University of Novi Sad
Faculty of Sciences
Department of Mathematics and
Informatics

Maja Jović

Rumor spreading in complex social networks

Master's thesis

Mentor
Dr Dušan Jakovetić

Novi Sad, 2023

I would like to express my sincere gratitude to my mentor Dr Dušan Jakovetić for providing guidance and feedback during the writing of this thesis. I am also grateful to the board, Prof. Danijela Rajter-Ćirić and Prof. Sanja Rapajić for reviewing this thesis and providing valuable advice.
I owe special thanks to Ivan for his continuously provided encouragement throughout the studies.
I wish to extend my thanks and express deepest gratitude to my family for their support during my studies.

Contents

Chapter 1

Introduction

Rumors are unverified or unconfirmed information that circulates informally within social circles. They manifest in diverse formats, ranging from hearsay, gossip or speculations, to conspiracy theories, advanced propaganda and marketing material. They are disseminated verbally, through writing or digital media platforms, and rapidly, particularly within today's interconnected global landscape which is amplified by the presence of social media platforms and Internet in general.

Rumors represent a natural and universal phenomenon within human societies. The examination of rumors is valuable for several reasons. Rumors possess the capacity to affect person's decisions since individuals frequently base their choices and actions upon the information they encounter, whether or not they are actually true. So, the systematic analysis of rumors plays a pivotal role in unraveling the intricate interplay between misinformation and its potential consequences on choices and outcomes. Also, amidst times of crises, the rapid dissemination of rumors can create widespread panic and confusion. The goal of organizations and emergency response units is to confront misinformation, preserve the trust of the public and provide effective crisis communication strategies. In the specific context of health crises, rumors wield the potential to contribute to vaccine hesitancy and disapproval with vital public health guidelines. That is why this analysis is important for epidemiologists, healthcare practitioners and public health agencies in order to develop effective strategies for addressing and countering health-related rumors. Furthermore, there can exist potential risks and consequences by spreading false information that can impact national security during military actions or threats. In addition, in the era of fast and broad dissemination of online information, it is imperative to understand the role of algorithms and the influence of echo chambers and filter bubbles, an environ-

ment in which a person encounters only beliefs or opinions that coincide with their own, so that their existing views are reinforced and alternative ideas are not considered. In conclusion, the research on rumors is significant as it provides comprehension of human behavior and the impact of information in various domains, from psychology and sociology to healthcare and national security. Understanding how rumors function and spread allows individuals and organizations to respond effectively to the challenges posed by misinformation and rumors in today's information-rich world.

There exist various approaches to analyzing rumor spreading, each offering a unique perspective on this complex phenomenon. The psychological approach revolves around comprehending the cognitive, emotional and behavioral mechanisms that induce the creation, propagation and reception of rumors. Within this framework, psychologists delve into the intricacies of how individuals process and interpret information, the role of emotions such as fear in shaping the belief and propagation of rumors, and the various motivations that drive rumor creation, such as the pursuit of social status. On the other hand, the sociological approach explores the manner in which rumors operate within social collectives and interpersonal connections. This perspective recognizes that social norms and cultural dynamics wield substantial influence over the creation, interpretation and dissemination of rumors. The mathematical perspective on this phenomenon, which is covered in this thesis, emphasizes modeling and quantitative analysis to elucidate the intricate dynamics of rumor propagation within networks. Researchers create models of complex networks to simulate rumor spreading and use computer simulations and mathematical equations to study how rumors spread, considering parameters like network structure and transmission rates. It is necessary to mention that these approaches need not remain mutually exclusive, and researchers can draw from multiple disciplines to obtain a complete knowledge of rumor diffusion.

This thesis is organized into 8 chapters, where the first and the last provide an introduction and conclusion. In the second chapter, we delve into the very essence of complex networks and their properties. Chapter 3 explores the four prominent network models and explains the difference in their characteristics. In Chapter 4 we introduce the notion of rumor propagation and explain how knowledge of network models offers insight into different propagation patterns. The processes of rumor spreading and epidemic spreading are very similar, so in the fifth chapter we mention three models that belong to the category of deterministic compartmental models that represent the fundamental dynamic framework for infectious diseases. Even though the spread of rumors reflects certain aspects

of the spread of infectious diseases, they are not completely equivalent. Hence, in Chapter 6 we explain the fundamental rumor spreading models derived by drawing insights from the study of infectious disease transmission. In the seventh chapter, the work and results of the work "Theory of rumor spreading in complex social networks", by M. Nekovee et. al is discussed. In their work, they introduce a new model of rumor propagation on complex networks that provides a more realistic description of this process and explains the dynamics of the model on complex social networks.

Chapter 2

Complex social networks

The scientific exploration of networks, encompassing computer networks, social networks and biological networks, has become a subject of extensive scientific interest in recent decades. The widespread use of the Internet and its accessibility has facilitated the collection and analysis of network data on a large scale. Furthermore, the emergence of diverse theoretical frameworks has enabled the extraction of new possibilities for gaining insights from various types of networks. The study of networks transcends disciplinary boundaries, with significant advancements witnessed in fields such as mathematics, physics, computer and information sciences, biology and the social sciences.

Therefore, network science is a multidisciplinary field that delves into the study of connectivity and networks in various manifestations. Whether it is a simple connection between two nodes or intricate networks comprising millions of members, network scientists examine these structures, uncovering meaningful patterns between network attributes and their eventual results. By carefully observing and investigating these networks, valuable theories are formulated to clarify the underlying mechanisms governing their behavior and provide insights into optimizing their use for practical applications.

In this work, the focus is on social networks. To many individuals, the term "social network" commonly relates to online social networking services such as Facebook and Twitter. They can actually be seen as how people or groups of people are connected to each other through social interactions, where vertices in these networks represent individuals or groups, while edges indicate the existence of social ties. It is crucial to acknowledge that social networks can be defined in various ways, particularly in terms of their edges. The specific definition of edges used depends on the research questions being addressed. Edges could

represent friendships, professional relationships, exchange of goods or money, communication patterns, romantic relationships, or other types of connections [4].

Over the past few decades, various quantities and measures of complex networks have been proposed and studied extensively. The advent of powerful technologies has revolutionized our ability to explore the structure of complex networks. In the past, analyzing networks with millions of nodes would have required specialized facilities that were beyond the reach of most researchers. However, this is no longer the case. Now, even networks with millions of nodes can be analyzed with relative ease, opening up new possibilities for research and discovery in this exciting field.

2.1 Characteristics of complex networks

The anatomy of a network is of utmost importance because structure and function are inherently linked. In other words, the way a network is structured can significantly impact how it functions. For instance, the topology of social networks can affect the spread of information and disease, with densely connected networks facilitating faster propagation.
Analyzing certain fundamental characteristics of a complex network is the initial phase in comprehending its structure. Nonetheless, a certain number of exceptional concepts, such as the size and density of a network, average path length, clustering coefficient and degree distribution - have been involved in advancing the theory of complex networks. So, each of these properties will be covered in more detail [9,12].

2.1.1 Network size

The size of a network can be typically defined in terms of the number of nodes, denoted by n, or less commonly, the number of edges, denoted by m. For a connected simple graph, the number of edges can range from $n-1$ in the case of a tree to maximum number of edges, which is $\binom{n}{2}$, in the case of a complete graph. On the contrary, in the presence of self-connected vertices and multiple edges between a pair of vertices, the utmost number of edges is considered to be infinite.
Moreover, the complexity of a network grows exponentially with its size.

2.1.2 Network density

The density of a network is a normalized measure, ranging between 0 and 1, that evaluates the ratio between the actual number of edges and the total number of possible edges in a network consisting of n nodes. Network density serves as an indicator of the percentage of "optional" edges that are present within the network. Here, "optional" edges refer to the potential connections that could exist between nodes, regardless of whether they are currently established. The lower limit of network density corresponds to networks where no relationships exist, indicating a sparse or disconnected network. On the other hand, the upper limit represents networks where all possible relationships are present, indicating a dense or highly interconnected network. As the network density value approaches 1 , the network becomes denser, indicating a higher degree of cohesion among the nodes, leading to a tighter and more interconnected structure among the nodes. Also, information in dense networks can flow more easily than information in sparse networks.

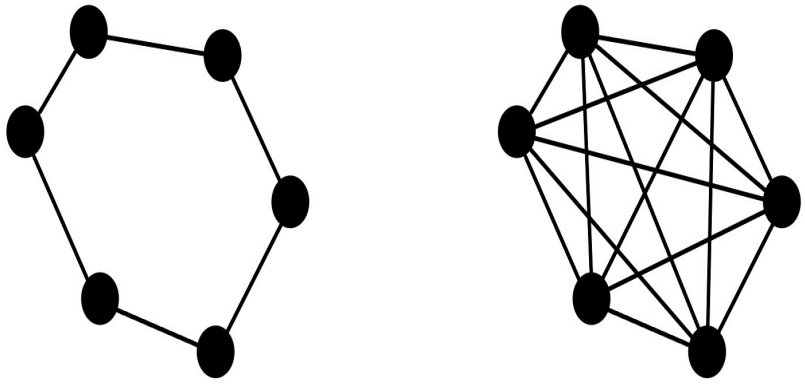

Figure 2.1: Sparse and dense networks

2.1.3 Average path length

Within a given network, the distance between any two nodes is described as the shortest path that connects them in terms of the amount of edges present. Correspondingly, the diameter of the network is established as the maximal distance

value among all the distances calculated between any two nodes present in the network. Additionally, the average path length L of the network is defined as the mean value of the distances between each pair of nodes in the network, calculated over all possible node pairs. This parameter L serves as a key determinant of the overall "size" of the network, representing the most common distance between two randomly chosen nodes present in the network. For instance, when analyzing a friendship network, the average path length represents the typical number of friends that any two individuals are connected by, and shows us how interconnected and closely knit these social circles can be.

2.1.4 Clustering coefficient

In a friendship network, it is quite common for two friends to also be friends with each other, which is known as clustering. To measure this property, the clustering coefficient C is used. To be more specific, the clustering coefficient is characterized as the average fragment of pairs of adjacent nodes for a given node that are also adjacent to one another, resembling a community of interconnected individuals sharing a common bond. It is a measure of the network's cohesion, indicating the strength of the relationships between nodes within the network.

To calculate the clustering coefficient C_i of a node i in a network, we examine all distinct pairs of vertices that are neighbours of i in the network. For each such pair, we count the number of links connecting them and divide by the total number of possible links between them, which is given by $\binom{k_i}{2}$, where k_i is the degree of i [2].

That is,

$$C_i = \frac{\text{number of pairs of neighbors of } i \text{ that are connected}}{\text{number of pairs of neighbors of } i}$$

This ratio gives us the local clustering coefficient C_i, which represents the average probability that two friends of i are also friends of each other. The clustering coefficient is an important measure of the local structure of a network, as it captures the tendency of nodes to form clusters or communities.

Local clustering plays a significant role in the study of networks due to its dependence on the degree. Generally, vertices with higher degrees exhibit a lower average local clustering coefficient. Another reason why local clustering is essential is its ability to detect "structural holes" in a network. Typically, especially in social networks, neighbors of a vertex in a network are connected among

themselves. However, this may not always be the case, leading to expected connections between neighbors being absent. Structural holes can be detrimental to the efficient spread of information or other traffic within a network, as they decrease the number of alternative routes available. However, for the central vertex i, whose friends lack connections, structural holes can be beneficial as it provides i with power over information flow between those friends. The local clustering coefficient measures the level of influence i has in this sense, with lower values indicating a higher number of structural holes in the network surrounding i.

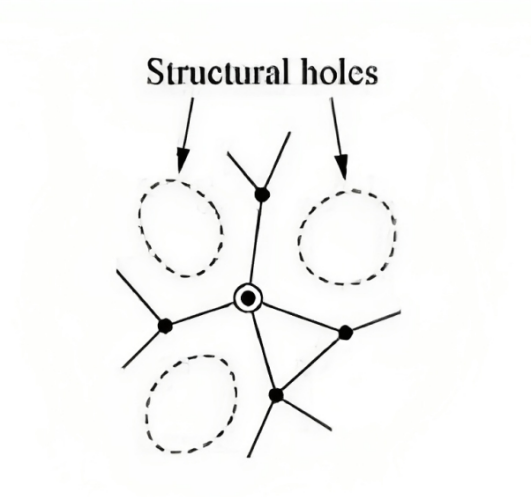

Figure 2.2: Structural holes. Source:[2]

In order to determine the clustering coefficient C of the entire network, we average the C_i values across all nodes i within the network.
Note that $0 \leq C \leq 1$. In a complete network, the clustering coefficient is equal to 1. On the other hand, in a completely random network consisting of n nodes, the clustering coefficient is approximately equal to $\frac{1}{n}$.

2.1.5 Degree distribution

A node's degree is its most basic, and, arguably, its most significant attribute. The degree of a single node, refers to the total number of connections it possesses, or equivalently, the total number of its nearest neighbors. Higher degrees signify greater node importance within the network. The network's average degree is denoted by $\langle k \rangle$ and can be given by $\langle k \rangle = \frac{2m}{n}$, where m is the number of edges and n is the number of vertices.

The distribution of node degrees in a network is represented by the probability function $P(k)$, which outlines the likelihood of a randomly selected node having exactly k edges. In a regular lattice network, where all nodes share the same number of edges, the degree sequence is straightforward, with the degree distribution plot displaying a single, concentrated spike known as a delta distribution.

However, randomness in the network causes this peak to broaden and the distribution shape to change. In the scenario of a completely random network, the degree sequence follows the Poisson distribution.

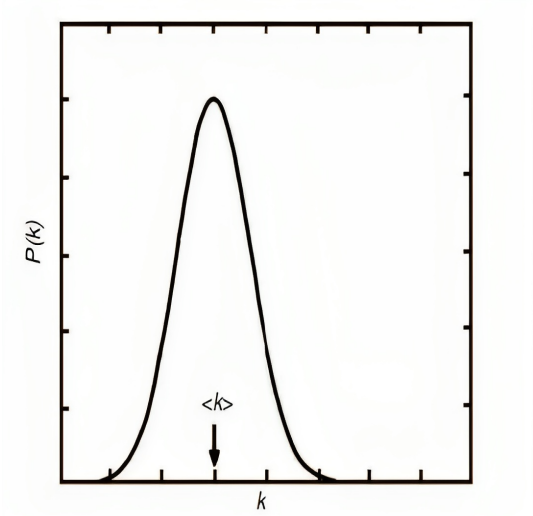

Figure 2.3: Poisson distribution. Source: [12]

In networks with a Poisson degree distribution, commonly known as exponential networks, the distribution curve is symmetrical and the shape of this distribution exhibits an exponential decrease from the peak value of $\langle k \rangle$. This means that discovering a node with a degree equal to k, for any k far greater than the average degree $\langle k \rangle$, is highly unlikely. A completely random graphs were studied first by Erdős and Rényi, about whom there will be more word later.

Recent empirical evidence has demonstrated that, for most large-scale real networks, the degree distribution differs significantly from the Poisson distribution. Instead, a power law of the form $\mathrm{P(k)} \sim k^{-\gamma}$, where γ is some exponent(known as the power law exponent), more accurately represents the degree distribution for certain networks.

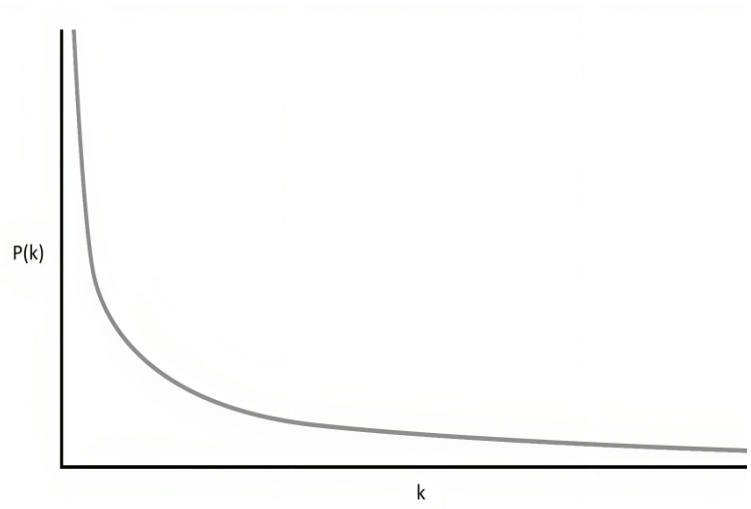

Figure 2.4: Power law distribution.

The indicated power law distribution displays a slower decline than the exponential distribution, which permits a few nodes with exceptionally large degrees to exist. Because power laws do not have any peculiar scale, a network that follows a power law degree distribution is known as a scale-free network. We will discuss them later on as well.

2.1.6 Degree correlation

Complex networks can be classified as assortative, disassortative or neutral.

In assortative networks, nodes with high degree tend to connect with other high-degree nodes, resulting in the formation of tightly knit communities or clusters. Conversely, small-degree nodes have a tendency to link with other small-degree nodes. Social networks are a classic example of such networks.

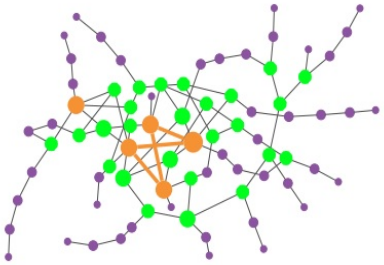

Figure 2.5: Assortative network. Source: [6]

In disassortative networks, on the other hand, high-degree nodes avoid each other, connecting instead to small-degree nodes, as observed in information and biological networks.

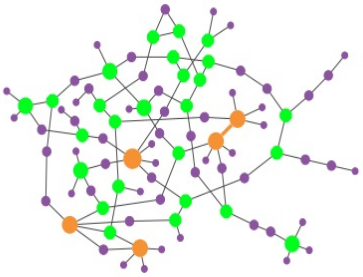

Figure 2.6: Disassortative network. Source: [6]

Neutral networks are characterized by a random linking pattern between nodes, that is, showing no significant preference for node degree in their connections, resulting in a lack of correlation and structural organization in the network. This randomness in the linking pattern is reflected in the symmetric density of links around the average degree, indicating that nodes are as likely to be linked to nodes with similar degrees as they are to nodes with different degrees.

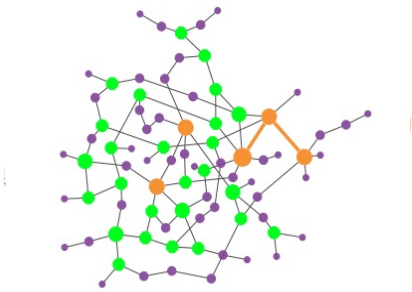

Figure 2.7: Neutral network. Source: [6]

Sociologists have long observed and discussed these divisions. Individuals tend to form social connections with others based on various attributes such as age, nationality, language, income and education level. This is known as assortative mixing, which is characterized by the tendency of individuals to connect with others who share similar features.
However sometimes, although rarely, people also form connections with those who possess different characteristics from them, which is disassortative mixing. Disassortative mixing is predominantly seen in the context of gender-based mixing in marriages, where partnerships between individuals of opposite genders are more prevalent compared to same-gender partnerships.

If the number of links between high and low-degree nodes deviates significantly from what would be expected by chance, then a network is said to display degree correlations. Classical random graphs do not exhibit degree correlations, but growing networks naturally display them.

Chapter 3

Network models

One of the fundamental aspects of network science is the study of network models, which are mathematical abstractions that capture essential characteristics of real-world networks. These models help researchers understand and analyze various properties and phenomena exhibited by real networks. Network analysis often involves calculating various attributes, mentioned in the previous chapter, that provide insights into the characteristics of a network. These properties play a crucial role in defining network models and can be utilized to compare and contrast different models, highlighting their distinct behaviors.

Four prominent network models widely studied in network science are random graphs, scale-free graphs, regular coupled networks and small-world networks, which will be explored in more detail. The first two will be of great importance to us later on in this paper.

3.1 Random graphs

In 1959, Paul Erdős and Alfréd Rényi, two outstanding mathematicians from Hungary, proposed a model to describe networks observed in fields like communications and life sciences. Their concept involved the random connection of nodes through links, offering a simple yet powerful approach. This approach brought new vitality to the realm of graph theory. As a result, a specialized area of mathematics emerged, focusing on the analysis and exploration of random networks.

In general, a random graph is a type of model network that has certain parameters with fixed values, but is random in other aspects. There are two

main constructions of a random graph model with a fixed number of nodes that were proposed.
The first construction involves a network where we only set the number of vertices n and the number of edges m. This means we have n vertices, and we randomly place m edges among them. Specifically, we select m pairs of vertices randomly from all possible pairs and connect them using an edge. It is often required that the network should be a simple graph, meaning it should not have self-edges or multiple edges. Therefore, the position of each edge should be selected only among distinct pairs that are not already connected. This model is commonly known as $G(n, m)$ in mathematical terms.
Another way to define the model is to state that the network is created by randomly choosing from the set of all simple graphs that have exactly n vertices and m edges.
Both constructions yield statistically equivalent sets of graphs.

It is important to note that the random graph model should not be defined in terms of a single randomly generated network, but as a collection of networks.

The majority of mathematical research has primarily focused on a relatively modified model known as $G(n, p)$, which offers significantly greater ease of analysis. Within this particular model, the emphasis shifts from fixing the number of edges to fixing the probability of edges between vertices. In this revised framework, we maintain a network of n vertices, where each distinct pair has an independent probability p of being connected by an edge. Unlike previous models, the number of edges in this network is not predetermined. It is feasible for the network to contain zero edges or even have edges connecting every possible pair of vertices, although these scenarios are less likely for most values of p.

Once again, it is essential to note that the technical definition of a random graph relates not to an individual network, but rather to an ensemble over all conceivable networks.

The advent of the $G(n, p)$ model brought forth significant contributions from renowned mathematicians Erdős and Rényi, who conducted a series of groundbreaking studies on this model during the late 1950s and early 1960s. Their invaluable insights led to the model being commonly referred to as the "Erdős-Rényi model" or the "Erdős-Rényi random graph" as a tribute to their influential work. Notably, when discussions revolve around a random graph without specifying a particular variant, $G(n, p)$ invariably serves as the fundamental reference point.

In our exploration of the $G(n, p)$ random graph, we commence with a straightforward yet crucial calculation, which is the determination of the expected number

of edges within our model network. Even though the number of edges in this model is not fixed, we can derive its expectation value through a systematic analysis. By examining the number of graphs featuring precisely n vertices and m edges, we can discover the various ways in which edge positions can be selected from $\binom{n}{2}$ distinct vertex pairs. Consequently, the total probability of selecting a graph with m edges from our ensemble can be expressed with a standard binomial distribution as:

$$P(m) = \binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m} \tag{3.1}$$

In addition, the expected number of edges connecting any given pair of vertices is equivalent to the probability p of an edge existing between those vertices. That being so, the expected total number of edges in the network is equal to the probability p connecting any pair of vertices, multiplied by the total number of pairs, that is

$$E(m) = \sum_{m=0}^{\binom{n}{2}} mP(m) = \binom{n}{2} p. \tag{3.2}$$

Now, since the average degree of a graph with precisely m edges is $k = \frac{1}{n}m$, the average degree in $G(n,p)$ is

$$< k >= \sum_{m=0}^{\binom{n}{2}} \frac{1}{n} mP(m) = \frac{1}{n}\binom{n}{2} p = (n-1)p. \tag{3.3}$$

The computation of the degree distribution of $G(n,p)$ is slightly more challenging. Each vertex in the graph is independently connected to every other vertex with a probability of p, excluding itself. Consequently, the probability of being connected to exactly k other vertices while not being connected to any of the remaining $n-1-k$ vertices is given by $p^k(1-p)^{n-1-k}$. With $\binom{n-1}{k}$ ways for k vertices to be selected, the total probability of being connected to precisely k vertices follows again the binomial distribution, which is

$$P(k) = \binom{n}{k} p^k (1-p)^{n-1-k}. \tag{3.4}$$

Hence, in essence, $G(n,p)$ is characterized by a binomial degree distribution. However, in numerous cases, our focus is directed towards the properties exhibited by expansive networks, where the value of n can be considered sufficiently large.

Moreover, many networks showcase a mean degree that remains relatively stable as the network size expands. For instance, the average number of friends an individual possesses is not heavily dependent on the total global population.

From Eq.(3.3) we can see that, as $n \to \infty, p = \frac{<k>}{n-1}$ becomes vanishingly small, which enables us to write

$$\ln\left[(1-p)^{n-1-k}\right] = (n-1-k)\ln\left(1-\frac{<k>}{n-1}\right) \simeq -(n-1-k)\frac{<k>}{n-1} \simeq -<k>$$

where we have expanded the logarithm as a Taylor series, and the equalities become exact as $n \to \infty$. By exponentiating both sides of the equation, we gain

$$(1-p)^{n-1-k} = e^{-<k>}. \tag{3.5}$$

Furthermore, for large n, we obtain

$$\binom{n-1}{k} = \frac{(n-1)!}{k!(n-1-k)!} \simeq \frac{(n-1)^k}{k!}. \tag{3.6}$$

Therefore, we can express $P(k)$ as

$$\begin{aligned} P(k) &= \frac{(n-1)^k}{k!}p^k e^{-<k>} = \frac{(n-1)^k}{k!}\left(\frac{<k>}{n-1}\right)^k e^{-<k>} \\ &= e^{-<k>}\frac{<k>^k}{k!}. \end{aligned} \tag{3.7}$$

Hence, for large n, we gain that $G(n,p)$ has the Poisson distribution, which is why this specific model is also sometimes called the "Poisson random graph".

Now, using that $\langle k \rangle$ is the average degree in a random graph, we know that a node in this network has on average:

- $< k >$ nodes at distance one $(d = 1)$
- $< k >^2$ nodes at distance two $(d = 2)$
- $< k >^3$ nodes at distance three $(d = 3)$
- $\ldots$
- $< k >^d$ nodes at distance d

The expected number of nodes up to distance d from the starting node is

$$N(d) = 1 + < k > + < k >^2 + < k >^3 + \ldots + < k >^d$$
$$= \frac{< k >^{d+1} - 1}{< k > - 1}. \tag{3.8}$$

Let L_{rand} represent the average path length of a random network. Intuitively, approximately $< k >^{L_{\text{rand}}}$ nodes within the random network are situated at a distance L_{rand} or in close proximity to it. Consequently, $n \sim < k > L_{\text{rand}}$, implying that $L_{\text{rand}} \sim \frac{\ln n}{\ln < k >}$. Due to the slow increase of $\ln n$ with respect to n, the average path length can remain relatively small even within relatively large networks.

This model $G(n, p)$ produces networks that are statistically homogeneous, even though the model is inherently random. This means that most nodes in the network have roughly the same number of links, around the average degree $(< k >)$. The connectivity in this model follows a Poisson distribution with a bell shape, with a strong peak at $\langle k \rangle$. This suggests that the likelihood of finding a node with a very high degree decreases exponentially.

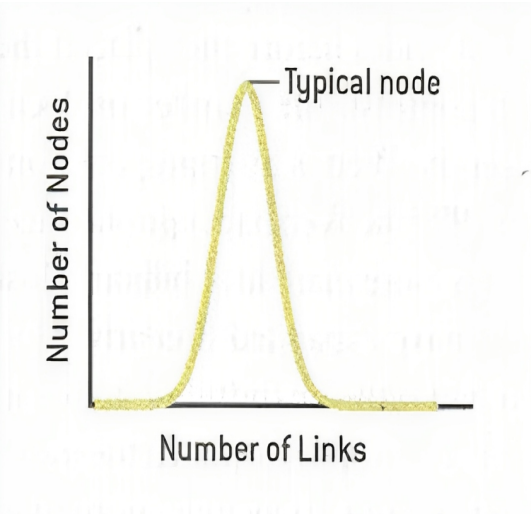

Figure 3.1: Bell curve distribution of node linkages. Source: [11]

When examining the Poisson random graph, one can easily compute an essential network metric - the clustering coefficient. As already stated, this coefficient quantifies the transitivity within a network by representing the probability that two neighboring vertices are also connected to each other. In the case of a random graph, the probability of any two vertices being neighbors is identical, and

all these probabilities are equal to $p = \frac{<k>}{n-1}$, where $< k >$ represents a constant value.
Accordingly,

$$C = \frac{< k >}{n-1}.$$

As the limit n approaches infinity with a fixed average degree, we can see from the above equation that the clustering coefficient tends to zero in the random graph. One noteworthy disparity between the random graph and real-world networks lies in their clustering coefficients. While real-world networks frequently exhibit high clustering coefficients, the random graph follows a different pattern.

3.2 Scale-free networks

For over four decades, the scientific community regarded complex networks as entirely random. Therefore, when scientists Albert-László Barabási, Eric Bonabeau, Hawoong Jeong and Réka Albert initiated their ambitious project to map the World Wide Web in 1998, they anticipated discovering a random network [11]. Their reasoning was straightforward: individuals link their web documents based on their distinctive interests, and with the vast variety of these interests and the sheer number of available web pages, one would expect the resulting connections to exhibit a fairly random pattern.
However, the measurements they obtained during their investigation contradicted their initial assumptions. Using specialized software, they unleashed a virtual robot that traversed the web, meticulously collecting all the links it encountered. Although the robot could only explore a diminutive fraction of the entire web, the resulting map it constructed unveiled that a select few highly connected pages played a vital role in holding the entire World Wide Web together. Strikingly, over 80% of the pages on the map possessed fewer than four links, while an extraordinarily minuscule minority(less than 0.01% of all nodes) had an astonishing number of over 1000 links. The magnitude of this phenomenon became even more apparent in a subsequent web survey, which uncovered a single document that had amassed references from an astounding two million other pages.
A careful analysis of the web pages, taking into account the number of links they possessed, revealed that the distribution followed a power law. This particular distribution demonstrated an intriguing characteristic whereby the probability of a node being connected to exactly k other nodes was $\frac{1}{k^n}$. In simpler terms,

nodes with a small number of links were plentiful, whereas nodes with a large number of links were relatively rare. The value of n exhibited an approximate value of 2 for incoming links. So, a node was roughly four times more likely to have half the number of incoming links compared to another node. This power law distribution stood in stark contrast to the familiar bell-shaped distributions commonly observed in random networks, thus emphasizing the distinctive and captivating nature of complex networks.

Unlike a bell curve, which exhibits a peak, a power law distribution is characterized by a continuously decreasing function. When represented on a double-logarithmic scale, a power law takes the form of a straight line.

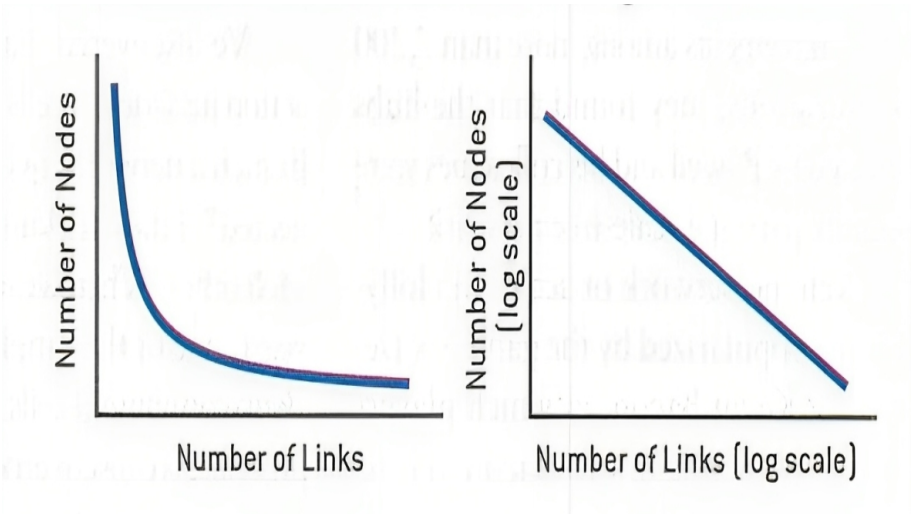

Figure 3.2: Power law distribution of node linkages. Source: [11]

Notably, power laws defy the distribution of links observed in random networks, instead characterizing systems where a select few dominant hubs (nodes that are highly connected to other nodes in the network), like for instance Google, exert significant influence. Random networks, on the other hand, do not enable the emergence of such hubs.

Anticipating to encounter nodes that follow a bell-shaped distribution similar to the distribution of human heights, the scientists found nodes that defied conventional understanding, analogous to the discovery of individuals with height of 30 meters, which prompted them to introduce the term "scale-free" to capture the essence of this extraordinary phenomenon.

The investigation conducted later on by the three Faloutsos brothers, who are computer scientists, focused on examining the routers interconnected through

optical or other types of communication lines. Their findings discovered that the network topology also exhibits a scale-free nature.
Researchers have further uncovered the presence of scale-free properties in select social networks. Among these revelations is the investigation which revealed that a network depicting relationships among individuals in Sweden is adhered to a power law distribution. While the majority of individuals had only a few lifelong partners, a small group of hubs had a considerably larger number of partners. Another study provided evidence of a scale-free network structure among individuals connected through email communication. Additionally, scientific papers, connected through citations, also exhibit a power law distribution. Furthermore, collaborations among scientists in various fields, including physicians and computer scientists, have scale-free characteristics in these networks.
Besides all of this, scale-free networks can occur in protein regulatory network, Hollywood actor's network, cellular metabolism and many more, which scientists discovered the more they studied networks.

An essential question arises concerning the inadequacy of random network theory in explaining the presence of hubs and the power law degree distribution within networks. To elucidate this, Barabási and Albert (BA) proposed another network model, through the closer examination of the original work by Erdős and Rényi (ER) which uncovers two key reasons behind this. Firstly, the ER model assumed a predetermined inventory of nodes prior to establishing connections, whereas the number of documents on the Web, for instance, is far from constant. The Web, which began with a solitary page in 1990, has since grown, surpassing a trillion pages. Similar growth patterns can be observed in most networks. For example, the Hollywood network initially consisted of a mere handful of actors, but as new individuals entered the industry, the network flourished, with newcomers connecting to established veterans, with each new addition linking to the existing network. Consequently, due to the dynamic nature of real networks, older nodes were afforded greater opportunities to establish connections.
The second crucial factor revolves around the inequality of nodes within networks. When individuals decide where to place their web page links, they face over a billion potential destinations. However, our familiarity is limited to only a tiny fraction of the entire web, with a tendency to favor well-connected sites that are easier to discover. By linking to these nodes, individuals unintentionally contribute to and strengthen a bias toward them. This phenomenon, referred to as "preferential attachment" is observed in various contexts. In the domain of Hollywood, for instance, more connected actors are more likely to be selected for new roles, thereby reinforcing their prominence. Similarly, well-established

companies naturally attract a higher number of alliances, enhancing their desirability for future partnerships. Likewise, highly cited articles in scientific literature attract more researchers to read and cite them, maintaining their influence.
Hence, the presence of hubs can be clarified by these two mechanisms: growth and preferential attachment. As the network expands and new nodes come into play, they exhibit a natural inclination to connect to the more connected sites that have already gathered attention. As a result, these popular locations accumulate a growing number of links over time, surpassing their less connected neighbors. This process, often characterized as "rich get richer," predominantly favors the early nodes, as they are more likely to evolve into hubs in due course.

Scientists have utilized computer simulations and precise calculations to demonstrate that a growing network with preferential attachment inevitably exhibits a scale-free nature, characterized by a distribution of nodes that adheres to a power law. While this theoretical model is simplistic in nature and necessitates adjustments to suit individual circumstances, it undeniably verifies the explanation for the ubiquitous presence of scale-free networks in real-world scenarios.

Numerical findings have provided evidence that, when compared to a random graph of the same size and average degree, the scale-free model showcases a marginally smaller average path length, alongside a significantly higher clustering coefficient. This observation suggests that the presence of a few "big" nodes, characterized by exceptionally large degrees serves as crucial role in bringing the remaining nodes in the network into closer proximity. Nevertheless, currently, no analytical predictive formula exists to determine the average path length and clustering coefficient for the scale-free model.

The intuitive understanding suggests that the breakdown of a considerable number of nodes would lead to the fragmentation of a network, resulting in isolated and non-communicating segments. This notion holds true for random networks, wherein the removal of a critical fraction of nodes causes the system to disintegrate into disconnected islands. However, simulations conducted then on scale-free networks reveal a different outcome. Surprisingly, even if up to 80% of Internet routers are arbitrarily selected and fail, the remaining nodes manage to form a cohesive cluster, maintaining a path between any two nodes. Similarly, disrupting a cell's protein-interaction network proves challenging, but the measurements indicate that despite introducing a high level of random mutations, the unaffected proteins continue to collaborate effectively.

Scale-free networks, in their essence, demonstrate a remarkable ability to withstand accidental failures, a quality deeply ingrained within their inhomogeneous topology. When nodes are randomly removed, it is the smaller ones that

are predominantly affected due to their abundance compared to hubs. The elimination of these small nodes does not significantly disrupt the overall network topology as they possess fewer connections in contrast to the hubs, which serve as connectors to a substantial portion of the network. However, the dependence on hubs also poses a notable vulnerability to deliberate attacks, which is often referred to as the "Achilles' heel" of such systems. This term draws an analogy from Greek mythology, where the legendary warrior Achilles possessed extraordinary strength and invulnerability except for a single weak spot in his heel.

Through a series of simulations, researchers have discovered that the elimination of a few pivotal hubs from the Internet has a detrimental effect, causing the system to fragment into isolated groups of routers with no viable connections between them. Similarly, experiments conducted on yeast have revealed that removing proteins with higher degrees of connectivity poses a significantly higher risk of organism fatality compared to the deletion of other nodes.

In conclusion, the significance of relying on hubs can vary depending on the specific system under consideration, offering both advantages and disadvantages.

3.3 Regular coupled networks

In the domain of graph theory, there exists another structure where each vertex has an identical number of neighbors, that is, has the same degree, and it is called a regular graph. Real-world networks rarely exist in isolation, intertwining through various modalities such as dependency relationships, multi-support and inter-connection patterns. Intuitively, a globally coupled network manifests the intriguing attributes of possessing the smallest average path length and the largest clustering coefficient.

An extensively examined, sparse and regular network model of notable interest is the nearest-neighbor coupled network, commonly referred to as a lattice. This type of regular graph exhibits nodes linked solely to a limited number of their neighbors. In this context, a lattice does not exclusively imply a two-dimensional square grid but can rather adopt diverse geometries. A minimal lattice, for instance, constitutes a straightforward one-dimensional structure akin to a row of individuals holding hands. On the other hand, a nearest-neighbor lattice with a periodic boundary condition encompasses of n nodes arranged in a ring-like formation, where each node i is adjacent to neighboring nodes, $i = 1, 2, \cdots, K/2$, with K representing an even integer. When K assumes a large value, such a network showcases substantial clustering. In fact, the clustering coefficient of

the nearest-neighbor coupled network approximates $C = \frac{3}{4}$.

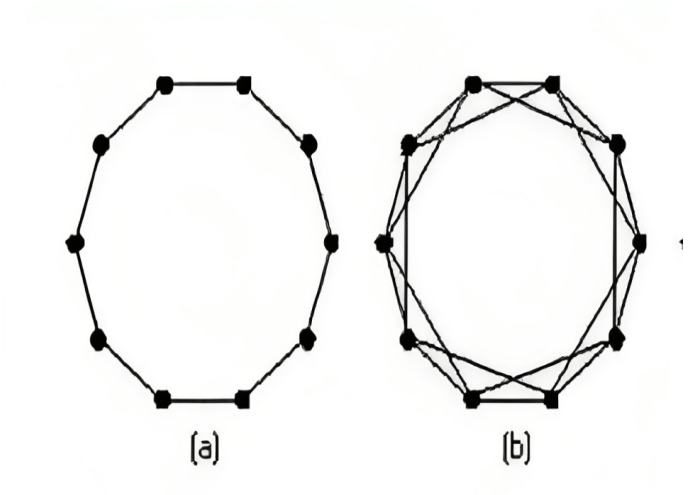

Figure 3.3: (a) The simplest nearest-neighbor coupled network. (b) Nearest-neighbor coupled network with four nearest-neighbor coupled nodes. Source: [16]

The nearest-neighbor coupled network exhibits a significantly large average path length that diverges as $n \to \infty$. This characteristic poses a challenge in achieving effective global coordination, such as synchronization, within the confines of this locally connected network. Nevertheless, there exists a regular network that combines sparsity, clustering and a small average path length called the star-shaped coupled network. In this particular network, a central node interconnects with all other $n-1$ nodes, while the peripheral nodes remain unconnected amongst themselves. Surprisingly, as n tends to infinity, the average path length converges to 2 along with the clustering coefficient approaching to 1. This star-shaped network model embodies the sparse, clustering and other properties observed in many real-world networks, rendering it a more fitting model than the regular lattice for several well-known real networks. Notably, although most real networks do not exhibit an exact star shape, this model captures essential characteristics of real-world systems.

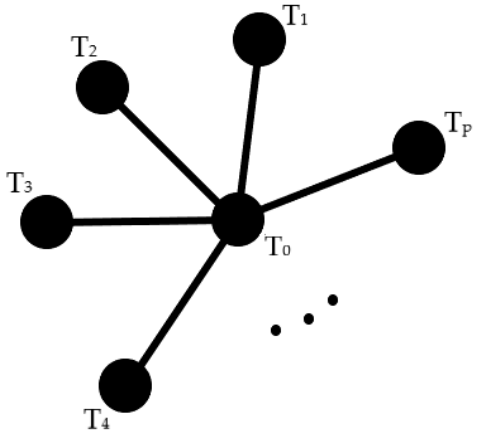

Figure 3.4: Star-shaped network.

3.4 Small-world networks

The small-world phenomenon captures the prevalence of interconnections within human networks, often referred to as "six degrees of separation". This principle underscores the abundance of short paths in social graphs composed of individuals, linked by acquaintances. Its origins trace back to an American social psychologist, Stanley Milgram in the 1960s, who conducted experiments by asking the participants to forward letters through personal connections in the United States. Each participant could forward a letter to solely one person with the aim of the letter reaching a "target person" in Boston. Milgram unveiled that the average chain length was six. This intrigued the exploration of compact pathways within networks.

The succeeding studies by eminent applied mathematicians Duncan Watts and Steve Strogatz proposed a distinctive interpretation of networks possessing the small-world attribute. They conceptualized such networks as a superposition: a densely interconnected sub-network encompassing nodes' local connections, along with a set of randomized long-range connections facilitating the creation of concise pathways. Beyond empirical analyses of social, technological and biological networks, Watts and Strogatz explored this fundamental model. This model commences with a d-dimensional lattice network, augmented by a small amount of extended links from each node, randomly chosen with uniform distribution. The resultant network demonstrates localized clustering and short paths,

mirroring prevalent real-world networks. (Refer to Figures 3.5 and 3.6 for visual representation.)

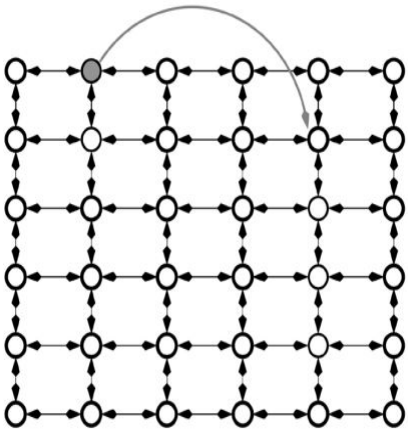

Figure 3.5: Two-dimensional grid with a single random shortcut superimposed. Source: [19]

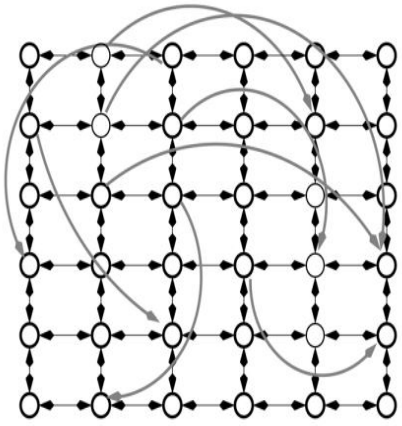

Figure 3.6: Two-dimensional grid with many random shortcuts superimposed (as in the Watts-Strogatz model). Source: [19]

Conventionally, network connection topology adheres to either total regularity or complete randomness. However, a shift has transpired with the advent of "small-world" networks, introducing a middle ground. In these networks, regular structures undergo progressive rewiring, introducing a growing degree of disorder. It becomes a type of graph in which most nodes are not neighbors of one another,

but most nodes can be reached from every other by a small number of steps. Strikingly, these networks embody the dual nature of high clustering similar to regular lattices, coupled with small characteristic path lengths similar to random graphs.

To transition between regular and random networks, Watts and Strogatz employ the following random rewiring protocol. Commencing with a ring lattice consisting of n vertices and k edges per vertex, they randomly rewire each edge with a probability of p. In this context, "rewiring" refers to relocating one end of a link to a randomly selected node within the entire network, following the conditions that distinct nodes cannot possess multiple connections between them and no node can be connected to itself. This procedure establishes connections between nodes that would otherwise belong to separate local clusters. With this approach, the graph is gradually adjusted from regularity $(p = 0)$ to randomness $(p = 1)$, thereby exploring the intermediate zone where $0 < p < 1$.

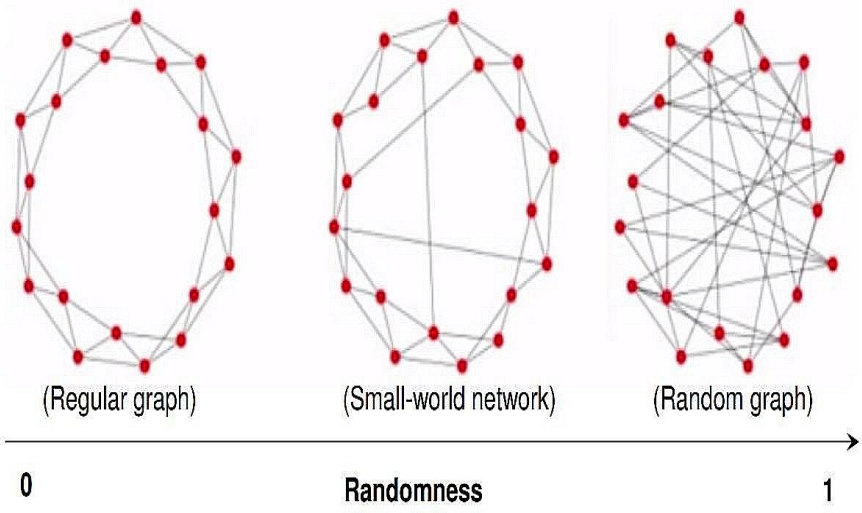

Figure 3.7: Three graphs depending on the value of p. Source: [20]

In the WS small-world model, both clustering coefficient $C(p)$ and average path length $L(p)$ can be regarded as functions of the rewiring probability p. At the beginning, when $p = 0$, a regular ring lattice has high clustering $C(0)$, but a long average path length $L(0)$. When the rewiring probability is low, the local network features remain similar to the original regular network, maintaining clustering coefficient close to the initial value $(C(p) \approx C(0))$. Yet, the average path length decreases significantly and approaches that of random networks $(L(p) >>$

$L(0)$). This is expected since a few random rewirings can substantially reduce the average path length while keeping local clustering properties relatively unchanged.

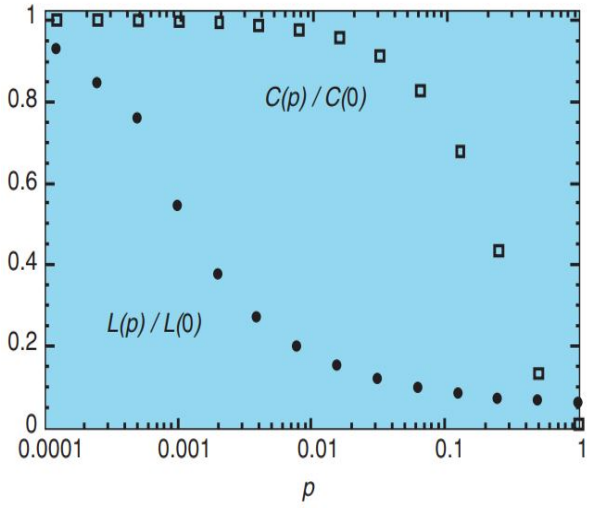

Figure 3.8: Average path length and clustering coefficient of the WS small-world model as a function of the rewiring probability p. Both are normalized to their values for the original regular lattice $(p = 0)$. Source: [12]

The study of WS small-world networks has led to an exploration of new models of complex networks, which includes certain adaptations of the WS model. One well-known variation is the NW small-world model, introduced by Newman and Watts [21]. In the NW model, existing connections between nearest neighbors remain intact, and new connections are introduced between node pairs with probability p. A node's connectivity also cannot involve multiple links to another node or self-connections. When p is set to 0, the NW model mirrors the original nearest-neighbor coupled network, and if p is 1, it transforms into a globally coupled network. Analyzing the NW model is comparatively simpler than the original WS model, as it avoids the emergence of isolated clusters, a phenomenon that may be present in the WS model. When p is sufficiently small and n is adequately large, the NW model becomes essentially indistinguishable from the WS model. Presently, these two models are collectively referred to as small-world models for the sake of conciseness.

The origin of small-world models lies in social networks, where individuals usually have friendships with nearby neighbors, like those living on the same street

or coworkers in the same workplace. However, many individuals also maintain a handful of friendships through significant distances, such as friends in different countries. These distant relationships are symbolized by the extended edges introduced through the rewiring process in the WS model or by the connection-addition process in the NW model.

Chapter 4

Rumor spreading

In the age of information, with massive amounts of data being generated and circulated without any monitoring, gossip can spread like a wildfire. In today's world, it is more important than ever to be cautious about the accuracy and reliability of the information we consume and share.

Rumor spreading refers to the process of dissemination of unverified or false information among individuals or groups in a social network. Research has shown that the spread of rumors is influenced by a variety of factors, including psychological and social factors, such as the need for information, the emotional appeal of the rumor and social identity dynamics.

Misinformation, being incorrect information, and disinformation, being deliberately misleading false information, have always existed in human communication, propagating through word-of-mouth transmission. Hence, the informational content of rumors varies widely, spanning from basic gossip to developed propaganda and marketing material. However, contemporary research utilizing modern analytical techniques, facilitated by the Internet, has shed new light on this phenomenon. In the past, the spread of rumors was limited to interpersonal interactions, but the emergence of social media has ushered in a new era of rapid and extensive rumor propagation which has become more destructive.

The study of rumor propagation in social networks is vital for various reasons. Firstly, it helps develop strategies to limit misinformation, especially during crises or sensitive occurrences. Secondly, examining rumor diffusion patterns yields insights into human behavior, cognitive biases and information processing in digital communities. Moreover, identifying influential nodes (individuals with significant rumor propagation potential) is crucial for focused intervention and precise information dissemination. Analyzing rumor spreading also helps in detecting and

tracking the origin and evolution of specific rumors, thus aiding in building more effective fact-checking mechanisms.
Furthermore, social network analysis benefits researchers in comprehending the network's structure that either facilitates or impedes rumor diffusion. Recognizing vulnerable regions in the network guides the implementation of measures to limit the rapid spread of misinformation. Additionally, machine learning and data mining techniques can model the dynamics of rumor propagation using large datasets, enabling the development of predictive models for forthcoming rumor outbrakes.
Better insight into the mechanics of rumor diffusion in complex social networks and its consequences can relieve the establishment of a more wholesome and educated online realm. This, in turn, can promote critical thought, digital media awareness and the advancement of technologies for identifying and countering the issues presented by misinformation in the digital era.

Knowledge about network models has implications for understanding the spread of computer viruses, diseases, information and rumors. The connection between these graph structures and rumor spreading lies in the different propagation patterns they enable. In random graphs, for instance, due to the more homogeneous connectivity, rumors tend to spread more slowly and less extensively. The lack of highly connected hubs restricts the ability of rumors to propagate rapidly across the network. Consequently, random graphs generally exhibit a slower and localized diffusion of rumors. On the other hand, scale-free graphs provide an environment convenient for rapid and widespread rumor spreading. The presence of highly connected hubs in scale-free networks enables rumors to quickly propagate from a small set of influential nodes to a significant portion of the network. These hubs serve as information spreaders, capable of disseminating rumors to numerous connections efficiently. Once a hub has heard the rumor, it will pass it to numerous other nodes, eventually compromising other hubs, which will then spread the rumor through the entire network. As a result, scale-free graphs often exhibit a faster and more extensive spread of rumors, with the potential for global reach.

Understanding the connection between various network models and their impact on rumor spreading dynamics is crucial in various fields, such as social network analysis and epidemiology. By studying how information or rumors spread through these different network structures, researchers can gain insights into the mechanisms underlying the propagation of ideas, diseases or opinions in real-world networks.

Chapter 5

Epidemics on networks

The dynamics of information and rumor diffusion in complex networks are often modeled as contagion processes, resembling the spread of epidemics among neighboring individuals. The study of social networks has gathered substantial attention from the scientific community, particularly due to their association with disease transmission. Diseases propagate through networks of interpersonal contacts, where airborne illnesses like influenza or tuberculosis can be transmitted through shared air in enclosed spaces, while contagious diseases and parasites spread through physical contact. These contact patterns can be represented as networks, and extensive empirical research has been dedicated to understanding the structure of these networks. In this section, we delve into the relationship between network structure and disease dynamics.

The biological processes that unfold when an individual, known as a "host" in epidemiological terms, acquires an infection are complex. Upon infection, the pathogen generally undergoes replication within the host's body, while the immune system initiates a response to counteract its presence, often leading to the manifestation of symptoms. Ultimately, either the pathogen or the immune system gains the upper hand, although there are instances where neither prevails, leading to outcomes such as the individual's recovery, mortality or the establishment of a chronic infection state.

5.1 Compartmental models of the disease spread

To completely understand how diseases disseminate throughout populations would, in theory, require the understanding of all the intricate biological aspects in-

volved. Thankfully, in practice, there exist more manageable approaches that employ simplified models of disease spread, which offer valuable insights into disease dynamics in numerous scenarios. Infectious disease modeling has proven invaluable in unraveling the intricate mechanisms governing the transmission of diseases, enabling us to gain insights into their propagation, anticipate the trajectory of outbreaks and assess the effectiveness of epidemic control strategies. Within this realm, the SI, SIR and SIS models stand as fundamental dynamical frameworks for infectious diseases. The aforementioned models fall under the category of deterministic compartmental models presented in the form of ordinary differential equations. Deterministic models operate under the assumption that the observed population precisely behaves as the model's description, thus excluding any random events within the population. In compartmental models individuals within a population are allocated into distinct subgroups or compartments, each representing a particular stage of the epidemic. In this chapter, our focus lies on these tractable approaches.

Historical background

Mathematical modeling in the realm of epidemics has a rich historical background that predates the study of networks. The foundations of mathematical modeling in the context of disease transmission can be traced back to the groundbreaking work of Daniel Bernoulli in 1760. Bernoulli, a trained physician, embarked into the realm of mathematics to construct a model that aimed to defend the practice of smallpox vaccination. Through his calculations, Bernoulli demonstrated that widespread vaccination against smallpox had the potential to increase life expectancy from 26 years and 7 months to 29 years and 9 months.

In 1906, William Hubert Hamer made a significant contribution to the understanding of infection spread by proposing that it depends on the number of susceptible and infective individuals [23]. This idea outlined in Hamer's work introduced the concept of a so-called mass-action law to describe the rate of new infections. Since then, this fundamental concept has formed the basis of compartmental models in epidemiology.

An illustrative example highlighting the power of mathematical modeling in epidemiology is the revolutionary work of doctor Ronald Ross on malaria. Dr. Ross's significant contributions earned him the prestigious Nobel Prize in Medicine in 1902, recognizing his remarkable explanation of the intricate dynamics underlying the transmission of malaria between mosquitoes and humans. At

that time, it was widely believed that complete removal of malaria was impossible as long as mosquitoes persisted within a population. However, Dr. Ross's work presented a concise compartmental model in 1911, incorporating both mosquitoes and humans, that challenged this prevailing notion [24]. His model demonstrated that reducing the mosquito population below a critical threshold could indeed lead to the elimination of malaria. Field trials subsequently provided empirical support for Dr. Ross's conclusions, paving the way for remarkable achievements in malaria control efforts. These triumphs stand as a testament to the remarkable effectiveness of applying mathematical models to combat infectious diseases. In addition, also dating back to the early twentieth century, the innovative efforts of Anderson McKendrick, a military medical practitioner and self-taught mathematician, laid the foundation for the emergence of compartmental models. The theoretical frameworks developed by McKendrick and his contemporaries constitute the core principles of traditional mathematical epidemiology, a well-established and extensively explored discipline.
It is noteworthy that the groundwork for the entire approach to epidemiology using compartmental models was laid not by mathematicians, but by esteemed public health physicians such as Sir R. Ross, W.H. Hamer and A. McKendrick between 1900 and 1935.

5.1.1 The SI model

In the conventional mathematical depiction of an epidemic, the intricate dynamics occurring within individual hosts are reduced to transitions between a limited number of fundamental disease states. In its simplest form, the model comprises two states: susceptible and infected. Individuals in the susceptible state have not yet contracted the disease but are susceptible to infection upon contact with an infected individual. On the other hand, individuals in the infected state have contracted the disease and have the potential to transmit it if they come into contact with susceptible individuals. While this binary classification overlooks many detailed biological complexities, it captures some fundamental characteristics of disease dynamics and serves as a valuable simplification, particularly when our focus is primarily on the interplay between networks and populations rather than the internal dynamics of individual host organisms.

The traditional approach to studying disease spread takes a rather simplified perspective, avoiding any explicit discussion of contact networks. It adopts a fully mixed or mass-action approximation. *The Law of Mass Action* is a term widely applicable in the realm of chemistry, and it is an assumption that ensures

that the rate of a chemical reaction is directly proportional to the concentration of the reactants.
In the context of population dynamics and epidemic modeling, the Law of Mass Action is used to describe how individuals move between two interacting categories within a population and presumes that people within a population interact homogeneously which in fact means that every individual has an equal chance, over a given time period, of encountering any other individual. In addition, the frequency of interaction between two distinct subgroups of the population is directly linked to the product of the respective sizes of those subgroups. In this model, people interact in a random and unrestricted manner. However, it is important to acknowledge that this representation does not accurately reflect the complexities of the real world. In reality, people have limited contact with only a fraction of the global population, and these interactions are far from random. This is precisely why the study of networks becomes pivotal in understanding the transmission of diseases. Despite its lack of realism, this assumption enables the use of reached conclusions about the spread of a disease in one population to predict the spread of that disease in another population, regardless of their size. So, it is still worthwhile to explore the principles underlying the traditional approaches, as they provide a foundation for our investigation into network epidemiology.

Let us imagine a scenario where a disease is spreading throughout a population of individuals. In this context, we define $S(t)$ as the number of individuals who are susceptible to the disease at time t, and $X(t)$ as the number of individuals who are infected at time t. However, it is important to note that due to the random nature of disease transmission, these numbers are not fixed and can vary. If we were to repeat the disease spreading process multiple times under similar conditions, we would likely obtain different numbers each time. To address this variability, we can purify our understanding of S and X as the average or expected numbers of susceptible and infected individuals. This means that we consider running the disease spreading process multiple times under identical conditions and calculate the average values of S and X based on the outcomes.
Naturally, the number of individuals who become infected increases when susceptible individuals contract the disease from infected individuals. Assume a scenario where people randomly interact and make contacts, which are sufficient for the spread of the disease. We can presume that the rate at which these contacts occur is β per individual, meaning that each person has an average of β contacts with randomly chosen individuals per unit time. To visualize the transitions between different states, we can use flow charts. Here is a simple flow chart

representing the SI model:

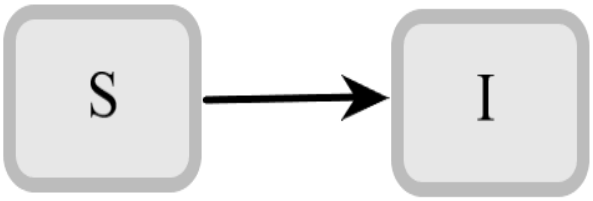

Figure 5.1: SI model flow chart.

The transmission of the disease occurs only when an infected person comes into contact with a susceptible individual. If the total population consists of n people, then the average probability of encountering a susceptible person at random is $\frac{S}{n}$. Therefore, an infected person has contact with an average of $\beta\frac{S}{n}$ susceptible individuals per unit time. Considering that there are, on average, X infected individuals in the population, the overall average rate of new infections can be expressed as $\beta\frac{SX}{n}$. We can now formulate a differential equation to describe the rate of change with respect to time t of the number of infected individuals X:

$$\frac{dX}{dt} = \beta\frac{SX}{n} \tag{5.1}$$

At the same time, the number of susceptible individuals S decreases at the same rate as new infections occur. Therefore, we can write the corresponding differential equation for the rate of change of S with respect to time t:

$$\frac{dS}{dt} = -\beta\frac{SX}{n} \tag{5.2}$$

To simplify the analysis, it is commonly preferred to introduce variables that represent the fractions of susceptible and infected individuals. As such, we express the following relations:

$$s = \frac{S}{n}, \quad x = \frac{X}{n}. \tag{5.3}$$

Hence the equations Eq.(5.1) and (5.2) can be written as

$$\frac{dx}{dt} = \beta s x, \tag{5.4}$$

$$\frac{ds}{dt} = -\beta s x \tag{5.5}$$

The mathematical model described, where individuals are either susceptible or infected, is known as the fully mixed susceptible-infected (SI) model. In the SI model, there is no recovery or removal of individuals from the infected state. Once an individual becomes infected, they remain infected throughout the course of the epidemic. Therefore, the total population size remains constant, so we can express the relationship between the susceptible and infected individuals using the equation $S + X = n$, or equivalently $s + x = 1$. By substituting $s = 1 - x$ in the Eq.(5.4), we can eliminate the variable s and express the dynamics of the system solely in terms of the variable x, which represents the proportion of infected individuals. This simplification allows us to focus on the dynamics of the infection without explicitly considering the susceptible individuals. We gain:

$$\frac{dx}{dt} = \beta(1 - x)x. \tag{5.6}$$

The previous equation, known as the logistic growth equation, is encountered in various fields such as biology, physics and other domains. By employing conventional mathematical techniques, the equation can be solved, resulting in the expression

$$x(t) = \frac{x_0 e^{\beta \mathrm{t}}}{1 - x_0 + x_0 e^{\beta \mathrm{t}}}, \tag{5.7}$$

where x_0 represents the initial value of x at time $t = 0$. The resulting pattern, as depicted in Figure 5.2, takes the form of an S-shaped "logistic growth curve", representing the proportion of infected individuals. During the initial phase, the curve experiences exponential growth, signifying the high susceptibility of the population. However, as the number of susceptibles diminishes over time, the disease encounters increasing difficulties in finding new victims, causing the curve to reach a saturation point.

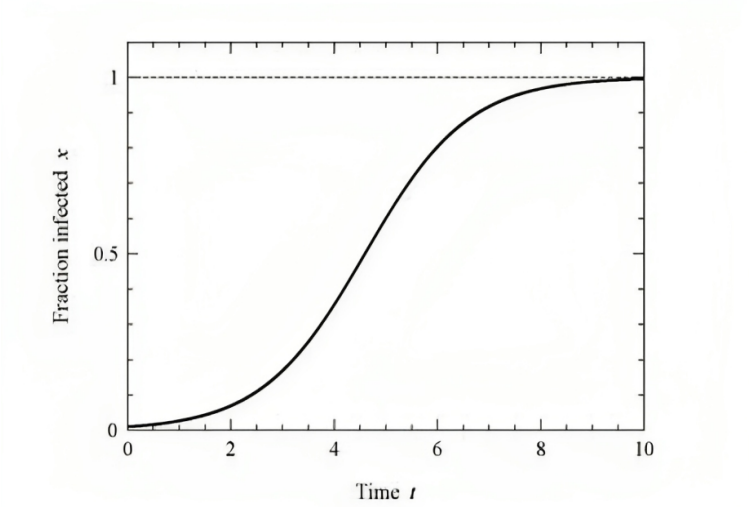

Figure 5.2: The classic logistic growth curve of the SI epidemic model. Source: [2]

5.1.2 The SIR model

The SI model represents the most elementary approach of infection modeling. Nevertheless, to increase realism or to better capture the characteristics of specific diseases, various extensions can be employed. One common extension addresses the concept of recovery from the disease. In the SI model, once individuals become infected, they remain in the infected state indefinitely and continue to spread the infection. However, in reality, many diseases involve a recovery period during which the immune system successfully eliminates the pathogen, leading to the restoration of health. Moreover, individuals often develop immunity to the disease following recovery, making them resistant to reinfection. To incorporate these aspects into the model, a third disease state is introduced, typically referred to as "recovered" and denoted by R. This extended framework is known as the susceptible-infected-recovered (SIR) model.

In certain diseases, individuals do not experience recovery but instead succumb to the infection after a certain interval and pass away. Although this outcome stands in direct contrast to the notion of recovery in human terms, it holds little significance in the field of epidemiology. From an epidemiological perspective, whether an individual is immune or deceased, they are effectively eliminated as potential hosts for the disease. In our model, both recovery and death are represented by the R state. Diseases characterized by mixed outcomes,

where individuals can either recover or die, can also be accommodated within this framework. Mathematically speaking, it is inconsequential whether individuals in the R state are classified as "recovered" or "dead". Hence, some individuals opt to interpret the R state as "removed" to encompass both possibilities, leading to the usage of the susceptible-infected-removed (SIR) model.

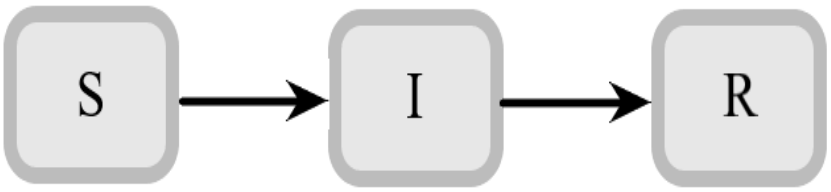

Figure 5.3: SIR model flow chart.

In the fully mixed susceptible-infected-recovered (SIR) model, the dynamics unfold in two stages. In the initial stage, susceptible individuals become infected through contact with infected individuals. The rate of contact between individuals remains β per person, as discussed earlier. In the subsequent stage, infected individuals recover (or pass away) at a constant average rate γ. Also, once some individual recovers or dies, it will never get infected again.

By knowing the value of γ, we can determine the expected duration of infection, denoted as τ, for an infected individual prior to their recovery. The probability of recovering within a small time interval $\delta\tau$ is $\gamma\delta\tau$, while the probability of not recovering within that interval is $1 - \gamma\delta\tau$. This information allows us to determine the probability of an individual still being infected after a total time τ, as shown in the upcoming equation:

$$\lim_{\delta\tau\to 0}(1-\gamma\delta\tau)^{\frac{\tau}{\delta\tau}} = e^{-\gamma\tau} \tag{5.8}$$

Furthermore, the probability, $p(\tau)d\tau$, that the individual remains infected for a duration between τ and $\tau + d\tau$, and then recovers, is obtained by multiplying this previous probability by $\gamma d\tau$, as expressed in equation which corresponds to a standard exponential distribution:

$$p(\tau)d\tau = \gamma e^{-\gamma\tau}d\tau \tag{5.9}$$

Consequently, although an infected person is most likely to recover shortly after infection, there exists the possibility of remaining in the infected state for an extended period, potentially exceeding the mean infectious time of $\frac{1}{\gamma}$.

Contrary to the exponential distribution assumed in the SIR model, the actual distribution of infection durations in most diseases exhibits a narrow peak centered around an average value. Nevertheless, these behaviors do not accurately capture the dynamics observed in most real diseases. In real-world scenarios, the duration of infection for the majority of individuals tends to group around a specific time period, such as a week or a month. Deviations from the average duration are relatively rare. Despite this limitation, we will continue employing this model for the sake of mathematical convenience.

The SIR model equations, which describe the fractions of individuals in the susceptible (s), infected (x), and recovered (r) states, can be expressed as follows:

$$\frac{ds}{dt} = -\beta s x, \tag{5.10}$$

$$\frac{dx}{dt} = \beta s x - \gamma x, \tag{5.11}$$

$$\frac{dr}{dt} = \gamma x. \tag{5.12}$$

These equations capture the transitions between the different states, reflecting the transmission of the disease from susceptible to infected individuals, the recovery or death of infected individuals, and the acquisition of immunity, respectively. It is through these equations that the SIR model elucidates the dynamics of infectious diseases.

In addition, the total population size remains invariant. Hence, the three variables satisfy

$$s + x + r = 1. \tag{5.13}$$

We can obtain the value of s from these equations, by first eliminating x from equations (5.10) and (5.12) and gaining

$$\frac{1}{s}\frac{ds}{dt} = -\frac{\beta}{\gamma}\frac{dr}{dt},$$

and then integrating both sides with respect to t to obtain

$$s = s_0 e^{-\frac{\beta r}{\gamma}} \tag{5.14}$$

where s_0 is the value of s at the initial time $t = 0$, and the constant of integration was chosen so that there are no individuals in the recovered state at $t = 0$. Other options are possible but this one will be used for now.

Now using that $x = 1 - s - r$ and placing it in the equation (5.12) with the help of (5.14), we get

$$\frac{dr}{dt} = \gamma \left(1 - r - s_0 e^{-\frac{\beta r}{\gamma}}\right). \tag{5.15}$$

If this equation is solved for r, and s is found from Eq.(5.14) and x from Eq.(5.13), the solution is given by

$$t = \frac{1}{\gamma} \int_0^r \frac{du}{1 - u - s_0 e^{-\frac{\beta u}{\gamma}}}. \tag{5.16}$$

Unfortunately, in practice we cannot evaluate the integral in closed form, but we can however evaluate it numerically, for which an example is shown in the figure below. Here, the three curves demonstrate the fractions of the population in the susceptible, infected and recovered states as a function of time. The parameters are $\beta = 1, \gamma = 0.4, s_0 = 0.99, x_0 = 0.01$ and $r_0 = 0$.

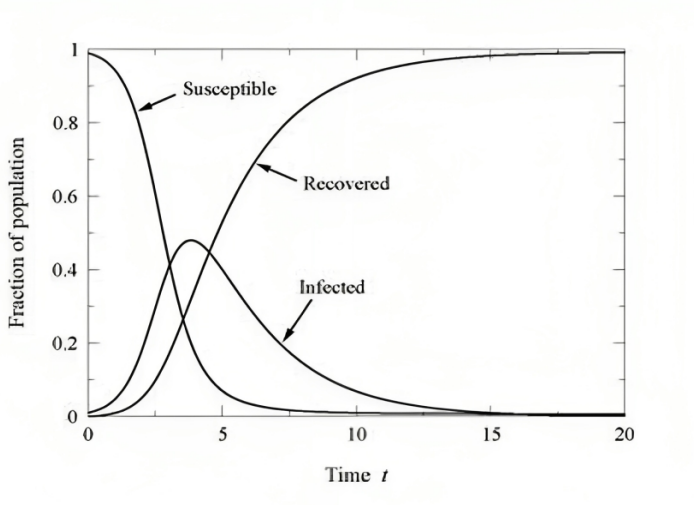

Figure 5.4: Time evolution of the SIR model. Source: [2]

Several noteworthy observations can be made regarding this figure. Firstly, the fraction of susceptible individuals in the population decreases monotonically as more susceptibles become infected, while the fraction of recovered individuals monotonically increases. Conversely, the fraction of infected individuals initially rises as infections occur, then declines as individuals recover, ultimately approaching zero as the time approaches infinity.
Importantly, the susceptible population does not reach zero. The curve representing $s(t)$ concludes slightly above the axis. This is due to the absence of infected individuals $(x \to 0)$, resulting in no further transmission to the remaining susceptibles. Individuals who manage to avoid infection for a prolonged period are unlikely to contract the disease at all, symbolizing their fortunate survival throughout the outbreak. Similarly, the fraction of recovered individuals fails to reach one as time progresses.

The asymptotic value of r represents the total number of individuals who become infected throughout the entire epidemic duration, serving as a measure of the outbreak's overall magnitude. This value can be derived by solving equation (5.15) at the point where $\frac{dr}{dt} = 0$, resulting in

$$r = 1 - s_0 e^{-\frac{\beta r}{\gamma}} \tag{5.17}$$

When determining the initial conditions for the model, there are different approaches to consider, but the most common one assumes that the disease originates from either a single infected individual or a small group of c individuals, while the remaining population is in the susceptible state. In other words, the initial values of the variables are

$$s_0 = 1 - \frac{c}{n}, x_0 = \frac{c}{n}, r_0 = 0$$

In the scenario of a substantial population size $(\mathrm{n} \to \infty)$, we can simplify the notation to $s_0 \approx 1$, leading to the expression $r = 1 - e^{-\frac{\beta r}{\gamma}}$ for the final value of r.

If $\beta \leq \gamma$, the size of the epidemic gradually diminishes until it reaches zero. This occurs because when $\beta \leq \gamma$, the recovery rate of infected individuals surpasses the rate at which susceptible individuals become infected preventing the disease from establishing a strong presence in the population. As a result, the number of infected individuals, which starts at a small value, decreases instead of increasing, resulting in the extinction of the disease rather than its spread.

The fundamental epidemiological metric is the basic reproduction number, denoted as R_0, which signifies the potential for an epidemic's development and its

speed. R_0 is calculated by considering the early stages of a disease outbreak when there are only a few cases, and the majority of the population is susceptible. This is called a naive population. It represents how many more people, on average, one infected person will infect before recovering. For instance, if each infected person infects two others on average, $R_0 = 2$. If half of them infect one person, and the rest infect none, $R_0 = \frac{1}{2}$, and so on.

- In the case of $R_0 = 2\,(R_0 > 1)$, each infected person, on average, infects two others. This means that the number of new cases doubles in each iteration, resulting in exponential growth and epidemic.
- Conversely, if $R_0 < 1$, the disease experiences exponential decline, eventually dying out. The pivotal point is $R_0 = 1$, serving as the threshold between these growth and decline scenarios. Below $R_0 = 1$, the disease shrinks, while above it, the disease spreads.
- If $R_0 = 1$, it indicates that one infected person, on average, infects one susceptible person. In this context, the disease becomes *endemic*, that is, always present in the population with a relatively unchanging number of infected individuals.

To calculate R_0 in our model, we consider that if an individual remains infectious for a duration τ, they will come into contact with approximately $\beta\tau$ other individuals during this period. R_0 is defined for a naive population where all individuals in contact are susceptible, so $\beta\tau$ represents the total number of individuals the infected person will infect. We find the average R_0 by taking the τ distribution into account, as described in Eq. (5.9).

$$R_0 = \beta\gamma \int_0^\infty \tau e^{-\gamma\tau} d\tau = \frac{\beta}{\gamma}$$

5.1.3 The SIS model

An alternative variation of the SI model explores the concept of reinfection, where individuals can be infected multiple times due to diseases that either do not provide immunity for their victims after recovery or offer limited immunity. This model is useful for simulating the transmission of diseases caused by bacteria. In these situations, individuals do not acquire immunity to further infections caused by the same bacterium.

Hence, this leads to the development of the SIS model, a simplified framework consisting of two states: susceptible and infected. In this model, individuals transition back to the susceptible state upon recovery.

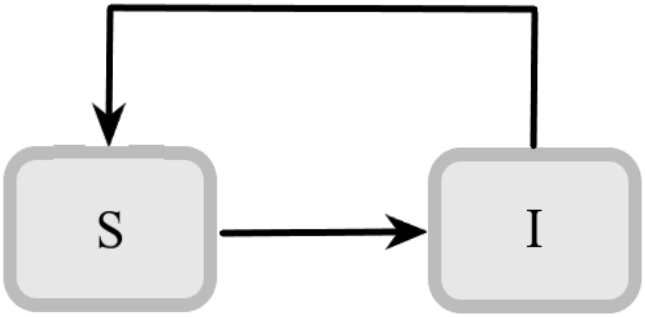

Figure 5.5: SIS model flow chart.

The mathematical differential equations managing the dynamics of the SIS model are as follows:

$$\frac{ds}{dt} = \gamma x - \beta s x, \tag{5.18}$$

$$\frac{dx}{dt} = \beta s x - \gamma x. \tag{5.19}$$

Once more, the rate of contact between individuals is β per person. Once an individual recovers, there exists a distinct probability, γ, for them to be reinfected, which is different from the probability of initial infection for susceptible individuals. Additionally, the total population size remains constant, leading to the equation $s + x = 1$.
By substituting $s = 1 - x$ into equation (5.19), we derive

$$\frac{dx}{dt} = (\beta - \beta x - \gamma)x \tag{5.20}$$

for which the solution is

$$x(t) = \left(1 - \frac{\gamma}{\beta}\right) \frac{Ce^{(\beta-\gamma)t}}{1 + Ce^{(\beta-\gamma)t}}, \tag{5.21}$$

where the integration constant C is fixed by the initial value of x to be

$$C = \frac{\beta x_0}{\beta - \beta x_0 - \gamma}. \tag{5.22}$$

In scenarios where the population size is large and the number of initial disease carriers is small, leading to x_0 approaching zero and $C = \frac{\beta x_0}{\beta - \gamma}$, we can derive a simpler solution expressed as

$$x(t) = x_0 \frac{(\beta - \gamma)e^{(\beta-\gamma)t}}{\beta - \gamma + \beta x_0 e^{(\beta-\gamma)t}}. \tag{5.23}$$

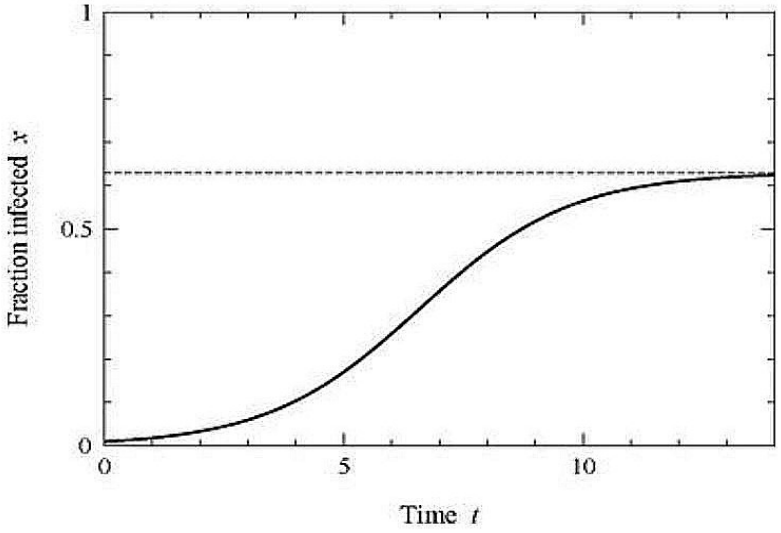

Figure 5.6: Fraction of infected individuals in the SIS model. Source: [2]

If β is greater than γ in the SIS model, a logistic growth curve similar to the basic SI model is observed in the figure above. However, a key distinction is that the entire population does not become infected by the disease. Instead, the system reaches a stable state in the long run, where the infection and recovery rates are perfectly balanced. Consequently, a constant fraction of the population remains infected at any given time. It is important to note that the identity of the infected individuals changes over time as some recover and others become infected.

By setting the derivative $\frac{dx}{dt}$ to zero in Equation (5.20), the fraction of infected individuals can be obtained as

$$x = \frac{\beta - \gamma}{\beta}. \tag{5.24}$$

In the field of epidemiology, as mentioned, this stable state is commonly referred to as an *endemic disease state*.

An interesting observation in the SIS model is that as the rate β approaches rate γ, the fraction of infected individuals in the endemic state tends towards zero. Conversely, when β is smaller than γ, the disease is projected to diminish exponentially, as predicted by Eq. (5.23).

In the vast landscape of epidemiology, numerous other models have emerged to capture the dynamics of specific diseases. These models go beyond the basic frameworks and introduce additional states to account for various aspects of disease transmission. For instance, the SIRS model, where individuals experience recovery from infection and acquire immunity, similar to the SIR model. However, this immunity is not permanent, and individuals eventually lose it after a specified time, rendering them susceptible once more. Also, an "exposed" state may be incorporated to represent individuals who have contracted the disease but have yet to reach the infectious stage in order to transmit it to the rest. Alternatively, an initial immune state can be included prior to the susceptible state to reflect the immunity passed on from mothers to newborns. Furthermore, there exist models that account for population dynamics, allowing for the arrival of new individuals through birth or immigration. Additionally, certain models make a distinction between individuals who fully recover from the disease and those who become carriers capable of transmitting the infection to others. However, we will not pursue these models further since the models already seen will be sufficient for our subject of interest.

5.2 Epidemic models on networks

In the preceding section, we explored the standard approach to epidemic modeling, which assumes full mixing of the population. This assumption implies that each individual has the potential to come into contact with any other individual, and these contacts, with a transmission probability of β per unit time, contribute to the spread of the disease.

However, in reality, it is unreasonable to assume that any two individuals have the potential to cross paths. Instead, most individuals have a defined circle of acquaintances, family, friends, neighbors and coworkers, with whom they interact

regularly. Consequently, the majority of the world's population can be safely disregarded.
To capture the essence of these interpersonal dynamics, the concept of a contact network emerges, shedding light on the underlying structural complexities that shape the trajectory of disease transmission throughout the population.

When it comes to modeling disease dynamics, network models function in a manner similar to the fully mixed models we have previously encountered, albeit with a key difference - the utilization of a network of potential contacts instead of assuming universal contact within the entire population. The population's contact network is portrayed as a graph. Here, individuals are nodes, and the edges denote contact links that enable infection transmission. Within the graph, each node symbolizes an individual within the population. The degree of a node denotes the count of connections associated with that individual. In this context, we define the transmission rate or infection rate as the probability per unit time that infection will be transmitted between two individuals who are connected by an edge in the network, one being susceptible and one infected. To put it differently, it represents the rate at which sufficient contact for disease transmission occurs between any two individuals linked by an edge. It is important to note that the transmission rate in the fully mixed case differs from its counterpart in the network case. In the fully mixed models, it represents the rate of contacts between an infected individual and all others in the population. In contrast, in the network models, it denotes the rate of contacts with only one other individual. Despite this distinction, we employ the same notation β to facilitate comparison and comprehension.

The transmission rate, as a crucial factor in disease dynamics, encompasses both disease-specific characteristics and social and behavioral parameters of the population. On one hand, some diseases are inherently more transmissible than others, resulting in distinct transmission rates. On the other hand, it is important to recognize that the transmission rate is not solely determined by the disease itself. The social and behavioral norms within a population also play a significant role. For instance, in certain countries, it is customary for individuals with minor respiratory infections like colds to wear surgical face masks as a preventive measure. By adopting such practices, the transmission of the disease can be reduced. In contrast, other countries may lack such conventions, leading to a disparity in transmission rates.

By assigning a specific value to the transmission rate, it becomes possible to develop models that depict the spread of disease within a network. The models previously introduced in the initial section of this chapter can be extended to

accommodate network scenarios. In the network variant of the SI model, for example, the network's vertices represent n individuals, the majority of whom are initially in the susceptible state at time $t = 0$. Only a small fraction x_0, or even a solitary vertex, is in the infected state. Through a probability β per unit time, infected nodes transmit the disease to their susceptible neighbors, resulting in the gradual dissemination of the disease throughout the network. In the context of this model, it is evident that as $t \to \infty$, every individual who is susceptible to the disease will eventually become infected. This is due to the fact that infected individuals remain infectious indefinitely. Consequently, their susceptible neighbors will inevitably contract the disease, regardless of the transmission rate (as long as it is not zero). The only requirement for infection is that a vertex must have at least one path connecting it to an infected individual in the network, enabling the disease to reach and affect them.

A graph is considered connected if there exists a sequence of edges connecting any two nodes. In reality, most contact networks are described by disconnected graphs, which consist of components. These components exhibit the property that nodes from different components lack a connection, yet nodes within the same component are mutually reachable. In real-world networks, a common observation is that one substantial component predominates, encompassing more than half of the total graph, while numerous smaller components exist.
In the scenario where an outbreak originates from a single infected individual, over an extended period, only individuals within the same component of the network will be affected. Other distinct components remain unaffected by the outbreak. Consequently, in the asymptotic limit of prolonged durations, the disease will spread from each initial carrier to infect all vertices that are reachable within the carrier's component. In the simplest instance, where the disease initiates with a solitary infected carrier, only one component will experience infection.

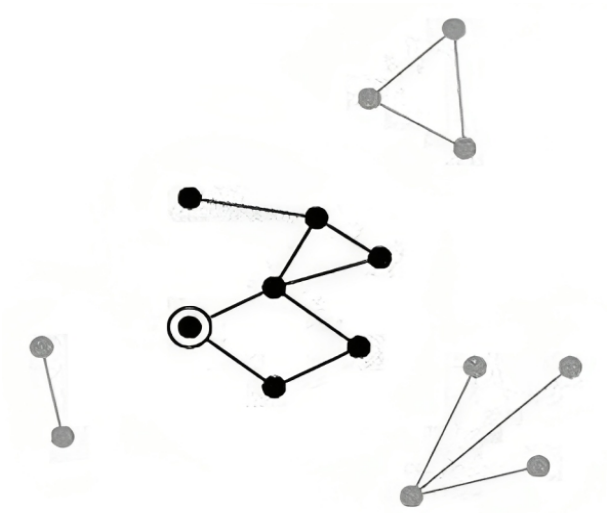

Figure 5.7: Network components with a single infected individual. Source: [2]

Accordingly, the magnitude of the disease's influence on the population is directly determined by two key factors: the underlying network structure and the initial location of the infected individual within the network. In the event that the infected individual is situated within a large component of the network, it is highly probable that the disease will propagate extensively, infecting a significant fraction of the population. Conversely, if the individual is positioned within a small component, the disease may have a limited reach, initially infecting only a few individuals before eventually diminishing and dying out.

Besides the structure of the network, its connectivity, contact patterns meaning how frequently individuals interact and with whom they interact, the starting point of the transmission, and the transmission rate β along with social and behavioral changes, there are numerous factors that influence the spread, both for epidemic and information. One that can also significantly alter the dynamics is the implementation of interventions, such as vaccination campaigns, quarantine measures or information dissemination strategies. In some cases, the physical location of individuals within a network or geographic factors can also have an affect on the spread. Furthermore, real-world processes often exhibit stochastic behavior, with probabilistic transmission contributing to variability in epidemic outcomes. In addition, the ability of a network to withstand and recover from disruptions, such as targeted interventions or the removal of influential nodes, is a critical aspect worth evaluating.

5.2 Epidemic models on networks

In order to predict and control the epidemic and information dissemination precisely, it is crucial to gain an understanding of how these factors combine in specific network models. Researchers utilize mathematical models, computational simulations and data analysis in their investigations, offering valuable insights for shaping public health strategies or adjusting marketing campaigns, depending on the context of the study.

Chapter 6

Rumor propagation models

The examination of rumors found its early motivation and influence within the realm of epidemic study. During the period spanning the 1940s to the 1960s, it was commonly viewed as a component of the extensive investigation into epidemics. Although often coupled in discussion, the research of rumors is significantly less progressed than the study of epidemics. This divergence is partly due to the relatively intricate nature of the rumor dynamics. This complexity primarily originates from differences in the way removal occurs, which makes the analysis distinct and more challenging than the one in epidemics. In epidemic scenarios, individuals attain immunity through mortality or recovery. Conversely, in rumors, those who spread the rumor can transition to not spreading the rumor in two distinct ways: through interactions between people who transmit the rumor, or through interactions with people who do not transmit the rumor.

Rumors can be viewed as an "infection of the mind", and within the interpersonal connections, the diffusion of rumors mirrors certain aspects of infectious disease spread, particularly in propagation principles and population classification. Consequently, this has led several scholars to analyze rumor propagation processes by drawing insights from the study of infectious disease transmission. However, they are not entirely equivalent. The earliest exploration into parallels and distinctions between rumors and epidemics was initiated in 1965 by scientists D.J. Daley and D.G. Kendall. Following this, the study of rumor propagation has evolved as an independent phenomenon, distinct from the realm of epidemics.

6.1 Stochastic models

The models elucidating the dissemination of information or the propagation of rumors serve as analytical tools to enhance our comprehension of social phenomena. Distinct mathematical models are typically constructed as stochastic models, from which deterministic counterparts are derived. This represents the logical methodology for developing mathematical models in discrete populations, whether they are finite or infinite. It is worth noting that, under appropriate conditions within the model, differential equations are used to approximate the stochastic formulation, rather than the other way around.

"Stochastic", a term originating from Greek, signifies randomness or chance. Stochastic models are mathematical representations used to predict outcomes of experiments or random events unfolding over time, their probabilities considered. These models have broad applicability, spanning fields like geophysics, social science, earthquake size and location prediction, pressure pulsations, atmospheric temperature, environmental problem-solving, stock prices in finance, purchasing behavior, many examples in service and control theory, and manufacturing systems. This chapter particularly focuses on the extensive use of stochastic modeling in examining the dissemination of information or rumors within populations of individuals.

The deterministic approach represents the process as a system of differential equations while the stochastic approach describes the process as either a discrete or continuous finite state Markov process. Frequently, it becomes necessary to resort to a deterministic model to facilitate numerical solution procedures. Nonetheless, it is essential to recognize that the deterministic model is fundamentally an approximation.

6.1.1 The Daley-Kendall model

Suppose a closed, homogeneously mixing population of total size S (where $S = N + 1$). In other words, suppose a constant collection of the same set of individuals who are equally likely to interact with each other in a given (small) interval of time.
At any given time, individuals from this population can be categorized into three distinct, non-overlapping categories consisting at time $t > 0$ of:

(A) $X(t)$ individuals who remain uninformed about the rumor,

(B) $Y(t)$ individuals actively involved in disseminating the rumor, and

(C) $Z(t)$ individuals who have encountered the rumor but have discontinued their involvement in its dissemination.

Initially, $X(0) = N, Y(0) = 1$ and $Z(0) = 0$, while the overall population size remains constant, ensuring that the equation

$$X(t) + Y(t) + Z(t) = N + 1 \tag{6.1}$$

holds true at all times.
This mathematical relation offers a means for eliminating any mention to the variable Z.

The classes labeled as (A), (B) and (C) whose members are called ignorants, spreaders and stiflers, respectively, may initially feature resemblance to the groups of susceptibles, infectious and removed commonly encountered in epidemic SIR model.

The propagation of the rumor throughout the population relies on interactions between ignorants and spreaders, following the Law of Mass Action. To be specific, in any pairwise encounter involving a spreader, an attempt is made to transmit the rumor to the other person. This other individual could be an ignorant, a spreader or a stifler. In the first scenario, the ignorant becomes a spreader, while in the other two scenarios, those engaged in the encounter realize that the rumor is already known, prompting them to refrain from further dissemination, thus joining the stiflers as a result of this "stifling experience". In the context of pairwise meetings happening at rate β, there are three types of encounters that occur at any given time t with rates relative to the populations of each category, that is:

- ignorant $\leftrightarrow$ spreader, with rate $\beta X(t)Y(t)$
- spreader $\leftrightarrow$ spreader, with rate $\beta\frac{1}{2}Y(t)[Y(t) - 1]$
- stifler $\leftrightarrow$ spreader, with rate $\beta Y(t)Z(t)$

To conveniently express the changes over a small time interval $(t, t + h)$, we can use the following format:

$$\Delta_h(X, Y)(t) \equiv (X, Y)(t + h) - (X, Y)(t) \tag{6.2}$$

Hence, for the sake of an illustration, the interaction between an ignorant and a spreader yields $\Delta_h(X, Y)(t) = (-1, 1)$.

The transition rates governing state changes in the population are determined by the frequencies of these meetings. From there, we derive relationships while making approximations that ignore terms with minimal impact $(o(h))$:

$$\begin{aligned}
\Pr\{\Delta_h(X,Y)(t) = (-1,1) \mid (X,Y)(t) = (x,y)\} &= \beta xyh, \\
\Pr\{\Delta_h(X,Y)(t) = (0,-2) \mid (X,Y)(t) = (x,y)\} &= \beta\frac{1}{2}y(y-1)h, \\
\Pr\{\Delta_h(X,Y)(t) = (0,-1) \mid (X,Y)(t) = (x,y)\} &= \beta yzh, \\
\Pr\{\Delta_h(X,Y)(t) = (0,0) \mid (X,Y)(t) = (x,y)\} &= 1 - \beta y\left(N - \frac{1}{2}(y-1)\right)h.
\end{aligned} \tag{6.3}$$

with all other transitions having probability $o(h)$.
It is evident that β functions as a time constant, so through an appropriate selection of time units, we can set β to 1.

We will now present two straightforward variations of the basic Daley-Kendall model, maintaining our focus on models where rumor transmission occurs through pairwise interactions. In the first variant, referred to as the "k-fold stifling model", we suppose that a spreader continues to disseminate the rumor until encountering k stifling experiences, where k is a positive integer $(k \geq 1)$. In essence, a spreader persists until they have had "unsuccessful tellings" k times. By categorizing the $Y(t)$ spreaders into one of k exclusive groups, each representing individuals with j stifling experiences $(j = 0, \ldots, k-1)$, we can maintain a Markovian description of the process. This involves employing an augmented $(k+1)$ -dimensional vector $(X, Y_1, \ldots, Y_k)(t)$, where $(Y_1 + \cdots + Y_k)(t) = Y(t) = N + 1 - X(t) - Z(t)$, just like before.
Of course, more intricate models can be developed, such as introducing randomness into the critical number, k, of rebuffs.

We recognize that the original Daley-Kendall model oversimplifies certain aspects. Specifically, it assumes that

(1) every encounter between two spreaders inevitably leads to an attempt to propagate the rumor, and

(2) a spreader transforms into a stifler solely after one stifling experience.

Therefore, in the second variant, known as the (α, p)-probability variant, consider a scenario where, during a pairwise meeting, a spreader attempts to disseminate the rumor with a probability p. Furthermore, suppose that upon such

an attempt, each spreader involved independently decides, with a probability α to become a stifler. In the original model, both p and α are set to 1. However, we now introduce variations where $0 < p \leq 1$ and $0 < \alpha \leq 1$. In this context, when $(X, Y)(t) = (x, y)$, several events can unfold within the time interval $(t, t + h)$. These include (neglecting terms of order $o(h)$):

- (X, Y) meeting resulting in $\Delta_h(X, Y) = (-1, 1)$ with probability of $pxyh$, or
- (Y, Y) meeting resulting in $\Delta_h(X, Y) = (0, -1)$ with probability of $p(2 - p)\alpha(1- \alpha)y(y - 1)h$, or
- (Y, Y) meeting resulting in $\Delta_h(X, Y) = (0, -2)$ with probability of $p(2 - p)\alpha^2\frac{1}{2}y(y-1)h$, or
- (Y, Z) meeting resulting in $\Delta_h(X, Y) = (0, -1)$ with probability of $payzh$, or
- no change with probability $1-p\left[(1 + \alpha)x + \left(1 - \frac{1}{2}p\right)\alpha(2 - \alpha)(y - 1)\right]yh$.

6.1.2 The Maki-Thompson model

Another classic model for rumor propagation was presented by scholars D.P. Maki and M. Thompson in 1973. Continuing in our established manner, we categorize and quantify the population as (X, Y, Z). The propagation of the rumor occurs through directed interactions involving the $Y(t)$ spreaders at time t and others within the population. These interactions result in two fundamental outcomes:

(a) when a specific spreader encounters an ignorant individual, the ignorant becomes a spreader, and

(b) when the spreader interacts with another spreader or a stifler, the initiating spreader becomes a stifler.

There are no other micro-level transitions considered in this model.

Consequently, the model encompasses merely two elementary transitions at infinitesimal rates: $\Delta_h(X, Y) = (-1, 1)$ at rate $XYh + o(h)$ and $\Delta_h(X, Y) = (0, -1)$ at rate $Y(Y - 1 + Z)h + o(h)$.

Accordingly, disregarding terms $o(h)$:

$$\begin{aligned} \Pr\{\Delta_h(X, Y)(t) = (-1, 1) \mid (X, Y)(t) = (x, y)\} &= xyh, \\ \Pr\{\Delta_h(X, Y)(t) = (0, -1) \mid (X, Y)(t) = (x, y)\} &= y(y - 1 + z)h, \\ \Pr\{\Delta_h(X, Y)(t) = (0, 0) \mid (X, Y)(t) = (x, y)\} &= 1 - Nyh. \end{aligned} \tag{6.4}$$

It is worth highlighting that due to the spreader having directed contact, the term $\frac{1}{2}y(y-1)$ found in Eq. (6.3) is replaced by $y(y-1)$ in Eq.(6.4). Apart from this change, the model remains akin to the Daley-Kendall model, as it still operates on the premise that a spreader transitions into a stifler solely through interactions between the initiating spreader and other members of the population.

The SI, SIR and SIS epidemic models have also been proposed as potential models for explaining how rumors circulate [29], [33], [35]. The SIR epidemic model involves two possible fundamental transitions over a small time interval of length h, specifically $\Delta_h(X,Y)(t) = (-1,1)$ or $(0,-1)$. While the first transition can arise from an (X,Y) contact, the second does not necessitate any interaction between individuals in the population at all. It is sufficient for the cessation of rumor spreading to happen solely due to a spreader permanently "forgetting" to share the rumor with those they encounter in the population.

When it comes to modeling factors that contribute to the decline in the spread of a rumor, we might consider things like forgetfulness or simply not feeling like sharing the rumor anymore. However, it does not seem very realistic to describe human behavior by making forgetfulness the sole reason for a rumor to stop spreading, just as it would be unreasonable to rely entirely on the stifling mechanism. Diseases come in different strengths, so rumors can also vary in how urgently people want to spread them. Hence, it might make sense to use both stifling and forgetfulness mechanisms. In that case, the relative rates of pairwise contact (for spreading and stifling) and individual forgetfulness would likely impact how far the rumor spreads.

6.2 Deterministic analysis of rumor models

The models we have outlined to describe how rumors spread can be mathematically represented as continuous-time Markov processes. Analyzing them using explicit algebra is not straightforward, especially when it comes to finding the transition probabilities for the various parameter values. Particularly, our primary concern revolves around the fraction f within the initial ignorant population $X(0) = N$ who eventually become aware of the rumor, so that

$$\lim_{t\to\infty} X(t) = N - Nf, \quad \lim_{t\to\infty} Y(t) = 0, \quad \lim_{t\to\infty} Z(t) = 1 + Nf \tag{6.5}$$

For this reason, we use deterministic versions of these models for our analysis.

In the deterministic version of Eq. (6.2), we introduce continuous functions $x(t), y(t)$ and $z(t)$, each of which is differentiable with respect to time t. These functions share identical initial values: $x(0) = X(0) = N, y(0) = Y(0) = 1, z(0) = Z(0) = 0$. It is imperative that they adhere to the equation

$$x(t) + y(t) + z(t) = N + 1 \tag{6.6}$$

which holds for a closed population.
The differential equations are given as:

$$\dot{x} = \frac{dx}{dt} = \sum_{\text{simple transitions}} (\text{change in } x) * (\text{rate of change}) = (-1)xy$$
$$\dot{y} = \frac{dy}{dt} = (+1)xy + (-2)\frac{1}{2}y(y-1) + (-1)yz = y(x - y + 1 - z) = y(2x - N) \tag{6.7}$$

A point of significance here is that, because of our use of Eq. (6.6) to eliminate z, N emerges as a coefficient within the differential equations. Additionally, N is a consequence of its distinct appearance in the initial values.

The last two equations provide us with:

$$\frac{dy}{dx} = -2 + \frac{N}{x}. \tag{6.8}$$

We will not delve into specifics, but they used the solution derived from their deterministic model's ordinary differential equations to establish a constant of motion, $\lambda(x, y)$, and replaced it with a random variable, $\lambda(x(t), y(t))$. An important outcome was that, under the given initial conditions with N initial ignorants, one initial spreader and no initial stiflers, the proportion of the population who never heard the rumor asymptotically approaches 0.203188. For further details, refer to [31], [32].

In a similar manner, the generalized (α, p)-probability variant of the fundamental Daley-Kendall model results in the following system of differential equations:

$$\begin{aligned}\dot{x} &= (-1)pxy \\ \dot{y} &= (+1)pxy + (-1)\left[p(2-p)2\alpha(1-\alpha)\frac{1}{2}y(y-1) + p\alpha yz\right] + (-2)p(2-p)\alpha^2\frac{1}{2}y(y-1) \\ &= py[x - \alpha(z + (2-p)(y-1))]\end{aligned} \tag{6.9}$$

Now these two equations yield

$$\frac{dy}{dx} = -(1+\alpha) + \frac{\alpha(N-1+p)}{x} + \frac{\alpha(1-p)y}{x} \tag{6.10}$$

In the limit as N approaches infinity, it can be deduced that the proportion of a sizable population that remains uninformed about the rumor at the point of its cessation is approximately 0.284668. For comprehensive information, refer to [31], [32].

In addition, the deterministic version of the k-fold stifling model introduces a set of equations for which now $y_i(t)$, for $i = 1, \ldots, k$, tells how many spreaders have met other spreaders or stiflers on different $i-1$ occasions. By summing them all up, we obtain $y = y_1 + \cdots + y_k$, and so the equations take the form

$$\begin{gathered}
\dot{x} = -xy, \\
\dot{y}_1 = xy - y_1\left(y_2 + \cdots + y_k + z\right) - 2\frac{1}{2}y_1\left(y_1 - 1\right) = xy - y_1(N-x), \\
\dot{y}_i = y_{i-1}\left(N - x - y_i\right) - y_i\left(N - x - y_{i-1}\right) = \left(y_{i-1} - y_i\right)(N-x), i = 2, \ldots, k,
\end{gathered} \tag{6.11}$$

The set of differential equations (6.11) seems to lack a fully solved closed-form solution for $k \geq 2$, thus necessitating numerical integration. For those interested, more in-depth exploration is available in [31], [32].

The deterministic examination of Maki and Thompson's model is outlined as follows:

$$\begin{aligned}
\dot{x} &= (-1)xy \\
\dot{y} &= (+1)xy + (-1)y(N-x) = y(2x-N)
\end{aligned} \tag{6.12}$$

In this case,

$$\frac{dy}{dx} = -2 + \frac{N}{x}$$

from which we can observe is the same as Eq.(6.8), so the percentage of individuals who remain uninformed will be 0.203188 , mirroring the same proportion as the Daley-Kendall model.

6.3 Model variations

A notable limitation of the class of models mentioned above is that they either disregard the topology of the underlying social interaction networks through which rumors propagate, assuming a homogeneous population, or employ overly simplistic network models. While these simplistic models may adequately capture the dynamics of rumor spread within small-scale social networks, primarily driven by person-to-person communication, they are unsuitable when applied to the dynamics of rumor dissemination in large social interaction networks, particularly those facilitated by the Internet. These extensive networks often consist of tens of thousands to millions of interconnected nodes with an exceptionally high number of social connections that exhibit intricate and distinctive connectivity patterns.

Up to this point, a multitude of researchers have delved into the examination of rumor transmission dynamics. They have extended and adapted the two classical rumor propagation models in order to gain a more profound understanding of how rumors spread. Here are some of the new directions and modifications that researchers have investigated:

- Heterogeneous networks: A notable development in this field involves the exploration of heterogeneous networks. In real-world scenarios, individuals often exhibit diverse levels of connectivity and influence. Researchers have analyzed how the dissemination of rumors varies within networks where certain individuals exhibit more extensive connections or influence compared to others.

- Multi-layer networks: Rumors can propagate through different communication channels, such as social networks, mass media and face-to-face interactions. This approach aims to capture the intricate nature of real-world information diffusion.

- Behavioral factors: Scholars have embarked on investigations regarding the influence of behavioral aspects, encompassing cognitive biases and emotional reactions, in the context of rumor transmission. These factors exert substantial influence on how individuals perceive and respond to rumors.

- Stochastic and agent-based approaches: Stochastic models and agent-based simulations provide researchers with the means to introduce randomness and individual-level behavior enhancing comprehension of uncertainty and unpredictability in the transmission of rumors.

- Geographical and spatial aspects: The emphasis on spatial models is particularly pronounced in contexts involving disease-related rumors or emergency situations. So, investigations have been undertaken to discern the impact of geography and spatial variables on the dynamics of rumor dissemination.
- Information integrity and the combat against misinformation: The studies of information quality discuss the differentiation between accurate information and falsehoods. They explain the mechanisms through which misinformation proliferates and explore strategies for countering its dissemination.
- Interventions and Control Measures: A segment of scholarly investigations is dedicated to the formulation of comprehensive methodologies designed to control and reduce the diffusion of rumors. This involves the strategic identification of nodes with significant influence, facilitating precise interventions, along with the optimization of the timing of information releases.

The aforementioned cases represent just a fraction of the ongoing research concerning the dynamics of rumor diffusion. Researchers are committed to continuous exploration and refinement of models to account for the complexity and diversity of real-world circumstances. In the upcoming chapter, our focus will be on an in-depth investigation of a more realistic model concerning the dissemination of rumors.

Chapter 7

A rumour spreading model for complex social networks

In this Chapter, I will present the review and results of the paper "Theory of rumour spreading in complex social networks" by M. Nekovee et. al [1].

The rumor model

Consider a population comprising N individuals who are categorized into three groups: ignorants, spreaders and stiflers. Building upon the work of Maki and Thompson, suppose that the rumor spreads via directed interactions initiated by spreaders with other individuals within the population. However, these interactions are restricted to occur exclusively along the edges of an undirected social interaction network denoted as $G = (V, E)$ where, naturally, V represents the network's vertices and E represents its edges. The dynamics of interactions between spreaders and the remaining population follow a set of specific rules:

1. In the event of a spreader interacting with an ignorant, the ignorant transitions to become a spreader at a rate λ.

2. In the event of a spreader interacting with another spreader or a stifler, the initiating spreader transitions to become a stifler at a rate α.

The first rule can be interpreted as a probabilistic representation of individuals' tendency to accept a rumor. This probability, loosely speaking, is influenced by factors like the rumor's urgency or credibility. On the other hand, the second rule

models the phenomenon where individuals tend to lose interest in propagating a rumor upon learning, through interactions with others, that the rumor has grown outdated or is incorrect.
As stated in the previous chapter, the discontinuation of rumor dissemination can occur through the stifling mechanism as well as the spreader's forgetfulness to disseminate the rumor or they lose their motivation to continue doing so. With regard to this crucial mechanism, individuals may cease rumor propagation spontaneously, without any external contact, at a rate denoted as δ. The propagation process initiates when one or more individuals gain awareness of a rumor and concludes when there are no longer any spreaders within the population.

Formulation of the model

The writers of this paper present an elucidation of this model in terms of Interacting Markov chains. Within this framework, they proceed to derive mean-field equations governing the dynamics of rumor propagation on complex networks characterized by arbitrary degree correlations.

An interacting Markov chain(IMC) comprises N interacting nodes, each having a state that evolves over time according to an internal Markov chain. Distinct from traditional Markov chains, the corresponding internal transition probabilities of an IMC depend not only on the individual node's current state but also on the states of all nodes to which it is linked. Consequently, the overall system's evolution adheres to a global Markov Chain, whose state space dimension is the product of the states characterizing each individual node. Given the presence of interactions between nodes within expansive networks, the task of solving the dynamics associated with IMCs becomes exceptionally challenging. While Monte Carlo simulations often provide viable solutions to such problems, there persists a preference for analytical solutions, even if they necessitate approximations. When it comes to the rumor model, each internal Markov chain can be in one of three states - ignorant, spreader or stifler.
Within the domains of physics and probability theory, mean-field theory is dedicated to the examination of the behavior exhibited by complex high-dimensional stochastic models. It accomplishes this by examining a simplified model that serves as an approximation to the original. This theory simplifies the representation of the actual contagion process by replacing specific local variables of the dynamics with their corresponding global mean values. In the following, we proceed to derive a set of coupled rate equations specifically tailored to this scenario. These equations, functioning on a mean-field level, effectively capture the

intricate dynamics of the Interacting Markov chains.

Let j be a node which in time t is in the ignorant state. Define p_{ii}^j as the probability that this node remains in the ignorant state during the time interval $[t, t + \Delta t]$, while p_{is}^j represents the complementary probability that it undergoes a transition to the spreader state, that is, $p_{is}^j = 1 - p_{ii}^j$. Consequently, we have the following relationship:

$$p_{ii}^j = (1 - \lambda \Delta t)^g, \tag{7.1}$$

where $g = g(t)$ corresponds to the number of neighboring nodes connected to node j that are in the spreader state at time t.
To simplify the microdynamics of our system, all nodes in the network are divided into different classes based on their degrees, and statistical averages of the transition probability discussed above are computed, grouping them by degree classes.

Supposing node j possesses k links, g can be modeled as a stochastic variable, following a binomial distribution characterized by

$$\Pi(g, t) = \left(\begin{array}{c} k \\ g \end{array} \right) \theta(k, t)^g (1 - \theta(k, t))^{k-g}, \tag{7.2}$$

where θ denotes the probability at time t that an edge originating from an ignorant node with k links connects to a spreader node. This quantity can be expressed as

$$\theta(k, t) = \sum_{k'} P\left(k' \mid k\right) P\left(s_{k'} \mid i_k\right) \approx \sum_{k'} P\left(k' \mid k\right) \rho^s\left(k', t\right). \tag{7.3}$$

Within this equation, $P\left(k' \mid k\right)$ represents the degree-degree correlation function, $P\left(s_{k'} \mid i_k\right)$ denotes the conditional probability that a node possessing k' links resides in the spreader state, given its connection to an ignorant node with a degree of k, and $\rho^s\left(k', t\right)$ stands for the density of spreader nodes at time t that fall within the connectivity class of k'.

The transition probability $\bar{p}_{ii}(k, t)$ averaged over all possible values of g can be expressed as:

$$\begin{aligned} \bar{p}_{ii}(k, t) &= \sum_{g=0}^{k} \left(\begin{array}{c} k \\ g \end{array} \right) (1 - \lambda \Delta t)^g \theta(k, t)^g (1 - \theta(k, t))^{k-g} \\ &= \left(1 - \lambda \Delta t \sum_{k'} P\left(k' \mid k\right) \rho^s\left(k', t\right) \right)^k \end{aligned} \tag{7.4}$$

Similarly, we can derive an expression for the probability $\bar{p}_{ss}(k,t)$ that a spreader node with k links remains in the spreader state within the time interval $[t, t+\Delta t]$. In this scenario, we must also compute the expected value of the stifler neighbors of the node at time t. Following previous procedures, we obtain

$$\bar{p}_{ss}(k,t) = \left[1 - \alpha\Delta t \sum_{k'} P\left(k' \mid k\right)\left(\rho^{s}\left(k',t\right) + \rho^{r}\left(k',t\right)\right)\right]^{k}\left(1-\delta\Delta t\right) \quad (7.5)$$

As for the corresponding probability of transitioning from the spreader to the stifler state, $\bar{p}_{sr}(k,t)$ it is defined as $\bar{p}_{sr}(k,t) = 1 - \bar{p}_{ss}(k,t)$.

The transition probabilities discussed above enable the formulation of a system comprising coupled Chapman-Kolmogorov equations, designed to describe the probability distributions concerning the populations of ignorants, spreaders and stiflers within each connectivity class. However, by disregarding fluctuations around their expected values, we can alternatively derive a set of deterministic rate equations governing the expected values of these quantities. Let $I(k,t), S(k,t)$ and $R(k,t)$ signify the expected values of node populations belonging to connectivity class k which, at time t, occupy the ignorant, spreader or stifler state, respectively. Consider the scenario where an ignorant node within class k undergoes a transition to the spreader state within the interval $[t, t+\Delta t]$. This event can be modeled as a Bernoulli random variable with a success probability of $1 - p_{ii}(k,t)$. Since it involves the summation of independent and identically distributed random Bernoulli variables, the total count of successful transitions throughout this time span follows a binomial distribution, with an expected value equivalent to $I(k,t)\left(1 - p_{ii}(k,t)\right)$. Accordingly, the rate of change in the expected population of ignorant, spreader and stifler nodes, respectively, within class k is expressed as

$$I(k,t+\Delta t) = I(k,t) - I(k,t)\left[1 - \left(1 - \lambda\Delta t \sum_{k'} P\left(k' \mid k\right)\rho^{s}\left(k',t\right)\right)^{k}\right] \quad (7.6)$$

$$\begin{aligned} S(k,t+\Delta t) = {} & S(k,t) + I(k,t)\left[1 - \left(1 - \lambda\Delta t \sum_{k'} P\left(k' \mid k\right)\rho^{s}\left(k',t\right)\right)^{k}\right] \\ & - S(k,t)\left[1 - \left(1 - \alpha\Delta t \sum_{k'} P\left(k' \mid k\right)\left(\rho^{s}\left(k',t\right) + \rho^{r}\left(k',t\right)\right)\right)^{k}\left(1-\delta\Delta t\right)\right] \end{aligned} \quad (7.7)$$

$$R(k,t+\Delta t) = R(k,t) + S(k,t)\left[1-\left(1-\alpha\Delta t\sum_{k'}P\left(k'\mid k\right)\left(\rho^{s}\left(k',t\right)+\rho^{r}\left(k',t\right)\right)\right)^{k}\left(1-\delta\Delta t\right)\right] \tag{7.8}$$

Within the context discussed above, $\rho^{i}(k,t)$, $\rho^{s}(k,t)$ and $\rho^{r}(k,t)$ denote the fractions of nodes belonging to class k that occupy the ignorant, spreader and stifler states, respectively. Also, the normalization condition holds, that is $\rho^{i}(k,t)+\rho^{s}(k,t)+\rho^{r}(k,t)=1$. As $\Delta t\to 0$, we have

$$\frac{d\rho^{i}(k,t)}{dt} = -k\lambda\rho^{i}(k,t)\sum_{k'}P\left(k'\mid k\right)\rho^{s}\left(k',t\right), \tag{7.9}$$

$$\begin{aligned}\frac{d\rho^{s}(k,t)}{dt} &= k\lambda\rho^{i}(k,t)\sum_{k'}P\left(k'\mid k\right)\rho^{s}\left(k',t\right)\\ &- k\alpha\rho^{s}(k,t)\sum_{k'}P\left(k'\mid k\right)\left(\rho^{s}\left(k',t\right)+\rho^{r}\left(k',t\right)\right)-\delta\rho^{s}(k,t),\end{aligned} \tag{7.10}$$

$$\frac{d\rho^{r}(k,t)}{dt} = k\alpha\rho^{s}(k,t)\sum_{k'}P\left(k'\mid k\right)\left(\rho^{s}\left(k',t\right)+\rho^{r}\left(k',t\right)\right)+\delta\rho^{s}(k,t). \tag{7.11}$$

It is worth noting that when it comes to the equations we have been discussing, the information about the underlying network is tied solely to the degree-degree correlation function. This means that in the upcoming analytical and numerical studies, there is no need to generate an actual network. The sole prerequisites are either an analytical expression for $P\left(k'\mid k\right)$ or a numerical representation of this quantity.

Analytical results

Now we will present the analytical findings for homogeneous social networks. Homogeneous networks consist of nodes that serve identical roles and perform the same functions within the network. On the contrary, heterogeneous networks encompass networks characterized by the presence of two or more distinct node classes, differentiated by both their function and utility, thus introducing greater complexity into the analysis. Hence, we will refrain from delving into the intricacies of inhomogeneous networks and the associated analytical findings. Those

seeking a deeper understanding of this subject matter are encouraged to turn to reference [1] for more detailed information.

To gain insight into fundamental aspects of our rumour model, we examine the scenario of homogeneous networks. In such networks, degree fluctuations are minimal and degree correlations are absent. As a result, the rumour equations simplify to

$$\frac{d\rho^i(t)}{dt} = -\bar{k}\lambda\rho^i(t)\rho^s(t), \tag{7.12}$$

$$\frac{d\rho^s(t)}{dt} = \bar{k}\lambda\rho^i(t)\rho^s(t) - \bar{k}\alpha\rho^s(t)\left(\rho^s(t) + \rho^r(t)\right) - \delta\rho^s(t), \tag{7.13}$$

$$\frac{d\rho^r(t)}{dt} = \bar{k}\alpha\rho^s(t)\left(\rho^s(t) + \rho^r(t)\right) + \delta\rho^s(t) \tag{7.14}$$

with $\bar{k}$ signifying the network's constant degree distribution (or the average value, in cases where the probability of encountering a node with differing connectivity decays exponentially).

The system of equations outlined above can be integrated analytically using a standard approach. The scholars concluded that, in the infinite time limit, when there are no spreaders remaining, the final fraction of nodes that have encountered the rumor, commonly referred to as the final size of the rumor, is obtained from the subsequent transcendental equation:

$$R = \rho^r(\infty) = 1 - e^{-\varepsilon \mathrm{R}}, \tag{7.15}$$

where

$$\varepsilon = \frac{(\alpha + \lambda)\bar{k}}{\delta + \alpha\bar{k}}.$$

Equation (7.15) possesses a non-zero solution solely when $\varepsilon > 1$. In cases where $\delta \neq 0$, this condition holds true as long as $\frac{\lambda}{\delta}\bar{k} > 1$.
However, in the specific scenario where $\delta = 0$ (indicating the absence of the forgetting mechanism), we have $\varepsilon = 1 + \frac{\lambda}{\alpha} > 1$, thus ensuring that Equation (7.15) always accommodates a non-zero solution.

The outcome above elucidates that the introduction of a forgetting mechanism results in the emergence of a finite threshold in the rate of rumor propagation, beneath which rumors cannot spread within homogeneous networks. Notably, this threshold value remains unaffected by α which represents the stifling mechanism. Consequently, in the initial stages of the spreading process,

characterized by $\rho^s \approx 0$ and $\rho^r \approx 0$, the influence of stifling is insignificantly minor in comparison to that of forgetting.

Numerical results

We now transition to the exploration of exact numerical solutions, enabling us to analyse both the steady-state and the time-dependent behavior of the model across various social network models. Initially, we focus on Erdős-Rényi (ER) random graphs and uncorrelated scale-free networks(SF), followed by an examination of scale-free networks with assortative degree correlations.

It is important to recall that ER random graphs, as we discussed in Chapter 3, exhibit a Poisson degree distribution for large values of N. Specifically, this distribution is represented by

$$P(k) = e^{-<k>}\frac{< k >^k}{k!},$$

which attains its peak at an average value $\langle k \rangle$ and exhibits minor fluctuations around this value.

Scale-free networks are characterized by a power law degree distribution:

$$P(k) = \begin{cases} Ak^{-\gamma}, k_{\min} \leq k \\ 0, \text{ otherwise} \end{cases}$$

Within the aforementioned equation, $k_{\min}$ represents the minimum degree within these networks, while A serves as a normalization constant.

Their investigations were conducted employing the aforementioned forms of $P(k)$ to represent ER and SF networks, with $\gamma = 3$, respectively. The network size remained fixed at $N = 10^6$, with a consistent average degree of $\langle k \rangle = 7$. For each network variant, they generated a sequence of N random integers, distributed in accordance with the degree distribution. Subsequently, in order to numerically solve the coupled set of differential equations (7.9)-(7.11), they employed a numerical technique for solving differential equations called the standard finite difference scheme. Also, numerical convergence with respect to the step size was evaluated numerically.

Throughout this paper the initiation of the rumor occurred from a randomly selected initial spreader. The outcomes are based on averaging results obtained from 300 runs, each commencing with different initial spreaders.

In the initial series of computations, the value of δ was set at 1, and the dynamics were examined as functions of the rumour spreading rate λ and the stifling rate α. The initial focus was directed towards assessing the influence of network topology on the final size of the rumor R, which for inhomogeneous networks is derived from

$$R = \sum_k \rho^r \left(k, t_\infty\right),$$

where t_∞ represents a sufficiently extended period at which the propagation process attains its steady state characterized by the complete absence of spreaders within the network.

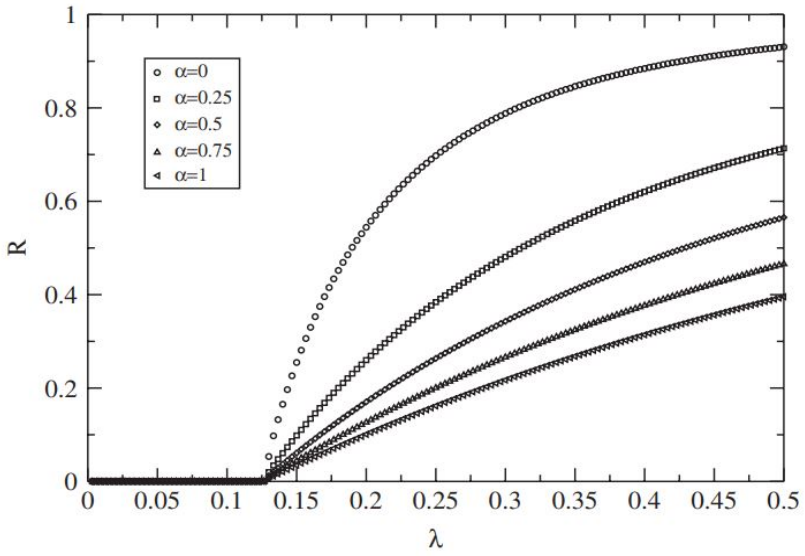

Figure 7.1: R for the ER network plotted as a function of λ for several values of α. Source: [1]

In Figure 7.1, final size of the rumor corresponding to the ER network is plotted as a function of λ showcasing various values of the stifling parameter α. The key observation is the presence of a critical threshold λ_c, which is 0.12507, below which the dissemination of a rumor is not possible. Intriguingly, akin to the findings witnessed in homogeneous networks, the value of this threshold remains unaffected by variations in α.

As far as the results for the scale-free network are concerned, the subsequent Figure presents the outcomes concerning the final size of the rumor.

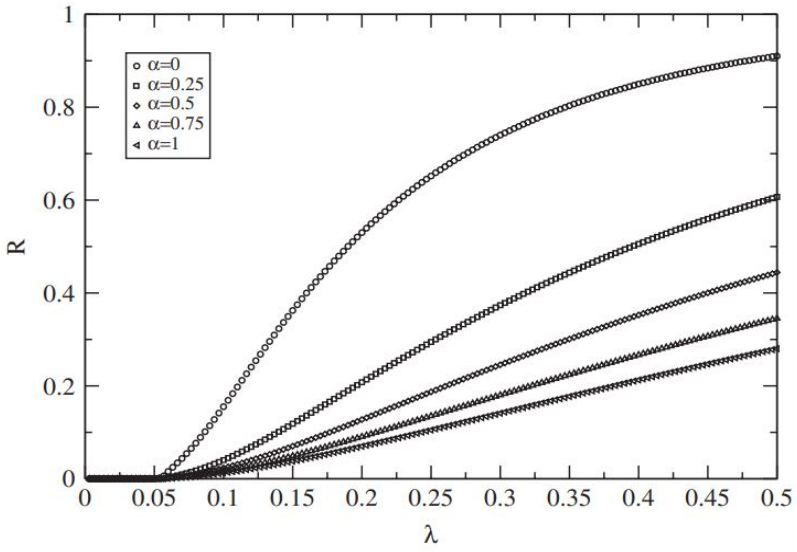

Figure 7.2: R for the SF network plotted as a function of λ for several values of α. Source: [1]

In this scenario as well, the rumor threshold remains independent of α, although it manifests at considerably lower spreading rates when compared to the ER network.

Besides assessing the influence of network topology on the steady-state characteristics of the model, attention is also paid around comprehending the impact of topology on the time-dependent behavior of the model. Temporal evolution of the overall proportions of stiflers and spreaders within both ER and SF network are observed, where the visualizations are constructed under the premise of $\lambda = 1$ and two distinct sets of cessation parameters:

- $\{\alpha = 0; \delta = 1\}$, where parameter set signifies a spreading process in which cessation arises exclusively due to the spontaneous forgetting of a rumor by spreaders or their disinclination to propagate the rumor further.
- $\{\alpha = 1; \delta = 0\}$, where parameter set corresponds to a situation in which individuals persist in spreading the rumor until they transition to stiflers as a consequence of their interactions with other spreaders or stiflers within the network.

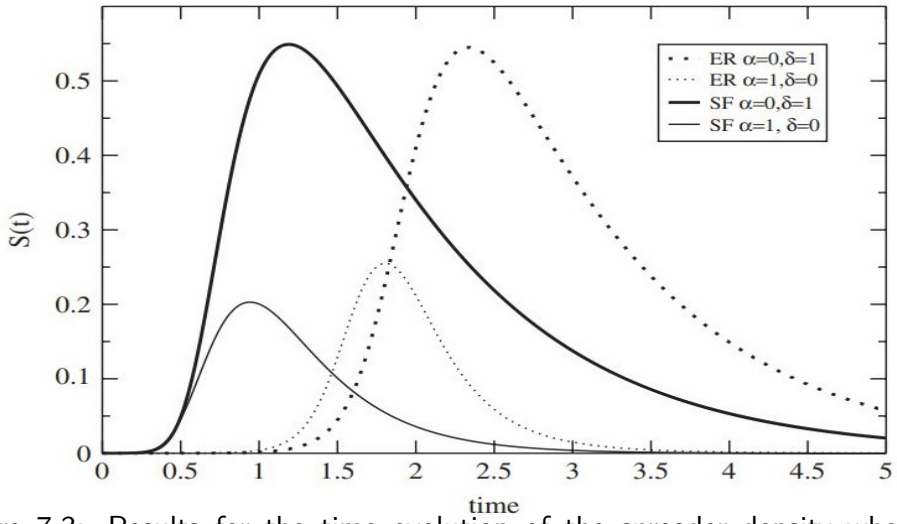

Figure 7.3: Results for the time evolution of the spreader density where ER network is represented with dashed lines and SF network with solid lines, for two distinct parameter sets. Source: [1]

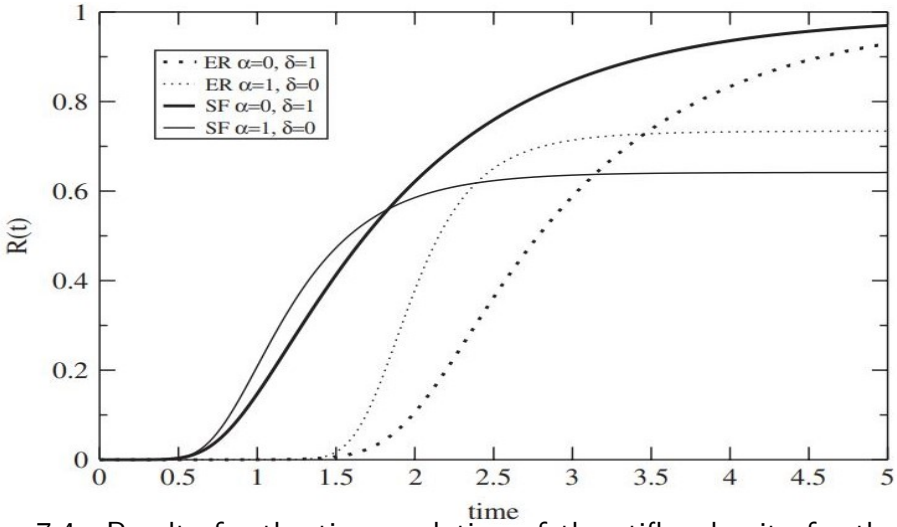

Figure 7.4: Results for the time evolution of the stifler density for the same networks. Source: [1]

As demonstrated in Figures above, within the first spreading scenario, the initial propagation rate of a rumor across the SF network substantially surpasses that of the ER network. In fact, findings indicate that the timeframe necessary for the rumor to encompass 50% of the nodes in the ER random graph is nearly twice as extended as the corresponding duration observed within the SF networks. This substantial contrast in the propagation rate is expected and attributed to the presence of highly connected nodes, hubs, within the SF network. These hubs possess considerable influence, greatly accelerating the process of rumor propagation once they become informed. Furthermore, not only does the rumor propagate more rapidly during its initial phase within SF networks, but it also ultimately encompasses a larger fraction of nodes by the culmination of the dissemination process.
Within the second scenario of propagation, where stifling represents the sole cessation mechanism, the initial rate of dissemination within the SF network once again surpasses that witnessed within the ER network. However, in contrast to the prior scenario, the final size of the rumor is immenser on the ER network. This intricate behavior arises from the divergent roles played by hubs when the stifling mechanism is activated. Initially, the presence of hubs expedites the propagation, but as these hubs transition into stiflers, they effectively prevent further dissemination of the rumor.

Our attention now turns to the examination of scale-free networks featuring assortative degree correlation. Investigations have unveiled that social networks tend to exhibit assortative degree correlations, signifying that highly connected nodes rather establish connections with other highly connected nodes [8]. To explore the influence of such correlations on the dynamics inherent to the proposed model a specific approach concerning the correlation parameter $\beta, 0 \leq \beta < 1$ is introduced.
Once again, equations (7.9) - (7.11) were solved with a specific focus on a SF network characterized by a degree exponent $\gamma = 3$ and an average degree of $\langle k \rangle = 7$. The network size remained constant at $N = 10^5$, while using two distinct values for β, specifically $\beta = 0.2$ and $\beta = 0.4$. Figure 7.5 serves as the visual representation of the outcomes, illustrating the behavior of the final size of the rumor as a function of λ, while considering values of α at $0.5, 0.75$ and 1. Also, again, the value of δ was held at 1 during this analysis.

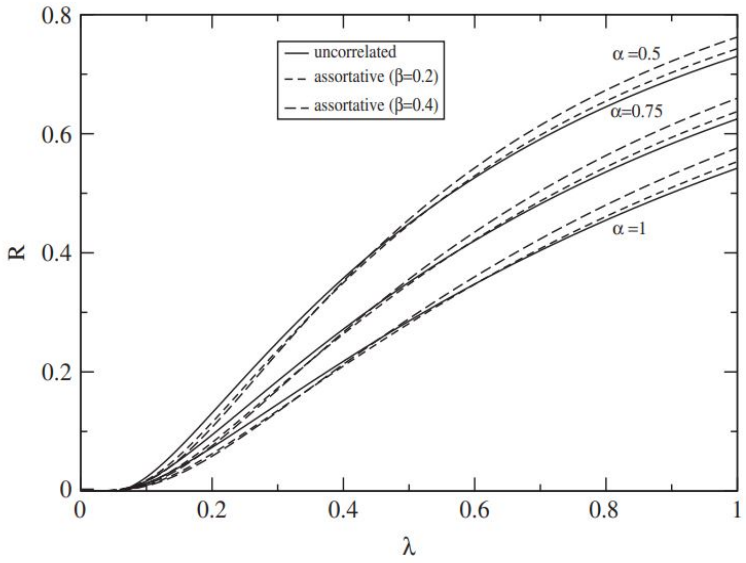

Figure 7.5: R for the SF network plotted as a function of λ for several values of α. The outcomes are showcased both in the absence of assortative degree-degree correlations (represented by solid lines) and in their presence. The correlation parameter $\beta = 0.2$ is indicated by short dashed lines and $\beta = 0.4$ by long dashed lines. Source: [1]

Observations reveal that when λ is approximately less than 0.5, a rumor tends to reach a somewhat smaller fraction of nodes within correlated networks compared to uncorrelated ones. However, as λ surpasses this threshold, the behavior reverses. In this scenario, the final extent of the rumor within assortatively correlated networks surpasses that observed in uncorrelated networks. This observation leads us to the conclusion that the qualitative impact of degree correlations on the ultimate rumor size is highly dependent on the rate of rumor propagation.

In a concluding exploration, the investigators wanted to understand the impact of assortative correlations on the speed of rumor propagation. In Figure 7.6, the results are presented, showcasing the time evolution of the total fraction of spreaders, $S(t)$. These observations were made within scale-free networks, each composed of $N = 10^5$ nodes. The correlation strength, β, spanned from 0 to 0.4 in this analysis. Throughout these calculations, λ was kept constant at a value of 1, and two distinct values of α, specifically 0 and 1, were considered.

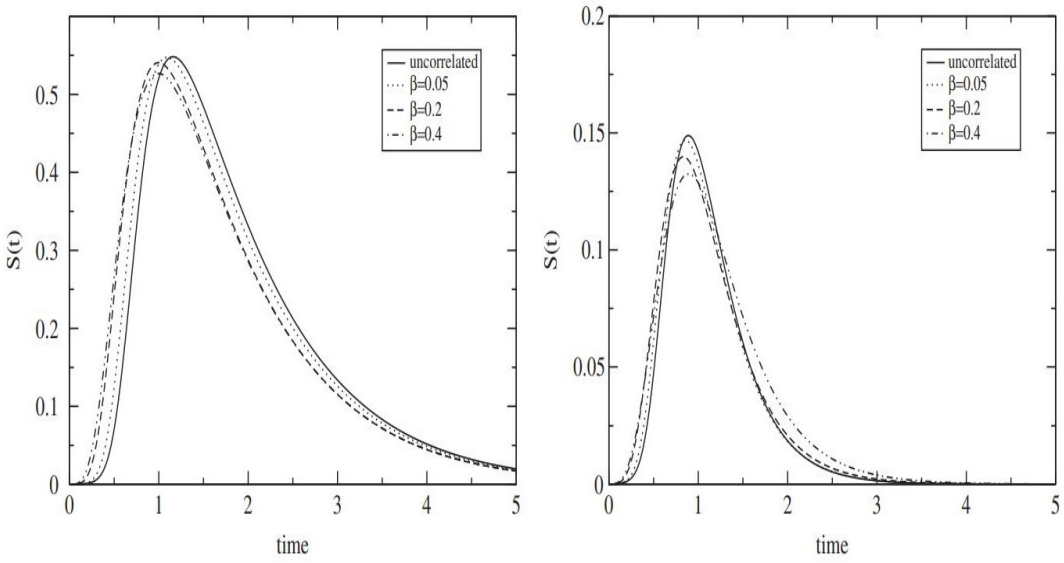

Figure 7.6: The influence of assortative correlations on the time dynamics of rumor spreaders. The left panel showcases the outcomes when $\alpha = 0$, while the right panel illustrates the results when $\alpha = 1$. Source: [1]

The graphical representation reveals that the initial rate of rumor dissemination increases proportionally with the strengthening of assortative correlations, regardless of the chosen value for α. However, when $\alpha = 1$ the rumor's lifespan also diminishes more rapidly when such correlations are stronger.

Chapter 8

Conclusion

In conclusion, the conducted analysis of rumor propagation within complex social networks has unveiled the intricate interplay between network structure and rumor dissemination. As discussed throughout this thesis, understanding the anatomy of a network is of the greatest importance since the structure of the network seriously affects its functioning, which includes the spread of rumors, information and even diseases.

In initial chapters, we provided insight into the fundamental attributes characterizing complex networks, which laid the foundation for comprehending network's structure. Subsequently, we discussed four prominent network models, illuminating the distinct characteristics that differentiates them and elucidating their implications for information dissemination. Following this, we pursued a comparative analysis of rumor propagation and epidemic transmission. The utilization of infectious disease models, SI, SIR and SIS, unveiled parallels between these seemingly distinct processes. This exploration underscored the role of compartmental models in capturing the essence of propagation patterns.
Acknowledging rumors as a form of "mind infection", our inquiry led us from the realm of disease modeling into the realm of developing models for the dissemination of rumors. The pioneering models, for which Daley, Kendall, Maki and Thompson are credited to, explained the dynamics of rumor transmission and established a basis for more sophisticated investigations. In the final chapter, the groundbreaking work of Nekovee et. al, introduced an innovative, realistic model for the propagation of rumors across complex networks, which incorporates two distinct mechanisms that cause cessation of a rumour, stifling and forgetting. Their research reported the existence of a critical threshold in the rate of rumor propagation within ER networks, below which the dissemination of a rumor re-

mains unattainable. This threshold has proven to be immune of the influence of the stifling mechanism. In parallel, a similar threshold, although smaller in value, has surfaced within the finite-size SF networks examined.
Additionally, their research marked a disparity in the initial rumor dissemination rates between scale-free networks and ER random graphs. This phenomenon can be attributed to the role assumed by hubs within scale-free networks, serving as highly effective agents in the rapid propagation of rumors, once they attain knowledge of the rumor. Moreover, an exploration was carried out to examine the impact of assortative degree correlations on the speed of rumor propagation within SF networks. It follows that these correlations speed up the initial rate of propagation within such networks. Nevertheless, the magnitude of their influence on the eventual proportion of nodes which hear the rumor depends on the rate at which the rumor itself propagates.
Hence, their findings further emphasize the impact of network topology on rumor spread, stressing the interconnectedness of network structure and function.

This study, despite its valuable insights, must be acknowledged to possess certain limitations. The ever-evolving area of communication and technology continually presents new challenges and promising opportunities for research within this sphere. Future investigations may delve deeper into the underlying network which, throughout this study, remains static, representing a fixed and unchanging network topology. In reality, many social and communication networks are often highly dynamic. For instance, consider the dynamic nature of Internet chatrooms, where individuals establish new social connections and cease existing ones on a continuous basis. The endeavor of modeling the propagation of information on such dynamic networks becomes an exceedingly intricate task, which could represent the forthcoming research.
In addition to that, within a single network, a multitude of spreading mechanisms can exist. For example, consider a scenario in which two distinct rumors, labeled "rumor 1" and "rumor 2", exhibit varying probabilities of acceptance as they disseminate throughout the network [40]. This conceptual framework extends the classical epidemic SIS model and introduces a level of complexity. Furthermore, in traditional rumor spreading models, the transition probability from ignorants to spreaders is typically treated as a fixed parameter. However, when examining real-world scenarios, it becomes evident that an individual's susceptibility to infection by a neighboring spreader is contingent upon the strength of the interpersonal ties between them. To address this, one can introduce a stochastic epidemic model for rumor diffusion, where the infectious probability is explicitly defined as a function of these tie strengths [41].

Furthermore, one can also challenge the conventional notion that, in the diffusion of rumors across networks, an individual exclusively acquires information from their immediate neighbors. With the advent of the Internet and the increase of popularity of online social networks, network participants now possess convenient access to a diverse array of information sources, including television, Facebook, newspapers, and more. Among these individuals, a subset may receive a rumor from alternative channels and subsequently disseminate it within their network, all without necessarily obtaining the rumor from their immediate neighbors within that network. This phenomenon transforms them into independent spreaders, individuals engaged in multiple networks with access to diverse information sources. Moreover, the scenario unfolds where participants engaged in one network concurrently maintain membership in various other networks. Addressing this complexity necessitates the development of an innovative model in which independent spreaders emerge, propagating rumors within a network while not relying on information from network neighbors [42].

In closing, it is evident that the examination of rumor propagation continually evolves through innovative research and a growing understanding of the factors that govern how rumors spread. This field provides valuable insights into the broader landscape of social networks and human interactions. Nevertheless, I embrace the the idea that interdisciplinary approaches are key to unraveling the intricacies of this phenomenon.

As responsible members of an information-rich world, we all have a role to play in combating the spread of misinformation and nurturing a culture that values critical thinking. By doing so, we can navigate the complex realm of rumor propagation with greater resilience and wisdom.

Bibliography

[1] Nekovee M., Moreno Y., Bianconi G., Marsili M., Theory of rumour spreading in complex social networks, Physica A 374, 457-470, 2007

[2] Newman M. E. J., Networks – An introduction, Oxford University Press, 2010

[3] Turenne N., The rumour spectrum, PLOS ONE, 2018

[4] Bondy, J. A., Murty, U. S. R., Graph Theory, Springer, 2008

[5] Albert R., Barabási A.-L., Statistical mechanics of complex networks, Reviews of Modern Physics 74, 47-97, 2002

[6] Barabasi A.-L., Network Science, Cambridge University Press, 2015

[7] Molontay R., Nagy M., Two decades of network science–as seen through the co-authorship network of network scientists, 1–6, 2019

[8] Newman M. E. J., The structure and function of complex networks, SIAM Rev. 45 167, 2003

[9] Strogatz S. H., Exploring complex networks, Nature, vol. 410, p. 268, 2001

[10] Zhou S., Parametrising and Modelling the Internet Topology, 2023

[11] Barabasi A.-L., Bonabeau E., Scale–Free Networks, Scientific American 288, 50-59, 2003

[12] Wang X. F., Chen G., Complex Networks: Small–World, Scale-Free and Beyond, IEEE Circuits and Systems Magazine, First Quarter, 2003

[13] Barabási A.-L., Albert R., Emergence of scaling in random networks, Science 286, 509–512, 1999

[14] Watts D. J., Strogatz S. H., Collective dynamics of small-world networks, Nature, vol. 393, 440–442, 1998

[15] Telesford Q. K., Joyce K. E., Hayasaka S., Burdette J. H., Laurienti P. J., The Ubiquity of Small-World Networks, Brain Connect, 2011

[16] Lu J., Yu X., Chen G., Cheng D., Characterizing the Synchronizability of Small-World Dynamical Networks, IEEE Transactions on Circuts and Systems I: Regular Papers 51, 787–796, 2004

[17] Yoon J.-H., Dias S., Hahn S., A method for assessing robustness of the results of a star-shaped network meta-analysis under the unidentifiable consistency assumption, BMC Medical Research Methodology, 2021

[18] Milgram S., The small-world problem, Psychology Today 2, 60-67, 1967

[19] Kleinberg J., The Small-World Phenomenon and Decentralized Search, SIAM News, vol. 37, 2004

[20] Meghanathan N., Small World Networks, Lecture Notes, Jackson State University, Jackson, MS, USA

[21] Newman M. E. J., Watts D. J. Renormalization group analysis of the small-world network model, Phys. Lett. A, vol. 263, 341-346, 1999

[22] Brauer F., Mathematical epidemiology: Past, present, and future, Infectious Disease Modelling, 113-127, 2017

[23] Hamer W.H., The Milroy Lectures on Epidemic disease in England – The Evidence of Variability and of Persistence of Type, The Lancet 1, 733–738, 1906

[24] Ross R. A., The prevention of malaria (2nd edition, with Addendum), London: John Murray, 1911

[25] Brauer F., Driessche P.V.D., Wu J., Mathematical Epidemiology, Springer Science and Business Media, 2008

[26] Sun H., Research on SI , SIS, SIR epidemic models and deeper investigation on diverse conditions, Southeast University, China, 2016

[27] Babić D., Matematički modeli u epidemiologiji(Master's thesis), Faculty of Sciences, University of Novi Sad, Novi Sad, Serbia, 2017

[28] Hethcote H. W., Three basic epidemiological models, In: Levin S.A., Hallam T.G. and Gross L.J. (eds.), Applied Mathematical Ecology, Biomathematics vol. 18, Berlin: Springer, 119-144, 1989

[29] Goffman W., Newill V. A., Generalization of Epidemic Theory–An application to the transmission of ideas, Nature, vol. 204, 225-228, 1964

[30] Belen S., The behaviour of stochastic rumours(Doctoral thesis), School of Mathematical Sciences, University of Adelaide, Adelaide, South Australia, 2008

[31] Daley D. J., Gani J., Epidemic Modelling, Cambridge University Press, 2000

[32] Daley D. J., Kendall D. G., Stochastic Rumours, J. Inst. Math. Appl. 1, 42-55, 1965

[33] Dunstan R., The rumour process, J.Appl.Prob. 19, 759-766, 1982

[34] Watson R., On the size of a rumour, Stochastic Processes and their Applications 27, 141-149, 1988

[35] Wang X., Zhao T., Model for multi-messages spreading over complex networks considering the relationship between messages, Communications in Nonlinear Science and Numerical Simulation 48, 63-69, 2017

[36] Maki D.P., Thompson M., Mathematical Models and Applications, with Emphasis on Social Life and Management Sciences, Prentice-Hall, Englewood Cliffs, NJ, 1973

[37] Zhu L., Yang F., Guan G., Zhang Z., Modeling the dynamics of rumor diffusion over complex networks, Information Sciences, vol. 562, 240-258, 2021

[38] López-Pintado D., Diffusion in complex social networks, Instituto Valenciano de Investigaciones Económicas, S.A, 2004

[39] Moreno Y., Nekovee M., Pacheco A. F., Dynamics of Rumor Spreading in Complex Networks, Phys. Rev. E 69, 2004

[40] Trpevski D., Tang W. K. S., Kocarev Lj., Model for rumor spreading over networks, Phys. Rev. E 81, 2010

[41] Cheng J.-J., Liu Y., Shen B., Yuan W.-G., An epidemic model of rumor diffusion in online social networks, The European Physical Journal B, vol. 86, no. 29, 2013

[42] Li W., Tang S., Pei S., Yan S., Jiang S., Teng X., The rumor diffusion process with emerging independent spreaders in complex networks, Physica A, vol. 397, 121-128, 2014

Biography

Maja Jović was born on the 14^{th} of June 1998 in Novi Sad, Serbia. In her hometown, she attended "Svetozar Marković Toza" elementary school and "Svetozar Marković" high school. In 2017, she enrolled at University of Novi Sad, Faculty of Sciences, Department of Mathematics and Informatics, where she finished her Bachelor studies in Mathematics in 2021, with a GPA 8.48. In the same year, she continued with her Master studies in Applied Mathematics: Data Science at the same faculty and passed all exams in 2023 with a GPA of 9.25.

UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI FAKULTET
KLJUCNA DOKUMENTACIJSKA INFORMACIJA

Redni broj:
RBR

Identifikacioni broj:
IBR

Tip dokumentacije: Monografska dokumentacija
TD

Tip zapisa: Tekstualni štampani materijal
TZ

Vrsta rada: Master rad
VR

Autor: Maja Jović
AU

Mentor: dr Dušan Jakovetić
MN

Naslov rada: Širenje glasina u kompleksnim društvenih mrežama
NR

Jezik publikacije: Engleski
JP

Jezik izvoda: Engleski
JI

Zemlja publikovanja: Republika Srbija
ZP

Uže geografsko područje: Vojvodina
UGP

Godina: 2023.
GO

Izdavač: Autorski reprint

IZ

Mesto i adresa: Novi Sad, Prirodno-matematički fakultet, Departman za matematiku i informatiku, Trg Dositeja Obradovića 4
MA

Fizički opis rada: 8 poglavlja, 79 strana, 42 lit. citata, 28 slika
FO

Naučna oblast: Matematika
NO

Naučna disciplina: Primenjena matematika
ND

Ključne reči: Širenje glasina, kompleksne društvene mreže, nasumično generisane mreže, nesrazmerne mreže, struktura i funkcija, širenje epidemije, kompartmentalni modeli, Daley-Kendall model, Maki-Thompson model
KR

Univerzalna decimalna klasifikacija:
UDK

Čuva se: Biblioteka Departmana za matematiku i informatiku, Prirodno-matematičkog fakulteta, u Novom Sadu
ČU

Važna napomena:
VN

Izvod: U ovom radu istražujemo kako struktura mreže utiče na njenu funkciju. U prvom poglavlju pružen je uvid u osnovne atribute koji karakterišu složene mreže, što je postavilo osnovu za razumevanje strukture mreže. Nakon toga, objasnili smo različite karakteristike četiri istaknuta mrežna modela. Zatim je sprovedena komparativna analiza širenja glasina i prenosa epidemije. Ovo istraživanje je naglasilo ulogu kompartmentalnih modela u hvatanju suštine obrazaca širenja. Priznajući glasine kao oblik „infekcije uma“, istraživanje nas je odvelo do razvoja modela za širenje glasina, gde smo se susreli sa osnovnim Daley-Kendall i Maki-Thompson modelima. U poslednjem poglavlju uvodimo inovativan, realističan model za širenje glasina kroz složene mreže. Na njemu upoređujemo početnu stopu, ponašanje granice nakon koje se glasina više ne

širi i dinamiku modela kod nasumično generisanih mreža, nesrazmernih mreža i nesrazmernih mreža sa asortativnim stepenom korelacije. Utvrđeno je da topologija mreže ima važan uticaj na širenje glasina.
IZ

Datum prihvatanja teme od strane NN veca: 7.9.2023.
DP

Datum odbrane:
DO

Članovi komisije:
KO
Predsednik: dr Danijela Rajter Ćirić, redovni profesor, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
Mentor: dr Dušan Jakovetić, vanredni profesor, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
Član: dr Sanja Rapajić, redovni profesor, Prirodno-matematički fakultet, Univerzitet u Novom Sadu

UNIVERSITY OF NOVI SAD
FACULTY OF SCIENCES
KEY WORDS DOCUMENTATION

Accession number:
ANO

Identification number:
INO

Document type: Monograph type
DT

Type of record: Printed text
TR

Contents code: Master's thesis
CC

Author: Maja Jović
AU

Mentor: Dr Dušan Jakovetić
MN

Title: Rumor spreading in complex social networks
TL

Language of text: English
LT

Language of abstract: English
LA

Country of publication: Republic of Serbia
CP

Locality of publication: Vojvodina
LP

Publication year: 2023.
PY

Publisher: Author's reprint

PU

Publ. place: Novi Sad, Department of Mathematics and Informatics, Faculty of Science and Mathematics, University of Novi Sad, Trg Dositeja Obradovića 4
PP

Physical description: 8 chapters, 79 pages, 42 references, 28 figures
PD

Scientific field: Mathematics
SF

Scientific discipline: Applied mathematics
SD

Key words: Rumor spread, complex social networks, random networks, scale-free networks, structure and function, epidemic spread, compartmental models, Daley-Kendall model, Maki-Thompson model
KW

Universal decimal classification:
UDC

Holding data: The Library of the Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad
HD

Note:
N

Abstract: In this paper, we investigate how the structure of the network affects its function. In the first chapter, an insight into the basic attributes that characterize complex networks was provided, which laid the foundation for understanding the structure of the network. After that, we explained the different characteristics of the four prominent network models. Then a comparative analysis of rumor spreading and epidemic transmission was conducted, where parallels between these apparently different processes were revealed. This research emphasized the role of compartmental models in capturing the essence of diffusion patterns. Recognizing rumors as a form of "mind infection", our research led us to develop models of rumor spreading, where we encountered the basic Daley-Kendall and Maki-Thompson models. In

the last chapter, we introduce an innovative, realistic model for spreading rumors through complex networks. We compare the initial rate, the behavior of the limit after which the rumor stops spreading, and the dynamics of the model for random networks, scale-free networks and scale-free networks with assortative degree correlation. It has been established that network topology has a significant impact on the spread of rumors.
AB

Accepted by the Scientific Board on: 7.9.2023.
ASB

Defended on:
DE

Thesis defend board:
DB
Chairperson: Dr Danijela Rajter Ćirić, full professor, Faculty of Sciences, University of Novi Sad
Mentor: Dr Dušan Jakovetić, associate professor, Faculty of Sciences, University of Novi Sad
Member: Dr Sanja Rapajić, full professor, Faculty of Sciences, University of Novi Sad