HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

EFFECTS OF SUPERSPREADERS IN SPREAD OF EPIDEMIC

MATH MODELING

Supervisor: Dr. Bui Quoc Trung

Authors: Luong Huu Thanh
Student ID: 20225458
Class: Data Science & AI 03 K67

Hanoi, June 2025

Contents

1 Introduction

The unpredictable dynamics of infectious disease outbreaks are often driven by a small number of individuals known as superspreaders, who are responsible for a disproportionately large number of secondary infections. The 2003 SARS outbreak, where a handful of individuals caused rapid, global transmission, exemplifies this phenomenon. Traditional SIR models, which assume uniform transmission rates, fail to capture the heterogeneity introduced by such individuals, limiting their explanatory and predictive capabilities.

Building on the work presented in "Effects of Superspreaders in Spread of Epidemic" by Fujie and Odagaki (2007) [1], this study examines the impact of superspreaders within a spatially structured SIR framework. The authors propose two distinct models: one in which superspreaders exhibit intrinsically higher infectiousness within a fixed spatial range, and another—termed the hub model—in which superspreaders possess a broader range of contact, simulating individuals with extensive social connectivity. Through Monte Carlo simulations, key epidemic characteristics such as percolation probability, propagation speed, epidemic curves, and secondary infection distributions are analyzed across varying population densities and superspreader fractions. The findings indicate that superspreaders significantly enhance disease spread, with the hub model closely aligning with empirical data from the 2003 SARS outbreak in Singapore. These results highlight the importance of incorporating heterogeneity in contact patterns and spatial dynamics when modeling infectious disease transmission and formulating control strategies.

2 Modeling Approach

This study adopts a spatially structured SIR (Susceptible–Infected–Recovered) framework to explore how superspreaders influence the spread of infectious diseases. Individuals are placed within a two-dimensional continuous space, where infection probability depends on the distance between individuals, allowing the model to account for spatial transmission dynamics. Each individual occupies a fixed position and exists in one of three possible states: susceptible (S), infected (I), or recovered (R).

Two distinct models of superspreaders are considered. In the strong infectiousness model, superspreaders are characterized by a higher intrinsic ability to infect others within a fixed local range. In contrast, the hub model represents superspreaders as individuals who interact with a broader range of contacts, reflecting greater social connectivity rather

than increased infectiousness.

The simulation begins with a single infected individual, and a defined fraction of the population is designated as superspreaders. Disease transmission and recovery are simulated over discrete time steps using Monte Carlo methods. By varying population density and the proportion of superspreaders, the study analyzes key epidemic characteristics such as outbreak size, speed of spread, secondary infection patterns, and the probability of large-scale transmission.

This modeling approach provides a framework for comparing how different types of superspreading behavior impact epidemic outcomes and supports evaluation against real-world data, such as the 2003 SARS outbreak in Singapore.

3 Mathematical Model

3.1 Foundational Elements

- **Simulation Space:** A two-dimensional continuous square domain of size $L \times L$ with periodic boundary conditions is used to eliminate edge effects.
- **Population:** A total of N individuals are uniformly and randomly distributed across the domain. One individual is initially infected (state I), and the remaining individuals are initially susceptible (state S). A fixed fraction λ of the population is randomly designated as superspreaders.
- **SIR States:** Each individual can be in one of three discrete epidemiological states: susceptible (S), infected (I), or recovered (R).
- **Transmission Mechanism:** Infection occurs probabilistically. During each time step, an infected individual can transmit the disease to nearby susceptible individuals. The probability of transmission depends on the spatial distance r between the two individuals and is denoted as $w(r)$.

3.2 Strong Infectiousness Model

This model explores the hypothesis that superspreaders are biologically more infectious, meaning they transmit the disease more efficiently to individuals in close proximity.

- **Normal Individual:** The transmission probability for a normal infected individual decreases with distance and is limited to a maximum range r_0:

$$w_n(r) = \begin{cases} w_0 \left(1 - \frac{r}{r_0}\right)^2, & 0 \le r \le r_0 \\ 0, & r > r_0 \end{cases}$$

- **Superspreader:** Superspreaders in this model have the same interaction radius as normal individuals but transmit the disease at a constant (maximum) probability within that range:

$$w_s(r) = \begin{cases} w_0, & 0 \le r \le r_0 \\ 0, & r > r_0 \end{cases}$$

In this formulation, superspreaders are not socially distinct but are significantly more potent at transmitting the disease to close contacts. As a result, the epidemic is expected to form tight, densely infected clusters that expand gradually.

3.3 Hub Model

This model represents the alternative hypothesis that superspreaders are defined by their extensive social connectivity, enabling them to spread the disease across larger spatial scales.

- **Normal Individual:** Transmission dynamics for normal individuals remain unchanged from the Strong Infectiousness Model:

$$w_n(r) = \begin{cases} w_0 \left(1 - \frac{r}{r_0}\right)^2, & 0 \le r \le r_0 \\ 0, & r > r_0 \end{cases}$$

- **Superspreader:** Superspreaders in the hub model share the same functional form of $w(r)$ as normal individuals but have a larger effective contact radius $r_n = \sqrt{6} r_0$:

$$w_s(r) = \begin{cases} w_0 \left(1 - \frac{r}{r_n}\right)^2, & 0 \le r \le r_n \\ 0, & r > r_n \end{cases}$$

In this case, superspreaders are not inherently more infectious at close range but are ca-

pable of reaching a wider population. This leads to more rapid and spatially discontinuous spread, with new infection clusters potentially forming far from existing ones.

4 Solving Algorithm

The epidemic simulation is conducted using a Monte Carlo approach within a Python class framework. This algorithm models the stochastic nature of infection transmission and recovery over discrete time steps in a 2D spatial environment. It incorporates the influence of superspreaders under two distinct models: the strong infectiousness model and the hub model. The simulation employs probabilistic rules based on spatial distances and individual characteristics to capture the dynamics of an epidemic.

4.1 Initialization

At the start of each simulation, the following steps are performed:

- **Individual Placement:** N individuals are randomly positioned within an $L \times L$ two-dimensional space, where $L = 10 \times r_0$. **Periodic boundary conditions** are applied, meaning the space is treated as a torus where interactions can occur across the edges. This is crucial for eliminating boundary effects and is calculated using the formulae:

$$r = \sqrt{(\min(|x_1 - x_2|, L - |x_1 - x_2|))^2 + (\min(|y_1 - y_2|, L - |y_1 - y_2|))^2}$$

- **Superspreader Assignment:** A fraction λ of individuals are randomly designated as superspreaders through a uniform random draw.

- **Patient Zero:** The initial infected individual (patient zero) is placed at a randomly chosen position along the bottom edge of the simulation space.

- **State Initialization:** All other individuals begin in the susceptible state. Each individual's state is tracked in an array, with `0` representing susceptible, `1` representing infected, and `2` representing recovered.

- **Data Structures:** Various data structures are initialized to record positions, infection times, the infection tree, counts of secondary infections, new infections per time step, and the maximum geographical spread of the infection.

4.2 Simulation Loop

The simulation progresses in discrete time steps, up to a `max_steps` limit or until no infected individuals remain. In each step:

1. **Identify Infected Individuals:** Identify all individuals currently in the infected state.
2. **Infection Transmission:** For each infected individual:
    - The periodic distance to every susceptible individual is calculated using the `periodic_distance` method.
    - The infection probability $w(r)$ is determined using the `infection_probability` method. This probability is dependent on the distance r, the infector's superspreader status, and the chosen model type.
    - Susceptible individuals become infected with probability $w(r)$. If infected, their state is updated to `1`, their infection time is recorded, and the infection tree is updated to reflect the transmission. The count of secondary infections for the infector is incremented.
3. **Recovery Process:** Each infected individual has a probability γ of recovering and transitioning to the recovered state `2`. This probability is drawn independently for each infected individual.
4. **Metrics Updates:** After each step, the number of new infections that occurred is recorded, and the maximum distance of any infected individual from the initial patient zero's position is calculated.

4.3 Termination

The simulation concludes when either the `max_steps` limit is reached or when there are no longer any infected individuals in the simulation.

Finally, the simulation returns a dictionary containing all relevant data collected during the run, including the final positions and states of all individuals, their superspreader statuses, infection times, the complete infection tree, a summary of secondary infections, the progression of new infections per step, and the maximum spread of the epidemic.

Algorithm 1 RunSimulation

```
procedure RunSimulation(simulation_data, max_steps, model_type)
    for step ← 0 to max_steps − 1 do
        infected_indices ← FindIndices(simulation_data.states == INFECTED)
        if Length(infected_indices) == 0 then
            break                                   ▷ No more infected individuals
        end if
        infected_positions ← GetPositions(simulation_data.states > SUSCEPTIBLE)
        if Length(infected_positions) > 0 then
            distances ← CalculateDistances(infected_positions, initial_pos, L)
            Append(simulation_data.max_distances, Max(distances))
        else
            Append(simulation_data.max_distances, 0)
        end if
        new_infections_this_step ← 0
        for all infector_idx ∈ infected_indices do
            infector_pos ← simulation_data.positions[infector_idx]
            infector_super ← simulation_data.is_superspreader[infector_idx]
            for target_idx ← 0 to N − 1 do
                if simulation_data.states[target_idx] == SUSCEPTIBLE then
                    target_pos ← simulation_data.positions[target_idx]
                    distance ← CalculatePeriodicDistance(infector_pos, target_pos, L)
                    prob ← CalculateInfectionProbability(distance, infector_super, model_type)
                    if Random() < prob then
                        simulation_data.states[target_idx] ← INFECTED
                        simulation_data.infection_times[target_idx] ← step + 1
                        simulation_data.infection_tree[target_idx] ← infector_idx
                        simulation_data.secondary_infections[infector_idx] ← +1
                        new_infections_this_step ← +1
                    end if
                end if
            end for
        end for
        for all idx ∈ infected_indices do
            if Random() < γ then
                simulation_data.states[idx] ← RECOVERED
            end if
        end for
        Append(simulation_data.new_infections_per_step, new_infections_this_step)
    end for
    return simulation_data
end procedure
```

5 Simulation Result

The Monte Carlo simulation results provide insights into the spread of epidemics under the influence of superspreaders, comparing the strong infectiousness model and the hub model. The simulations were conducted on an $L \times L$ continuous space with N individuals, where $L = 10r_0$, and initial values for w_0 and γ were fixed at 1. Each result represents an average over 1000 Monte Carlo runs with different initial positions.

Percolation Probability

The percolation probability is defined as the ratio of trials where the infection, starting at the bottom of the system, reaches the top.

In both the strong infectiousness model and the hub model, a percolation transition is observed for all values of λ. At low densities, the infection does not percolate, meaning the disease stops before spreading throughout the system. Conversely, at high densities, the infection percolates, and the disease spreads across the entire system.

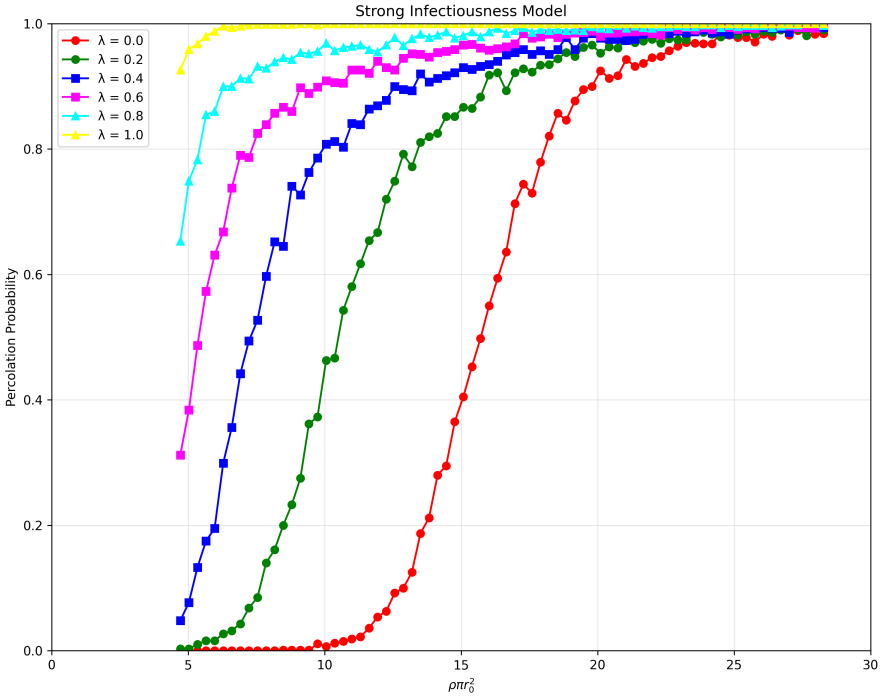

Figure 1: Percolation probability vs. population density ($\rho\pi r_0^2$) for different λ in the strong infectiousness model

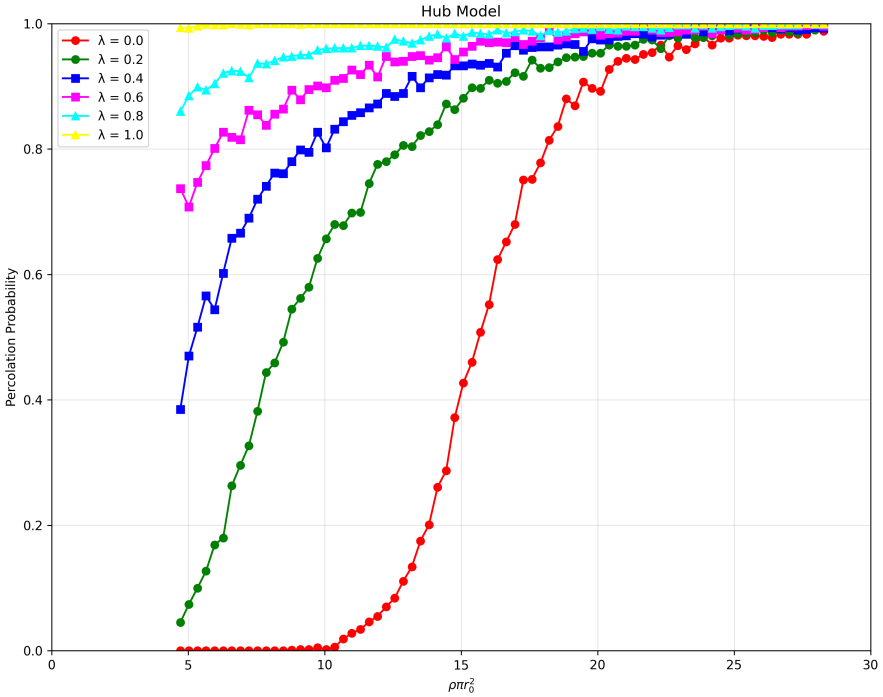

Figure 2: Percolation probability vs. population density ($\rho\pi r_0^2$) for different λ values in the hub model

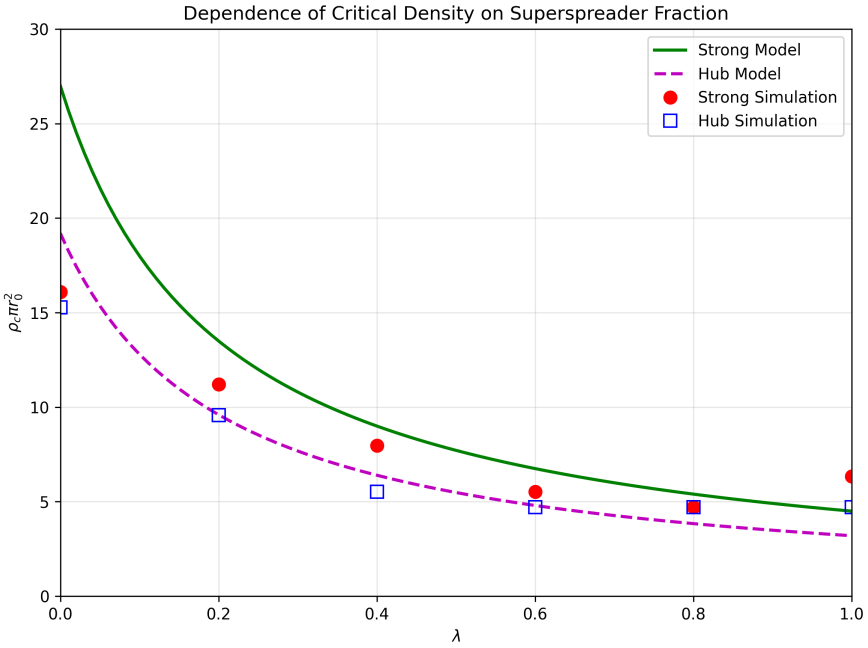

Figure 3: Critical density ($\rho_c\pi r_0^2$) vs. superspreader fraction (λ). Circles/solid line: strong infectiousness model; squares/dashed line: hub model. Percolation occurs above the curves

Propagation Speed

Propagation velocity is defined as the velocity of the front line of the infection. The time evolution of the distance between the initial infected individual and the furthest infected individual (r_f) shows that as λ increases, the distance covered at a given time step is larger, and the plateau is reached faster.

The propagation velocity is an increasing function of λ. For any $\lambda > 0$, the velocity in the hub model is greater than in the strong infectiousness model.

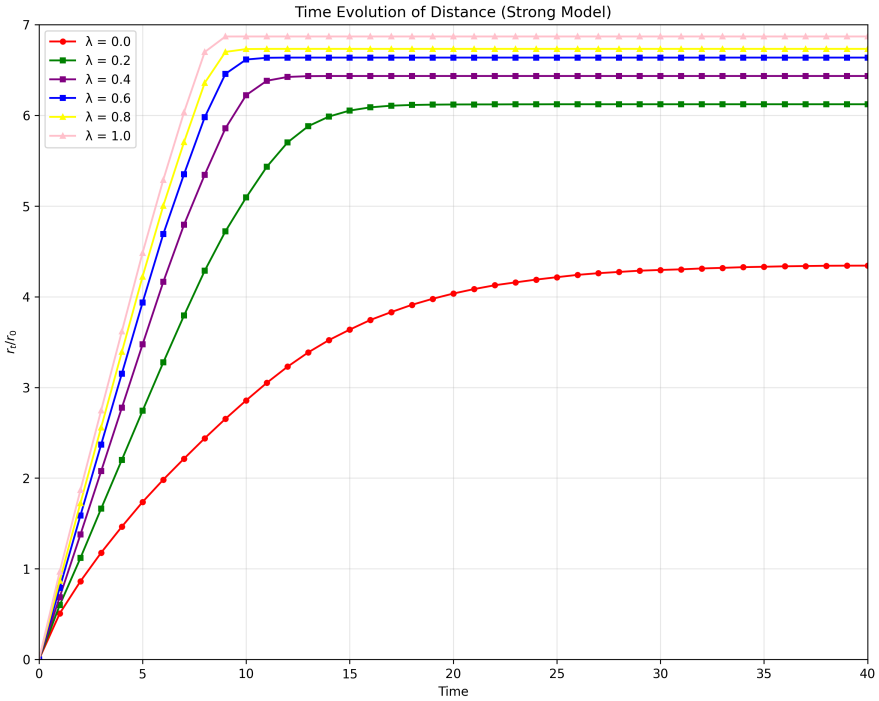

Figure 4: Time evolution of distance from initial infection to propagation front for different λ in the strong infectiousness model

Epidemic Curves

Epidemic curves illustrate the number of newly infected individuals over time. The results show that when superspreaders are present in the hub model, the number of newly infected individuals increases rapidly, and the peak appears earlier compared to the strong infectiousness model or cases without superspreaders.

Distribution of Secondary Infected

The network of infection routes reveals how the disease spreads. In the strong infectiousness model, some superspreaders infect a large number of individuals locally, while

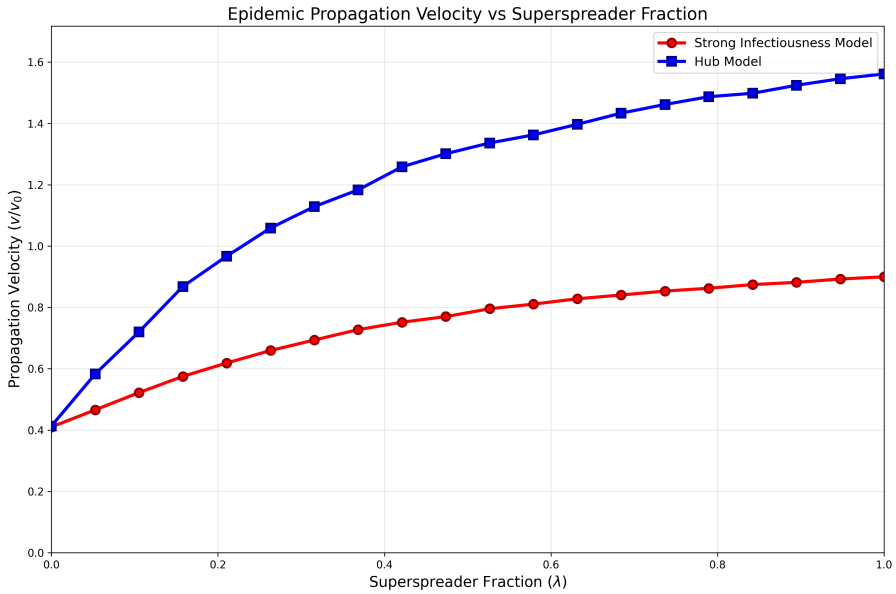

Figure 5: Propagation velocity vs. λ for both models. Circles: strong infectiousness; squares: hub model

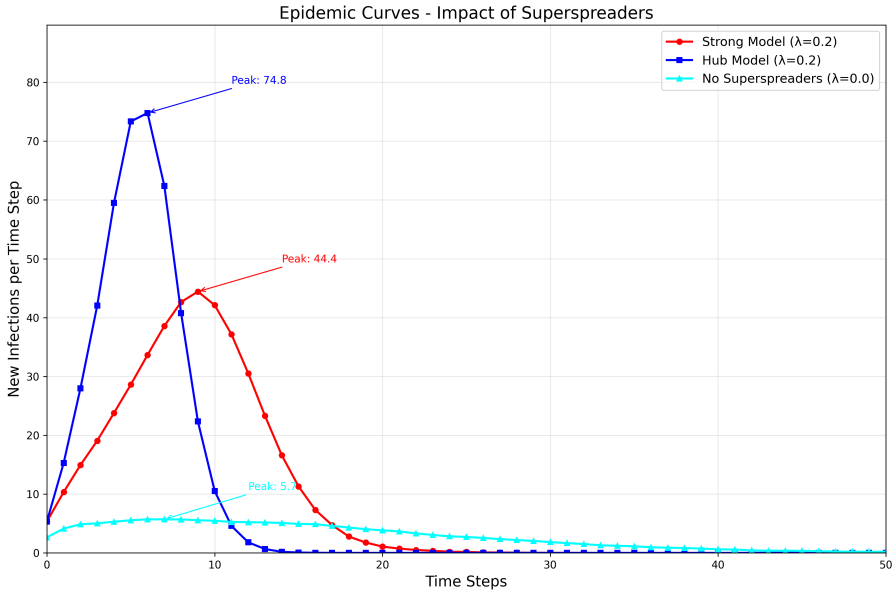

Figure 6: Epidemic curves for different models. Circles: strong infectiousness; squares: hub model ($\lambda = 0.2$); triangles: no superspreaders ($\lambda = 0.0$), with $\rho\pi r_0^2 = 15.7$.

most normal individuals infect very few (at most two or three individuals). In the hub model, superspreaders create long infection paths, which contribute to higher propagation velocities. When no superspreaders are present, the infection often stops before percolating due to low density. The distribution of links in these networks corresponds to the distribution of secondary infections.

When superspreaders are mixed into the system, both models yield similar distributions. These distributions are characterized by a much larger number of links at zero compared to other numbers of links, and they exhibit long tails. This contrasts with the distribution without superspreaders, where the distribution at zero is smaller and the tail is shorter.

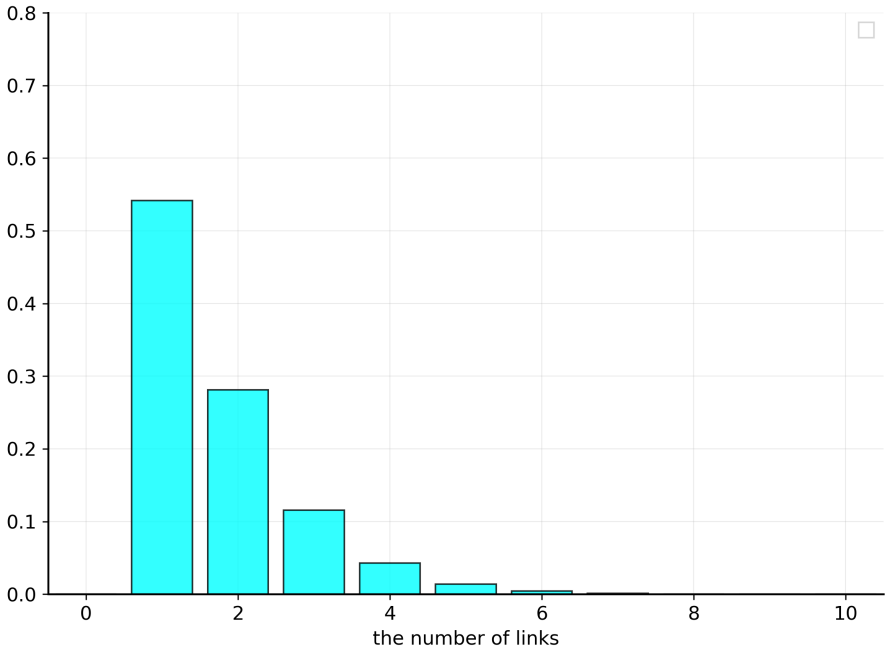

Figure 7: Link distribution after influence for $\lambda = 0.0$ (no superspreaders) and $\rho\pi r_0^2 = 15.7$.

Comparison with SARS Data

When comparing the epidemic curve of SARS in Singapore (February 13–June 13, 2003) with the simulation results, the hub model, with parameters $N = 500$ ($\rho\pi r_0^2 = 15.7$), $\lambda = 0.4$, and $\gamma = 1.0$, shows a strong resemblance to the observed SARS data. This suggests that SARS spread in a manner consistent with the hub model, indicating that the superspreader in the SARS outbreak was a person having many social connections.

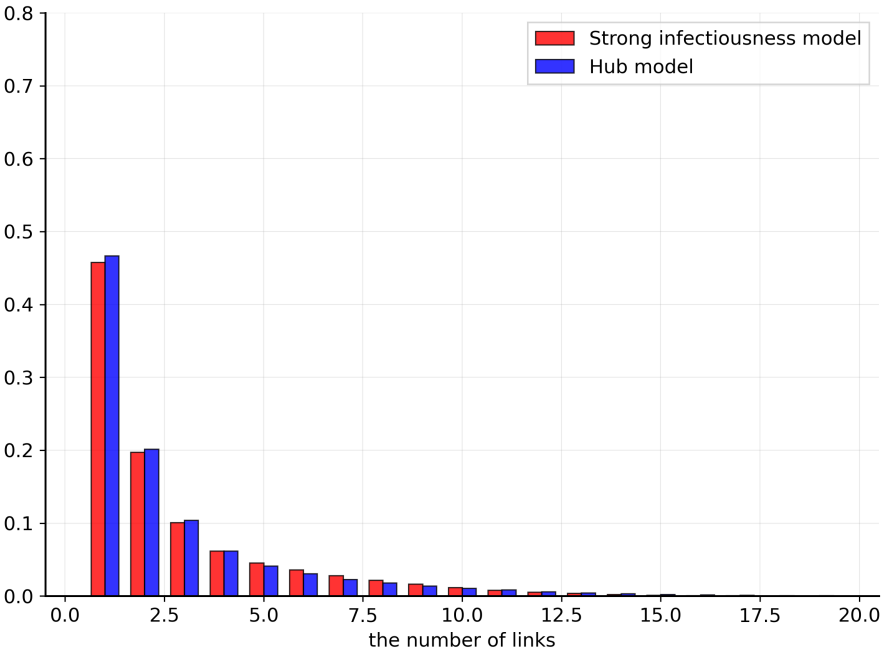

Figure 8: Link distribution after influence for superspreaders in both models, with $\lambda = 0.2$ and $\rho\pi r_0^2 = 15.7$.

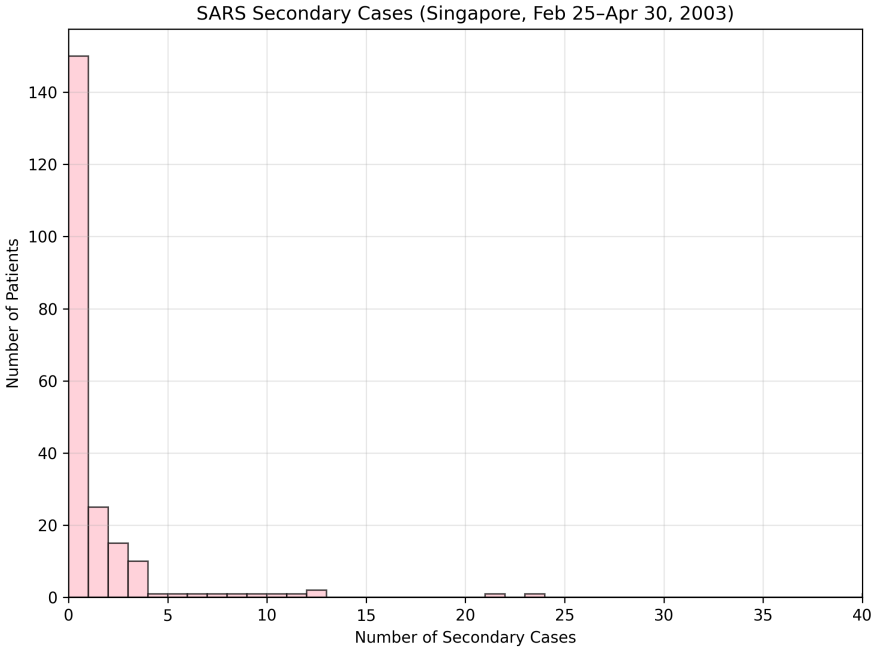

Figure 9: Number of direct secondary patients from SARS (Singapore, February 25-April 30, 2003).

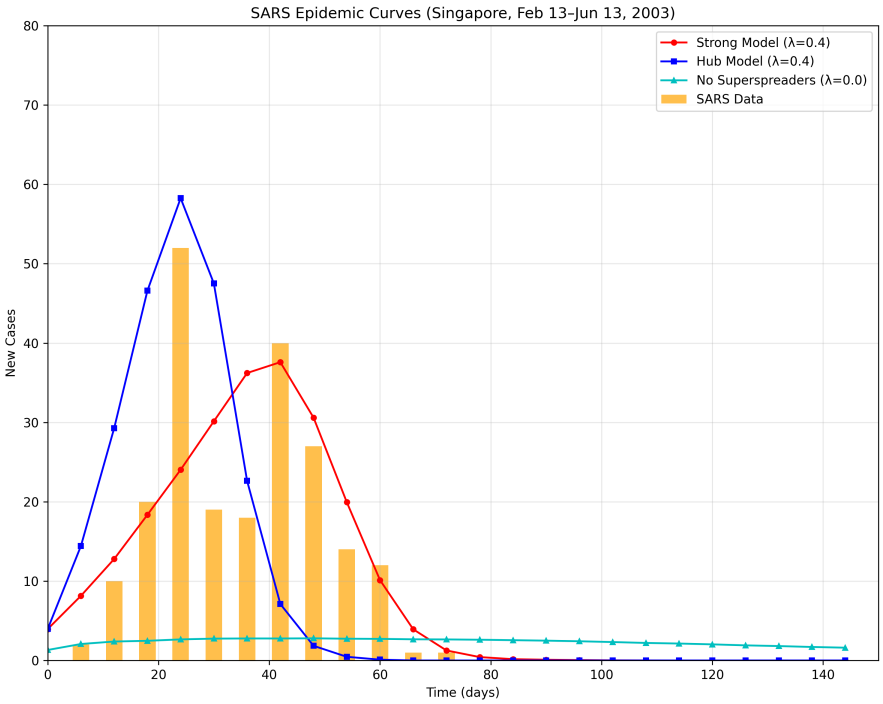

Figure 10: Epidemic curves of SARS (Singapore, Feb 13–Jun 13, 2003) compared with model results. Circles: strong infectiousness model; squares: hub model ($\lambda = 0.4$); triangles: no superspreaders ($\lambda = 0.0$). Initial population: $N = 500$ ($\rho\pi r_0^2 = 15.7$); 1 timestep = 6 days.

6 Discussion

6.1 Sentivity Analysis

The simulation results demonstrate a strong sensitivity to the fraction of superspreaders, λ. For both the strong infectiousness and hub models, increasing λ significantly lowers the critical density for percolation, making the disease spread more easily and widely even in less dense populations. This suggests that even a small proportion of superspreaders can drastically alter the epidemic's trajectory. Furthermore, the propagation velocity is an increasing function of λ, indicating that higher fractions of superspreaders lead to faster disease dissemination. The epidemic curves also highlight this sensitivity, showing that with the presence of superspreaders, particularly in the hub model, the peak of new infections occurs earlier and is often sharper. This rapid increase in cases could overwhelm healthcare systems more quickly.

The model also shows sensitivity to the underlying model of superspreader behavior (strong infectiousness vs. hub model). While both models exhibit percolation transitions and increased propagation with λ, the hub model consistently shows a lower critical density and higher propagation velocities for any $\lambda > 0$. This indicates that the "social connec-

tion" aspect of superspreading (as modeled by the hub model) is a more potent driver of epidemic spread than simply increased individual infectiousness.

6.2 Robustness of the model

The robustness of the models is supported by their qualitative consistency across various simulated parameters and their ability to reproduce key features of real-world epidemic data, specifically SARS. The observed percolation transitions, the dependence of critical density on λ, the increasing propagation velocity with λ, and the characteristic shapes of the epidemic curves are all consistent and logical responses to the introduced superspreader mechanisms.

The strong correlation between the simulated distribution of secondary cases and the actual SARS secondary case data provides a compelling argument for the models' ability to capture the underlying transmission dynamics of real-world outbreaks with superspreaders. Furthermore, the hub model's ability to closely reproduce the epidemic curve of SARS in Singapore is a significant validation, suggesting that the model captures the essential mechanism of spread during that specific outbreak – namely, the importance of social connections in superspreading events. While the models are simplified, their ability to replicate observed phenomena with distinct parameter-driven behaviors (e.g., strong infectiousness vs. hub) indicates a fundamental robustness in their conceptualization of superspreader effects. However, further validation with diverse epidemic data sets would strengthen the assessment of their generalizability.

7 Conclusion

This study demonstrates that superspreaders significantly influence epidemic dynamics. Both the strong infectiousness and hub models show that even a small fraction of superspreaders can greatly increase disease spread and speed. Among them, the hub model better replicates real-world outbreaks like SARS in Singapore, emphasizing the impact of social connectivity. These findings highlight the need to consider individual heterogeneity in modeling and suggest that targeting highly connected individuals could improve epidemic control strategies.

References

[1] R. Fujie and T. Odagaki. Effects of superspreaders in spread of epidemic. *Physica A: Statistical Mechanics and its Applications*, 374(2):843–852, 2007. ISSN 0378-4371. doi: https://doi.org/10.1016/j.physa.2006.08.050. URL https://www.sciencedirect.com/science/article/pii/S0378437106008703.

[2] M. Lipsitch, T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman, and M. Murray. Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300(5627):1966–1970, 2003. doi: 10.1126/science.1086616. URL https://www.science.org/doi/abs/10.1126/science.1086616.

[3] World Health Organization. Severe acute respiratory syndrome (sars). https://www.who.int/health-topics/severe-acute-respiratory-syndrome, 2003.