Epidemic Models On Complex Network

A Project Report Submitted
in Partial Fulfillment of the
Requirements for the Degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING

by

Ayush Rathor (CS21B1004)

ಭಾರತೀಯ ಮಾಹಿತಿ ತಂತ್ರಜ್ಞಾನ ಸಂಸ್ಥ ರಾಯಚೂರು
भारतीय सूचना प्रौद्योगिकी संस्थान रायचूर
Indian Institute of Information Technology Raichur

to

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING INDIAN INSTITUTE OF INFORMATION TECHNOLOGY RAICHUR-784135, INDIA

April, 2025

DECLARATION

I, Ayush Rathor (Roll No: CS21B1004), hereby declare that, this report entitled **"Epidemic Models On Complex Network"** submitted to Indian Institute of Information Technology Raichur towards partial requirement of Bachelor of Technology/Bachelor of Technology(Hon)/Master of Science in Computer Science and Engineering is an original work carried out by me under the supervision of Priodyuti Pradhan and has not formed the basis for the award of any degree or diploma, in this or any other institution or university. I have sincerely tried to uphold the academic ethics and honesty. Whenever an external information or statement or result is used then, that have been duly acknowledged and cited.

Raichur-584135
April 2025

Ayush Rathor

CERTIFICATE

This is to certify that the work contained in this project report entitled **"Epidemic Models On Complex Network"** submitted by [Ayush Rathor (Roll No: CS21B1004) to the Indian Institute of Information Technology Raichur towards partial requirement of Bachelor of Technology/Bachelor of Technology (Hon) in Computer Science and Engineering has been carried out by him under my supervision and that it has not been submitted elsewhere for the award of any degree.

Raichur-584135
May 2025

Prof. Priodyuti Pradhan
Project Supervisor

Abstract

Epidemic models on complex networks provide critical insights into the spread of infectious diseases by capturing heterogeneity in individual interactions. Classical models such as SI, SIR, and SIS describe basic contagion mechanisms but assume homogeneous mixing, overlooking real-world network structures. Incorporating network science reveals how features like degree distributions and network topology impact transmission dynamics, with key parameters such as the spectral radius influencing epidemic thresholds. Graph-based modeling approaches describe time-dependent infection processes more accurately, highlighting degree-dependent spread and eigenmode-driven growth patterns. Data-driven methods using Graph Neural Networks (GNNs) overcome the limitations of classical models by learning local contagion dynamics directly from temporal data, without strong prior assumptions. A GNN architecture with attention mechanisms and recurrent layers enables modeling both structural dependencies and temporal evolution, achieving strong generalization across different network types such as Erdős–Rényi and Barabási–Albert graphs. Results demonstrate accurate prediction of both simple and complex contagion behaviors, illustrating the effectiveness of network-aware, learning-based approaches in epidemic forecasting.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Understanding the spread of infectious diseases is a fundamental problem in epidemiology and public health. Classical epidemic models, such as SI, SIR, and SIS, provide a mathematical framework for analyzing disease dynamics by partitioning populations into health-related compartments. However, these models typically assume homogeneous mixing, where every individual interacts uniformly with every other, an assumption that rarely holds true in real-world social structures. The emergence of network science has revolutionized

the study of epidemics by introducing the idea that individuals interact through complex, heterogeneous networks. These structures, characterized by features such as degree distributions, hubs, and communities, significantly influence how diseases propagate. Modeling epidemics on networks reveals new dynamics, including modified outbreak thresholds, super-spreader effects, and network-driven acceleration of contagion. In addition to structural complexity, time-dependent properties play a critical role in epidemic spread.

Temporal analysis uncovers patterns such as early exponential growth, saturation effects, and the formation of persistent endemic states. Traditional models often miss these temporal signatures, highlighting the need for more detailed approaches. Recent advances in machine learning, particularly Graph Neural Networks

(GNNs), offer a new paradigm for learning contagion dynamics directly from data. GNNs exploit the underlying graph structure to model both local interactions and global spread patterns without relying on fixed transition rules. By integrating GNNs with recurrent architectures, it becomes possible to capture both the spatial and temporal evolution of disease spread in complex networks.

This work focuses on combining classical epidemic models, network-based frameworks, and GNN-based learning to provide a comprehensive view of contagion dynamics.

1.1 Motivation

Traditional epidemic models have been instrumental in shaping early theories of disease spread. However, their simplifying assumptions limit their predictive accuracy in complex social systems. In modern society, where human interactions are highly structured and dynamic, incorporating network topology into epidemic models becomes essential for understanding real-world outbreaks. The COVID-19 pandemic has further emphasized the necessity of accounting for network effects, super-spreading events, and heterogeneous contact patterns.

1.2 Classical vs Network-based Modeling

Classical models treat populations as well-mixed systems, leading to uniform disease transmission rates across individuals. In contrast, network-based models represent individuals as nodes and interactions as edges, allowing a realistic description of how diseases propagate through diverse contact structures. Key phenomena such as degree heterogeneity, clustering, and community formation profoundly affect epidemic thresholds and dynamics. Moreover, data-driven methods such as Graph Neural Networks (GNNs) enable learning complex contagion patterns directly from empirical data, improving both prediction and generalization across different network topologies.

Chapter 2

Epidemic Models

Epidemic models help us understand how diseases spread through populations by representing people and their interactions mathematically. These models are particularly important for studying infections that travel through direct contact between individuals, such as flu viruses or sexually transmitted diseases. The network approach recognizes that not all contacts between people are equal - some individuals have many more social connections than others, and these differences can significantly affect how quickly a disease spreads. By using network science, we can create more accurate predictions about disease outbreaks and develop better strategies to control them.

2.1 The SI Model

The SI model is the simplest way to represent disease spread, dividing the population into just two groups. Susceptible individuals (S) are those who could potentially catch the disease, while infected individuals (I) are those who currently have the disease and can pass it to others. In this basic model, once people become infected, they remain infected indefinitely, continuing to spread the disease to susceptible people they meet.

The mathematical formulation of the SI model describes how individuals transition between these two states. Let S(t) be the number of susceptible individuals at time t and X(t) be the number infected (we use X instead of I to avoid confusion with indices later). The total population size is n=S+X.

The model makes these key assumptions[1]:

- The disease spreads through random contacts between individuals
- Each infected individual has β contacts per unit time that could transmit the disease
- The probability a random contact is with a susceptible is $\frac{S}{n}$
- Once infected, individuals remain infected forever (no recovery)

This leads to the differential equations:

$$\frac{dX}{dt} = \beta \frac{SX}{n} \tag{2.1}$$

$$\frac{dS}{dt} = -\beta \frac{SX}{n} \tag{2.2}$$

These equations state that:

- The infected population grows at rate $\beta\frac{SX}{n}$.
- The susceptible population decreases at exactly the same rate (Eq. 2)

We can simplify by working with fractions of the population:

$$s = \frac{S}{n}, \quad x = \frac{X}{n} \tag{2.3}$$

The equations then become:

$$\frac{ds}{dt} = -\beta sx \tag{2.4}$$

$$\frac{dx}{dt} = \beta sx \tag{2.5}$$

Since $s + x = 1$, we can eliminate s to get a single equation for x:

$$\frac{dx}{dt} = \beta(1 - x)x \tag{2.6}$$

This is the *logistic growth equation*, whose solution is:

$$x(t) = \frac{x_0 e^{\beta t}}{1 - x_0 + x_0 e^{\beta t}} \tag{2.7}$$

where x_0 is the initial fraction infected at t=0.

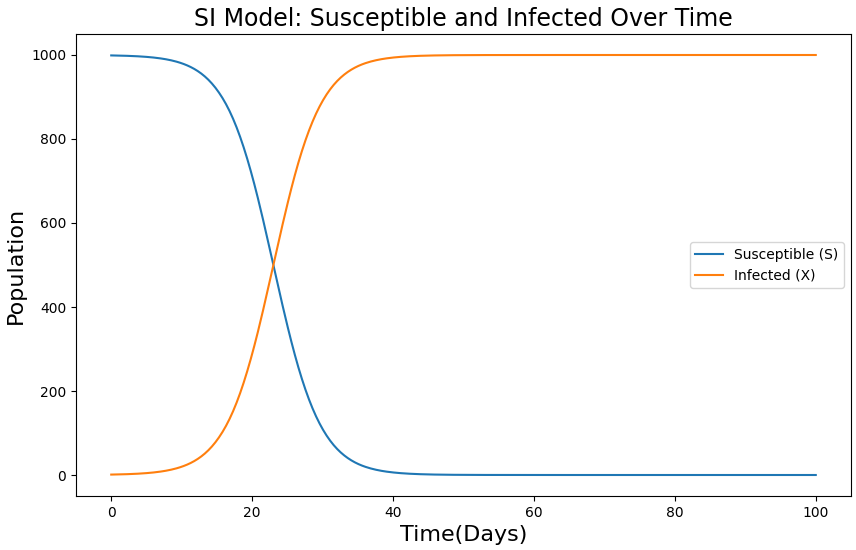

Figure 2.1: [1]The characteristic S-shaped logistic growth curve of the SI model, showing the fraction of infected individuals over time. The curve has exponential growth initially when most of the population is susceptible, then saturates as the supply of susceptibles is exhausted. Parameters: $\beta = 0.5$, $n = 100000$ $s_0 = 0.99$, $x_0 = 0.01$.

The solution produces an S-shaped logistic growth curve (Figure 2.1) with three characteristic phases:

1. **Slow initial growth**: When few are infected (x$\ll$1), growth is roughly exponential $x(t) \approx x_0$

2. **Rapid spread**: As more become infected, the growth accelerates
3. **Saturation**: When most are infected ($x \approx 1$), growth slows as few susceptibles remain.

Key limitations of the SI model:

- No recovery mechanism (infected individuals remain infectious forever)
- Predicts entire population eventually gets infected
- Assumes homogeneous mixing (all contacts equally likely)

2.2 The SIR Model

The SIR model provides a more realistic representation of many diseases by adding a third group - recovered individuals (R). In this model, infected people can eventually recover from the disease and gain permanent immunity, meaning they can't get infected again or spread the disease to others. The recovery process happens at a rate γ, which determines how long people typically remain infectious. The key insight from this model is the concept of the basic reproduction number R_0, which is the ratio of the infection rate to the recovery rate (β/γ). This important number tells us whether a disease will spread through the population or die out. If each infected person spreads the disease to more than one other person on average ($R_0 > 1$), the disease will grow into an epidemic. If each infected person spreads to less than one other ($R_0 < 1$), the disease will eventually disappear. The model shows that diseases don't necessarily infect everyone - there's often a portion of the population that never gets infected because the disease runs out of susceptible people to infect before reaching them.

This model captures three key processes[1]:

- Susceptible individuals becoming infected through contact
- Infected individuals recovering with immunity
- The possibility of disease dying out before reaching everyone

The model is governed by these coupled differential equations:

$$\frac{ds}{dt} = -\beta sx \tag{2.8}$$

$$\frac{dx}{dt} = \beta sx - \gamma x \tag{2.9}$$

$$\frac{dr}{dt} = \gamma x \tag{2.10}$$

where:

- $s(t)$, $x(t)$, $r(t)$ are fractions of susceptible, infected, and recovered individuals
- β is the infection rate (contacts per person per time)
- γ is the recovery rate (recoveries per infected per time)
- The constraint $s + x + r = 1$ always holds

We can find an analytical solution through these steps:

1. **Eliminate variables**: Combine equations (1) and (3):

$$\frac{1}{s}\frac{ds}{dt} = -\frac{\beta}{\gamma}\frac{dr}{dt} \tag{2.11}$$

2. **Integrate** to find the relationship between s and r:

$$s = s_0 e^{-\beta r/\gamma} \tag{2.12}$$

where s_0 is the initial susceptible fraction.

3. Substitute into the r equation:

$$\frac{dr}{dt} = \gamma(1 - r - s_0 e^{-\beta r/\gamma}) \tag{2.13}$$

4. The solution can be expressed as:

$$t = \frac{1}{\gamma}\int_0^r \frac{du}{1 - u - s_0 e^{-\beta u/\gamma}} \tag{2.14}$$

While this integral cannot be solved in closed form, it can be evaluated numerically. Figure 2.2 shows a typical numerical solution.

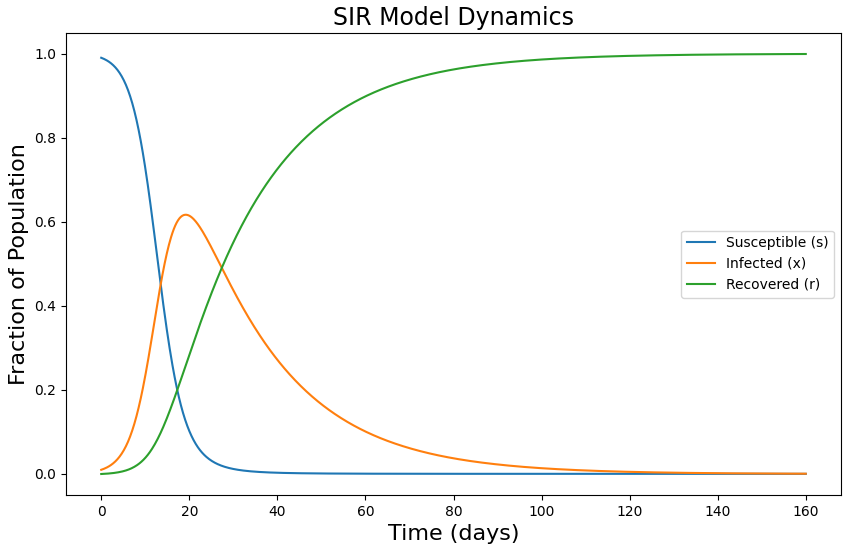

Figure 2.2: [1]Time evolution in the SIR model showing fractions of susceptible (S), infected (I), and recovered (R) individuals. The infected fraction peaks then declines as recovery creates herd immunity. Parameters: $\beta = 0.5$, $\gamma = 0.4$, $n = 100000$, $s_0 = 0.99$, $x_0 = 0.01$.

The model reveals a critical threshold condition:

$$R_0 = \frac{\beta}{\gamma} \tag{2.15}$$

where R_0 is the *basic reproduction number*.

- When $R_0 > 1$: The disease spreads (epidemic)
- When $R_0 \leq 1$: The disease dies out

While powerful, the SIR model has several limitations:

- Assumes homogeneous mixing (all contacts equally likely)
- Uses exponential recovery times (unrealistic for many diseases)
- Ignores network structure of contacts
- Assumes permanent immunity after recovery

2.3 The SIS Model

Some diseases don't provide permanent immunity after infection, meaning people can get sick again after recovering. The SIS model handles this situation by allowing recovered individuals to return to the susceptible state. This creates a continuous cycle where people move between being susceptible and infected. Unlike the SIR model where diseases eventually die out or infect part of the population, the SIS model can reach a steady state where the disease remains present indefinitely at a constant level in the population. The balance point depends on the relationship between the infection rate and recovery rate. When the infection rate is high compared to recovery, a large fraction of the population will be infected at any given time. When recovery happens quickly relative to new infections, the disease maintains only a small presence in the population. This model is particularly useful for understanding diseases like the common cold or certain bacterial infections where people can be reinfected multiple times throughout their lives.

The SIS model features two states with continuous cycling between them[1]:

- Susceptible (S) $\rightarrow$ Infected (I) via contact with infected individuals
- Infected (I) $\rightarrow$ Susceptible (S) through recovery without immunity

The model is governed by these coupled differential equations:

$$\frac{ds}{dt} = \gamma x - \beta s x \tag{2.16}$$

$$\frac{dx}{dt} = \beta s x - \gamma x \tag{2.17}$$

with the constraint:

$$s + x = 1 \tag{2.18}$$

where:

- β = infection rate (transmission probability $\times$ contact rate)
- γ = recovery rate (inverse of average infectious period)
- s, x = fractions of susceptible and infected populations

Substituting $s = 1 - x$ into equation (18) gives:

$$\frac{dx}{dt} = (\beta - \gamma - \beta x)x \tag{2.19}$$

This logistic-type equation has solution:

$$x(t) = \frac{(1 - \gamma/\beta)Ce^{(\beta-\gamma)t}}{1 + Ce^{(\beta-\gamma)t}} \tag{2.20}$$

where the constant C is determined by initial conditions:

$$C = \frac{\beta x_0}{\beta - \gamma - \beta x_0} \tag{2.21}$$

For small initial infection ($x_0 \to 0$) in large populations:

$$x(t) = \frac{x_0(\beta - \gamma)e^{(\beta-\gamma)t}}{\beta - \gamma + \beta x_0 e^{(\beta-\gamma)t}} \tag{2.22}$$

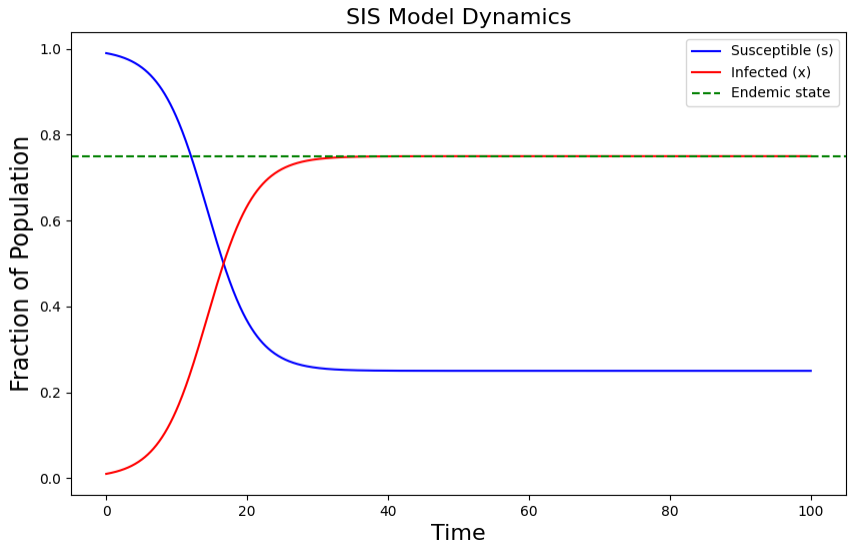

Figure 2.3: [1]Dynamics of the SIS model showing convergence to endemic equilibrium when $R_0 > 1$ ($\beta > \gamma$). Initialy susceptible converts to infected exponentially then an equilibrium reaches as infected are also converted to suseptible. Parameters: $\beta = 0.5$, $\gamma = 0.3$, $n = 100000$, $s_0 = 0.99$, $x_0 = 0.01$.

When $\beta > \gamma$, the system reaches a stable endemic state where:

$$x^* = \frac{\beta - \gamma}{\beta} = 1 - \frac{1}{R_0} \tag{2.23}$$

Characteristics of this equilibrium:

- Constant fraction infected despite individual recoveries
- Balance between new infections and recoveries
- Specific individuals change status over time

The basic reproduction number remains:

$$R_0 = \frac{\beta}{\gamma} \tag{2.24}$$

Critical behaviors:

- $R_0 > 1$: Disease persists at endemic level
- $R_0 \leq 1$: Disease dies out exponentially

Limitation of SIS Model:

- Assumes homogeneous mixing population
- Exponential distributions for infectious periods
- No demographic turnover (births/deaths/migration)
- Constant parameters over time

2.3.1 Comparison of SI, SIR and SIS Models

Table 2.1: Feature comparison of basic epidemic models

Characteristic	SI Model	SIR Model	SIS Model
Recovery	No	Yes (with immunity)	Yes (no immunity)
End state	All infected	Some recover	Endemic or die out
Epidemic threshold	None	$R_0 = 1$	$R_0 = 1$
Typical applications	Incurable diseases	Chickenpox, measles	Colds, STDs

Chapter 3

Network Effects on Disease Spread

While the basic models assume everyone has an equal chance of meeting everyone else, real-world social networks show that disease transmission follows more complex patterns. Some individuals, known as "super-spreaders," have many more social connections than average and can infect large numbers of people. Other people may belong to tightly-knit groups where diseases can spread quickly within the group but have limited connections to outside. Network models capture these realities by representing individuals as nodes and their interactions as connections between these nodes. This approach reveals that diseases can spread much faster through real social networks than the simple models predict, especially in the early stages when the disease reaches well-connected individuals. Understanding these network effects helps public health officials design more effective intervention strategies, such as targeted vaccination programs that focus on the most connected individuals rather than random members of the population.

Table 3.1: Comparison of Key Epidemic Models

Model Name	**Population Groups**	**Key Characteristics**	**Critical Value**
SI Model	Susceptible, Infected	No recovery, all get infected	None
SIR Model	Susceptible, Infected, Recovered	Permanent immunity after recovery	$R_0 = \beta/\gamma$
SIS Model	Susceptible, Infected	Temporary immunity, reinfection possible	$R_0 = \beta/\gamma$

3.1 Time-Dependent Properties on Networks

Traditional epidemic analysis often focuses on final outcomes, but understanding temporal progression is crucial for real-world applications. On networks, the dynamics differ fundamentally from homogeneous mixing models due to structural constraints. The infection front propagates through connected nodes, creating degree-dependent spreading patterns where well-connected nodes accelerate early transmission. The network structure also induces characteristic temporal signatures: rapid early spread followed by slower percolation through peripheral nodes, and potential resurgence in models with reinfection. These features are inaccessible for static analysis but critical for assessing the timing and effectiveness of the intervention.

3.1.1 Time-Dependent Properties of the SI Model on Networks

[1]The temporal dynamics of epidemics on networks require analyzing the evolution of vertex-state probabilities. For repeated outbreaks on the same network, we consider $s_i(t)$ and $x_i(t)$ as the probabilities that

vertex i is susceptible or infected at time t, governed by the adjacency matrix A_{ij}. Focusing on the giant component (where epidemics ultimately reach all vertices), the dynamics obey:

$$\frac{ds_i}{dt} = -\beta s_i \sum_j A_{ij} x_j = -\beta s_i \sum_j A_{ij}(1 - s_j) \tag{3.1}$$

$$\frac{dx_i}{dt} = \beta s_i \sum_j A_{ij} x_j = \beta(1 - x_i) \sum_j A_{ij} x_j \tag{3.2}$$

These coupled nonlinear equations use initial conditions $x_i = c/n$ and $s_i = 1 - c/n$ for uniformly random initial infections, simplifying to $x_i = 0$, $s_i = 1$ in the large-n limit.

For early stages ($x_i \ll 1$), we linearize Eq. 3.2:

$$\frac{dx_i}{dt} = \beta \sum_j A_{ij} x_j \tag{3.3}$$

Expressed in matrix form with $\mathbf{x}$ as the infection probability vector:

$$\frac{d\mathbf{x}}{dt} = \beta A \mathbf{x} \tag{3.4}$$

The solution decomposes into adjacency matrix eigenmodes:

$$\mathbf{x}(t) = \sum_{r=1}^{n} a_r(t) \mathbf{v}_r \tag{3.5}$$

where $\mathbf{v}_r$ are eigenvectors with eigenvalues κ_r. Substituting into Eq. 3.4 gives:

$$\frac{da_r}{dt} = \beta \kappa_r a_r \tag{3.6}$$

with solutions:

$$a_r(t) = a_r(0) e^{\beta \kappa_r t} \tag{3.7}$$

The dominant behavior comes from the largest eigenvalue κ_1:

$$\mathbf{x}(t) \sim e^{\beta \kappa_1 t} \mathbf{v}_1 \tag{3.8}$$

This reveals that:

- Early growth depends on $\beta\kappa_1$ rather than $\beta\langle k \rangle$
- Infection probabilities align with eigenvector centrality $\mathbf{v}_1$

3.1.2 Time-Dependent Properties of the SIR Model on Networks

[1]The temporal dynamics of the SIR model on networks reveal crucial differences from both the SI model and homogeneous mixing assumptions. For a network with adjacency matrix A_{ij}, we track three time-dependent probabilities for each vertex i:

$$\begin{aligned} &s_i(t) : \text{Susceptible probability} \\ &x_i(t) : \text{Infected probability} \\ &r_i(t) : \text{Recovered probability} \end{aligned} \tag{3.9}$$

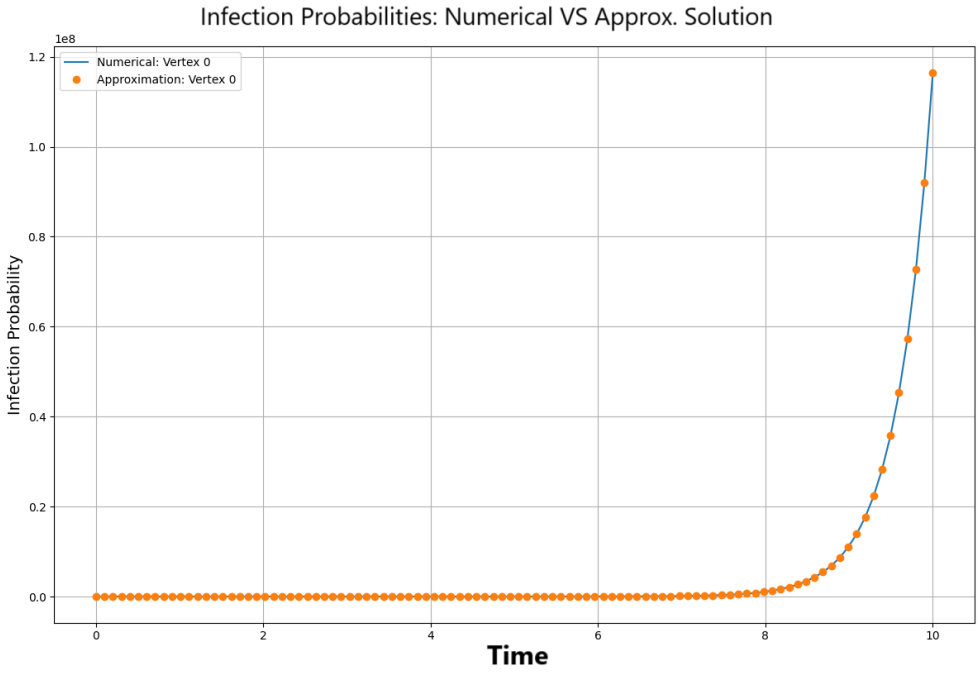

Figure 3.1: Comparison of first-order approximation (dashed) with simulations (dots) on a configuration model network. It shows that Infection Probability obtained by Numerical solution and Approximate solution are same. Parameters: $\beta = 0.5$, $n = 100000$, $s_0 = 0.99$, $x_0 = 0.01$.

The exact evolution follows these coupled nonlinear equations:

$$\frac{ds_i}{dt} = -\beta s_i \sum_{j=1}^{n} A_{ij} x_j \quad \text{(Infection process)} \tag{3.10}$$

$$\frac{dx_i}{dt} = \beta s_i \sum_{j=1}^{n} A_{ij} x_j - \gamma x_i \quad \text{(Infection minus recovery)} \tag{3.11}$$

$$\frac{dr_i}{dt} = \gamma x_i \quad \text{(Recovery process)} \tag{3.12}$$

with conservation law $s_i(t) + x_i(t) + r_i(t) = 1$ and initial conditions:

$$s_i(0) = 1 - \frac{c}{n}, \quad x_i(0) = \frac{c}{n}, \quad r_i(0) = 0 \tag{3.13}$$

For $t \to 0$ with $x_i \ll 1$ and $s_i \approx 1$, we linearize Eq. 3.11:

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} (\beta A_{ij} - \gamma \delta_{ij}) x_j \tag{3.14}$$

This can be expressed in matrix form using the modified adjacency matrix:

$$\mathbf{M} = \mathbf{A} - \frac{\gamma}{\beta}\mathbf{I} \tag{3.15}$$

yielding the matrix differential equation:

$$\frac{d\mathbf{x}}{dt} = \beta \mathbf{M}\mathbf{x} \tag{3.16}$$

The solution decomposes into adjacency matrix eigenmodes:

$$\mathbf{x}(t) = \sum_{r=1}^{n} a_r(0)\mathbf{v}_r \exp\left[(\beta\kappa_r - \gamma)t\right] \tag{3.17}$$

where $\{\kappa_r\}$ and $\{\mathbf{v}_r\}$ are the eigenvalues and eigenvectors of $\mathbf{A}$. The dominant behavior comes from the principal eigenmode:

$$\mathbf{x}(t) \approx a_1(0)\mathbf{v}_1 e^{(\beta\kappa_1 - \gamma)t} \tag{3.18}$$

The outbreak criterion requires at least one growing mode:

$$\beta\kappa_1 - \gamma > 0 \quad \Rightarrow \quad \frac{\beta}{\gamma} > \frac{1}{\kappa_1} \tag{3.19}$$

This reveals three key insights:

- The epidemic threshold depends on the spectral radius κ_1
- Networks with larger κ_1 facilitate disease spread
- The critical point differs from the homogeneous mixing case ($\beta/\gamma > 1$)

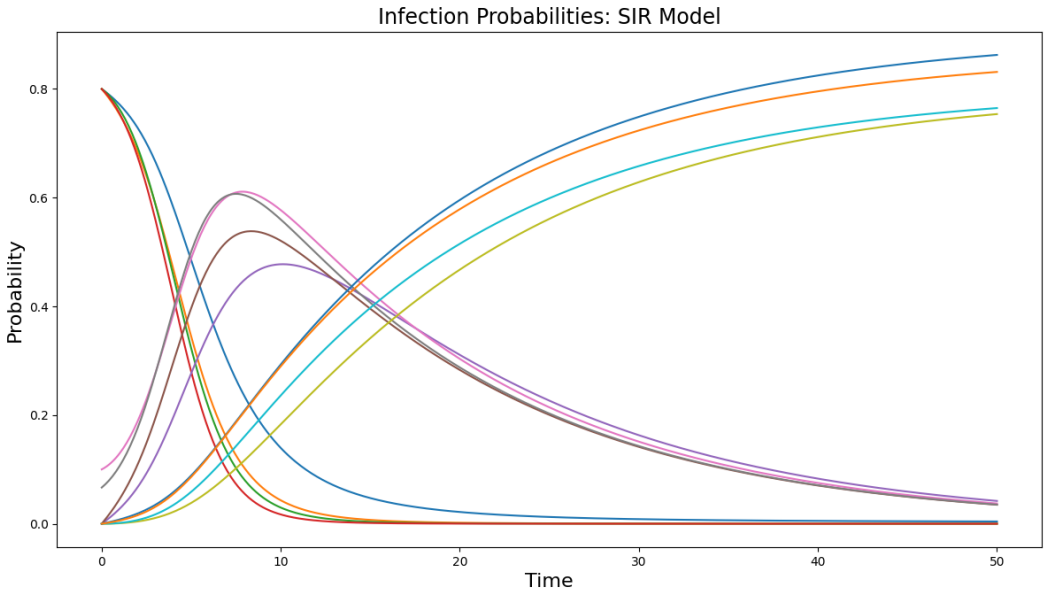

Figure 3.2: Fractions of susceptible, infected, and recovered vertices of various degrees in the SIR model. The fraction of vertices of degree k that are susceptible, infected, and recovered as a function of time for k = 1, 2, 4, ... on a network with an exponential degree distribution. Parameters: $\beta = 0.5$, $\gamma = 0.4$, $n = 100000$, $s_0 = 0.99$, $x_0 = 0.01$.

3.1.3 Time-Dependent Properties of the SIS Model on Networks

[1]The temporal dynamics of the SIS model in networks captures the essential features of diseases with reinfection. The model is governed by the following master equations for each vertex i:

$$\frac{ds_i}{dt} = -\beta s_i \sum_j A_{ij} x_j + \gamma x_i \quad \text{(Susceptible dynamics)} \tag{3.20}$$

$$\frac{dx_i}{dt} = \beta s_i \sum_j A_{ij} x_j - \gamma x_i \quad \text{(Infected dynamics)} \tag{3.21}$$

with the constraint $s_i + x_i = 1$ for all times. These equations share the same initial conditions as previous models: $x_i(0) = c/n$, $s_i(0) = 1 - c/n$ for c initial infections.

Linearizing Eq. 3.21 for small x_i yields:

$$\frac{dx_i}{dt} = \beta \sum_j A_{ij} x_j - \gamma x_i \tag{3.22}$$

Remarkably, this is identical to the SIR model's early dynamics (Eq. 3.14), leading to:

$$\mathbf{x}(t) \approx \sum_{r=1}^{n} a_r(0) \mathbf{v}_r e^{(\beta\kappa_r - \gamma)t} \tag{3.23}$$

The epidemic threshold consequently matches the SIR case:

$$\frac{\beta}{\gamma} > \frac{1}{\kappa_1} \tag{3.24}$$

At late times ($t \to \infty$), setting $dx_i/dt = 0$ in Eq. 3.21 gives the self-consistent equation:

$$x_i^* = \frac{\sum_j A_{ij} x_j^*}{\gamma/\beta + \sum_j A_{ij} x_j^*} \tag{3.25}$$

This nonlinear equation admits two important limiting cases:

1. **Near Threshold ($\beta/\gamma 1/\kappa_1$)** When just above the epidemic threshold:

$$x_i^* \approx \frac{\beta}{\gamma} \sum_j A_{ij} x_j^* \approx \kappa_1 x_i^* \tag{3.26}$$

implying the solution is proportional to the principal eigenvector:

$$x_i^* \propto [\mathbf{v}_1]_i \tag{3.27}$$

2. **Strong Epidemic ($\beta/\gamma \gg 1/\kappa_1$)** When transmission dominates recovery:

$$x_i^* \approx 1 - \frac{\gamma}{\beta \sum_j A_{ij}} \approx 1 \tag{3.28}$$

Chapter 4

GNN Model: A Modern Approach to Complex Network Dynamics

Graph Neural Networks (GNNs) have emerged as a transformative modeling framework for learning and predicting contagion dynamics on complex networks. Classical mechanistic models—such as SIS (Susceptible-Infected-Susceptible) and SIR (Susceptible-Infected-Recovered)—are based on predefined, often simplistic, transition rules that assume homogeneity and independence among individuals. While these models have significantly contributed to our understanding of epidemic spreading, they face critical limitations in representing the nuanced, nonlinear, and heterogeneous interactions present in real-world networks. They often ignore important factors like node and edge attributes, higher-order neighborhood effects, or dynamic interactions over multiplex networks. Consequently, their predictive power is limited, especially when applied to networks or dynamics that deviate from idealized assumptions.

GNNs, in contrast, are data-driven models capable of learning complex contagion mechanisms directly from observational time series data, without requiring a priori assumptions about the underlying processes. Each node in a GNN learns a local function that determines its future state based on its current state, attributes, and the aggregated information from its neighbors. This aggregation is performed using message passing and attention mechanisms, which ensure that the model respects the network's topological structure and accounts for edge attributes, such as weights or types of interactions. The GNN architecture is inherently *permutation-invariant*, *locally contextual*, and *inductive*, allowing it to generalize effectively to unseen networks with similar local patterns. This stands in stark contrast to classical models, which often lack such flexibility and generalization capabilities.

Another advantage of GNNs is their robustness and expressivity. They can model both stochastic and deterministic dynamics, handle both discrete and continuous state spaces, and encode metadata at both node and edge levels. Furthermore, attention mechanisms within GNNs introduce a layer of interpretability by quantifying the influence of neighboring nodes on the outcome of a target node, offering valuable insights into the nature of the spreading process. These features make GNNs highly adaptable to the challenges of real-world contagion modeling, where networks are often dynamic, multi-scale, and only partially observable.

4.1 Modeling Network Dynamics with Graph Neural Networks[2]

In this approach, we consider an unknown dynamical process M that evolves on a known network structure or ensemble of networks, denoted as $G = (V, E, \Phi, \Omega)$. The node set is $V = \{v_1, \ldots, v_N\}$, and the edge set is $E = \{e_{ij} \mid v_j \text{ is connected to } v_i\}$. The network includes metadata: node attributes $\Phi_i = (\varphi_1(v_i), \ldots, \varphi_Q(v_i))$ and edge attributes $\Omega_{ij} = (\omega_1(e_{ij}), \ldots, \omega_P(e_{ij}))$, where $\varphi_q : V \to R$ and $\omega_p : E \to R$. These attributes can encode properties such as node characteristics or edge weights. We denote the full node and edge attribute matrices as $\Phi = (\Phi_i)_{v_i \in V}$ and $\Omega = (\Omega_{ij})_{e_{ij} \in E}$, respectively.

We assume that the process M generates a time series dataset $D = (X, Y)$ on the network G, where $X = (X_1, \ldots, X_T)$ represents the node states at each time step, and $Y = (Y_1, \ldots, Y_T)$ the corresponding outcomes. Formally, the dynamic is described as:

$$Y_t = M(X_t, G) \tag{4.1}$$

where $X_t \in S^{|V|}$, $Y_t \in R^{|V|}$, and S, R are the spaces of possible node states and outcomes, respectively. For node v_i, we denote $x_i(t) = (X_t)_i$ and $y_i(t) = (Y_t)_i$.

When S is discrete and finite, $y_i(t)$ becomes a probability vector over S, representing the conditional transition probabilities. Since we typically do not observe these directly, we define the observed outcome vector $\tilde{y}_i(t) \in \{0, 1\}^{|S|}$ as:

$$(\tilde{y}_i(t))_m = \delta(x_i(t + \Delta t), m), \quad \forall m \in S \tag{4.2}$$

where δ is the Kronecker delta function and Δt is the timestep size.

The model assumes *locality*: each node's next state depends only on its own state and the states of its neighbors. Thus, the true function governing the process M at a node v_i is:

$$y_i = f(x_i, \Phi_i, x_{N_i}, \Phi_{N_i}, \Omega_{iN_i}) \tag{4.3}$$

Here, N_i denotes the neighbors of v_i, x_{N_i} their states, Φ_{N_i} their attributes, and Ω_{iN_i} the edge attributes from v_i to its neighbors.

Our goal is to construct an approximate model $\hat{M}$, parameterized by a Graph Neural Network (GNN) with weights Θ, that learns to mimic the true process M:

$$\hat{M}(X'_t, G', \Theta) \approx M(X'_t, G') \tag{4.4}$$

for any network G' and node states X'_t. At the node level, the GNN learns a function $\hat{f}$ to predict the outcome:

$$\hat{y}_i = \hat{f}(x_i, \Phi_i, x_{N_i}, \Phi_{N_i}, \Omega_{iN_i}, \Theta) \tag{4.5}$$

To train the GNN, we define a global loss function $L(\Theta)$ as the average of local losses over the dataset. However, because data may be imbalanced (e.g., some node configurations appear more frequently), we introduce weights $w_i(t)$ and define the loss as:

$$L(\Theta) = \sum_{t \in T'} \sum_{v_i \in V'(t)} \frac{w_i(t)}{Z'} \mathcal{L}(y_i(t), \hat{y}_i(t)) \tag{4.6}$$

where Z' is a normalization constant. The weights are chosen to upweight rare input configurations:

$$w_i(t) \propto \rho(k_i, x_i, \Phi_i, x_{N_i}, \Phi_{N_i}, \Omega_{iN_i})^{-\lambda} \tag{4.7}$$

Here, ρ is the empirical frequency of the input pattern, k_i is the degree of node v_i, and $\lambda \in [0, 1]$ is a hyperparameter controlling the level of importance sampling.

This framework is first validated using a simple synthetic contagion model—specifically, the Susceptible-Infected-Susceptible (SIS) dynamics. In this model, nodes are either in state S or I and the infection probability depends on the number of infected neighbors ℓ, via a function $\alpha(\ell)$, while recovery occurs at a constant rate β. A key assumption here is that infections occur independently across neighbors, making it an ideal baseline for evaluating the GNN's ability to learn from simple, interpretable transition rules.

Thus, this formulation lays the mathematical and architectural foundation for modeling local, temporal contagion dynamics using GNNs—allowing us to build models that are inductive, data-driven, and capable of generalizing across heterogeneous networks.

4.2 GNN Architecture and Training Details

The GNN model is structured in three main stages: input encoding via MLPs, attention-based aggregation of neighborhood information, and final prediction through output MLPs. The architecture is expressive, modular, and capable of incorporating both node and edge features. By modifying the classical GAT attention into a flexible weighted summation form, the model gains increased ability to learn complex dynamic behaviors on graphs, which is further demonstrated through experiments on various synthetic contagion processes.

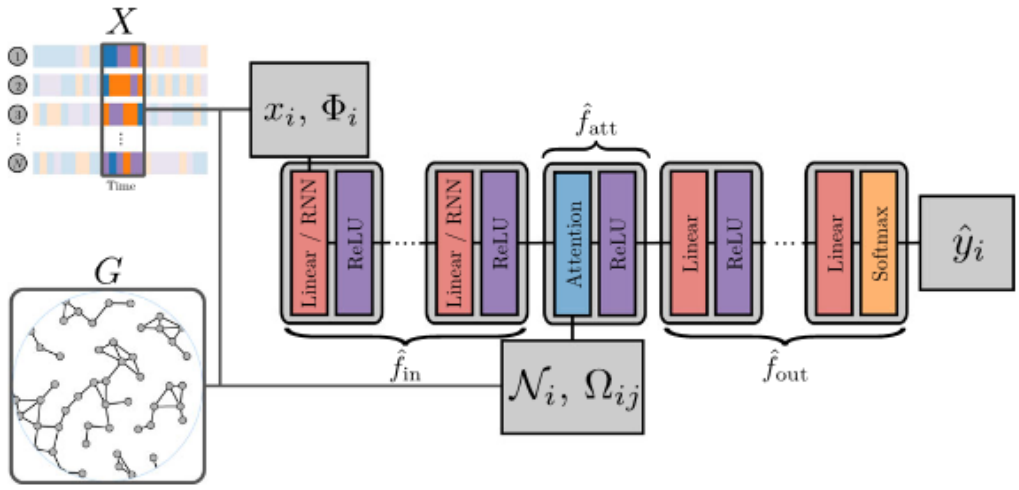

Figure 4.1: Visualization of the GNN architecture[2]. The blocks of different colors represent mathematical operations. The red blocks correspond to trainable affine transformation parametrized by weights and biases. The purple blocks represent activation functions between each layer. The core of the model is the attention module, which is represented in blue. The orange block at the end is an activation function that transforms the output into the proper outcomes.

4.2.1 Architecture

The Graph Neural Network (GNN) architecture proposed in this work (4.1) is designed to approximate complex contagion dynamics by learning node-level transition functions in a fully data-driven and locally structured manner. Each node v_i begins with a state x_i, which is transformed into a feature vector using a shared multilayer perceptron (MLP) $\hat{f}_{\text{in}}$:

$$\xi_i = \hat{f}_{\text{in}}(x_i) \tag{4.8}$$

where $\hat{f}_{\text{in}} : S \to R^d$ if node attributes are not used. When node attributes $\Phi_i \in R^Q$ are available, they are concatenated with the state x_i, and the input MLP becomes $\hat{f}_{\text{in}} : S \times R^Q \to R^d$. The resulting vector ξ_i encodes a learned representation of the node state and attributes.

Attention mechanism. To aggregate information from the neighborhood, the model applies a modified attention mechanism $\hat{f}_{\text{att}}$, which considers both node features and, when available, edge attributes $\Omega_{ij} \in R^P$. These edge features are first passed through an MLP:

$$\psi_{ij} = \hat{f}_{\text{edge}}(\Omega_{ij}), \quad \hat{f}_{\text{edge}} : R^P \to R^{d_{\text{edge}}} \tag{4.9}$$

The attention coefficient a_{ij}, which measures the influence of node v_j's state on node v_i, is calculated using three affine transformations $A, B : R^d \to R$ and $C : R^{d_{\text{edge}}} \to R$:

$$a_{ij} = \sigma\left(A(\xi_i) + B(\xi_j) + C(\psi_{ij})\right) \tag{4.10}$$

where $\sigma(x) = \frac{1}{1+e^{-x}}$ is the logistic sigmoid function. This constrains $a_{ij} \in (0, 1)$, meaning that when $a_{ij} \approx 0$, v_j has minimal influence on v_i, and when $a_{ij} \approx 1$, its influence is maximal.

The aggregated node feature vector ν_i is then computed as:

$$\nu_i = \hat{f}_{\text{att}}(\xi_i, \xi_{N_i}) = \xi_i + \sum_{v_j \in N_i} a_{ij}\xi_j \tag{4.11}$$

Unlike the normalized attention of the original GAT architecture—where $\sum_{j \in N_i} a_{ij} = 1$—this formulation allows general weighted sums. This enhances expressivity and makes the model capable of representing a broader class of network dynamics.

Finally, the prediction for each node is produced by another MLP $\hat{f}_{\text{out}}$:

$$\hat{y}_i = \hat{f}_{\text{out}}(\nu_i) \tag{4.12}$$

which maps the aggregated feature vector to the output space, such as a probability distribution over future states.

4.2.2 Training Settings

The model is trained using the cross-entropy loss between the ground truth outcome y_i and the predicted vector $\hat{y}_i$:

$$\mathcal{L}(y_i, \hat{y}_i) = -\sum_m y_{i,m} \log \hat{y}_{i,m} \tag{4.13}$$

where $y_{i,m}$ is either a transition probability or a one-hot encoded observed outcome, depending on whether the underlying dynamics is stochastic or deterministic. For stochastic processes like SIS, observed outcomes $\tilde{y}_i$ are used as labels instead of the true transition probabilities due to practical limitations.

4.2.3 Performance Measure

To evaluate performance, the Pearson correlation coefficient r is used, defined between true outcomes Y and predicted outcomes $\hat{Y}$ as:

$$r = \frac{E[(Y - E[Y])(\hat{Y} - E[\hat{Y}])]}{\sqrt{E[(Y - E[Y])^2] \cdot E[(\hat{Y} - E[\hat{Y}])^2]}} \tag{4.14}$$

where $E[\cdot]$ denotes the expectation. A perfect correlation corresponds to $r = 1$, and deviations from this ideal are measured by the global error $1 - r$.

4.3 Importance of Graph Attention Networks (GAT) and Advantage of Combining GAT with RNN

In the study of contagion dynamics on networks, a central challenge is to accurately model how a node's state evolves based on the states of its neighbors and the underlying network structure. While recurrent neural networks (RNNs) are effective for modeling temporal sequences, they lack the inherent capability to encode spatial dependencies or topological information about the network. In contrast, Graph Neural Networks (GNNs) naturally handle graph-structured data by propagating information across nodes and edges.

Among various GNN architectures, the **Graph Attention Network (GAT)** is particularly well-suited for contagion modeling because it enables each node to dynamically assign importance to its neighbors through an attention mechanism. Instead of treating all neighbors equally (as in standard Graph Convolutional Networks), GAT learns an attention coefficient a_{ij} for each neighbor j of node i, computed as:

$$a_{ij} = \sigma\left(A(\xi_i) + B(\xi_j) + C(\psi_{ij})\right) \tag{4.15}$$

where ξ_i and ξ_j represent node features, ψ_{ij} denotes edge features, and A, B, C are trainable affine transformations. The logistic sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$ ensures that attention weights lie between 0 and 1.

This attention mechanism allows the model to flexibly learn *context-sensitive contagion rules*, where different neighbors can have varying levels of influence depending on their state and the edge characteristics. Such flexibility is crucial for accurately modeling heterogeneous and complex spreading behaviors in real-world networks.

Why RNN + GAT Outperforms RNN Alone

While GATs effectively capture spatial and relational information at each time step, they do not model the temporal evolution of node states over time. On the other hand, RNNs are designed to model temporal dependencies but operate on sequences, without a natural way to incorporate graph topology.

By combining the two, the model leverages the strengths of both architectures:

- **GAT** layers compute a feature representation for each node at each time step, incorporating the influence of neighboring nodes in a learned, adaptive manner.
- **RNN** layers then model the temporal evolution of these graph-based features, learning how node states change over sequences of time steps.

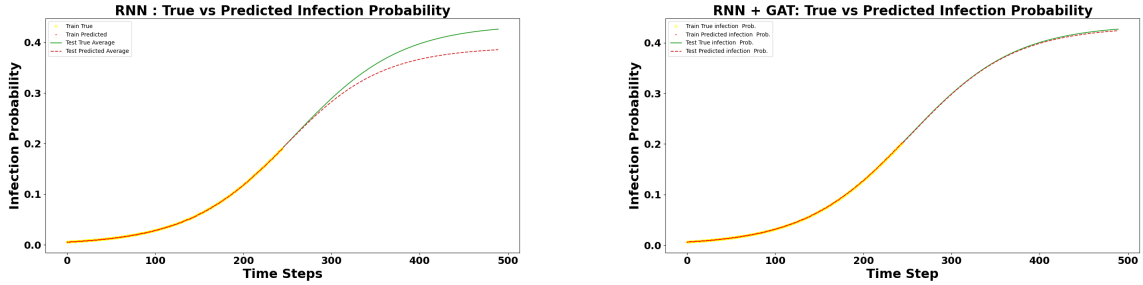

Figure 4.2: Predictions of RNN and GNN trained on a SIS Model dynamics. (a).Show prediction of SIS model data using RNN model. (b).Show prediction of SIS model data using GNN(RNN+GAT) model. Clearly GNN perform better prediction as it also consider neighbor interactions. Parameters: $\beta = 0.5$, $n = 1000$, $X(windowsizefortimeseriesdata) = 10timesteps$.

This hybrid architecture enables the model to:

- Capture both *local structural dependencies* and *global temporal dynamics*.
- Adapt to varying network topologies, improving generalization across unseen graphs.
- Handle complex contagion processes where infection probability depends nonlinearly on the number and influence of infected neighbors.

Empirical results from the study confirm that the RNN+GAT model achieves higher predictive accuracy than RNN alone, particularly in:

- **Complex contagion dynamics** where nonlinear reinforcement and inhibition are present.
- **Transfer learning scenarios** where models are tested on networks different from those seen during training.
- **Sparse or noisy data settings** where the attention mechanism helps prioritize influential neighbor interactions.

4.4 Behavior of Infection Probabilities: Simple vs. Complex Contagion

Figure 4.3 provides a comparative visualization of the transition probability functions for both the *simple contagion* (SIS) and *complex contagion* dynamics. These curves illustrate how the probability of a node becoming infected varies as a function of the number of infected neighbors ℓ. The figure compares three approaches: the ground truth (GT), predictions from the GNN model, and the maximum likelihood estimator (MLE) calculated directly from data.

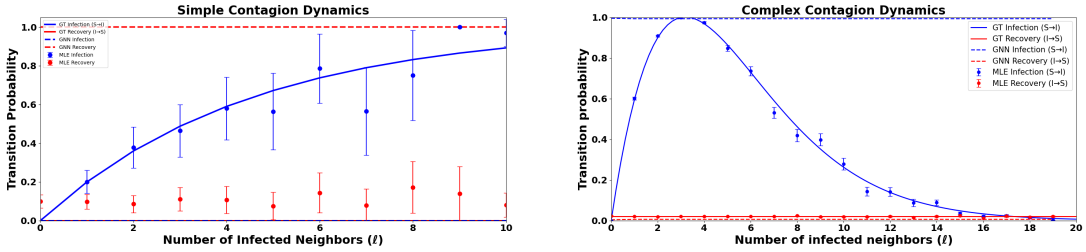

Figure 4.3: Predictions of GNN trained on a Barabási-Albert random network 72 (BA). (a). Transition probabilities of the simple contagion dynamics. (b). Transition probabilities of the complex contagion dynamics. The solid and dashed lines correspond to the transition probabilities of the dynamics used to generate the training data (labeled GT for "ground truth"), and predicted by the GNN, respectively. Symbols correspond to the maximum likelihood estimation (MLE) of the transition probabilities computed from the dataset D. The colors indicate the type of transition: infection (S $\rightarrow$ I) in blue and recovery (I $\rightarrow$ S) in red. The standard deviations, as a result of averaging the outcomes given ℓ, are shown using a colored area around the lines (typically narrower than the width of the lines) and using vertical bars for the symbols

Simple Contagion Dynamics

In Figure 4.3(a), the infection probability under the SIS model is governed by an exponential function:

$$\alpha(\ell) = 1 - (1 - \gamma)^{\ell} \tag{4.16}$$

This function increases monotonically with ℓ, reflecting the assumption that each infected neighbor contributes independently to the infection risk of a susceptible node. As ℓ grows, the probability saturates toward 1, producing a concave upward curve. The diminishing marginal effect of additional neighbors results from the multiplicative nature of independent infection events. The recovery process in the SIS model remains constant at a fixed probability β, regardless of ℓ, which is shown as a horizontal line in the figure.

Complex Contagion Dynamics

Figure 4.3(b) shows the infection probability for a complex contagion model, defined by the function:

$$\alpha(\ell) = \frac{\ell^3}{z(\eta)(e^{\ell/\eta} - 1)} \tag{4.17}$$

where $z(\eta)$ is a normalization factor ensuring that the function reaches a maximum value of 1, and η controls the location of the peak ℓ^*. This curve is non-monotonic and resembles the shape of the Planck radiation law. Initially, the infection probability increases with ℓ, capturing social reinforcement effects (e.g., peer pressure, collective behavior). However, after a critical threshold ℓ^*, the probability begins to decrease, modeling

the saturation or inhibitory effect where too many infected neighbors may trigger avoidance behaviors or reduce marginal influence. This results in a bell-shaped curve, which is qualitatively distinct from the simple contagion case.

Model Performance and Behavior

In both figures, solid lines denote the true analytical probability functions (ground truth), dashed lines show the GNN predictions, and markers represent the MLE estimates derived from empirical data. The GNN predictions closely match the ground truth curves, indicating the model's ability to learn accurate, smooth approximations of the transition dynamics from sparse and noisy data. In contrast, the MLE estimates are more irregular, especially at higher values of ℓ, due to limited sample counts for rare configurations. The GNN benefits from its parameter-sharing and inductive architecture, allowing it to generalize across neighborhood configurations and smooth over sampling noise. This capacity is especially evident in the complex contagion setting, where the GNN accurately recovers the non-monotonic shape of the transition curve that MLE fails to capture reliably.

4.5 Evaluating Generalization Across Network Structures: ER vs. BA Analysis

Motivation for Using ER and BA Networks

In the study of contagion dynamics on networks, it is essential to evaluate whether a model trained on one type of network can generalize to others with significantly different topological properties. Real-world networks exhibit a wide variety of structural features—such as degree distributions, clustering, and community structure—that influence the spread of contagion. To simulate these variations and stress-test the generalization capabilities of the GNN, the authors evaluate performance on two canonical network models:

- **Erdős–Rényi (ER)** networks, which are random graphs where each pair of nodes is connected independently with the same probability. These networks exhibit Poisson-like degree distributions and relatively uniform topology.
- **Barabási–Albert (BA)** networks, which are scale-free and built through a preferential attachment mechanism. These networks have a heavy-tailed (power-law) degree distribution, resulting in the emergence of high-degree hub nodes.

Evaluating the model on both ER and BA networks provides insight into how well the learned local rules generalize to unseen network structures with different spreading characteristics. While ER networks are homogeneous and easier to model, BA networks present challenges due to their high degree variability and localized hub influence.

Figure 4.4(a) and Figure 4.4(b) report the Pearson correlation coefficient r between the GNN-predicted infection probabilities and the true outcomes for ER and BA networks, respectively. These experiments test the ability of the GNN model, trained on synthetic contagion dynamics, to generalize to new, structurally distinct graphs.

In Figure 4.4(a), the model is evaluated on ER networks. The results show a high Pearson correlation, indicating that the GNN accurately predicts infection outcomes on these networks. The uniform structure of ER networks aligns well with the locality assumption in the GNN, where each node's dynamics depend only on its neighborhood. Since ER graphs lack hubs and have a relatively symmetric degree distribution, the infection process is smooth and consistent, making it easier for the GNN to generalize.

In contrast, Figure 4.4(b) evaluates the model on BA networks. Although the Pearson correlation remains strong, it is slightly reduced compared to the ER case. This difference is attributed to the heterogeneous structure of BA networks. In BA graphs, high-degree nodes (hubs) play a disproportionate role in spreading contagion, introducing local dynamics that may not have been fully captured if the training data lacked

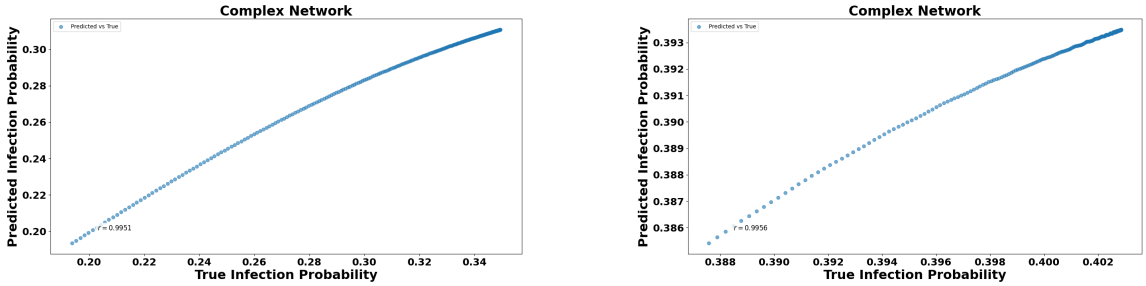

Figure 4.4: (a)GNN trained on Erdős–Rényi networks (ER). (b)GNN trained on Barabási-Albert networks (BA).The errors are obtained from the Pearson coefficients computed from subsets of the prediction-target pairs where all nodes have degree k. Parameters: $n = 1000$, $X(windowsizefortimeseriesdata) = 10timesteps$, $r = PearsonCorrelationCoefficient$.

similar structures. Despite this, the GNN still performs well, highlighting its inductive learning capabilities and its ability to handle both homogeneity and heterogeneity in node connectivity.

Reason for the Observed Behavior

The observed linear trends in both figures stem from the alignment between predicted and actual infection probabilities across a wide range of node states and degrees. In ER networks, the regular structure promotes consistent prediction quality across nodes, leading to a high correlation. For BA networks, the presence of hubs introduces variability—nodes connected to hubs experience different contagion pressures than isolated ones—which slightly disturbs the correlation.

Nevertheless, the GNN demonstrates strong generalization ability. Its use of local message-passing and attention mechanisms enables it to learn neighborhood-level patterns that apply across a broad set of network structures. The fact that it maintains high accuracy on both ER and BA networks supports the claim that the model learns transferable, structure-invariant contagion dynamics.

Chapter 5

Conclusion and Future Scope

5.1 Conclusion

In this project, we explored the dynamics of epidemic spread on complex networks through both classical compartmental models (SI, SIR, SIS) and modern learning-based methods like Graph Neural Networks (GNNs). Classical models offer foundational insights but assume homogeneous mixing, which fails to reflect real-world heterogeneity. Incorporating network theory allowed us to model the structure-driven variations in disease propagation, highlighting the role of degree distributions, network topology, and spectral properties.

The introduction of GNNs provided a data-driven approach capable of learning local contagion mechanisms from time series data. With attention mechanisms and temporal recurrence, GNNs demonstrated strong performance in capturing simple and complex contagion behaviors, and generalizing across different network structures such as Erdős–Rényi and Barabási–Albert graphs.

5.2 Future Scope

To further enhance the modeling and predictive capabilities of epidemic simulations on networks, the following directions are proposed:

1. **Integration with Real-World Datasets**:
   Future work can focus on validating the proposed models using real-world epidemiological data, such as COVID-19 or influenza outbreaks across regions. By incorporating features specific to real datasets (e.g., mobility patterns, demographic attributes, contact tracing), the models can be calibrated to reflect actual population behaviors and improve predictive realism.

2. **Feature-Based Model Customization**:
   As real-world data often come with multiple attributes—such as age, occupation, and vaccination status—models can be extended to include node and edge features. Using these features, we can model heterogeneous transmission risks, simulate targeted interventions, and personalize predictions at finer granularities[2].

3. **Incorporating Quantum Computing Approaches**:
   Inspired by the paper *"Studying the effect of lockdown using epidemiological modelling of COVID-19 and a quantum computational approach using the Ising spin interaction"* [4], there is significant promise in leveraging quantum computation for epidemic modeling. Quantum circuits derived from Ising Hamiltonians can simulate infection dynamics efficiently, offering potential advantages in time and space complexity. As epidemic simulations grow in complexity due to high-dimensional state spaces and network interactions, quantum algorithms may provide faster and more scalable solutions compared to classical counterparts.

4. **Hybrid Classical-Quantum Modeling**:
   A promising avenue is the combination of classical GNN-based predictors with quantum components for sub-tasks like parameter optimization, probabilistic state simulation, or encoding complex interactions. This could lead to novel hybrid epidemic models that exploit the strengths of both paradigms.

5. **Policy Modeling and Intervention Simulation**:
   Future studies could simulate various intervention strategies—such as lockdowns, vaccinations, and travel restrictions—on dynamic networks using the proposed frameworks. Combining real-time data assimilation with predictive modeling will help in proactive decision-making during public health crises.

Bibliography

[1] M. E. J. Newman, *Networks: An Introduction*, University of Michigan and Santa Fe Institute, 2010.

[2] Charles Murphy, Edward Laurence, Antoine Allard, *Deep learning of contagion dynamics on complex networks*, (2021).

[3] Kiss, I. Z., Miller, J. C., Simon, *Mathematics of Epidemics on Networks p. 598 https://doi.org/10.1007/978-3-319-50806-1* , (Springer, 2017).

[4] Anshuman Padhi, Sudev Pradhan, Pragna Paramita Sahoo, Kalyani Suresh, Bikash K. Behera, Prasanta K. Panigrahi, *Studying the effect of lockdown using epidemiological modelling of COVID-19 and a quantum computational approach using the Ising spin interaction*, (2020).

[5] M. H. A. Biswas, L. T. Paiva, MdR de Pinho, *A SEIR MODEL FOR CONTROL OF INFECTIOUS DISEASES WITH CONSTRAINT*, (2024).

[6] Lasko Basnarkov, *SEAIR Epidemic spreading model of COVID-19*, (2021).