2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Spatial Modeling and Analysis of Human Traffic and Infectious Virus Spread in Community Networks

Siqi Zhang and Hui Yang*

***Abstract*—The use of network models to study the spread of infectious diseases is gaining increasing interests. They allow the flexibility to represent epidemic systems as networks of components with complex and interconnected structures. However, most of previous studies are based on networks of individuals as nodes and their social relationships (e.g., friendship, workplace connections) as links during the virus spread process. Notably, the transmission and spread of infectious viruses are more pertinent to human dynamics (e.g., their movements and interactions with others) in the spatial environment. This paper presents a novel network-based simulation model of human traffic and virus spread in community networks. We represent spatial points of interests (POI) as nodes where human subjects interact and perform activities, while edges connect these POIs to form a community network. Specifically, we derive the spatial network from the geographical information systems (GIS) data to provide a detailed representation of the underlying community network, on which human subjects perform activities and form traffics that impact the process of virus transmission and spread. The proposed framework is evaluated and validated in a community of university campus. Experimental results showed that the proposed simulation model is capable of describing interactive human activities at an individual level, as well as capturing the spread dynamics of infectious diseases. This framework can be extended to a wide variety of infectious diseases and shows strong potentials to aid the design of intervention policies for epidemic control.**

***Index Terms*—Epidemiology, network modeling, spatial network, human traffic, infectious disease, COVID-19**

I. INTRODUCTION

There is an increasing interest in network models for the characterization of virus spread in a pandemic. Because of the flexibility of network structures, nodes can be individuals, groups of individuals, and/or spatial locations at which individuals perform activities in a connected way. Links can be social contact patterns between individuals, interrelationships between groups of the population, and transportation connectivity among spatial locations.

Extensive studies have been conducted on network models that consider individuals as nodes to characterize the spread dynamics of infectious diseases. For example, Pastor-Satorras et al. [1] formulated an SIS spreading dynamics model in exponential and scale-free networks. In their model, infection occurs between a susceptible node and an infected node under a certain probability at each time-step. Gross et al.[2] proposed to study SIS dynamics in adaptive networks where the structure of connectivity between nodes dynamically responds to the change of human behaviors over time. Wu et al. [3] investigated the influence of the community structure in social networks on the dynamical process. A major drawback of these works is the lack of human mobility patterns in infectious disease modeling.

However, the pattern of human mobility is a critical factor of epidemic diffusion. From the Google and Apple mobility data, Cot et al. [4] observed a strong decreasing trend of COVID-19 infection rate occurring two to five weeks after the onset of human mobility reduction in Europe and in the United States. Zhang et al. [5] showed that changes in contact patterns are significantly shaping the COVID-19 spread dynamics in China. Hence, a number of frameworks have been developed to incorporate human mobility patterns in epidemic models. For example, the metapopulation method proposed by Watts [6] gained a lot of popularity as it allows the spatial migration of sub-population groups to be modeled in system dynamics models. This framework operates in a hierarchical level, in which local contexts assume homogeneous mixing of sub-population groups and are then embedded in a nested level. Note that this model sticks to the assumption of aggregated human behaviors. In other words, the heterogeneity of human activities tends to be underestimated in the spatial environment. There is an urgent need to develop network-based models that are capable of providing a detailed representation of human activities in order to capture the complexity and heterogeneity of epidemic dynamics in the community.

Daily activities of residents in the community are often interconnected through a number of key spatial points of interests (POI) (e.g., schools, grocery stores, shopping malls). The spatial data are readily available from geographical information systems (GIS) providers (i.e., US Census Bureau [7]). However, little has been done to extract the network structure from spatial data to model movements of people living in the community and capture spatiotemporal dynamics pertinent to the transmission and spread of infectious diseases. In this paper, we aim to fill the gaps and develop a novel network-based framework for the characterization of human behaviors and epidemic dynamics in community networks. Our contributions are summarized as follows:

- We derive the community network from the geographical information systems (GIS) data.
- We leverage the derived network to model human activities and infectious disease transmission at micro-level details during the virus spread process.

---

*Corresponding author: Hui Yang (email: huy25@psu.edu)
The authors are with Complex Systems Monitoring, Modeling and Control Lab, The Pennsylvania State University, University Park, PA 16802 USA

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

II. RESEARCH METHODOLOGY

In this section, we first discuss how to derive the spatial network of university communities from the spatial data. Then, we provide details about the human movement model in the spatial network. Finally, we discuss the infection model to characterize the spatiotemporal dynamics of the virus spread in community networks.

A. *Spatial Modeling of Community Networks*

Many real-world systems can be described by network models with a large number of nodes that are connected by edges/links. In a small-scale community, residents often travel through a list of key spatial locations that are connected by roads. However, little has been done to extract the network structure from roadway data to model individuals' activities and capture the dynamics of infectious disease in community networks of a spatial region.

In this investigation, we consider OpenStreetMap [8] as the main data source to derive the spatial network for the community of interest. The OpenStreetMap uses 2005 TIGER/Line roads as fundamental data source and enriches its platform through additional road features and attributes, such as pedestrian paths through parks, bike lanes, and trails. With the specified place name, OpenStreetMap's Overpass API can provide a detailed road map for the spatial area of interest. The place name can be an area that is within certain distances of an address, a university, and a city. Then, the road network can be constructed as follows.

Let $G(V, E)$ be an undirected graph for the community network, where V is the set of road junctions as nodes and E is the set of relations between two junctions as edges [9], [10]. In the downloaded road map, it is common that roads with long distances are segmented into several parts even when there are no junctions of multiple routes. In order to simplify the road network, these breakpoints are not considered. Then, through the search of the entire road map and the exclusion of these breakpoints, all nodes can be identified and spatial coordinates (x_i, y_i) for any given node i will be added as node attributes. An edge will be linked between two nodes when there is a road between them and the attributes of this road is also linked to the edge, such as the road id, the road type, the road length, the road geometry.

B. *Human Movement Model*

In real-world practices, the movement of human activities tends to form traffic flows in a connected way [11]. Hence, we design a human movement model to simulate detailed behaviors of human activities and interactions as human traffics in the spatial network. Note that most of existing studies assume that individuals have a fixed list of activities in certain spatial locations (e.g,, 2 hours in commuting and 8 hours in either workplaces or shopping malls). Instead, we leverage the structure of the spatial network to simulate individuals' movements in the spatial environment.

Figure 1 provides an illustration of the movement of a human subject i in the spatial network. Note that this network is the community of Penn State University Park campus. Because human movement patterns are highly heterogeneous, the population of the community is divided into different activity groups (e.g, high, medium, and low), which corresponds to the number of nodes they are going to visit. In other words, individuals with a lower level of activity tend to visit fewer nodes than those in higher groups on a daily basis. The blue line in Figure 1 shows the planned path for a human subject i, The planned path consists of a starting node, a destination node, and the route. The starting node is determined by the network-oriented sampling from nodes of the spatial network. Then, the destination is chosen based on the activity category and other factors (i.e. workplace settings). Also, we assume that individuals follow the shortest path to plan the route that can be determined by Dijkstra's algorithm [12]. During the simulation, individuals plan the route daily. The movement of human activities is scheduled in a day of 24 hours as follows, the daily activity is sparse before 8 AM, becomes busier at 8 AM, and slows down after 9 PM. The number of individuals actively moving and interacting in the spatial network is time-dependent and at each time slot, new individuals become active and join the human activity network. Because the daily activity slows down after 9 PM, no new human subjects will be added.

C. *Disease Transmission In the Community Network*

When individuals move and interact with each other, the virus quickly spreads in the community. As shown in the left dash circle of Figure 2, when susceptible individuals and virus carriers share the same environment, infection occurs. In this paper, we assume that infection mainly occurs in the nodes but rarely between nodes of the spatial network. The infection probability $f_{i,j}$ of an individual i getting infected by a virus carrier j depends on a number of factors, which can be expressed as

$$f_{i,j}(x_{i,j}, s_i, \rho) = \frac{1}{[1 + \rho e^{-(1-s_i)x_{i,j}}]^{1/\rho}}, 0 < s_i, \rho < 1 \quad (1)$$

where ρ is the disease transmissibility that describes how likely an individual gets infected by a disease on average, s_i is the susceptibility of an individual that is related to different risk factors (e.g., age, gender, and comorbidity) and $x_{i,j}$ is the exposure risk of a susceptible individual i from a surrounding virus carrier j that maps a virus carrier's capability to infect susceptible individuals to the risk value. We assume that it can be modeled with a logit function [13] with the infectivity level r_j of a surrounding virus carrier, that is

$$x_{i,j} = log\frac{r_j}{1 - r_j}, 0 < r_j < 1 \quad (2)$$

Some virus carriers, also called "super spreaders", may have higher infectivity levels than others. In addition, larger gathering events and imported cases are other major concerns for the virus spread process in the community. Based on either the geographical data or network structure, spatial entities/nodes of the network are categorized into different gathering levels. When a virus carrier and a susceptible

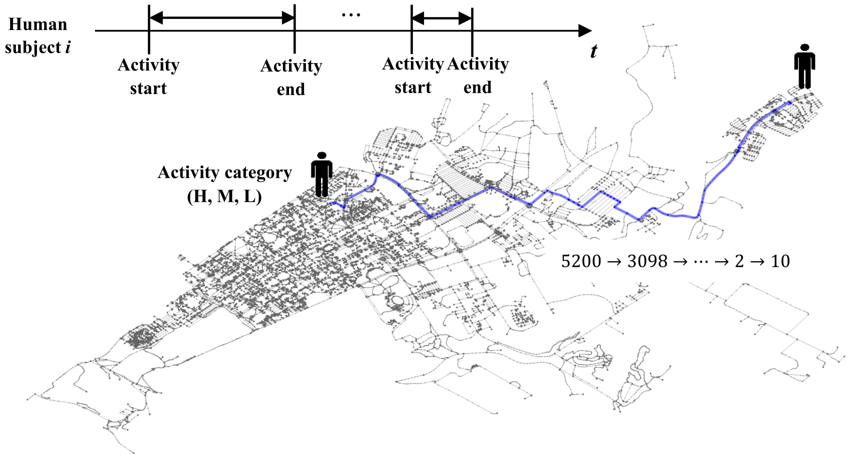

Fig. 1. Human activity and traffic in the spatial network (Note: H - high ; M - medium ; L - low.)

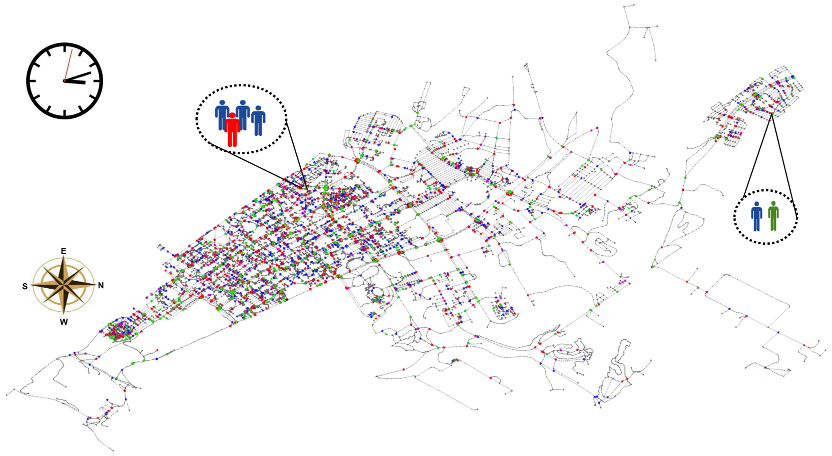

Fig. 2. An illustration of virus transmission and spread process in the community network (Youtube demo link: https://youtu.be/mKC_TXry3Lc)

individual come across a node of high gathering levels, the infection probability will be adjusted to a higher value. In addition, several nodes from the spatial network are considered as travel-related nodes to model imported cases in the spatial region. When individuals visit these nodes during their daily activities, they have a certain probability of getting infected from just visiting. On the other hand, when individuals have protective measures (e.g, mask-wearing and social distance of 6-feet), their probability of getting infected with surrounding virus carriers will be adjusted to a lower value.

III. EXPERIMENTS AND RESULTS

In the simulation, we consider the community of Penn State University Park campus where a population of 10000 people live, move, and interact with each other. There are 7903 nodes/spatial entities and 20533 edges/roads in the community network. In comparison with random, small-world, and scale-free networks, the derived community network has a smaller value of the average degree and tends to have much longer distances of shortest paths between nodes. As such, the probability of individuals interacting with each other becomes smaller given the network of the same size (i.e, number of nodes). In addition, nodes/POIs of the community network are categorized into five groups to represent different levels of gathering events (e.g., no gathering, virtual-only activities, smaller outdoor and in-person gatherings, medium-sized in-person gatherings, and large gatherings). Then, 100 nodes out of those in no gathering groups are randomly select to model travel-related risks outside the community network. In other words, when individuals visit nodes pertinent to high-risk travel locations, their infection risks will be increased.

Five levels of gathering events are considered to quantify different gathering events (e.) which are quantified by visiting different POIs/nodes of the community network. When individuals visit nodes pertinent to large gathering events, their infection risks will be increased. Similarly, the risk factor of travel history can be simulated by randomly assigning 100 POIs from no-gathering groups to represent different levels of travel-related risks.

Because such a community consists of students, faculties and staffs, there are four age groups: (1) 18-35 (2) 36-45 (3) 46-60 (4) 60+ and each group of which equals 70%, 15%, 12% and 3% of the entire population, respectively. The total simulation time is 60 days with a time step of 1 hour. In the human activity network, individuals are categorized into five activity groups to characterize the heterogeneity of human activities in the community. In the disease transmission module, we assume that the percentage of symptomatic individuals is about 50% of the infected population and 3.4% of symptomatic individuals need to be hospitalized after symptom onset. The transition time between exposed state and symptom onset takes about 5 days and the time between symptom onset and recovered state is about 12 days [14]. The Poisson distribution is considered to model the transition between these states. When individuals have protective measures (i.e., wearing masks), the infection probability $f_{i,n}$ will be reduced by 86%.

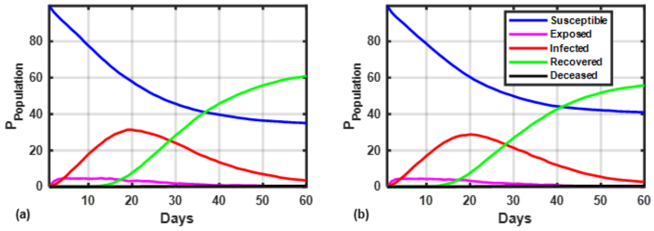

Fig. 3. Temporal characteristic curves of virus spread when (a) only 20% of the population and (b) more than 80% of the population in the community have protective measures

Figure 3(a) and 3(b) respectively shows the time evolution of infectious disease spread when only a few individuals (e.g, 20% of the population wears a mask for sure) and the majority of the population (e.g., over 80% of the population) take protective measures. The red curve in both plots reach the peak value on the 20th day but there is a decreasing value of 2.56% infection cases at this moment, which is equivalent to 256 out of 10000 people in the spatial community. Because the virus spread process is slowing down due to non-pharmaceutical interventions (e.g., masks, hygiene and social distancing) from more people living in the community, total cases (%) are reduced from 64.97% to 59.06% at the end of 60 days. In other words, protective measures are critical to mitigate the virus spread, reduce the total number of infections, and thereby save more lives in the community.

IV. CONCLUSIONS

Network models provide valuable frameworks to study the dynamical process of infectious disease spread in the human population. In the past, several research works have been done to understand the spread dynamics of infectious diseases in complex networks of which individuals are represented as nodes and their social connections as edges. However, little has been done to consider epidemic systems as networks of spatial entities to which individuals are connected and over which individuals dynamically move and interact with each other in the community. This paper presents a novel framework to characterize the network spread of infectious diseases in community networks. First, we derive the network structure of the community from the spatial data. Then, we design a human movement model to provide a detailed representation of networked human activities. Finally, we propose an infection model to capture the spread dynamics of infectious diseases in the community network. Experimental results show that the proposed framework is not only capable of providing detailed behaviors of individuals in the community network, but also characterizing the spread dynamics in the community network. The proposed simulation models can be further utilized for the design of clinical testing programs and the investigation of traffic control in the spatial network in order to mitigate the virus spread process.

ACKNOWLEDGMENT

The authors would like to acknowledge the NSF I/UCRC Center for Healthcare Organization Transformation (CHOT), NSF I/UCRC award IIP-1624727, and NSF RAPID grant IIP-2026875 for funding this research.

DISCLAIMER

This paper doesn't involve experimental procedures on either human subjects or animal models, but rather use the data available in the public domain. Thus, approval is not needed from the Institutional Review Board.

REFERENCES

[1] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics and endemic states in complex networks," *Physical Review E*, vol. 63, no. 6, p. 066117, 2001.

[2] T. Gross, C. J. D. D'Lima, and B. Blasius, "Epidemic dynamics on an adaptive network," *Physical review letters*, vol. 96, no. 20, p. 208701, 2006.

[3] X. Wu and Z. Liu, "How community structure influences epidemic spread in social networks," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 2-3, pp. 623–630, 2008.

[4] C. Cot, G. Cacciapaglia, and F. Sannino, "Mining google and apple mobility data: temporal anatomy for covid-19 social distancing," *Scientific reports*, vol. 11, no. 1, pp. 1–8, 2021.

[5] J. Zhang, M. Litvinova, Y. Liang, Y. Wang, W. Wang, S. Zhao, Q. Wu, S. Merler, C. Viboud, A. Vespignani *et al.*, "Changes in contact patterns shape the dynamics of the covid-19 outbreak in china," *Science*, vol. 368, no. 6498, pp. 1481–1486, 2020.

[6] D. J. Watts, R. Muhamad, D. C. Medina, and P. S. Dodds, "Multiscale, resurgent epidemics in a hierarchical metapopulation model," *Proceedings of the National Academy of Sciences*, vol. 102, no. 32, pp. 11 157–11 162, 2005.

[7] U. C. Bureau, "2015 tiger/line shapefiles," 2015.

[8] M. Haklay and P. Weber, "Openstreetmap: User-generated street maps," *IEEE Pervasive computing*, vol. 7, no. 4, pp. 12–18, 2008.

[9] H. Yang and G. Liu, "Self-organized topology of recurrence-based complex networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 23, no. 4, p. 043116, 2013.

[10] H. Yang, C. Kan, A. Krall, and D. Finke, "Network modeling and internet of things for smart and connected health systems—a case study for smart heart health monitoring and management," *IISE Transactions on Healthcare Systems Engineering*, vol. 10, no. 3, pp. 159–171, 2020.

[11] H. Yang, S. Zhang, R. Liu, A. Krall, Y. Wang, M. Ventura, and C. Deflitch, "Epidemic informatics and control: A holistic approach from system informatics to epidemic response and risk management in public health," in *AI and Analytics for Public Health - Proceedings of the 2020 INFORMS International Conference on Service Science*. Springer, 2021, ch. 1, pp. 1–46.

[12] F. B. Zhan and C. E. Noon, "Shortest path algorithms: an evaluation using real road networks," *Transportation science*, vol. 32, no. 1, pp. 65–73, 1998.

[13] H. Yang and E. K. Lee, *Healthcare analytics: from data to knowledge to healthcare improvement*. John Wiley & Sons, 2016.

[14] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R. L. Hesketh, L. Yang *et al.*, "Time course of lung changes on chest ct during recovery from 2019 novel coronavirus (covid-19) pneumonia," *Radiology*, 2020.