and Communication Processes

The study of network topologies provides interesting insights into the way in which the principles on which interconnected systems are constructed influence the dynamics of diffusion and communication processes in many kinds of socio-technical systems. Empirical research has shown that there are principles of construction similar to those of the laws of nature for social networks and their technical derivatives, like E-mail networks, the internet, publication co-authoring, or business collaboration. For decades, the paradigm of a randomly connected network has been used as a model for real world networks, in ignorance of the fact that they are only a poor fit for such networks. Apparently, all the above-mentioned networks share the same building blocks. They attach new members over time and the attachment prefers existing members that are already well connected. This principle of "preferential attachment" leads to interesting properties that have to be taken into consideration when analyzing and designing systems with some kind of network background.

What are called "scale-free" networks seems to be a better fit for the description of real world networks. They use preferential attachment as a construction principle to resample real world networks. Their behavior in terms of diffusion and communication processes is fundamentally different from that of random networks.

To illustrate the potential value of the discovery of scale-free networks for applications in information systems related research, an example will be used in this article to illustrate their usefulness for realistic network modeling. A scale-free communication network of security traders will show what impact network topology has on the dynamics of complex socio-technical systems.

Keywords: Random Networks, Scale-Free Networks, Communication Networks, Socio-Technical Networks, Diffusion

traders. It may be concluded that the network topology has a substantial influence on the outcome of our simulations results, while leading to significantly higher volatility of market prices if a scale-free network is assumed. The application of scale-free networks within simulation models is still in its infancy, which may be related to the lack of awareness of their existence. The authors hope to contribute to a better awareness and use of the above-mentioned network models within simulation models.

However, it is not only socio-economic systems like the trading model presented that are affected by the discovery of scale-free networks. Modeling network related systems is a widely used method in IS, and is also relevant for the investigation of the diffusion of software standards [StSW04] and the robustness of IT systems [DoMe03]. The prevalence of random networks is, in the absence of any better solution, a dangerous assumption. It might influence the simulation results in an unintended way. Therefore, the empirical observation of network properties, if possible, is a necessary step towards ensuring the absence of undesirable side effects. The step from random networks to scale-free networks is an important step in that sense. Scale-free networks show properties of real world networks and are often a better fit for modeling real world processes and will probably gain in importance in the field of IS in the future.

References

[AlBa02] *Albert, Reka; Barabasi, Albert-Laszlo:* Statistical mechanics of complex networks. In: Reviews of Modern Physics 74 (2002) 1, pp. 47–97.

[AlJB99] *Albert, Reka; Jeong, Hawoong; Barabasi, Albert-Laszlo:* Diameter of the World Wide Web. In: Nature 401 (1999), pp. 130–131.

[AHLP93] *Arthur, Brian W.; Holland, John; LeBaron, Blake; Palmer, R. G.; Taylor, Paul:* Asset Pricing Under Endogenous Expectations in an Artificial Stock Market. Working Paper #96-12-093 (1996).

[Bane92] *Banerjee, Abhijit V.:* A Simple Model of Herd Behavior. In: Quarterly Journal of Economics 107 (1992) 3, pp. 797–817.

[Bane93] *Banerjee, Abhijit V.:* The Economics of Rumours. In: The Review of Economic Studies 60 (1993) 2, pp. 309–327.

[Bara03] *Barabasi, Albert-Laszlo:* Linked. Plume, London 2003.

[BaAl99] *Barabasi, Albert-Laszlo; Albert, Reka:* Emergence of Scaling in Random Networks. In: Science 286 (1999), pp. 509–512.

[BaBo03] *Barabasi, Albert-Laszlo; Bonabeau, Eric:* Scale-Free Networks. Scientific American 2003.

[Cont01] *Cont, Rama:* Empirical properties of asset returns: stylized facts and statistical issues. In: Quantitative Finance (2001) 1, pp. 223–236.

[CoBo00] *Cont, Rama; Bouchaud, Jean-Philippe:* Herd behavior and aggregate fluctuations in financial markets. In: Macroeconomic Dynamics 4 (2000), pp. 170–196.

[DoMe03] *Dorogovtsev, Sergey; Mendes, J. Fernando:* Evolution of Networks, From Biological Nets to the Internet and WWW. Oxford University Press, Oxford 2003.

[EbMB02] *Ebel, H.; Mielsch, L.-I.; Bornholdt., S.:* Scale-free topology of e-mail networks. In: Phys. Rev. E 66 (2002), pp. 035103.

[ErRé59] *Erdös, Paul; Rényi, Alfred:* On random graphs. In: Publ. Math. Debrecen 6 (1959), pp. 290.

[HeSc05] *Hein, Oliver; Schwind, Michael:* Standardisierte Mikrosimulation eines Kapitalmarktes mit scale-free Netzwerken und Informationsansteckung. In: *Ferstl, O. K.; Sinz, E. J.; Eckert, S.; Isselhorst, T.* (Eds.): Wirtschaftsinformatik 2005, Physica-Verlag, Heidelberg 2005.

[HeSS05] *Hein, Oliver; Schwind, Michael; Spiwoks, Markus:* A Microscopic Stock Market Model with Heterogeneous Interacting Agents in a Scale-Free Communication Network. Workshop on Economic Heterogeneous Interacting Agents (WEHIA 05), University of Essex 2005.

[Keel99] *Keeling, Matthew J.:* The effects of local spatial structure on epidemiological invasions. Proceedings of the Royal Society of London B 266, 1999, pp. 859–867.

[KeMc27] *Kermack, William O.; McKendrick, Anderson G.:* A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London Series A 115, 1927, pp. 700–721.

[Leve02] *Leveille, Jasmin:* Epidemic Spreading in Technological Networks. Working Paper, HP, Bristol 2002.

[LeLS00] *Levy, Moshe; Levy, Haim; Solomon, Sorin:* Microscopic Simulations of Financial Markets. Academic Press, New York 2000.

[LlMa01] *Lloyd, Alun L.; May, Robert M.:* How Viruses Spread Among Computers and People. In: Science 292 (2001) 5520, pp. 1316–1317.

[Lope04] *López-Pintado, Dunia:* Diffusion in complex social networks. Working Paper, University of Alicante, Alicante, 2004.

[Lux00] *Lux, Thomas; Marchesi, Michele:* Volatility Clustering in Financial Markets: A Microsimulation of Interacting Agents. In: International Journal of Theoretical and Applied Finance 3 (2000) 4, pp. 675–702.

[Mand04] *Mandelbrot, Benoit B.:* The (Mis)Behavior of Markets, A Fractal View of Risk, Ruin, and Reward. Basic Books, NewYork 2004.

[MaLo01] *May, Robert M.; Lloyd, Alun L.:* Infection dynamics on scale-free networks. In: Phys. Review E 64 (2001), pp. 066112.

[MoPV02] *Moreno, Yamir; Pastor-Satorras, Romualdo; Vespignani, Alessandro:* Epidemic outbreaks in complex heterogeneous networks. In: The European Physical Journal B 26 (2002), pp. 521–529.

[PaVe01] *Pastor-Satorras, Romualdo; Vespignani, Alessandro:* Epidemic Spreading in Scale-Free Networks. In: Physical Review Letters 86 (2001) 14, pp. 3200–3203.

[PaVe02] *Pastor-Satorras, Romualdo; Vespignani, Alessandro:* Epidemics and immunization in scale-free networks. In: *Bornholdt., S.; Schuster, G.* (Eds.): Handbook of Graphs and Networks: From the Genome to the Internet. Wiley-VCH, Weinheim 2002.

[ScPr95] *Schmidt, Hartmut; Prigge, Stefan:* Börsenkursbildung. In: *Gerke, W.; Steiner, M.* (Eds.): Handwörterbuch des Bank- und Börsenwesens. Schäffer-Poeschl, Stuttgart 2001.

[Shil01] *Shiller, Robert:* Paradigmenwechsel in der Finanzmarktforschung. Neue Züricher Zeitung 2001, pp. 27.

[Shle00] *Shleifer, Andrei:* Inefficient Markets: An Introduction to Behavioral Finance. Oxford University Press, Oxford 2000.

[StSW04] *Stockheim, Tim; Schwind, Michael; Weiss, Kilian:* A Diffusion Model for Communication Standards in Supply Networks. In: Journal of IT Standards & Standardization Research 4 (2006) 2, pp. 24–42.

[VaFS02] *Valverde, Sergi; Ferrer-Cancho, Ramon; Sole, Richard V.:* Scale-free Networks from Optimal Design. In: Europhysics Letters 60 (2002), pp. 512–517.

[VaPV02] *Vazquez, Alexei; Pastor-Satorras, Romualdo; Vespignani, Alessandro:* Internet topology at the router and autonomous system level. cond-mat/0206084 2002.

[WaCh03] *Wang, Fan X.; Chen, Guanrong:* Complex Networks: Small-World, Scale-Free and Beyond. In: IEEE Circuits and Systems Magazine 1 (2003), pp. 6–20.

[WaSt98] *Watts, Duncan J.; Strogatz, Steven H.:* Collective dynamics of ‘small-world’ networks. In: Nature 393 (1998), pp. 440–442.