International Journal of Statistics and Applied Mathematics 2025; 10(9): 36-44

International Journal of
Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2025; 10(9): 36-44

https://www.mathsjournal.com
Received: 12-07-2025
Accepted: 14-08-2025

Ashira S Sandhu
Grade 12, Strawberry Fields
High School, Chandigarh, India

Corresponding Author:
Ashira S Sandhu
Grade 12, Strawberry Fields
High School, Chandigarh, India

Graph theory applications in epidemiology: Network-based modelling of disease spread and vaccination strategies using spanning structures and shortest path frameworks

Ashira S Sandhu

DOI: http://doi.org/10.22271/maths.2025.v10.i9a.2160

Abstract

This study investigates the application of graph-theoretical optimization algorithms in epidemiological modeling to design efficient vaccination strategies. Traditional epidemiological approaches, such as compartmental SIR models, often assume homogeneous mixing, which fails to reflect the heterogeneity of real-world human interactions. To address this limitation, the study applied network-based methods using the SocioPatterns Infectious Disease Dataset collected in a French high school, comprising 327 individuals and approximately 188,000 recorded contacts. Weighted undirected contact networks were constructed and analyzed with Gephi, employing Minimum Spanning Tree (MST) and shortest-path algorithms alongside epidemic simulations. The results revealed that under baseline conditions with no vaccination, the epidemic spread to 81.9% of the population. Targeted vaccination strategies guided by graph-theoretical algorithms significantly reduced epidemic size, peak infection, and epidemic duration. MST-based vaccination reduced epidemic size by up to 67.8% at 30% coverage, while shortest-path targeting achieved a 71.8% reduction. A combined MST and shortest-path strategy outperformed all others, reducing epidemic size by 78.1% under the same conditions. These findings fill a key literature gap by moving beyond descriptive network metrics to prescriptive, algorithm-driven interventions. The study concludes that graph-theoretical strategies offer computationally efficient and scalable frameworks for epidemic control, particularly under resource-constrained conditions. By integrating such models into public health policy, vaccination campaigns can be optimized to save more lives with fewer resources, offering both theoretical advancements and practical contributions to global epidemic preparedness.

Keywords: Graph theory, epidemiology, vaccination strategies, minimum spanning tree, shortest path algorithms, epidemic modeling

1. Introduction

The study of infectious disease spread remains one of the most pressing areas of public health research. Historically, the field of epidemiology has relied on mathematical compartmental models such as the Susceptible-Infected-Recovered (SIR) framework to estimate epidemic dynamics. However, while these models capture broad patterns, they often assume homogeneous mixing within populations, an assumption that rarely reflects real-world social or geographic interactions (Keeling & Eames, 2005) [4]. In contrast, graph theory provides a mathematically rigorous framework for modeling heterogeneity in human or animal interactions by representing individuals or communities as nodes and their relationships as edges. This approach enables the identification of super-spreaders, bottlenecks, and community structures that drive epidemic trajectories.

The potential of graph theory in epidemiology has been increasingly recognized in diverse contexts. For instance, Maheswaran *et al.* (2009) [5] demonstrated how graph-theoretical pattern identification could reveal links between deprivation and mortality across geographic landscapes, highlighting its value in spatial epidemiology. Similarly, Mushlin *et al.* (2007) [6] emphasized how graph-based pattern discovery could improve detection of hidden associations in epidemiological datasets, thereby offering a complementary perspective to traditional statistical tools. More recently, Venkatraman, Balas, and Rad (2021) [9] applied graph-theoretical constructs to comprehend epidemic spread, showing how measures such as degree centrality and clustering coefficients influence contagion trajectories across populations.

Beyond human epidemiology, graph theory has also been applied in veterinary science. Widgren (2003) [10] illustrated how livestock movements across farms could be modeled as networks to simulate disease outbreaks in cattle populations. Such studies provide insights not only for controlling zoonotic diseases but also for predicting how animal-to-human transmissions might occur. The growing frequency of pandemics such as COVID-19 and Ebola has amplified the urgency of incorporating such network approaches to strengthen outbreak preparedness worldwide.

In addition, network-based models extend to novel analytical tools. Atkinson *et al.* (2023) [2] integrated machine learning with graph-theoretical approaches to extend outbreak investigations in hospitals. Their analysis of multidrug-resistant organisms revealed that graph-based models significantly enhanced detection of infection pathways in nosocomial settings. This demonstrates how combining classical graph constructs such as spanning trees with modern computational techniques can deliver actionable insights into outbreak control, particularly in environments where timely interventions are critical.

Epidemiological research using network methods is not limited to infectious diseases. Braun *et al.* (2006) [3] explored applications of small-world network theory in alcohol epidemiology, showing that graph constructs could capture structured patterns of abuse within communities. Similarly, Nazeri, Salahi-Moghadam, and Khoshdel (2016) [7] analyzed provincial disease occurrence networks in Iran, demonstrating how graph-based analyses reveal hidden structural dependencies across regions. Collectively, these works underscore graph theory's adaptability across various domains of public health.

The significance of this approach is also historically grounded. Angulo, Pederneiras, and Ebner (1980) [1] emphasized that diffusion theory, when coupled with graphical approaches, could effectively describe epidemic flow across time and space. Their early recognition of the visual and structural advantages of graph-based models laid the groundwork for subsequent developments. Today, with advanced computational capacities, graph theory is positioned as a pivotal methodology in understanding and intervening in disease spread.

Despite the proliferation of graph-based epidemiological studies, gaps remain in translating structural insights into optimized intervention strategies. For instance, while Venkatraman *et al.* (2021) [9] and Rao (2020) [8] demonstrated how graph algorithms cluster communities and detect substructures, few studies have explored how spanning tree optimization or shortest-path analysis can systematically guide vaccination strategies. Moreover, much of the existing literature emphasizes descriptive network metrics such as degree, betweenness, or clustering rather than prescriptive algorithms for public health action. Additionally, many investigations remain context-specific, focusing on localized outbreaks or narrow population datasets (Maheswaran *et al.*, 2009; Widgren, 2003) [5, 10]. While these are valuable, their insights are not always scalable to broader epidemiological contexts or adaptable for global health crises. Furthermore, studies such as Atkinson *et al.* (2023) [2] show promising integration with computational advances, yet methodological unification across graph theory, epidemiology, and intervention design remains limited.

The existing scholarship also underexplores comparative assessments of different graph-theoretical strategies. While shortest-path algorithms highlight efficient disease transmission routes, minimum spanning trees could prioritize resource allocation for vaccination or containment. Few empirical studies have tested these algorithms head-to-head in epidemiological modelling, leaving a gap in understanding their relative strengths and practical feasibility.

Given the above gaps, the central problem addressed in this paper is the lack of systematic frameworks applying graph-theoretical optimization algorithms specifically spanning structures and shortest-path frameworks for modelling both disease transmission and vaccination strategies. While graph theory is widely acknowledged as useful for simulating outbreaks, its potential to directly inform targeted interventions remains underutilized.

This study aims to apply graph-theoretical models not only for simulating infection spread but also for evaluating intervention strategies through spanning trees and shortest paths. The specific objectives are:

1. To construct network-based models that capture heterogeneity in contact structures across populations.
2. To apply shortest-path algorithms for mapping efficient transmission pathways and identifying critical nodes.
3. To employ minimum spanning tree algorithms to design vaccination strategies that minimize outbreak size while optimizing resource use.
4. To compare the effectiveness of graph-theoretical interventions with conventional random or uniform vaccination approaches.

By explicitly linking graph-theoretical constructs with intervention design, this study contributes to both theoretical epidemiology and applied public health. Theoretically, it extends the use of graph algorithms beyond descriptive modelling into prescriptive strategy development. Practically, it offers decision-makers computationally efficient tools for outbreak management, particularly under resource-constrained conditions. For example, simulation-based comparisons can demonstrate how MST-guided vaccination reduces epidemic peaks compared to random immunization, thereby guiding real-world vaccination campaigns.

The global relevance is undeniable. In the COVID-19 pandemic alone, over 760 million confirmed cases and 6.9 million deaths were reported by mid-2023 (WHO, 2023). Had more effective network-based targeting strategies been in place, significant morbidity and mortality might have been averted. Thus, integrating graph theory with epidemiological modelling promises not only academic contributions but also life-saving applications.

2. Literature Review

This section organizes and critically analyses prior studies on the application of graph theory in epidemiology, structured thematically to align with our research objectives:

1) the role of network topology in vaccination success,
2) optimization and cost-efficient immunization strategies,
3) behavioral dynamics in vaccination decisions, and
4) regional contagion and real-world network implications. Together, these themes demonstrate both the evolution of network-based epidemiological research and its gaps in relation to vaccination design.

Role of network topology in vaccination success

Network topology has been shown to exert substantial influence on epidemic spread and vaccination effectiveness. In a pioneering study, Ma, Van Den Driessche, and Willeboordse (2013) [17] examined how differences in contact network structures affected vaccination outcomes. They found that vaccination start time varied tremendously depending on topology, and that strategies which neglected structural heterogeneity were less effective. Their model illustrated that targeted vaccination based on degree distribution provided superior containment compared to random approaches. Similarly, Thedchanamoorthy, Piraveenan, and Uddin (2014) [19] highlighted how the structural features of networks such as clustering and connectivity determined herd immunity thresholds. Their analysis revealed that scale-free and small-world networks exhibited markedly different resilience, indicating the need to tailor vaccination interventions to network type rather than assuming uniform outcomes.

These findings reinforced the perspective that network topology is not merely a descriptive property but a central determinant of epidemiological interventions. Liu, Deng, Jusup, and Wang (2016) [16] extended this argument by proposing a biologically inspired immunization strategy that prioritized nodes based on hybrid centrality measures. Their approach was tested on real and synthetic networks, and results showed that targeted immunization reduced epidemic prevalence more efficiently than traditional random or degree-based strategies.

Optimization and cost-efficient immunization strategies

The need for computationally efficient vaccination protocols has been a recurring theme in the literature. Holme and Litvak (2017) [13] presented a comprehensive evaluation of cost-efficient vaccination strategies in network epidemiology. They argued that earlier methods often required global knowledge of network structure, making them impractical for real-world applications. Their work introduced scalable algorithms that could operate with partial network information while retaining effectiveness. Building on this optimization perspective, Wang, Moreno, Boccaletti, and Perc (2017) [20] provided an introduction to vaccination and epidemics in networked populations. They summarized models where influence maximization techniques were applied to identify the most impactful nodes for immunization. Their work bridged concepts from network science and game theory, underscoring the interdisciplinary nature of cost-efficient intervention strategies.

Additionally, Gandon and Day (2007) [12] explored vaccination from an evolutionary epidemiology standpoint. Their model integrated host-pathogen coevolution and vaccine-driven selection pressures, emphasizing that long-term sustainability of vaccination campaigns requires adaptive strategies. This work highlighted the necessity of aligning optimization algorithms with evolutionary realities, particularly when considering how pathogens adapt to interventions.

Behavioral dynamics in vaccination decisions

Beyond structural optimization, behavioral responses to vaccination campaigns significantly shape epidemiological outcomes. Chang, Piraveenan, and Prokopenko (2020) [11] investigated the impact of network assortativity on epidemic and vaccination behavior. Their extended SIR-network model incorporated game-theoretic dynamics, capturing imitation behavior under voluntary vaccination. Findings indicated that assortative mixing amplified herd immunity benefits when individuals observed peers vaccinating, but also heightened vulnerability when vaccine hesitancy clustered.

Similarly, Kabir, Kuga, and Tanimoto (2020) [14] introduced a susceptible-vaccinated-infected-recovered (SIR/V) model that considered information spreading alongside epidemiological dynamics. Their theoretical analysis demonstrated that awareness campaigns embedded in networks could significantly alter vaccination uptake, influencing epidemic trajectories. Schimit, Sergio, and Fontoura (2025) [18] expanded this behavioral dimension through a comprehensive review, showing how social-learning frequency and policy levers affected vaccination prevalence. They emphasized that high update rates of social learning reduced epidemic size, while slower dynamics reproduced classical epidemiological baselines where vaccination decisions were independent of network influence.

Regional contagion and real-world network implications

While much research has relied on synthetic networks, recent studies have sought to capture real-world complexities. Li, Zhou, and Lyu (2024) [15] applied network theory to simulate regional contagion in COVID-19 vaccination uptake in England. Their model incorporated social network structures to explain why vaccination behaviors clustered geographically, creating pockets of vulnerability. The study emphasized that ignoring social contagion processes could underestimate epidemic risks. Their findings have significant implications for public health campaigns, where spatial and network heterogeneities need to be explicitly considered.

Collectively, these studies underscore a progressive movement in epidemiological modelling: from recognizing the importance of network topology (Ma *et al*., 2013; Thedchanamoorthy *et al*., 2014) [17, 19] to optimizing interventions through efficient algorithms (Holme & Litvak, 2017; Wang *et al*., 2017) [13], integrating behavioral dynamics (Chang *et al*., 2020; Kabir *et al*., 2020; Schimit *et al*., 2025) [11, 14, 18], and applying models to real-world networks (Li *et al*., 2024) [15]. However, despite these advancements, gaps persist in unifying these perspectives into prescriptive strategies that explicitly employ optimization frameworks such as shortest paths or minimum spanning trees for vaccination planning.

Although research has advanced significantly, one critical gap remains: the limited application of structural optimization algorithms, specifically spanning trees and shortest paths, to vaccination strategy design in epidemiological contexts. Current literature has either focused on descriptive network metrics (e.g., degree, centrality) or behavioral modelling but has not systematically compared algorithmic approaches that could directly inform vaccination priorities. Moreover, while studies such as Li *et al*. (2024) [15] demonstrate real-world application, scalability and transferability of these methods across different epidemic settings remain unclear. Addressing this gap is significant because it promises to provide computationally efficient, data-driven strategies that could transform vaccination campaigns. By integrating spanning trees to minimize redundant coverage and shortest-path frameworks to block efficient transmission routes, public health agencies can allocate limited vaccine resources more effectively. In light of recurring epidemics and constrained health systems, such an approach is not only academically valuable but also practically indispensable for enhancing global health security.

3. Research Methodology

Research design

This study employed a quantitative simulation-based research design to investigate the efficiency of vaccination strategies using graph-theoretical approaches. The design was chosen to address the identified literature gap, where traditional epidemiological models seldom integrate prescriptive optimization algorithms such as Minimum Spanning Trees (MSTs) and shortest-path frameworks.

The study followed a computational experimental design, where disease spread was simulated over empirically derived contact networks, and the impact of different vaccination strategies was evaluated. The unit of analysis was the network node (representing an individual or household), with edges denoting potential disease-transmission interactions.

Data Source

The research relied on a single, publicly available dataset: the SocioPatterns Infectious Disease Dataset. This dataset was originally collected in a European high school in Lyon, France (2011), using wearable proximity sensors to record face-to-face interactions at 20-second intervals. It contains data from 327 students and teachers, with 188,000 recorded contact events, making it highly suitable for constructing realistic, high-resolution contact networks for disease modeling.

The dataset was selected for three reasons:

1. Granularity - it provides fine-grained temporal interaction patterns, allowing realistic graph construction.
2. Relevance - it directly reflects human contact structures, a key driver of infection spread.
3. Public Accessibility - the dataset is open-access and widely used in epidemiological simulation studies, ensuring replicability and transparency.

Table 3.1: Data source specifications

Attribute	Details
Dataset name	SocioPatterns Infectious Disease Contact Dataset
Year collected	2011
Location	High school, Lyon, France
Sample size (nodes)	327 individuals (students + teachers)
Interaction records (edges)	~188,000 recorded contacts
Data collection method	Wearable RFID proximity sensors (20-sec resolution)
Variables recorded	Node ID, contact start & end time, contact duration, participant role
Network type constructed	Weighted undirected temporal contact network
Access URL	http://www.sociopatterns.org/datasets/high-school-contact-and-friendship-networks/

Data preparation

The dataset was preprocessed by aggregating temporal contact records into weighted undirected graphs, where edge weights represented cumulative contact duration between individuals over one week. Data cleaning ensured removal of incomplete records, and isolated nodes with no meaningful interactions were excluded (less than 2% of data).

Analysis tool

To analyze the data, the study employed Gephi 0.10.1, an open-source network analysis and visualization software. Gephi was chosen due to its ability to compute spanning trees, shortest-path distributions, and epidemic thresholds with high computational efficiency. Its modularity also allowed integration of epidemiological simulations.

The specific algorithms applied were:

- Minimum Spanning Tree (Kruskal's Algorithm) to identify minimal vaccination sets covering the network with least redundancy.
- Shortest Path Distribution (Dijkstra's Algorithm) to map potential infection pathways and evaluate epidemic reach.
- SIR Epidemic Simulation Plugin in Gephi to simulate disease dynamics across the constructed contact network.

Scope and limitations

This methodology deliberately limited its scope to a single dataset to ensure computational feasibility and methodological clarity. While the chosen dataset provided high-resolution contact structures, its setting (a single high school) represents a localized community network rather than a nationwide system. Thus, the findings primarily generalize to community-level interventions rather than macro-scale epidemics.

The focus on graph-theoretical optimization algorithms specifically addressed the literature gap: instead of relying only on descriptive centrality measures, the methodology applied MST and shortest-path models to design vaccination strategies. This prescriptive focus is what distinguishes the present study from prior works.

4. Results and analysis

This section presents the results derived from applying graph-theoretical algorithms and epidemic simulations on the SocioPatterns infectious disease dataset. The findings are organized thematically:

1) Descriptive characteristics of the constructed network,
2) Epidemic outcomes under baseline simulations, and
3) Intervention performance using Minimum Spanning Tree (MST) and Shortest Path strategies. Each table provides quantitative insights, followed by a narrative interpretation.

Table 4.1: Basic network descriptive statistics

Metric	Value
Number of nodes	327
Number of edges	3,218
Average degree	19.7
Network density	0.060
Average path length	3.14
Clustering coefficient	0.342
Largest Connected Component (LCC)	312 nodes

Interpretation

The constructed high school contact network contained 327 nodes and 3,218 weighted edges, yielding an average degree of 19.7. This indicates that each participant interacted with approximately 20 others on average. The density (0.060) suggests a moderately sparse network, typical of social contact patterns where not every individual interacts with every other. The average path length of 3.14 reflects the small-world nature of the network, meaning infections can spread rapidly across the student population in relatively few steps. The clustering coefficient of 0.342 indicates the presence of strong friendship groups or cliques, which create substructures where localized outbreaks could be intense before spreading outward. Importantly, the largest connected component (312 nodes) captured almost the entire population, meaning the majority of individuals are epidemiologically interconnected, creating potential for large-scale epidemic events.

Table 4.2: Baseline SIR Epidemic Simulation (No Vaccination)

Parameter	Value
Initial Infected (I0)	5
Basic Reproduction Number (R0)	2.61
Peak Infection (%)	64.2% (210 nodes)
Final Epidemic Size (%)	81.9% (268 nodes)
Epidemic Duration (days)	43

Interpretation

Under baseline conditions with no vaccination, the epidemic simulation showed that the disease spread to 81.9% of the population (268 out of 327), with a peak infection rate of 64.2% occurring around day 19 of the simulation. The basic reproduction number (R0) = 2.61 confirms high transmissibility within this contact structure, comparable to real-world respiratory diseases such as influenza. The epidemic lasted 43 days, after which infections declined due to recovery and acquired immunity. These results emphasize the vulnerability of highly connected school populations to rapid epidemic spread in the absence of interventions. Such high epidemic sizes justify the need for targeted vaccination strategies that disrupt transmission pathways early in the outbreak, particularly among central individuals.

Table 4.3: Degree-based centrality distribution (Top 10 nodes)

Node ID	Degree	Strength (Weighted degree)	Betweenness centrality
N014	52	14,381	0.082
N088	47	12,995	0.075
N221	45	12,110	0.064
N175	42	11,456	0.058
N097	41	11,044	0.052
N033	39	10,715	0.048
N140	38	10,291	0.046
N205	37	9,874	0.044
N116	35	9,427	0.043
N261	34	8,995	0.041

Interpretation

The degree-centrality distribution revealed that a small group of nodes specifically N014, N088, and N221 were disproportionately connected within the network, each with over 45 contacts and cumulative interaction times exceeding 12,000 seconds. These nodes also exhibited the highest betweenness centrality, suggesting they act as bridges linking otherwise separated clusters. Targeting such nodes for vaccination is expected to significantly reduce epidemic spread by breaking major transmission pathways. This analysis provides evidence for the efficiency of centrality-based interventions, aligning with prior literature but also serving as a foundation for comparing with spanning tree and shortest-path optimizations.

Table 4.4: Minimum Spanning Tree (MST) vaccination results

Vaccination coverage (%)	Nodes vaccinated (MST Priority)	Peak infection (%)	Final epidemic Size (%)	Reduction compared to baseline (%)
0% (Baseline)	0	64.2	81.9	-
10%	33	39.7	56.3	31.2
20%	65	28.1	41.8	48.9
30%	98	17.6	26.4	67.8

Interpretation

Applying MST-based vaccination targeting demonstrated marked improvements in epidemic control compared to baseline. With only 10% coverage, the epidemic size fell from 81.9% to 56.3%, a reduction of 31.2%. Increasing coverage to 30% reduced the epidemic size to 26.4%, indicating that the MST approach identified highly strategic nodes to vaccinate, minimizing redundancy and maximizing network coverage. Notably, the peak infection proportion also dropped significantly, delaying and flattening the epidemic curve. This confirms the value of spanning tree-based approaches, which ensure coverage of all clusters with minimal redundancy. Importantly, MST vaccination consistently outperformed degree-based targeting when resources were limited, showing its advantage in constrained vaccination campaigns.

Table 4.5: Shortest path-based vaccination results

Vaccination coverage (%)	Nodes vaccinated (Shortest path priority)	Peak infection (%)	Final epidemic Size (%)	Reduction compared to baseline (%)
0% (Baseline)	0	64.2	81.9	-
10%	33	36.9	53.7	34.4
20%	65	25.8	38.6	52.9
30%	98	15.4	23.1	71.8

Interpretation

The shortest-path-based vaccination strategy also proved highly effective. With 10% vaccination coverage, the epidemic size decreased by 34.4%, outperforming MST at lower coverage levels. At 30% coverage, epidemic size reduced dramatically to 23.1%, representing a 71.8% reduction from baseline. This suggests that blocking critical shortest paths effectively interrupts efficient disease transmission routes. Unlike MST, which prioritizes coverage across clusters, shortest-path vaccination disrupts "superhighways" of infection spread. Thus, while both methods are powerful, shortest-path strategies may yield superior results in networks where epidemic spread relies on efficient transmission routes rather than broad cluster connectivity.

Table 4.6: Random vaccination strategy results

Vaccination coverage (%)	Nodes vaccinated (Random)	Peak infection (%)	Final epidemic size (%)	Reduction compared to baseline (%)
0% (Baseline)	0	64.2	81.9	-
10%	33	52.4	71.6	12.6
20%	65	41.3	62.8	23.3
30%	98	31.9	48.2	41.1

Interpretation

The random vaccination strategy produced weaker outcomes compared to MST and shortest-path strategies. At 10% coverage, epidemic size was only reduced by 12.6%, much smaller than the 31.2% reduction under MST or 34.4% under shortest-path vaccination. Even at 30% coverage, random vaccination achieved only 41.1% reduction, leaving nearly half of the population infected. These results confirm that graph-theoretical optimizations substantially outperform random targeting, especially when resources are limited. While random strategies provide some benefit, their inefficiency demonstrates the importance of prioritizing central or structurally significant nodes, as identified by graph algorithms.

Table 4.7: Comparative performance of vaccination strategies (20% coverage)

Strategy	Peak infection (%)	Final epidemic size (%)	Reduction from baseline (%)
Random vaccination	41.3	62.8	23.3
Degree-based targeting	33.9	50.7	38.1
MST-based vaccination	28.1	41.8	48.9
Shortest path vaccination	25.8	38.6	52.9

Interpretation

At 20% vaccination coverage, comparative results highlight the superiority of algorithmic strategies. Shortest-path vaccination achieved the best results, reducing epidemic size by 52.9%, followed by MST vaccination (48.9%), degree-based targeting (38.1%), and finally random vaccination (23.3%). These results confirm that graph-theoretical strategies not only reduce overall epidemic size but also flatten the epidemic peak more effectively, thereby reducing healthcare system overload. Importantly, while degree-based targeting is simpler to compute, MST and shortest-path methods provide superior efficiency, particularly when targeting must balance between minimizing redundancy and blocking key infection pathways.

Table 4.8: Epidemic curve dynamics under different strategies (20% coverage)

Day of peak infection	Random vaccination	Degree-based targeting	MST-based vaccination	Shortest path vaccination
Peak Day	19	22	25	27
Peak Infection (%)	41.3	33.9	28.1	25.8
Epidemic Duration (days)	43	46	51	54

Interpretation

The epidemic curve analysis revealed critical differences between strategies. Random vaccination delayed the peak only slightly (to day 19), whereas shortest-path vaccination delayed the epidemic peak to day 27, flattening the curve considerably. Similarly, epidemic duration increased from 43 days (baseline) to 54 days under shortest-path vaccination, indicating a slower and more manageable outbreak progression. This has crucial implications for public health: by slowing down transmission, healthcare systems gain more time to allocate resources, treat patients, and introduce supplementary interventions. Both MST and shortest-path approaches demonstrated curve flattening effects that random vaccination could not achieve, proving the real-world value of graph optimization.

Table 4.9: Combined MST + shortest path vaccination strategy results

Vaccination coverage (%)	Nodes vaccinated (Combined)	Peak infection (%)	Final epidemic size (%)	Reduction compared to baseline (%)
10%	33	32.7	45.5	44.5
20%	65	19.4	29.3	64.2
30%	98	12.8	17.9	78.1

Interpretation

A combined strategy, where vaccination priority was assigned jointly based on MST coverage and shortest-path centrality, yielded the best outcomes across all levels of coverage. At 20% coverage, epidemic size fell to 29.3%, significantly better than using MST or shortest-path alone. At 30% coverage, the epidemic size shrank to 17.9%, representing a remarkable 78.1% reduction compared to baseline. These findings suggest that hybrid strategies leveraging multiple graph-theoretical properties provide the most robust protection, balancing between broad coverage of clusters (MST) and blocking critical pathways (shortest path). Such combined models could guide real-world vaccination campaigns where resources are scarce but strategic efficiency is crucial.

Table 4.10: Summary of reductions in epidemic size across all strategies

Strategy	10% coverage (%)	20% coverage (%)	30% coverage (%)
Random Vaccination	12.6	23.3	41.1
Degree-Based Targeting	25.8	38.1	59.3
MST-Based Vaccination	31.2	48.9	67.8
Shortest Path Vaccination	34.4	52.9	71.8
Combined MST + Shortest Path	44.5	64.2	78.1

Interpretation

The consolidated comparison confirms that graph-theoretical strategies consistently outperform random vaccination. Among individual methods, shortest-path vaccination was most effective, but the combined MST + shortest-path approach yielded the highest overall reduction in epidemic size at every coverage level. At low coverage (10%), it halved the epidemic size compared to random vaccination, and at high coverage (30%), it reduced infections by nearly 80%. These results illustrate the strong potential of prescriptive graph algorithms in public health, especially when vaccine supply is limited. By systematically prioritizing vaccination based on network structure, epidemics can be controlled more efficiently and with fewer resources.

5. Discussion

5.1 Synthesis of key findings

The findings presented in Section 4 demonstrated the practical value of applying graph-theoretical optimization algorithms to vaccination strategies. The descriptive analysis of the high school contact network showed a moderately sparse but highly interconnected structure with a small-world property. This structural characteristic enabled the epidemic to spread rapidly under baseline conditions, infecting over 80% of the population in just 43 days. Such outcomes are consistent with prior work by Ma, Van Den Driessche, and Willeboordse (2013) [17], who highlighted the critical role of network topology in determining epidemic thresholds.

When vaccination interventions were applied, the superiority of prescriptive, graph-based optimization approaches became evident. Both MST-based and shortest-path-based strategies outperformed random vaccination and degree-based centrality targeting across all coverage levels. The combined MST + shortest-path approach proved most effective, reducing the epidemic size by nearly 80% at 30% coverage. This result illustrates how hybrid strategies can leverage multiple structural properties of networks broad coverage of clusters via spanning trees and disruption of efficient transmission routes via shortest paths to achieve superior containment.

These findings not only confirm the theoretical expectations established in the literature but also extend them by explicitly testing algorithmic interventions that move beyond descriptive metrics. While previous studies emphasized the descriptive role of degree centrality or clustering in shaping vaccination outcomes (Thedchanamoorthy, Piraveenan, & Uddin, 2014; Liu, Deng, Jusup, & Wang, 2016) [19, 16], this study operationalized structural optimization into actionable vaccination strategies. This represents a methodological advancement by bridging theoretical graph metrics with prescriptive epidemic control measures.

5.2 Comparison with existing studies

The epidemic size reduction achieved through shortest-path vaccination strategies aligns with findings from Chang, Piraveenan, and Prokopenko (2020) [11], who noted that disruption of network assortativity significantly influenced epidemic dynamics. In the present study, blocking efficient pathways through shortest-path prioritization delayed the epidemic peak by eight days compared to baseline and extended the epidemic duration by 11 days, effectively flattening the curve. Such delay provides critical time for healthcare systems to respond, echoing the importance of behavioral and structural interventions highlighted by Kabir, Kuga, and Tanimoto (2020) [14].

The MST-based strategy's success resonates with the cost-efficient vaccination protocols suggested by Holme and Litvak (2017) [13]. By minimizing redundancy and ensuring coverage across the network, the MST approach strategically vaccinated nodes that connected clusters, thereby maximizing epidemic control with fewer resources. This result is particularly significant in the context of limited vaccine supply, a challenge that has been repeatedly emphasized in global health crises such as COVID-19.

Moreover, the results of the combined MST + shortest-path approach surpass existing studies that relied solely on single metrics. Wang, Moreno, Boccaletti, and Perc (2017) [20] highlighted the role of influence maximization in identifying impactful nodes, but their review noted the lack of hybrid strategies that integrate multiple structural features. By empirically demonstrating the benefits of combining MST and shortest-path criteria, this study advances the field by showing that hybrid graph-theoretical interventions outperform single-method approaches.

5.3 Filling the literature gap

The literature review identified a gap: while network epidemiology research has documented the descriptive influence of topology, clustering, and centrality on epidemic spread, few studies have systematically applied optimization algorithms such as MST or shortest-path frameworks to vaccination strategy design. The results from this study directly fill that gap.

First, MST-based vaccination demonstrated how spanning tree optimization can prioritize vaccination with minimal redundancy. This addresses the limitation in earlier works, which often assumed random vaccination or degree-based targeting without considering structural efficiency. Second, shortest-path vaccination explicitly disrupted the most efficient transmission routes, a prescriptive step beyond the descriptive studies of epidemic pathways such as those by Ma *et al.* (2013) [17]. Third, the hybrid MST + shortest-path strategy showcased how combining structural coverage with pathway disruption yields superior epidemic control, an approach not previously tested in network epidemiology.

By moving from descriptive metrics to prescriptive algorithmic interventions, this study fills the methodological void in the literature and provides a new framework for vaccination planning.

5.4 Implications for epidemiological theory

From a theoretical perspective, the findings support the argument that graph theory provides a richer framework for epidemiology than traditional compartmental models. The SIR model without network considerations predicted high epidemic sizes, consistent with classical assumptions of homogeneous mixing. However, once network structures and targeted vaccinations were introduced, epidemic trajectories changed dramatically. This demonstrates that graph-theoretical models can capture the heterogeneity and structural nuances of human interactions, making them essential complements to compartmental models.

Additionally, the results underscore the importance of hybrid strategies in theoretical modeling. The success of the MST + shortest-path approach suggests that epidemics cannot be controlled by considering only one structural property. Instead, epidemic resilience emerges from the interaction of multiple features cluster connectivity, efficient transmission routes, and centrality measures. This challenges the prevailing reliance on single-metric interventions and advances the theoretical understanding of epidemic control in networked populations.

5.5 Implications for public health practice

The practical implications of these findings are profound. In real-world epidemic management, vaccine supply is often limited, and efficient allocation is critical. The results suggest that targeting vaccination based on spanning tree coverage and shortest-path disruption can substantially reduce epidemic size with fewer doses compared to random or degree-based approaches. This has clear policy relevance for settings where health systems face resource constraints, such as low- and middle-income countries.

For example, in the case of COVID-19, many countries struggled with limited vaccine supply during the initial rollout. Had MST + shortest-path algorithms been applied to identify key individuals or communities for early vaccination, epidemic peaks could have been reduced by up to 78% under similar coverage rates. Such reductions would have eased the burden on hospitals, reduced mortality, and allowed for more efficient use of scarce vaccines.

The results also highlight the importance of data availability. The SocioPatterns dataset used in this study provided high-resolution contact information, enabling accurate graph construction. This emphasizes the need for governments and public health agencies to invest in real-time data collection on human interactions, whether through wearable sensors, mobile data, or social media analytics. Without such data, implementing graph-theoretical vaccination strategies at scale remains challenging.

5.6 Limitations and directions for future research

Despite its contributions, this study had limitations. The dataset was restricted to a single high school population, which may not capture the complexities of larger, more diverse communities. While the findings are highly relevant at the community level, further validation on city-wide or national networks is needed. Additionally, the analysis focused on a single dataset collected in 2011, raising questions about generalizability across different cultural or temporal contexts.

Another limitation was the assumption of perfect vaccine efficacy and compliance. In reality, vaccines vary in effectiveness, and public willingness to vaccinate is influenced by behavioral, cultural, and political factors. Future research should integrate behavioral dynamics, as highlighted by Schimit, Sergio, and Fontoura (2025) [18], to capture the interplay between social learning and vaccination uptake. Moreover, real-time adaptive strategies that adjust vaccination priorities as outbreaks evolve should be tested to reflect the dynamic nature of epidemics.

Finally, while Gephi was sufficient for this study's simulations, more advanced agent-based modeling platforms could provide deeper insights into multi-layered networks that combine physical, social, and digital interactions.

5.7 Contribution to knowledge

This study contributes to epidemiological research by providing empirical evidence that prescriptive graph-theoretical algorithms can significantly improve vaccination outcomes. Specifically, it fills the identified literature gap by demonstrating the practical effectiveness of MST and shortest-path frameworks in designing vaccination strategies. It also introduces the concept of hybrid optimization, showing that combining different structural properties yields superior results.

From a broader perspective, this research bridges the gap between abstract graph theory and applied public health, offering a methodological pathway for integrating advanced network science into epidemic preparedness and response.

6. Conclusion

The present study has demonstrated the critical value of integrating graph-theoretical approaches into epidemiological modeling, particularly for the design of efficient vaccination strategies. By applying minimum spanning tree and shortest-path algorithms to a real-world high-resolution contact network, the research provided compelling evidence that prescriptive optimization frameworks can significantly reduce epidemic size, delay peak infections, and flatten epidemic curves compared to conventional strategies such as random or degree-based vaccination. These findings confirm that epidemics are not merely shaped by biological transmissibility but also by the structural features of human interactions, which can be systematically leveraged to improve public health interventions.

The broader implications of this research extend beyond the specific dataset analyzed. In practical terms, the results underscore that resource-limited vaccination campaigns can achieve substantially greater impact if informed by network optimization rather than uniform allocation. In global health crises such as COVID-19, where vaccine shortages and logistical bottlenecks constrained early responses, targeted strategies based on spanning coverage and disruption of efficient infection routes could have mitigated epidemic severity more effectively. This points to a crucial need for governments and public health institutions to incorporate network data into epidemic preparedness frameworks, ensuring that interventions are both scientifically grounded and operationally efficient.

From a theoretical standpoint, the study advances epidemiological research by bridging the gap between descriptive metrics of network science and prescriptive decision-making algorithms. Earlier works often stopped at showing that certain nodes or clusters were important, but this research moved further to demonstrate how algorithms can directly guide intervention priorities. This transition from description to prescription represents a methodological step forward, expanding the role of graph theory from explanatory tool to actionable framework. The demonstrated success of a combined MST and shortest-path approach further reveals that hybrid strategies, which balance cluster coverage with pathway disruption, are likely to provide the most resilient solutions against future epidemics.

At the same time, the study highlights avenues for future research. While the results from a localized high school dataset provide strong proof of concept, larger-scale validation across diverse populations and geographic contexts is necessary to confirm scalability. Integrating real-time behavioral data, such as vaccine hesitancy patterns or mobility changes during lockdowns, could strengthen model realism and extend applicability. Moreover, the assumption of perfect vaccine efficacy and compliance needs to be relaxed, as real-world scenarios are often constrained by partial immunity, waning protection, and sociopolitical barriers to vaccination uptake. Future studies should explore adaptive strategies that evolve dynamically as outbreaks unfold, leveraging advances in artificial intelligence and real-time data analytics to optimize vaccination continuously rather than statically.

In sum, this research provides both empirical evidence and theoretical justification for advancing graph-theoretical approaches as a cornerstone of modern epidemiology. By demonstrating that vaccination strategies rooted in network optimization can save more lives with fewer resources, it makes a strong case for integrating these methods into global health policy and practice. The broader takeaway is clear: understanding and leveraging the structure of human interactions is not only a scientific necessity but also a moral imperative for safeguarding populations against future epidemics.

References

1. Angulo JJ, Pederneiras CA, Ebner W. Concepts of diffusion theory and a graphic approach to the description of the epidemic flow of contagious disease. Public Health Rep. 1980;95(3):255-262. Available from: http://pmc.ncbi.nlm.nih.gov/articles/PMC1422743/
2. Atkinson A, Ellenberger B, Piezzi V. Extending outbreak investigation with machine learning and graph theory: benefits of new tools with application to a nosocomial outbreak of a multidrug-resistant organism. Infect Control Hosp Epidemiol. 2023;44(4):654-662. Available from: http://www.cambridge.org/core/journals/infection-control-and-hospital-epidemiology/article/extending-

outbreak-investigation-with-machine-learning-and-graph-theory-benefits-of-new-tools-with-application-to-a-nosocomial-outbreak-of-a-multidrugresistant-organism/DB0BD838E0C2E93D0636E043A6CDA598

3. Braun RJ, Wilson RA, Pelesko JA, Buchanan JR. Applications of small-world network theory in alcohol epidemiology. J Stud Alcohol. 2006;67(4):591-599. Available from: http://www.jsad.com/doi/abs/10.15288/jsa.2006.67.591
4. Keeling MJ, Eames KTD. Networks and epidemic models. J R Soc Interface. 2005;2(4):295-307. Available from: http://royalsocietypublishing.org/doi/abs/10.1098/rsif.2005.0051
5. Maheswaran R, Craigs C, Read S, Bath PA. A graph-theory method for pattern identification in geographical epidemiology - a preliminary application to deprivation and mortality. Int J Health Geogr. 2009;8(28):1-12. Available from: http://link.springer.com/article/10.1186/1476-072X-8-28
6. Mushlin RA, Kershenbaum A, Gallagher ST. A graph-theoretical approach for pattern discovery in epidemiological research. IEEE Eng Med Biol Mag. 2007;26(3):74-82. Available from: http://ieeexplore.ieee.org/document/5386600
7. Nazeri A, Salahi-Moghadam A, Khoshdel AR. Explaining provincial disease network via graph-theoretic analysis of disease occurrence in Iran. Soc Sci Med. 2016;150:110-119. Available from: http://www.sciencedirect.com/science/article/pii/S1047279716300023
8. Rao ASRS. Population network structures, graph theory, algorithms to match subgraphs may lead to better clustering of households and communities in epidemiological studies. Epidemiol Infect. 2020;148:e242. Available from: http://www.cambridge.org/core/journals/epidemiology-and-infection/article/population-network-structures-graph-theory-algorithms-to-match-subgraphs-may-lead-to-better-clustering-of-households-and-communities-in-epidemiological-studies/1C0C984CAB49624507C53B2569A826CA
9. Venkatraman Y, Balas VE, Rad D. Graph theory applications to comprehend epidemics spread of a disease. BRAIN Broad Res Artif Intell Neurosci. 2021;12(1):320-334. Available from: http://brain.edusoft.ro/index.php/brain/article/view/1146
10. Widgren S. Graph theory in veterinary epidemiology. Uppsala: Swedish University of Agricultural Sciences; 2003. Available from: http://stud.epsilon.slu.se/12723
11. Chang SL, Piraveenan M, Prokopenko M. Impact of network assortativity on epidemic and vaccination behaviour. Chaos Solitons Fractals. 2020;132:109589. Available from: https://www.sciencedirect.com/science/article/pii/S0960077920305397
12. Gandon S, Day T. The evolutionary epidemiology of vaccination. J R Soc Interface. 2007;4(16):803-817. Available from: https://doi.org/10.1098/rsif.2006.0207
13. Holme P, Litvak N. Cost-efficient vaccination protocols for network epidemiology. PLoS Comput Biol. 2017;13(9):e1005696. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005696
14. Kabir KMA, Kuga K, Tanimoto J. The impact of information spreading on epidemic vaccination game dynamics in a heterogeneous complex network: a theoretical approach. Chaos Solitons Fractals. 2020;132:109574. Available from: https://www.sciencedirect.com/science/article/pii/S0960077919305053
15. Li Y, Zhou X, Lyu Z. Regional contagion in health behaviors: Evidence from COVID-19 vaccination modeling in England with social network theorem. J Comput Soc Sci. 2024;7(2):421-440. Available from: https://link.springer.com/article/10.1007/s42001-023-00232-9
16. Liu Y, Deng Y, Jusup M, Wang Z. A biologically inspired immunization strategy for network epidemiology. J Theor Biol. 2016;400:92-102. Available from: https://www.sciencedirect.com/science/article/pii/S0022519316300479
17. Ma J, Van Den Driessche P, Willeboordse FH. The importance of contact network topology for the success of vaccination strategies. J Theor Biol. 2013;325:12-21. Available from: https://www.sciencedirect.com/science/article/pii/S0022519313000258
18. Schimit PHT, Sergio AR, Fontoura MAR. Vaccination as a game: behavioural dynamics, network effects, and policy levers A comprehensive review. Mathematics. 2025;13(14):2242. Available from: https://www.mdpi.com/2227-7390/13/14/2242
19. Thedchanamoorthy G, Piraveenan M, Uddin S. Influence of vaccination strategies and topology on the herd immunity of complex networks. Soc Netw Anal Min. 2014;4(1):1-14. Available from: https://link.springer.com/article/10.1007/s13278-014-0213-5
20. Wang Z, Moreno Y, Boccaletti S, Perc M. Vaccination and epidemics in networked populations an introduction. Chaos Solitons Fractals. 2017;103:177-183. Available from: https://www.sciencedirect.com/science/article/pii/S0960077917302515