Social Network Analysis and Mining (2025) 15:89
https://doi.org/10.1007/s13278-025-01509-9

ORIGINAL ARTICLE

Leveraging social network analysis and mobility data for modeling epidemic spread in urban tourist destinations

Nitika Pathania[1] · **Brian Labus**[2] · **Shaikh Arifuzzaman**[1]

Received: 31 May 2025 / Revised: 16 July 2025 / Accepted: 22 July 2025
© The Author(s) 2025

Abstract
This paper introduces a novel framework that integrates agent-based mobility simulation with social network analysis to model epidemic diffusion, using Las Vegas as a high-stakes test bed. By mining real-world transportation, demographic, and geographic data, we construct a high-resolution, dynamic contact network that captures the heterogeneous interactions among key subpopulations. Applying graph theory and diffusion modeling to this evolving network demonstrates that epidemic risk is structurally determined by critical transmission hotspots (major resorts, transport hubs) and is highly sensitive to the initial seeding location, with an introduction at the airport leading to a far more explosive outbreak than one in a residential zone. We demonstrate how distinct population groups play unique network roles: tourists act as primary importation vectors, while hospitality workers serve as crucial bridging nodes between otherwise disconnected clusters. The model also shows how some agents achieve "immunity by circumstance" due to mobility patterns that isolate them from high-traffic areas. Our experiments prove that targeted interventions, such as capacity limits at identified hotspots, substantially outperform universal restrictions. The framework validated in Las Vegas provides a transferable tool for strategic epidemic preparedness, offering a new paradigm for data-driven public health policy applicable to any globally connected urban center.

Keywords Social network analysis · Epidemic modeling · Human mobility · Graph theory · Machine learning · Infectious disease spread · Urban networks

1 Introduction

Network modeling and analysis have emerged as critical tools in understanding complex systems across diverse domains–from biology and infrastructure to finance, communication, and public health (Sattar et al. (2023); Faysal et al. (2023); Arifuzzaman and Pandey (2019); Arifuzzaman et al. (2019); Eubank et al. (2004); Azad et al. (2017); Sattar et al. (2025); Faysal and Arifuzzaman (2019); Sattar and Arifuzzaman (2022); Adiga et al. (2015); Funk et al. (2009)). By representing interactions as graphs, social network analysis (SNA) offers a powerful framework to uncover patterns of connectivity, influence, and diffusion (Faysal and Arifuzzaman (2019); Sattar and Arifuzzaman (2022); Faysal et al. (2021)). In the context of epidemiology, network-based approaches enable researchers to simulate disease propagation across populations with unprecedented granularity, helping to identify vulnerable subgroups, predict outbreak dynamics, and design effective intervention strategies (Eubank et al. (2004); Adiga et al. (2015); Funk et al. (2009)).

Las Vegas, often celebrated as the *entertainment capital of the world*, draws over 40 million tourists each year and thrives on a dynamic interplay between transient and resident populations. While this economic model fuels local businesses and employment, it also creates unique vulnerabilities during infectious disease outbreaks Tizzoni et al. (2013). Highly dense attractions, frequent mass gatherings, and the constant flux of people make the city

✉ Shaikh Arifuzzaman
shaikh.arifuzzaman@unlv.edu

Nitika Pathania
pathania@unlv.nevada.edu

Brian Labus
brian.labus@unlv.edu

[1] Department of Computer Science, University of Nevada, Las Vegas (UNLV), 4505 S Maryland Pkwy, Las Vegas, NV 89154, USA

[2] Department of Epidemiology and Biostatistics, University of Nevada, Las Vegas (UNLV), 4505 S Maryland Pkwy, Las Vegas, NV 89154, USA

Published online: 24 August 2025

a critical case for studying epidemic dynamics in tourism-driven urban environments.

Despite significant advancements in urban epidemic modeling (Adiga et al. (2015); Eubank et al. (2004); Kellermann et al. (2022); Charu et al. (2017); Tizzoni et al. (2013); Marra et al. (2022); Balcan et al. (2009); Patil et al. (2021); Chatterjee et al. (2016); Rothengatter et al. (2021)), few studies focus on cities whose identity and infrastructure are defined by tourism. Prior works have demonstrated the effectiveness of combining synthetic population models with detailed mobility networks to understand transmission, but most of these efforts target conventional metropolitan settings (Balcan et al. (2009); Patil et al. (2021); Chatterjee et al. (2016); Rothengatter et al. (2021)). This leaves a critical gap in accounting for cities like Las Vegas, where tourist activity dramatically shapes mobility behaviors and public health risk.

Recent studies have demonstrated the utility of SNA in understanding how transportation and mobility networks contribute to the spread of infectious diseases. For instance, SNA metrics are used to uncover vulnerabilities in Shanghai's urban traffic systems by identifying influential intersections and bottlenecks in Chang et al. (2022). Subway-based transmission is explored for Beijing, showing that high-centrality nodes like major transfer stations act as superspreaders in Wang et al. (2024). How modular structures within networks can delay or localize outbreaks is highlighted in Salathé and Jones (2010). These insights affirm that epidemic diffusion is not only a function of individual mobility, but also the structure of the underlying contact or transportation network. Our work builds on this perspective by modeling epidemic propagation on a real-world vehicular network of Las Vegas, where SNA is applied to mobility traces to identify critical locations and assess temporal-spatial risk dynamics.

In this paper, we extend these foundational modeling frameworks to a real-world, tourism-intensive city by integrating high-resolution vehicle trajectory data from mobility simulation (Eclipse Foundation (2024)) with demographic and spatial data from sources such as OpenStreetMap (OpenStreetMap contributors (2025)), Regional Transportation Commission (RTC), and the U.S. Census Bureau (Zhang and Wang (2023)). Drawing inspiration from the field of Social Network Analysis and Mining (SNAM) (Arifuzzaman and Pandey (2019); Faysal and Arifuzzaman (2019); Beamer et al. (2015); Chakravarty and Arifuzzaman (2024)), our approach represents a fusion of individual-level mobility modeling and graph-based analysis of dynamic social contact networks. This enables us to explore how local interaction structures and temporal connectivity patterns shape transmission risk, echoing recent trends in SNAM research that emphasize structural embedding, network-driven diffusion, and activity-aware modeling (Kellermann et al. (2022), Charu et al. (2017)).

1.1 Our contribution

Our primary contributions in this paper are organized into a multi-faceted simulation framework designed to model epidemic spread in a dynamic urban tourist destination:

- *A novel agent-based simulation framework.* We developed a high-fidelity simulation framework that bridges the gap between urban mobility and epidemic dynamics, specifically using Las Vegas as a test bed. Our approach uses individual agents as the core unit, where each agent has an epidemiological status (e.g., *Susceptible*, *Infected*, or *Recovered*), while simulated vehicles serve merely as their mobility proxies. This crucial distinction allows for a more realistic representation of human movement and interaction (Ruan et al. (2015); Topîrceanu (2024)).
- *Realistic spatiotemporal transmission mechanism.* To realistically model disease spread, we implemented a transmission model based on the spatiotemporal co-location of agents. An infection event can occur probabilistically when a susceptible agent shares the same physical location (e.g., a casino, restaurant, or bus stop) with an infected agent for a significant duration (e.g., 8-10 min), capturing the dynamics of airborne or close-contact transmission in dense urban environments.
- *Integration of rich, multi-layered data.* The model is enhanced by integrating diverse datasets to reflect real-world complexities. This includes demographic attributes, transportation patterns from airport arrivals and public transit, crowd density metrics, and location-specific behavioral data such as typical dwell times in tourist hotspots. This multi-layered approach ensures our outbreak scenarios are grounded in realistic urban activity (Yang et al. (2009); Craig et al. (2020); Malik et al. (2022)).
- *An interactive tool for policy & intervention analysis.* We created a user-interactive tool that allows for the custom design and simulation of trips between various Points of Interest (POIs). This feature, coupled with the overall framework, provides actionable insights for public health officials. It enables the evaluation of various non-pharmaceutical interventions, such as mobility restrictions, social distancing mandates, and targeted crowd control measures. Furthermore, the framework's modular design ensures it is adaptable for use in other urban tourist destinations beyond Las Vegas.

2 Our modeling framework and methodology

We construct a comprehensive, multi-modal simulation framework that models individual mobility and interaction patterns in tourism-driven cities characterized by high population transience and dense entertainment infrastructure. Leveraging historical and real-time public transportation data, demographic distributions, and annotated points of interest (POI), we generate high-resolution mobility traces that serve as the foundation for constructing dynamic contact networks. These networks capture the temporal co-location and movement overlap among agents, enabling the application of graph-based epidemic models. By embedding individuals within evolving social-mobility graphs, our methodology facilitates deeper insights into how urban dynamics amplify health risks. This framework not only quantifies the structural factors contributing to potential disease spread but also supports the evaluation of targeted public health interventions under various what-if scenarios. The overview of our modeling framework is shown in Fig. 1 and is described in the following subsections.

2.1 Comprehensive mobility modeling

Our methodology begins with a detailed mobility simulation designed to map the complex movement patterns of a major tourist destination. The model is built upon data

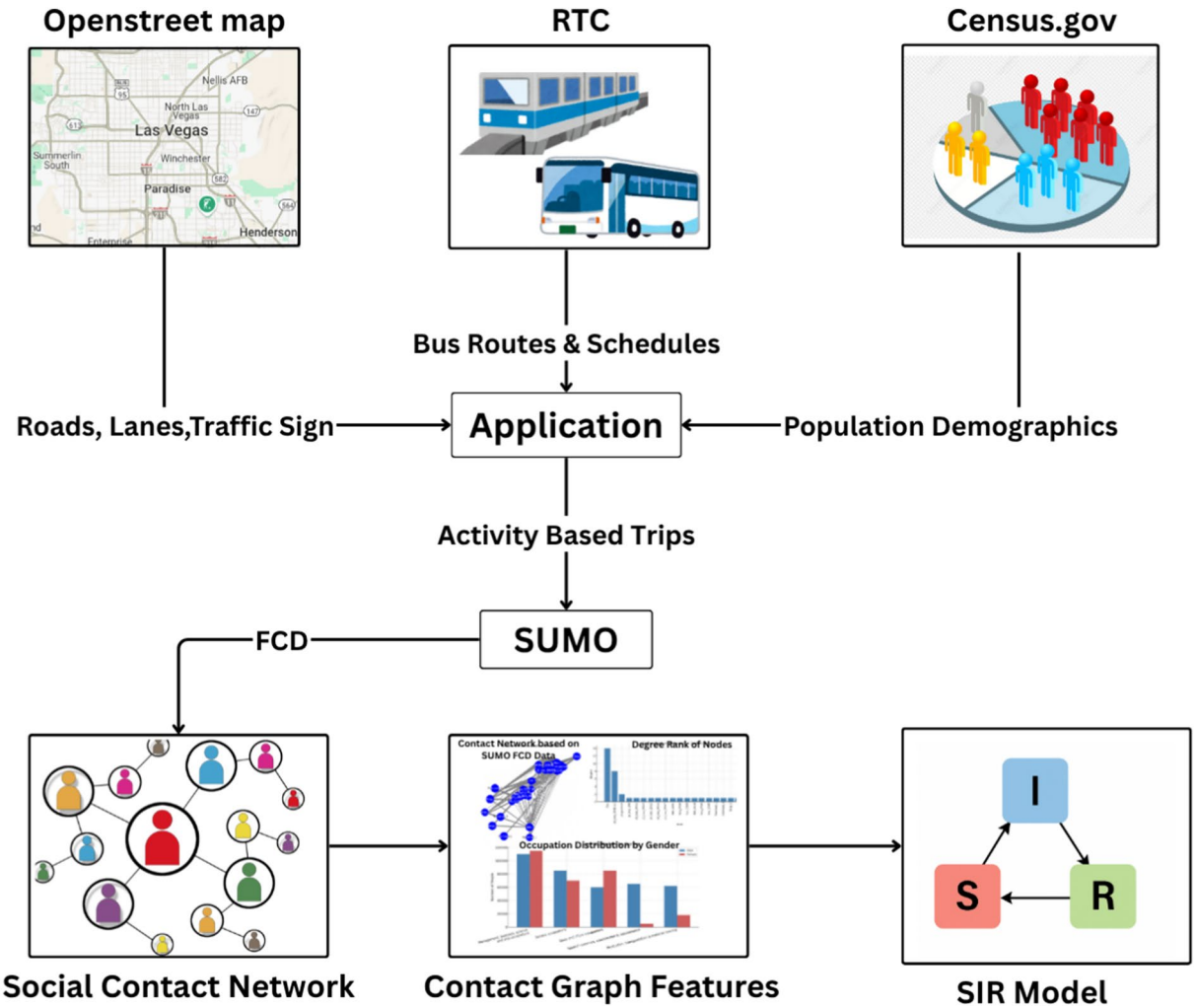

Fig. 1 Overview of the data-driven epidemic modeling framework. The system integrates geospatial road data from OpenStreetMap (OpenStreetMap contributors (2025)), public transit routes from RTC, and demographic data from U.S. Census Bureau into a unified application to generate activity-based mobility scenarios. These scenarios are simulated in SUMO (Eclipse Foundation (2024)), producing Floating Car Data (FCD), which is then used to construct a social contact network. The network enables the extraction of structural features, which are further used to simulate disease transmission dynamics using the SIR model

sourced from local transportation agencies, incorporating details such as public transit routes, schedules, and ridership statistics to ensure a realistic representation of urban travel.

2.1.1 Data collection

We extract Point of Interest (POI) information from OpenStreetMap (OSM) (OpenStreetMap contributors (2025)) that provides detailed spatial annotations for key urban locations such as casinos, hotels, transit stations, and entertainment venues. We also collect demographic data from the U.S. Census Bureau, which encompasses attributes such as race, education level, and occupational categories. Using these distributions, we can generate a synthetic population of the city, where individual agents are assigned demographic characteristics in proportion to real-world statistics. This probabilistic sampling approach ensures that the simulated population preserves the heterogeneity and diversity observed in the actual urban population.

2.1.2 POI extraction via named entity recognition

We employ spaCy's Named Entity Recognition model to extract Points of Interests (POIs) from unstructured text. Entities classified as GPE (Geo-Political Entity), LOC (Location), and ORG (Organization) were filtered to retain only relevant categories such as casinos, hotels, airports, schools, grocery stores, and convention centers. The resulting POIs were normalized for consistency and integrated as source and destination nodes in the SUMO trip-generation module, streamlining large-scale mobility scenario generation.

2.1.3 Passenger creation and activity assignment

We first create individual agents representing people in the simulation, each assigned demographic attributes such as age, occupation, and household status. Once individuals are created, their activities are assigned based on their demographic profiles. For example, students are assigned trips to school, workers are assigned trips to their workplaces, and elderly individuals are more likely to be assigned activities such as visiting parks or healthcare facilities. These assignments are informed by age, employment status, household structure, and other demographic indicators derived from U.S. Census data (Han et al. (2015); Burbano Lombana et al. (2022); Kumar et al. (2021)).

2.1.3.1 POI-based activity assignment using web application To support spatially accurate and dynamic activity assignment, we developed a custom-built web application that enables the interactive configuration of activity-based mobility scenarios. The user interface of this application is shown in Fig. 2. The purpose of this tool is to allow users to associate each synthetic agent with real-world Points of Interest (POIs) in a geographically meaningful manner. While demographic profiles (e.g., student, worker, retired) determine what type of activities an individual performs, this application determines where these activities take place within the simulated urban space. The application integrates several external data sources, as listed below.

- OpenStreetMap–used for road networks, POIs, lane and signal information.
- RTC Transit Data–includes bus routes, stop locations, and monorail schedules.
- U.S. Census Data–provides demographic distributions used to generate the synthetic population.

These datasets ensure that both spatial and behavioral aspects of mobility are grounded in real-world infrastructure and usage patterns.

2.1.3.2 User interface and interaction The simulation interface allows the users to perform certain activities and interactions with the tool.

- *User ID selection:* Users can begin by selecting a specific User ID, which corresponds to a particular agent within the simulation or dataset. This enables personalized data queries and analysis tied to individual travel patterns or behaviors.
- *Category filtering:* The interface allows users to choose distinct categories of Points of Interest (POIs), such as shop and sport. This feature helps narrow down the focus to relevant types of destinations, aligning with specific research or analytical objectives.
- *Specific POI type selection:* Within the chosen categories, users can further refine their selection by choosing particular POI types. For example, under the shop category, a user might select department store, while under sport, they might choose gym. This level of granularity ensures precise targeting of the types of locations being studied.
- *POI retrieval from OpenStreetMap:* Once the desired POI types are selected, the system automatically queries the OpenStreetMap (OSM) database to retrieve the relevant POI identifiers. These identifiers link to geospatial data that defines the location and metadata of each POI.
- *Interactive Map Visualization:* The selected POIs are then visualized on a dynamic, leaflet-based map embedded within the interface. This map allows users to explore the spatial distribution of the POIs, zoom into specific areas, and interact with individual markers for further information.

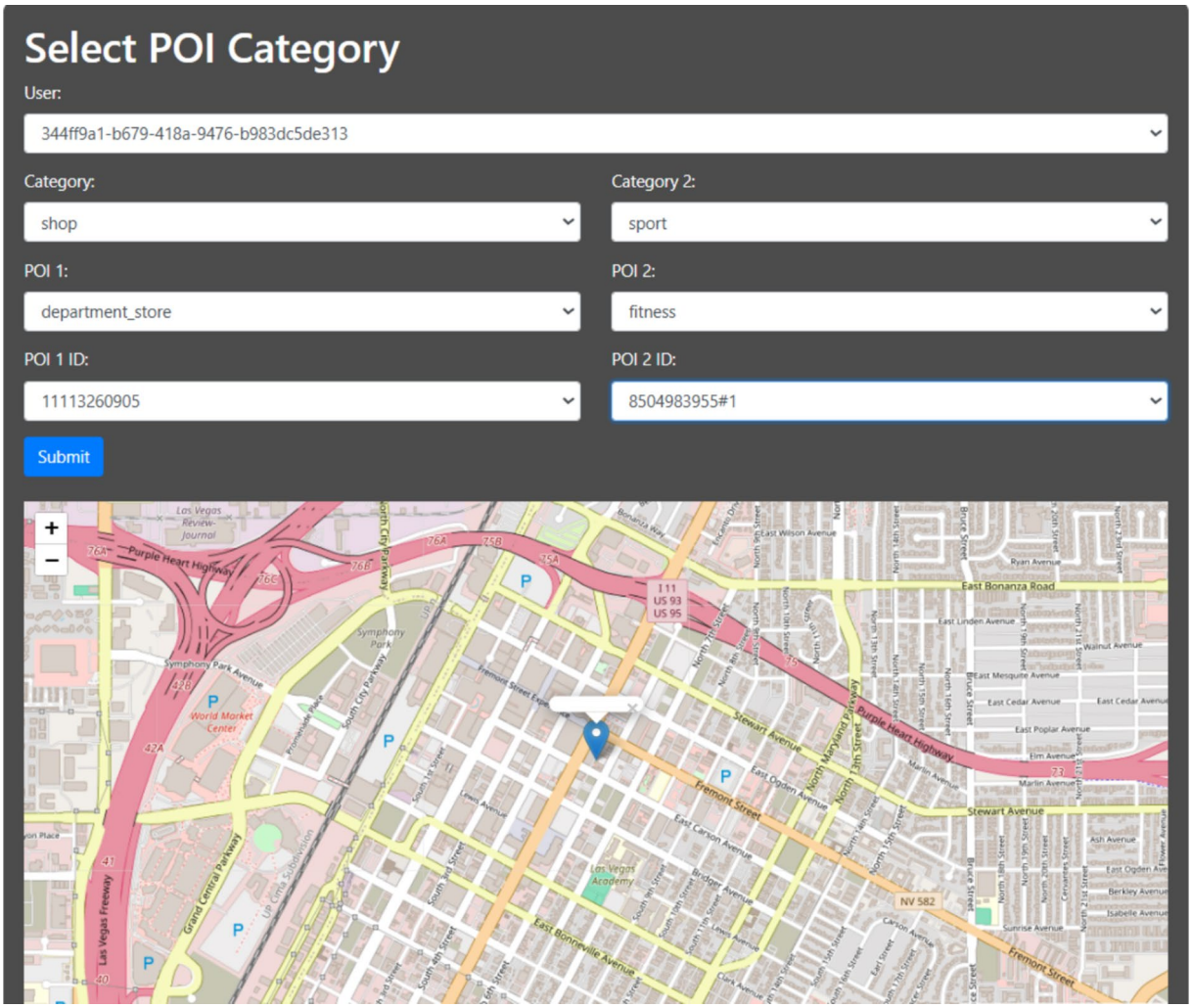

Fig. 2 Web-based application interface used for configuring activity locations. The interface allows users to select POI categories and subcategories, view real POI locations on a map, and assign activity destinations to individual synthetic agents based on real-world spatial data

The interface updates in real time, showing POI markers on the map to help the user verify spatial context, distribution, and proximity within the city network. Once POIs are selected, the application stores them as part of the agent's activity schedule. These POI-based selections become the basis for trip generation.

2.1.4 Trip generation strategy

With origin and destination POIs defined for each synthetic individual, the trip generation phase focuses on translating these activity pairs into actual vehicle movements on the road network. For each POI, the closest edge in the OpenStreetMap-derived road network is identified to ensure spatial accuracy and realistic anchoring of trip endpoints. These trip definitions are then processed using the mobility simulation tool's routing module (Eclipse Foundation (2024)), which computes the shortest or fastest paths through the network based on user-defined routing options and traffic rules.

2.2 Dynamic contact graph generation

The trips generated are executed in a mobility simulation tool, producing Floating Car Data (FCD). This FCD contains time-stamped location and movement data for every individual in the simulation. To simulate disease spread dynamics, a dynamic contact graph is constructed based on the spatial-temporal proximity of agents derived from the FCD. Contact events are identified by comparing vehicle positions at each timestep and detecting pairs of agents within a predefined spatial threshold (e.g., within 2m for at least 10s). These interactions are treated as edges in the dynamic contact graph, where nodes represent individuals and edges represent potential transmission events.

This graph is updated at each simulation step, allowing for the modeling of time-varying interactions. The contact graph serves as the foundational input for running the epidemiological model, capturing the nuanced movement and interaction behavior of the population in an urban setting like Las Vegas.

By basing the graph on realistic, activity-driven trajectories and vehicle-level interactions, this approach provides a higher-fidelity alternative to traditional static or grid-based contact estimation methods.

2.3 Using SIR model for disease spread

With the dynamic contact graph in place, we apply the classical SIR (Susceptible-Infected-Recovered) model to simulate the spread of an infectious disease through the urban population. Initially, a subset of agents is designated as infected, while the remainder are classified as susceptible. At each timestep, the model iterates through the contact graph, identifying which susceptible individuals are exposed to infected ones through prior contact events. Transition probabilities for infection and recovery are assigned based on configurable epidemiological parameters (Tizzoni et al. (2013)).

The model runs over multiple time steps to track the evolution of the disease, recording the number of susceptible, infected, and recovered individuals at each stage. This simulation enables the analysis of transmission dynamics in response to real-world mobility patterns and contact behaviors derived from both private and public transportation usage.

To create a realistic agent population and enhance model fidelity, our framework leverages demographic profiles from the U.S. Census Bureau to assign each agent specific attributes such as age, occupation, and residential location. This allows us to differentiate among population subgroups such as residents, tourists, and service workers-each with distinct mobility patterns and exposure risks. For example, service workers in the hospitality sector may experience higher interaction rates in dense urban areas like casinos and shopping centers, while tourists exhibit more spatial variability during major events and peak seasons.

Additionally, OpenStreetMap data is used to identify high-density locations such as hotels, entertainment venues, and key transit intersections. These zones are analyzed for high-contact potential, contributing to a more nuanced understanding of urban disease dynamics.

This integrated SIR modeling approach, grounded in realistic trip data and contact behavior, offers a powerful framework for exploring potential public health interventions, mobility restrictions, and policy impacts in complex urban environments like Las Vegas.

2.4 Forecast health risk and transmission patterns

Using the SIR model, this framework will identify areas where disease transmission is most likely, focusing on high-risk areas such as transportation hubs, tourist sites, and densely populated venues. The model will simulate interactions between individuals within these spaces, assessing how varying degrees of contact, event duration, and density impact the spread of infections. This allows us to forecast the transmission risk in crowded areas and assess how factors such as tourist influx, resident mobility, and use of public transport contribute to the spread of infectious diseases.

This information is particularly valuable for public health officials as it allows them to understand potential outbreak dynamics and prepare effective response plans for different scenarios (Pozzana et al. (2017); Schläpfer et al. (2014)). By identifying zones that are more susceptible to infection spikes, public health authorities can develop preventive interventions tailored to the specific needs of Las Vegas's urban infrastructure and population behavior.

2.5 Development of actionable insights for city officials

The insights generated from the simulation framework will address traffic congestion, public health concerns, and infrastructure needs. For example, by analyzing traffic flow and population density in specific areas, the model can suggest optimal placement of sanitation stations, increased transportation frequency during peak periods, or other interventions that can mitigate health risks. Furthermore, the model can help officials make data-driven decisions about potential restrictions, capacity limitations, or enhanced monitoring protocols in zones identified as high-risk.

Additionally, insights from the simulation can support long-term public health planning by identifying patterns that may persist beyond individual events. City officials can use these findings to design proactive policies that strengthen a city's resilience to future disease outbreaks. Data could also help urban planners consider mitigation of health risks in new infrastructure projects or transportation systems, ensuring the city is better prepared to handle similar risks in the future (Qian et al. (2021)).

2.5.1 Evaluation of preventive measures and mitigation strategies

The model further investigates targeted interventions tailored to high-risk locations. Measures such as reducing crowd density at transportation hubs, limiting occupancy in popular attractions, and implementing temporary closures during peak transmission periods are simulated to evaluate their comparative efficacy. In addition, we analyze the role

of proactive strategies like vaccination rollouts and rapid testing initiatives in suppressing infection rates under varying levels of compliance and resource availability (Nadini et al. (2020)). The results of these simulations are intended to inform adaptive policy design, enabling decision makers to implement context-sensitive evidence-based response plans. By identifying the most impactful and feasible mitigation strategies, this component of the study supports the development of dynamic public health responses capable of adjusting to evolving epidemiological conditions (Nyasulu et al. (2021)).

2.5.2 Optimization of health resource allocation

Efficient allocation of health resources, such as medical personnel, testing sites, and medical supplies, is crucial in responding effectively to disease outbreaks. Our simulation framework aims to optimize resource distribution by highlighting areas of the city where resources would be most impactful based on simulation results. By simulating different outbreak scenarios, the model can identify areas that may experience surges in infections and require additional support, such as hospitals, testing facilities, or vaccination sites. This objective seeks to create a data-driven approach for resource allocation, enabling the efficient deployment of resources to areas of highest need, thus minimizing response times and improving health outcomes during potential outbreaks.

Additionally, the model can help determine optimal locations for temporary health facilities or testing stations during peak tourist seasons or high-density events, ensuring that resources are easily accessible to both residents and visitors. This targeted approach to resource allocation aims to enhance the city's overall healthcare response, reducing the strain on existing infrastructure and ensuring readiness for future health crises.

2.5.3 Long-term public health strategy and preparedness

Beyond immediate disease outbreak scenarios, our simulation framework can contribute to the development of a long-term public health strategy for urban tourist environments. By identifying recurring patterns and potential health risks in areas of high population density, the model can guide the implementation of preventive measures that are sustainable over time. This will support the creation of long-term plans, such as incorporating disease transmission risk assessments into urban development projects, improving public health infrastructure, and fostering public awareness campaigns targeted at high-risk zones (Topîrceanu (2024)). With a data-driven understanding of mobility and health risks, city planners can work towards creating an environment that naturally reduces the risk of disease spread.

3 Simulating our framework using Las Vegas urban data

3.1 Creating realistic demographic profiles

To construct realistic agent profiles, we leveraged granular data from the U.S. Census Bureau. This dataset provided key demographic attributes for the Las Vegas population, including age distributions, occupation types, and common commuting patterns. A gender-wise occupation distribution for Las Vegas is shown in Fig. 3.

This information was essential for creating distinct agent archetypes within the simulation. We were able to model the unique mobility behaviors of different groups, such as:

- *Residents* following typical daily routines.
- Hospitality *workers* commuting to major employment zones like the Las Vegas Strip.
- *Tourists* exhibiting varied travel patterns between hotels, attractions, and other Points of Interest (POIs).

The synthesis of these demographic profiles with our transportation and geographic data layers creates a multi-layered, comprehensive model of the urban environment. This integration is crucial for ensuring the simulation accurately represents real-world conditions and produces valid, actionable results.

3.2 Integrating congestion and population density

To accurately simulate traffic flow and crowd dynamics, we built a detailed digital model of Las Vegas's physical infrastructure. Beyond the RTC transit data (Table 1), we constructed the simulation's foundational map using *OpenStreetMap (OSM)*, a collaborative, open-source mapping platform.

From OSM, we extracted two critical layers of data:

1. *A detailed road network:* This included the complete graph of streets and intersections, along with their properties like speed limits, lane counts, and traffic signal locations. This network forms the infrastructure upon which all vehicle movement is simulated.
2. *Geographic points of interest (POIs):* We precisely mapped high-traffic venues that act as major hubs for human activity, especially casinos, hotels, and entertainment spots along the Las Vegas Strip. See Fig. 4 for a visualization.

Table 1 Summary of selected RTC (Regional Transportation Commission) public bus routes in Las Vegas, detailing route IDs, start and end points, service frequency, average daily ridership, and peak operating hours. High-frequency routes such as the DEUCE and Tropicana (Route 201) serve major corridors with high daily ridership, indicating their significance in urban mobility and potential exposure hotspots for disease transmission.

Route ID	Route name	Start point	End point	Schedule (Freq)	Avg daily riders	Peak hours
201	Tropicana	Nellis Blvd	Rainbow Blvd	Every 15 mins	12,000	7:00 - 9:00AM, 5-7PM
109	Maryland Parkway	South Strip TC	Centennial Hills	Every 20 mins	9,800	6:30 - 8:30AM
DEUCE	Strip & Downtown	Mandalay Bay	Fremont Street	Every 10 mins	18,500	11:00AM - 9:00PM
SX	Sahara Express	Hualapai Way	Boulder Hwy	Every 12 mins	10,200	6:00 - 9:00AM
CX	Centennial Express	Downtown LV	Centennial Hills	Every 30 mins	5600	4:00 - 7:00PM
BHX	Boulder Hwy Express	Downtown LV	Henderson	Every 15 mins	6500	6:00 - 8:00AM

This integration is fundamental to our framework. The road network allows *congestion* to emerge organically in the simulation when traffic volumes exceed road capacity at bottlenecks. Simultaneously, the POIs serve as origins and destinations for agents, allowing the model to calculate dynamic *population density* based on how many agents are present at a specific location at any given time. By combining these elements, we can realistically model the formation of crowds and traffic jams, which are critical factors in epidemic spread (Fig. 5).

3.3 Simulating mobility dynamics

We developed a high-fidelity, agent-based mobility model for Las Vegas using the *Simulation of Urban MObility (SUMO)* platform (Eclipse Foundation (2024)). Within this framework, each individual is modeled as a distinct agent with unique behavioral patterns based on their demographic profile. We programmed two primary agent archetypes to capture the city's diverse movement dynamics:

- *Residents*, who follow structured daily routines such as commuting between residential areas and workplaces.
- *Tourists*, who exhibit more stochastic mobility patterns, traveling between hotels, casinos, entertainment venues, and other Points of Interest (POIs).

To generate realistic travel demand, we integrated multiple real-world data sources. Agent itineraries, created via a custom web application, were enriched with:

- *Intra-city tourist transit:* Las Vegas Monorail ridership data was used to accurately model high-frequency movement along the Resort Corridor.
- *Inter-city Ingress/Egress:* We incorporated FlixBus route data from major metropolitan hubs like Los Angeles and Phoenix to simulate the arrival and departure of visitors, providing a crucial vector for modeling potential external sources of infection.

Our simulation and analysis focused on key zones recognized for high population density and traffic, including the Las Vegas Strip, Harry Reid International Airport (LAS), and Downtown Las Vegas. Furthermore, the model's dynamism is enhanced by incorporating temporal factors, such as peak tourist seasons and large-scale events. This capability

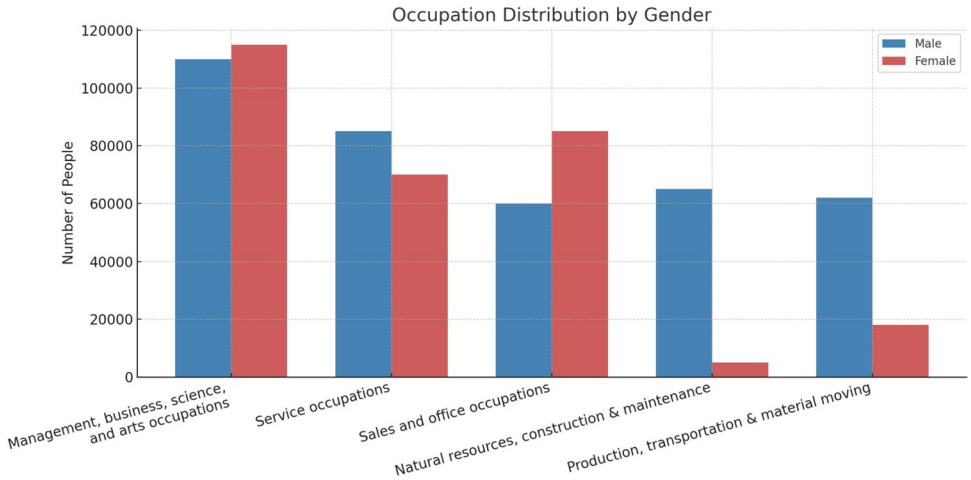

Fig. 3 Gender-wise occupation distribution in the modeled population based on U.S. Census occupational categories. The data distinguishes between male and female employment across five major occupational sectors: management/business/arts, service, sales/office, construction/maintenance, and transportation/material moving. Management and professional occupations dominate for both genders, while males are more concentrated in construction and transportation-related jobs

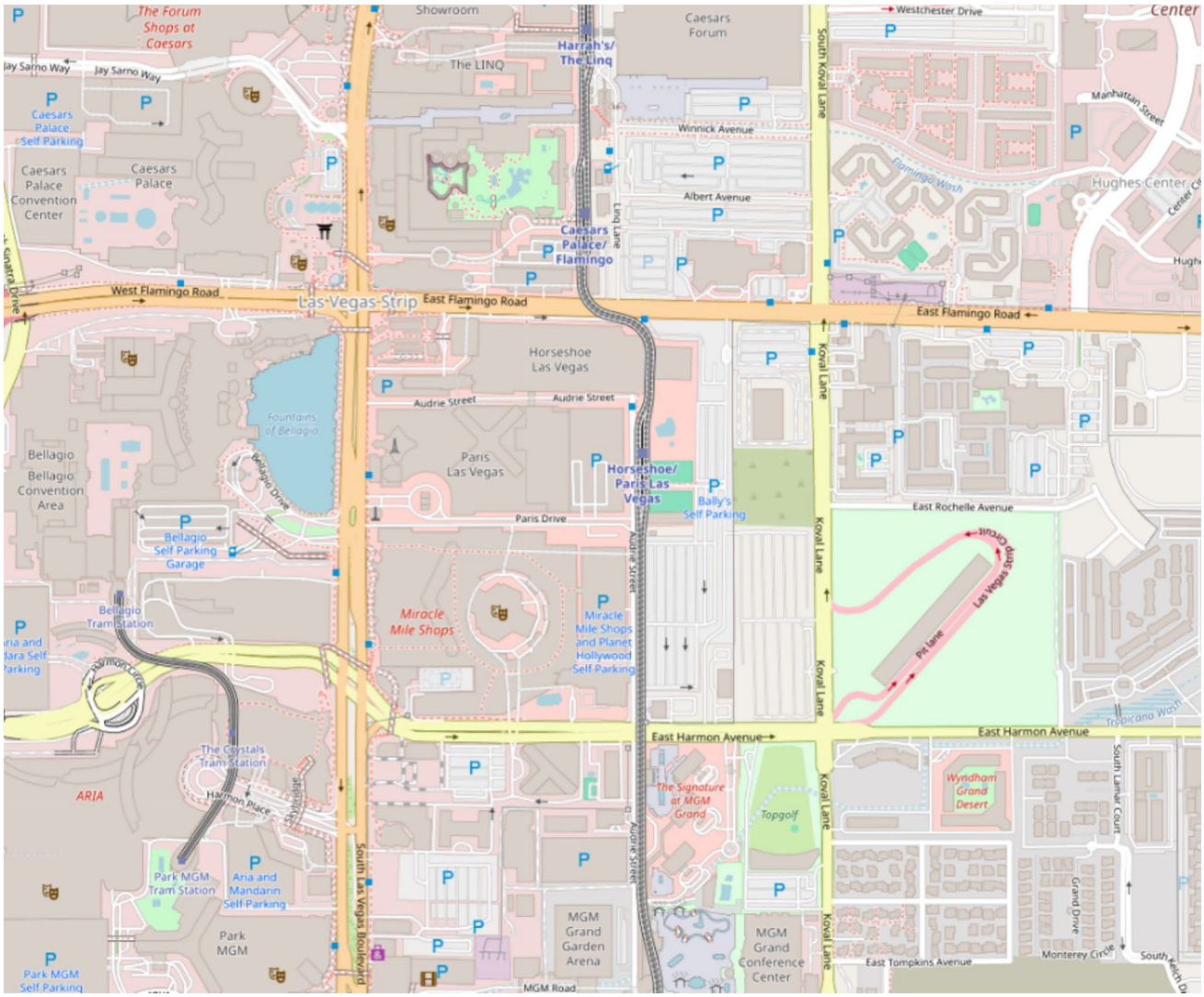

Fig. 4 Las Vegas strip with hotels, casinos and other points of interest shown, these POIs are major attractions of Las Vegas drawing majority of the foot traffic in the city and thus playing a key role in the activity assignment in our simulation

allows us to investigate how periods of heightened congestion and crowding amplify public health risks.

Ultimately, this process yields a rich, heterogeneous mobility landscape. This dynamic simulation serves as the essential foundation for generating contact networks and analyzing subsequent disease propagation scenarios.

3.4 Epidemiological modeling and diffusion

To simulate disease spread, we coupled our mobility framework with an agent-based epidemiological model. This health layer is built upon the classic *Susceptible-Infected-Recovered (SIR)* paradigm, a cornerstone of infectious disease modeling. Each agent in the simulation is assigned one of three health states:

- Susceptible (S): Healthy individuals who are vulnerable to infection.
- Infected (I): Individuals who have the disease and are capable of transmitting it.
- Recovered (R): Individuals who were previously infected and are now immune and non-contagious.

Disease transmission is not assumed but emerges directly from the agents' simulated interactions. An infection is modeled as a probabilistic event triggered by *spatiotemporal co-location*. Specifically, when a susceptible (S) agent shares a confined space (e.g., a casino floor, public bus, or restaurant) with one or more infected (I) agents for a significant duration, a transmission event may occur. The probability is weighted by factors such as proximity,

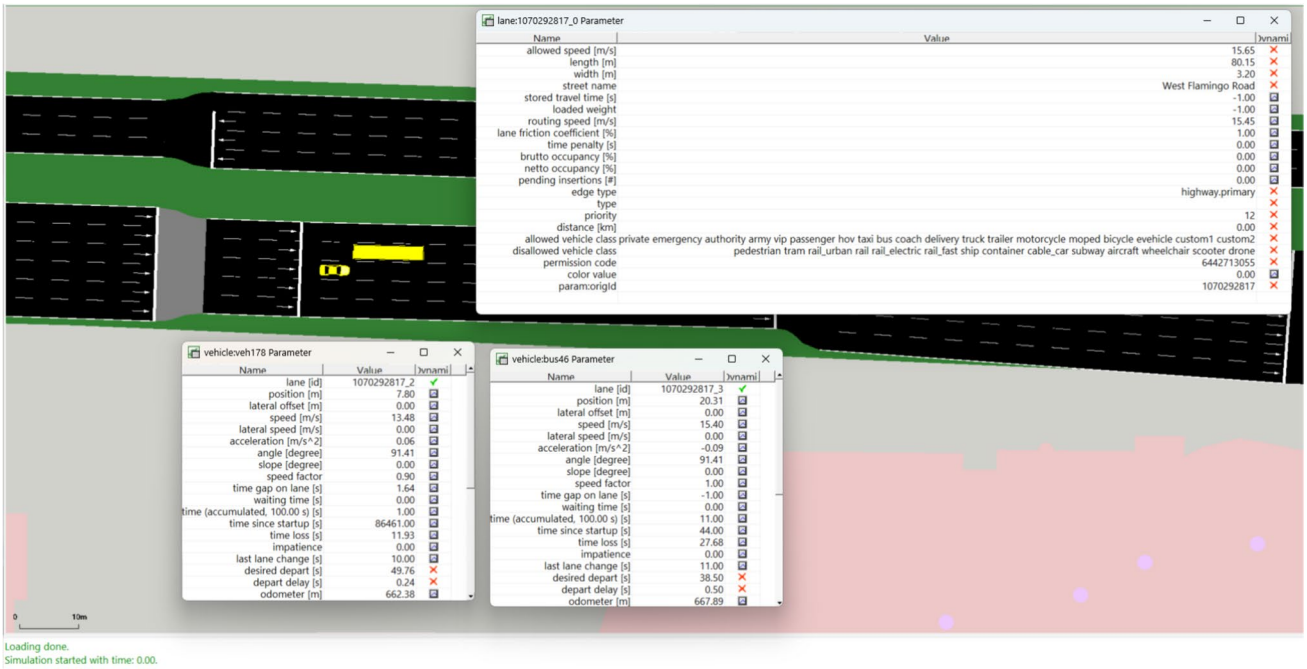

Fig. 5 A snapshot of the SUMO simulation environment. The main window displays vehicle agents on West Flamingo Road, while supplementary windows provide real-time data for selected entities, demonstrating the model's micro-simulation capabilities

duration of exposure, and location-specific attributes like crowd density.

By integrating SIR dynamics with fine-grained agent mobility, our framework can trace an epidemic's trajectory with high spatiotemporal resolution. This enables us to move beyond aggregate statistics to identify potential *super-spreader hotspots*-specific Points of Interest or transit routes that contribute disproportionately to transmission. Consequently, we can analyze how different mobility scenarios, such as a major convention or holiday weekend, influence overall infection rates and the geographic pattern of an outbreak across Las Vegas.

4 Simulation results and insights

4.1 General epidemic dynamics through agent-based modeling

Our integrated mobility and epidemiological model produced a classic epidemic trajectory, demonstrating the fundamental dynamics of disease spread within the simulated Las Vegas population. Figure 6 plots the evolution of the *Susceptible* (S), *Infected* (I), and *Recovered* (R) agent populations over a 100-day period, with contact events derived from the SUMO mobility simulation.

The simulation begins with a nearly entirely susceptible population. The outbreak exhibits an initial phase of exponential growth, with the number of infected agents peaking sharply at over 600 around day 18. This rapid spread directly corresponds to a steep decline in the susceptible pool, indicating widespread transmission during the early, unmitigated phase of the epidemic.

Following this peak, the infection rate slows and then reverses as the number of recovered individuals grows. The recovered curve follows a characteristic sigmoid shape, eventually plateauing as the majority of the population gains immunity. This dynamic illustrates the emergence of *herd immunity* within the model, which effectively starves the epidemic of new hosts and breaks the chain of transmission. The outbreak is fully resolved by approximately day 60, at which point the number of infected agents returns to zero.

4.2 Model calibration and real-world benchmarking

To ground our simulation in realistic epidemiological parameters, we set the transmission rate to $\beta = 0.02$ and the recovery rate to $\gamma = 0.01$. These values yield a basic reproduction number (R_0), a critical indicator of a pathogen's intrinsic transmissibility, calculated as:

$$R_0 = \frac{\beta}{\gamma} = \frac{0.02}{0.01} = 2.0$$

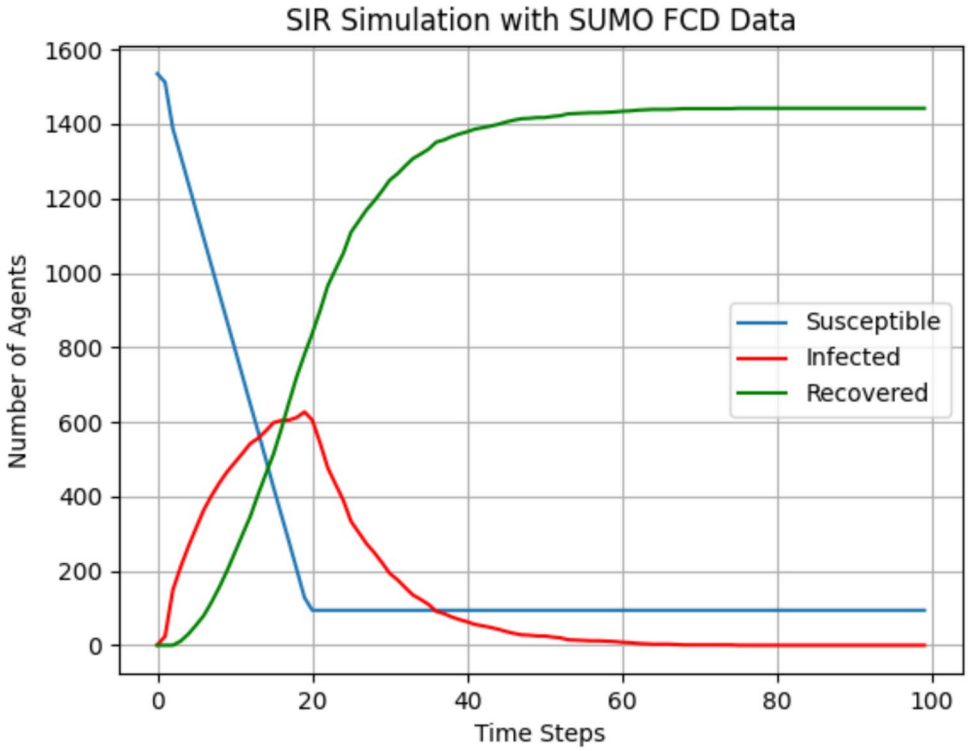

Fig. 6 Baseline Susceptible-Infected-Recovered (SIR) epidemic curve from the agent-based simulation. The graph illustrates the number of agents in each state over a 100-day period, with transmission dynamics driven by mobility data from SUMO's FCD output

We validated this parameterization by benchmarking our model's R_0 against empirical data. Public health reports from the initial, pre-intervention phase of the COVID-19 outbreak in Las Vegas (Spring 2020) estimated the local R_0 to be in the range of 2.2-−2.8.

While our framework models a generic respiratory pathogen, the close alignment of our simulated $R_0 = 2.0$ with these real-world figures provides strong validation for our model's core assumptions. It confirms that our simulation is calibrated to produce plausible epidemic dynamics consistent with a highly contagious disease spreading through a dense, tourism-driven urban landscape.

4.3 Residual susceptibility and mobility factors

A key insight from our simulation, visible in Fig. 6, is the persistence of a small but stable group of uninfected individuals. This is evidenced by the susceptible (S) curve plateauing at a non-zero value rather than dropping to zero. This outcome highlights a realistic feature of network-based epidemics that simpler, fully-mixed models often miss: transmission is not guaranteed to reach every corner of a population.

Our analysis attributes this residual susceptibility directly to *heterogeneous mobility patterns*. The agents who remained uninfected were not random; their "immunity by circumstance" was a direct product of their movement through the simulated environment. These agents typically exhibited routines that either:

- *Geographically isolated them* in low-traffic, peripheral zones of the network, far from transmission hotspots like the Las Vegas Strip.
- *Minimized spatiotemporal overlap* with contagious individuals, meaning their daily paths did not intersect with infected agents at the same time and place.

This finding is crucial because it demonstrates that an individual's infection risk is profoundly shaped by their unique mobility footprint. It underscores the model's capacity to reveal not only how diseases spread widely but also *why* certain segments of a population might be naturally shielded. This provides a granular foundation for designing and evaluating targeted, mobility-based interventions that can protect vulnerable populations without requiring universal restrictions.

4.4 Population-specific vulnerabilities and roles

The model provides a granular view of how different demographic groups contribute uniquely to the epidemic network. Each population segment plays a distinct role:

- *Tourists* act as the primary vectors for disease importation and long-range dissemination. Their high mobility across various hotels, casinos, and attractions ensures

rapid and widespread seeding of the virus across the city's most active hubs.

- *Hospitality Workers* function as critical *bridging nodes*. By interacting with a high volume of both tourists and local colleagues, they facilitate transmission between otherwise disconnected subpopulations, weaving disparate infection clusters into a single, cohesive epidemic.
- *Residents* sustain local, *community-level transmission*. Their infection patterns are largely governed by residential density and routine travel to workplaces and commercial centers, creating a baseline of infection that persists outside the main tourist corridors.

4.5 The critical impact of the initial seeding event

The overall trajectory of the outbreak is exceptionally sensitive to the timing and location of the initial infection. Our scenario analysis demonstrates that:

- An outbreak seeded in a *high-connectivity nexus*, such as Harry Reid International Airport or a major convention center, leads to an explosive and difficult-to-control epidemic. A single case in such a location gains immediate access to a large, highly mobile, and diverse population, maximizing its potential for rapid spread.
- Conversely, an initial infection occurring in a *low-density, residential area* results in a much slower, more contained initial growth phase. This provides a critical, albeit potentially brief, window of opportunity for public health interventions to be effective.

This finding quantifies the immense strategic value of early detection and rapid response measures, particularly surveillance focused on major travel hubs and mass gathering venues.

4.6 Scenario-based sensitivity analysis: public transit restriction

To demonstrate the framework's utility for policy evaluation, we simulated several "what-if" scenarios. This section details a high-impact intervention: the complete shutdown of public transportation, a key mode of transit known to concentrate individuals and increase transmission risk. In the simulation, we removed all public transit routes, forcing agents to either cancel trips or use alternative modes, thereby reducing overall mobility and contact density.

The results of this intervention are striking. Figure 7 shows a dramatically altered epidemic trajectory when compared to the uncontrolled baseline (Fig. 6). The impact is threefold:

- *Reduced infection peak:* The maximum number of concurrent infections is significantly suppressed, which

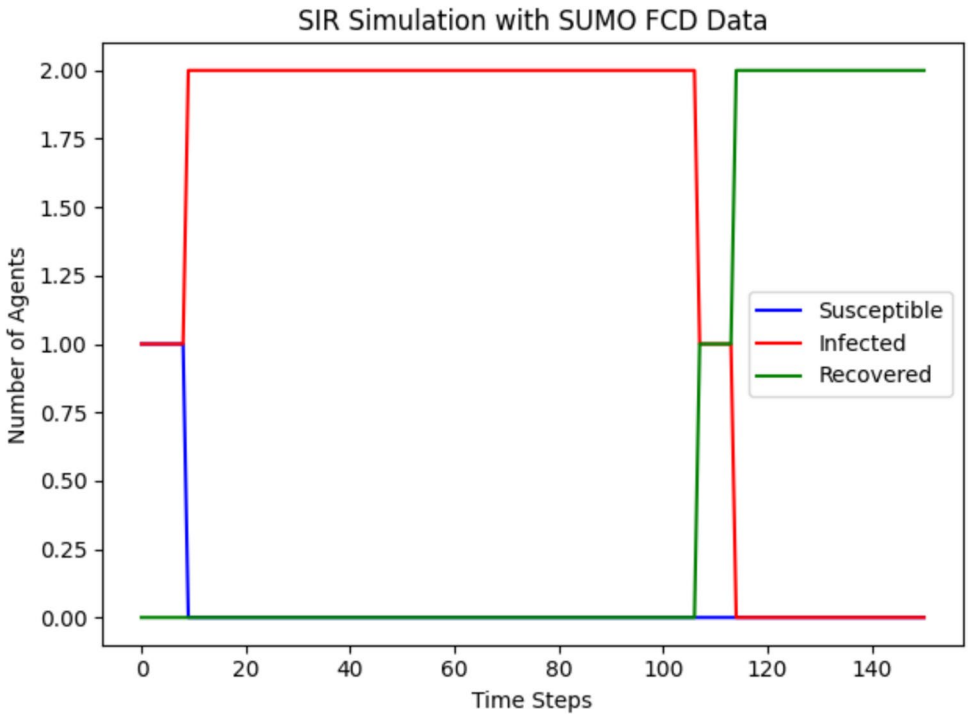

Fig. 7 SIR curve for the public transport restriction scenario. Compared to the baseline in Fig. 6, this intervention significantly flattens the infection curve (red), shortens the outbreak's duration, and preserves a larger susceptible population (blue). Time step = 1 day

would alleviate pressure on healthcare systems in a real-world scenario.

- *Faster epidemic resolution:* The outbreak is contained more rapidly, with the number of active cases returning to near-zero much earlier.
- *Lower attack rate:* A substantially larger portion of the population remains in the susceptible category, avoiding infection entirely.

This scenario quantitatively demonstrates how disrupting key mobility networks can starve a virus of transmission pathways. By limiting the opportunities for dense co-location at bus stops and within vehicles, the intervention effectively flattens the curve. While a full shutdown is an extreme measure, this analysis provides policymakers with a clear model of the public health gains that can be achieved through targeted mobility restrictions.

4.7 Identifying and characterizing transmission hotspots

A primary outcome of our simulation is the identification of specific geographic locations that function as *spatiotemporal hotspots*, driving a disproportionate amount of disease transmission. Rather than spreading uniformly, the epidemic's trajectory is dictated by these key amplifiers. By algorithmically analyzing the density and duration of high-risk contact events, our framework pinpoints the top ten epicenters of spread, as visualized in Fig. 8.

These hotspots are overwhelmingly concentrated in the city's economic and transit hearts: the Las Vegas Strip, Harry Reid International Airport, and the downtown corridor. The model reveals that these locations become transmission engines due to a potent combination of factors:

- *High agent density:* A large volume of individuals occupying a limited physical space.
- *Prolonged indoor co-location:* Significant agent dwell times within environments like casino floors, restaurants, and convention halls.
- *High population churn:* A constant influx and outflow of different individuals, maximizing the number of unique interactions.

Furthermore, this analysis captures the dynamics of *super-spreader events*. A large concert or convention hosted within one of these hotspots not only creates an intense local outbreak but also acts as a distribution hub, exporting cases across the wider community as attendees disperse. This granular, location-based risk assessment provides an

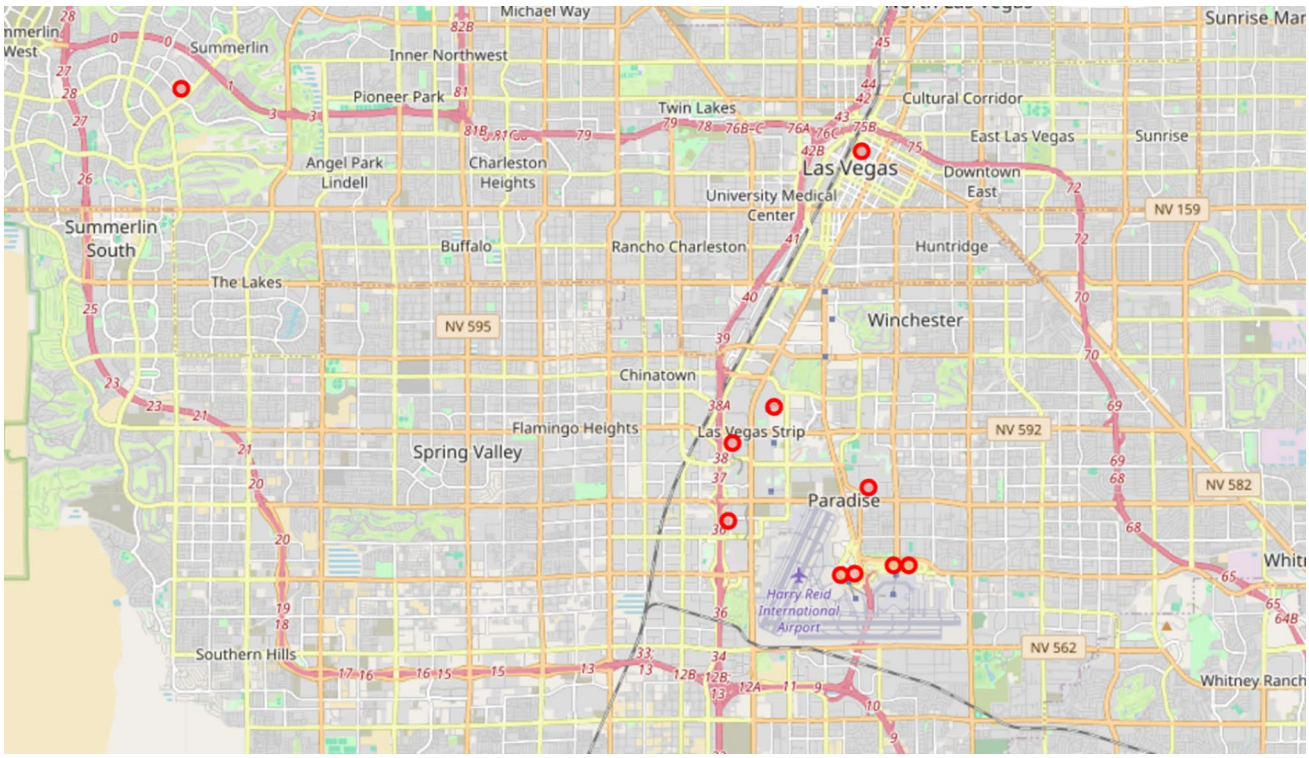

Fig. 8 Top ten transmission hotspots in Las Vegas identified by the agent-based model. Red circles indicate zones with the highest cumulative contact risk, algorithmically detected based on agent co-location density and duration. These areas represent critical nodes for disease amplification during an outbreak

empirical foundation for deploying targeted public health resources-such as enhanced surveillance or localized restrictions-precisely where they will have the greatest impact.

4.8 Assessing the impact of non-pharmaceutical interventions (NPIs)

Beyond forecasting an outbreak's natural course, our framework serves as a virtual laboratory for testing the efficacy of various Non-Pharmaceutical Interventions (NPIs). While policies like mask mandates are critical, our model excels at simulating interventions that directly manipulate mobility and density.

To illustrate this, we analyzed a *capacity restriction scenario*. In this simulation, all major POIs identified as hotspots (e.g., casinos, event venues, large restaurants) were limited to 75% of their normal occupancy. This was implemented by capping the number of agents allowed to enter a location at any given time, forcing others to alter their itineraries. The results were significant:

- Even this modest 25% reduction in capacity measurably slowed transmission by reducing the average number of close contacts per agent in high-risk indoor settings.
- The intervention's effectiveness was highly dependent on *timing*. Early implementation, before the exponential growth phase, not only lowered the peak number of infections but also delayed its arrival.

This "flattening of the curve" is a critical strategic outcome, as it provides invaluable time for healthcare systems to mobilize resources and avoid being overwhelmed. This analysis demonstrates the model's power to provide decision-makers with quantitative forecasts on the potential impact of specific, actionable public health policies, allowing for a data-driven approach to crisis management.

5 Discussion

Our integrated mobility and epidemiological model reveals that epidemic risk in a tourism-driven city like Las Vegas is not random but is systematically structured by the city's unique geography and the distinct mobility patterns of its subpopulations. The core finding is that a small number of *spatiotemporal hotspots*-overwhelmingly located along the Resort Corridor and at major transit hubs-drive a disproportionate share of transmission. These locations act as potent amplifiers only because *key population groups* (namely tourists and hospitality workers) create a constant flow of interactions within and between them.

The primary public health implication of this research is the strong validation of a *precision public health* approach. Instead of relying on universal, city-wide lockdowns that are socially and economically disruptive, our model provides an empirical basis for targeted, data-driven interventions. For example, our findings support strategies such as:

- Deploying rapid testing, enhanced ventilation, or prioritized vaccination resources specifically at the identified hotspot venues.
- Implementing dynamic capacity limits or mask mandates that are time-bound to large events or peak seasons, rather than being permanent.
- Focusing public health messaging on the hospitality workforce, recognizing their critical role as "bridging nodes" in the transmission network.

By revealing the underlying structure of risk, such models empower decision-makers to act with surgical precision, protecting public health while minimizing disruption to the lives and livelihoods that depend on a functioning urban center.

6 Conclusion

This paper presents a novel framework that bridges urban mobility modeling with *social network analysis*. Our primary contribution is a methodology that leverages agent-based mobility simulations to *mine and construct a dynamic, spatiotemporal contact network*. On this evolving network-where individuals are nodes and co-location events form weighted, time-varying edges-we can identify structural vulnerabilities for disease spread, moving beyond static or purely theoretical network models.

The limitations of this work point toward clear avenues for future research in network mining. Our current contact network is synthetically generated; the next generation of this research should focus on integrating real-world data streams (e.g., anonymized proximity data from mobile devices) to create an *empirically-grounded dynamic graph*. Furthermore, the SIR model represents a simple diffusion process; this could be extended to include more complex node states (e.g., asymptomatic carriers) or edge properties (e.g., varying transmission risk based on location type).

In conclusion, this study contributes to the field of social network analysis and mining by demonstrating a powerful application: using mobility data to generate otherwise unobservable contact networks for critical public health analysis. This approach-building a "digital twin" of a city's social fabric-offers a scalable, proactive paradigm for urban planning and crisis management. The ability to analyze how policy interventions reshape this contact network *in silico*

represents a critical step forward in building more resilient and healthier communities.

Acknowledgements This work is supported by a University of Nevada, Las Vegas (UNLV) Faculty Opportunity Award 2023. This material is also partially based upon work supported by the National Science Foundation under Grant Number 2323533.

Author contributions S.A. planned and designed the research. S.A. supervised the entire project. N.P. conducted the simulation, developed the web tool, and collected the simulation results. N.P. and S.A. interpreted the results and insights. B.L. verified the interpretations for domain relevance. N.P. and S.A. wrote the original draft of the manuscript. S.A. prepared the camera-ready version of the manuscript. All authors reviewed the manuscript.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

References

Adiga A, Agashe A, Arifuzzaman S, Barrett CL, Beckman RJ, Bisset KR, Chen J, Chungbaek Y, Eubank SG, Gupta S, et al (2015) Generating a synthetic population of the United States. Network Dynamics and Simulation Science Laboratory, Tech Rep NDSSL, 15–009

Arifuzzaman S, Pandey B (2019) Scalable mining, analysis and visualisation of protein-protein interaction networks. Int J Big Data Intell 6(3–4):176–187

Arifuzzaman S, Khan M, Marathe M (2019) Fast parallel algorithms for counting and listing triangles in big graphs. ACM Trans Knowl Discov Data (TKDD) 14(1):5–1534. https://doi.org/10.1145/3365676

Azad A, Buluc A, Pothen A (2017) Computing maximum cardinality matchings in parallel on bipartite graphs via tree-grafting. IEEE Trans Parallel Distrib Syst 28(1):44–59. https://doi.org/10.1109/TPDS.2016.2546258

Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A (2009) Multiscale mobility networks and the spatial spreading of infectious diseases. Proc Natl Acad Sci 106(51):21484–21489

Beamer S, Asanović K, Patterson D (2015) The gap benchmark suite. arXiv preprint arXiv:1508.03619

Burbano Lombana DA, Zino L, Butail S, Caroppo E, Jiang Z-P, Rizzo A, Porfiri M (2022) Activity-driven network modeling and control of the spread of two concurrent epidemic strains. Appl Netw Sci 7(1):66

Chakravarty UK, Arifuzzaman S (2024) Sentiment analysis of tweets on social security and medicare. Soc Netw Anal Min 14(1):91

Chang H, Li L, Huang J, Zhang Q, Chin K-S (2022) Tracking traffic congestion and accidents using social media data: a case study of Shanghai. Accid Anal Prev 169:106618

Charu V, Zeger S, Gog J, Bjørnstad ON, Kissler S, Simonsen L, Grenfell BT, Viboud C (2017) Human mobility and the spatial transmission of influenza in the United States. PLoS Comput Biol 13(2):1005382

Chatterjee A, Ramadurai G, Jagannathan K (2016) Contagion processes on urban bus networks in Indian cities. Complexity 21(S2):451–458

Craig BR, Phelan T, Siedlarek J-P, Steinberg J (2020) Improving epidemic modeling with networks. Econ Comment (2020-23)

Eclipse Foundation (2024) SUMO - Simulation of Urban MObility. Accessed: 2024-05-09. https://www.eclipse.org/sumo/

Eubank S, Guclu H, Anil Kumar V, Marathe MV, Srinivasan A, Toroczkai Z, Wang N (2004) Modelling disease outbreaks in realistic urban social networks. Nature 429(6988):180–184

Faysal MAM, Arifuzzaman S (2019) Distributed community detection in large networks using an information-theoretic approach. In: 2019 IEEE international conference on big data (big data), pp 4773–4782 . IEEE

Faysal MAM, Arifuzzaman S, Chan C, Bremer M, Popovici D, Shalf J (2021) Hypc-map: A hybrid parallel community detection algorithm using information-theoretic approach. In: 2021 IEEE high performance extreme computing conference (HPEC), pp 1–8. IEEE

Faysal MAM, Bremer M, Chan C, Shalf J, Arifuzzaman S (2023) Fast parallel index construction for efficient k-truss-based local community detection in large graphs. In: Proceedings of the 52nd international conference on parallel processing, pp 132–141

Funk S, Gilad E, Watkins C, Jansen VA (2009) The spread of awareness and its impact on epidemic outbreaks. Proc Natl Acad Sci 106(16):6872–6877

Han D, Sun M, Li D (2015) Epidemic process on activity-driven modular networks. XXPhys A 432:354–362

Kellermann R, Sivizaca Conde D, Robler D, Kliewer N, Dienel H-L (2022) Mobility in pandemic times: exploring changes and long-term effects of Covid-19 on urban mobility behavior. Transp Res Interdiscip Perspect 15:100668. https://doi.org/10.1016/j.trip.2022.100668

Kumar N, Oke J, Nahmias-Biran B (2021) Activity-based epidemic propagation and contact network scaling in auto-dependent metropolitan areas. Sci Rep 11(1):22665

Malik O, Gong B, Moussawi A, Korniss G, Szymanski BK (2022) Modelling epidemic spread in cities using public transportation as a proxy for generalized mobility trends. Sci Rep 12(1):6372

Marra AD, Sun L, Corman F (2022) The impact of Covid-19 pandemic on public transport usage and route choice: evidences from a long-term tracking study in urban area. Transp Policy 116:258–268

Nadini M, Zino L, Rizzo A, Porfiri M (2020) A multi-agent model to study epidemic spreading and vaccination strategies in an urban-like environment. Appl Netw Sci 5:1–30

Nyasulu JCY, Munthali RJ, Nyondo-Mipando AL, Pandya H, Nyirenda L, Nyasulu PS, Manda S (2021) Covid-19 pandemic in Malawi:

did public sociopolitical events gatherings contribute to its first-wave local transmission? Int J Infect Dis 106:269–275

OpenStreetMap contributors (2025) OpenStreetMap. https://www.openstreetmap.org. Accessed: August 14, 2025

Patil R, Dave R, Patel H, Shah VM, Chakrabarti D, Bhatia U (2021) Assessing the interplay between travel patterns and SARS-COV-2 outbreak in realistic urban setting. Appl Netw Sci 6:1–19

Pozzana I, Sun K, Perra N (2017) Epidemic spreading on activity-driven networks with attractiveness. Phys Rev E 96(4):042310

Qian X, Sun L, Ukkusuri SV (2021) Scaling of contact networks for epidemic spreading in urban transit systems. Sci Rep 11(1):4408

Rothengatter W, Zhang J, Hayashi Y, Nosach A, Wang K, Oum TH (2021) Pandemic waves and the time after Covid-19-consequences for the transport sector. Transp Policy 110:225–237

Ruan Z, Wang C, Ming Hui P, Liu Z (2015) Integrated travel network model for studying epidemics: interplay between journeys and epidemic. Sci Rep 5(1):11401

Salathé M, Jones JH (2010) Dynamics and control of diseases in networks with community structure. PLoS Comput Biol 6(4):1000736

Sattar NS, Arifuzzaman S (2022) Scalable distributed louvain algorithm for community detection in large graphs. J Supercomput 78 https://doi.org/10.1007/s11227-021-04224-2

Sattar NS, Buluc A, Ibrahim KZ, Arifuzzaman S (2023) Exploring temporal community evolution: algorithmic approaches and parallel optimization for dynamic community detection. Appl Netw Sci 8(1):64

Sattar NS, Ibrahim KZ, Buluc A, Arifuzzaman S (2025) DyG-DPCD: a distributed parallel community detection algorithm for large-scale dynamic graphs. Int J Parallel Prog 53(1):4

Schläpfer M, Bettencourt LM, Grauwin S, Raschke M, Claxton R, Smoreda Z, West GB, Ratti C (2014) The scaling of human interactions with city size. J R Soc Interface 11(98):20130789

Tizzoni M, Bajardi P, Decuyper A, King GKK, Schneider CM, Blondel V, Smoreda Z, González MC, Colizza V (2013) On the use of human mobility proxy for the modeling of epidemics. arXiv preprint arXiv:1309.7272

Topîrceanu A (2024) A spatial agent-based model for studying the effect of human mobility patterns on epidemic outbreaks in urban areas. Mathematics (2227-7390) 12(17)

Wang Z, Guo R, Zou L, Li T, Yao X (2024) Impact evaluation of Covid-19 on transit ridership: a case study of the Beijing subway. Urban Rail Transit, pp 1–20

Yang Y, Sugimoto JD, Halloran ME, Basta NE, Chao DL, Matrajt L, Potter G, Kenah E, Longini IM Jr (2009) The transmissibility and control of pandemic influenza a (H1N1) virus. Science 326(5953):729–733

Zhang J, Wang T (2023) Urban resilience under the Covid-19 pandemic: a quantitative assessment framework based on system dynamics. Cities 136:104265

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.