A Mobility-based Recommendation System for Mitigating the Risk of Infection during Epidemics

Gian Alix[§], Nina Yanin[§], Tilemachos Pechlivanoglou, Jing Li, Farzaneh Heidari and Manos Papagelis
Lassonde School of Engineering, York University, Toronto, Canada
{gcalix, tipech, jliellen, farzanah, papaggel}@eecs.yorku.ca, nina27@my.yorku.ca

***Abstract*—The relationship between human mobility and the spread of an infectious disease has been well documented. At the same time, availability of mobility data is growing due to advancements in digital contact tracing mobile applications and GPS-enabled devices. Motivated by these observations, we have designed and developed STRIPE (Safe Trips during Epidemics), a mobility-based recommendation system that can provide safer trip recommendations to individuals. The recommendation model considers the risk of infection of alternative trips between an origin and destination. It also considers the risk of infection of specific points of interests (POIs) that occur at the microscale. In this paper, we present a high-level architecture of the system, its main features and system use cases. The broader impact of our research is that by helping individuals making informed decisions, we promote more responsible behaviors in the community as a whole that could effectively alleviate the impact of the epidemic.**

***Index Terms*—mobility data, recommendation systems, risk maps, safe trips, epidemics**

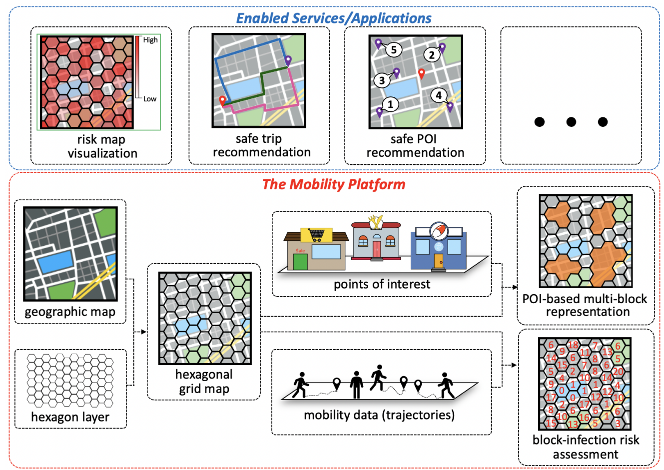

Fig. 1. High-level system overview of STRIPE.

I. INTRODUCTION

Motivation. The COVID-19 pandemic has led to a devastating social and economic disruption, ranging from a dramatic loss of human lives to unprecedented challenges to public health systems[1]. At the same time, the epidemic outbreak has been met by an unprecedented response by experts who simultaneously focused on pharmaceutical and non-pharmaceutical interventions to control the spread of the epidemic. Our work is motivated by recent advanced technological responses to the problem based on digital contact tracing that have claimed some success in controlling the epidemic [1].

Our approach. The focus of the current research is on utilization of GPS-enabled digital traces of individuals (i.e., mobility data or trajectories) to inform a more comprehensive analysis and modeling of disease spreading through methods of trajectory data mining. These methods can enable evidence-based data-driven models to support decision-making and inform public policy of targeted non-pharmaceutical mobility-related interventions [2]–[4]. In fact, it has been shown that human mobility has considerable impact on the spread of infectious diseases [5]. However, models have mostly focused on higher-level mobility patterns, such as flight connections and inter-city or inter-neighborhood mobility. Here, we focus on developing *microscopic models* to assess the risk of infection of individuals based on their mobility behavior and evaluate personalized strategies for mitigating that risk.

Contributions. Towards that end, we designed and developed a system that can assess the risk of infection based on mobility patterns. In particular, we present STRIPE (Safe Trips during Epidemics), a web-based proof-of-concept prototype that can provide several (alternative) risk-based personalized trip recommendations in a specified geographic area. The novelty of our data-driven recommendation system is multifold:

- it provides risk (of infection) assessment of a geographic area based on mobility dynamics;
- it provides risk (of infection) assessment of points of interests (POIs) that occur at the microscale;
- it incorporates the heterogeneity of the mobility patterns of individuals to provide safe trip recommendations.

Our work is based on the assumption that real-time mobility data of individuals is available, which can be enabled by mobile applications and location-based services [2].

Organization. The remainder of the paper is organized as follows. Section II provides a high-level overview of the system. Section III describes the user interface, and provides information on the demo scenarios. The related work is discussed in Section IV. We conclude in Section V.

This research has been partially supported by an NSERC Discovery Grant (#RGPIN-2017-05680) and by NSERC's Emerging Infectious Diseases Modelling (EIDM) initiative (#RGPID-560520-2020). This research is also partially supported by the Vector Scholarship in Artificial Intelligence, provided through the Vector Institute.

[§]Authors contributed equally.

[1]WHO COVID-19 Dashboard – https://covid19.who.int

II. System Overview

A high-level overview of our system can be seen in Fig. 1, where the lower part represents *a mobility platform* and the upper part represents *services/applications* enabled by the platform that have been developed. In this section, we further elaborate on these components of our recommendation system.

A. The Mobility Platform

Hexagonal grid map. Our STRIPE system is operating on top of an arbitrary geographic map that has been tessellated by a disjoint finite set of blocks $\mathcal{B} = \{b_i\}$ in the shape of *regular hexagons*. Hexagons are the most "circularly-shaped" polygons that are able to tessellate and form an evenly-spaced grid; they are also able to reduce sampling bias that is due to the edge effects of a grid map's shape. This is mostly due to the *circularity* of hexagons that enables a natural representation of curvatures in trajectory data, and as such is preferable over square or triangular grid tessellation when analyzing various properties of mobility data (e.g., movement paths, connectivity, etc.) [6]. Fig. 1 (leftmost part of the mobility platform layer) illustrates how a hexagonal grid overlays an arbitrary input map for the study of epidemics. The resultant *hexagonal grid map* can then be used in two ways: (*i*) as an input map where we seek to evaluate the risk of infection associated with each hexagon block, and (*ii*) as an input map providing the primitive units (hexagons) for forming hierarchical multi-blocks by grouping together contiguous hexagons.

Block infection risk assessment. We have developed methods which, given as input the hexagonal grid map of a geographic area and pedestrian mobility data (trajectories) on the same area for an observation period $[0, T]$, can assess the risk of infection of each hexagon block in the map as a function of time. The risk of infection of a hexagon $b \in \mathcal{B}$ at a time $t \in [0, T]$ depends on the number of individuals n_b^t that are found in block b at time t. Fig. 1 (lower right corner) shows an illustrative example where the value n_b^t of each block for a specific time is depicted in red. More specifically, the risk of infection in a single hexagon is related to the number of possible contacts with other individuals. If the size of the hexagon is small enough (as is the case in our epidemic-related analysis), then the potential number of *contacts* c_b^t (i.e., pairs of individuals) in a block b at time t is given by $c_b^t = n_b^t(n_b^t - 1)/2$. The larger the number of individuals found in a specific block at a specific time, the more the potential contacts, and thus the higher the infection risk in that block.

POI representation as multi-blocks. A *point of interest* (POI) in our context is any building that is visited often by individuals, such as grocery stores, restaurants, pharmacies, and so on. Given a set of POIs $\mathcal{P} = \{p_i\}$ within the specified geographic area, along with the tessellated hexagonal grid map (see upper right corner of the mobility platform of Fig. 1), we can form *multi-block* entities by grouping together several contiguous blocks that span the area of a POI. Note that some hexagons may not be part of any POI. Fig. 2 provides a more detailed example of the hierarchical multi-block representation

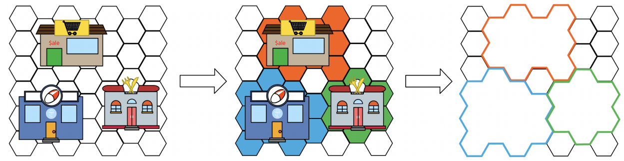

Fig. 2. Hierarchical multi-block representation of POIs.

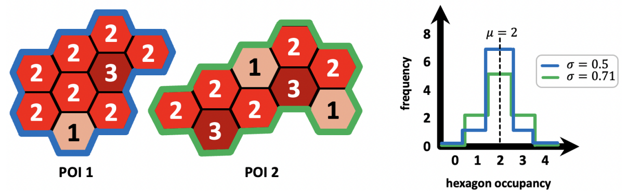

Fig. 3. An example of two POIs that span the same number of blocks (9 hexes) and have same occupancy (18 individuals), but exhibit different risk due to different statistical dispersion.

of POIs. In theory, this multi-block grouping scheme can be further extended to provide a more comprehensive hierarchy of entities that one can study. For instance, multiple POIs can belong to the same mall or city neighborhood, and so on.

POI infection risk assessment. The main purpose for representing POIs as multi-blocks is to be able to analyze the risk of infection dynamics within these higher-level entities (or the *POI risk*) based on some sensible aggregation of the risk of infection of block units (hexagons). Note however that the distribution of individuals to the blocks that make up a multi-block can have a significant impact on a POI's risk of infection. To better conceptualize this, imagine two POIs, say a grocery store and an open-concept mall, comprising the same number of hexagon blocks and having the same number of visitors. In the grocery store people are getting through narrow or restricted corridors, while in the open-concept mall people move around more freely. As the risk associated to a block b at time t has a quadratic relationship with the number of individuals n_b^t in it (as $c_b^t = n_b^t(n_b^t - 1)/2$), it can be concluded that these two POIs will be associated with a different risk of infection. Specifically, a higher risk of infection will be attributed to those POIs where more individuals are concentrated in fewer blocks, for the same amount of time. Fig. 3 shows an illustrating example, where POIs 1 and 2 are represented by the same number of blocks and they have the same total occupancy at time t. However, due to the different distribution of individuals in blocks, they are associated with a different risk of infection. More formally, we model this phenomenon, using the concept of *statistical dispersion* metric of individuals in POIs – in which the *standard deviation* is deemed as the natural choice of measure [7]. In particular, our system can evaluate the risk of infection of a specific POI for varying values of the standard deviation $\sigma = \{0.5, 1.0, 2.0, 3.0, \infty\}$ – where ∞ represents a uniform dispersion. Note that the uniform dispersion is what is usually captured by studies that evaluate risk based on building occupancy alone, without considering the distribution of individuals within the building. Our system provides quantitative evidence that POIs that lead to dense concentrations of individuals (i.e. lower values of dispersion) are associated with a higher risk of infection.

B. Enabled Services/Applications

Given the functionality provided by the mobility platform component of our STRIPE system, several enabled services and/or applications open up for use. Below, we elaborate on three important ones that have already been deployed.

Risk map visualization. This service provides a heatmap visual of the infection risk in a specific geographic area. Formally, given a specified map and mobility data defined on it, each block's risk is computed. Risk values then inform a heatmap, where the more red a hex is, the higher the risk.

Safe trip recommendation. This service provides personalized safe trip recommendations. Formally, given a query $q(s, d, t_{depart})$, where s and d are source and destination locations, and t_{depart} is the departure time, it recommends several safe alternative trips $\mathcal{T} = \{\tau_1, \tau_2, \tau_3, ...\}$ that minimizes a trip's infection risk. As shown in the services and applications layer of Fig. 1, safe (i.e., low risk) trips from the origin (red marker) to the destination (purple marker) can be suggested (blue, green and pink paths in the example).

Safe POI recommendation. This service provides personalized safe POI recommendations. Formally, given a query $q(s, r, k, t_{depart})$, where s is the source location, r is the radius distance around s, and t_{depart} is the departure time, it returns a ranked list of the top k safest POIs $\mathcal{P} = \{p_1, p_2, ..., p_k\}$ that minimizes the infection risk when visited. As illustrated in the services and applications layer of Fig. 1, safe POIs (i.e., low risk POIs) that are in the vicinity of the source (red marker) can be recommended. As our work does not intend on associating a specifically-named POI to a numerical risk value, we instead recommend safe POIs (purple markers) and display their rankings.

III. DEMONSTRATION

Two snapshots of the system's user interface are shown in Fig. 4 and Fig. 5. They are both comprised of *four panes*: (**A**) three tabs for choosing between service – *POI Recommender*, *Trip Recommender*, or *Searched Results*; (**B**) two tabs for choosing between third-party routing-services – *OSRM* or *GraphHopper*, which provide paths between two given points, along with their respective *distance* and *duration*; (**C**) based on the selection in (**A**), a customized form relevant to that query, and (**D**) visualization of the query result.

In the case of safe trip recommendation (Fig. 4), pane (**C**) depicts a form requesting information regarding the point of origin, destination and time of departure. Then, pane (D) provides the visual output, where the blue marker represents the origin, the red marker represents the destination, and possible trips are represented with semi-transparent blue; the suggested safest trip is represented with a bold blue color. The intensity of the red hexagons represent the relative risk of that area, where the more opaque the area is, the higher the risk.

In the case of safe POI recommendation (Fig. 5), pane (**C**) depicts a form requesting information regarding the origin, POI category (e.g., grocery store), number of results, distance from the origin (in km), sorting criteria (such as nearest, fastest

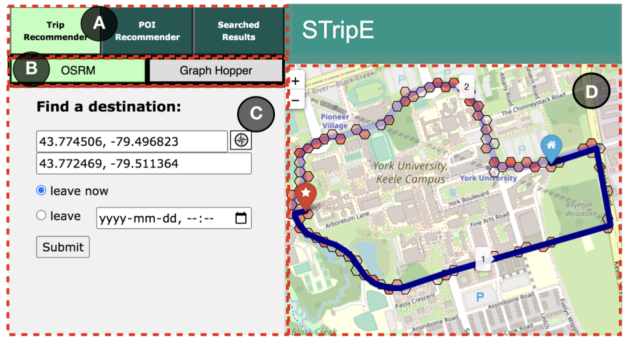

Fig. 4. A screenshot depicting safe trip recommendations.

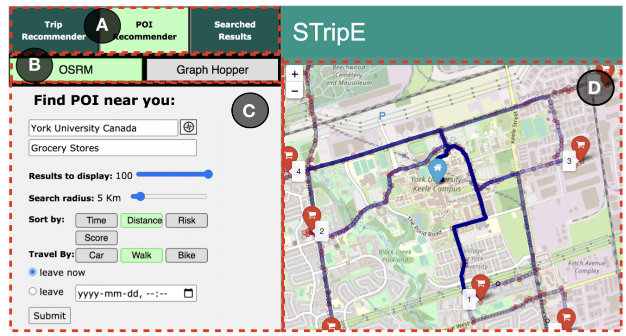

Fig. 5. A screenshot depicting safe POI recommendations.

or safest), mode of transit and time of departure. Then, pane (**D**) visualizes the results, where the blue marker is the origin and the red markers are the suggested POIs within r km from the origin. The red hexagons' intensity represents their relative risk, and the trips are represented with semi-transparent blue. The safest trip is represented with a bold blue color. The sorted results can be viewed in the *Searched Results* tab of pane (**A**).

Implementation details. Our work makes use of synthetic mobility traffic data of individuals generated by SUMO (Simulation of Urban MObility) [8], an open-source package that can simulate individual mobility. Routing information of trip recommendations rely on two third-party routing services, OSRM[2] and GraphHopper[3]. Our mobility platform uses visualization from Leaflet[4], an open-source JavaScript library that can display interactive visualizations of maps. Furthermore, our system makes use of OpenStreetMap[5] that contains free-to-use map data of the world under an ODbL license.

Demonstration scenarios. During the demonstration, end-users will be able to use our system, with preprocessed trajectories, risk map and POIs, in order to observe a risk map, and evaluate the risk associated with trips and POIs. In particular, they will be able to: (*i*) project the time-varying risk map on a pre-defined geographical area, (*ii*) query about safe trips from source to destination and be informed about their associated risk value, (*iii*) query about safest, nearest and fastest POIs under a varying set of constraints, such as place of origin, radius and departure time.

[2]http://project-osrm.org/
[3]https://www.graphhopper.com/
[4]https://leafletjs.com/
[5]https://www.openstreetmap.org/

IV. Related Work

Utilizing mobility data has been an entrancing research direction due to a plethora of useful applications [9]–[11]; as such we review some works relevant to this research direction.

Human mobility and epidemics. The relationship between mobility and the spread of infectious diseases has been well documented. For example, Hazarie et al. [12] studied 163 global cities and noted a positive correlation between COVID-19 infection risk and human mobility, where infection risk is higher in relatively denser places. Pechlivanoglou et al. [13] studied infectious disease spreading in mobility networks and suggested network targeted interventions that mitigate the spread without imposing strict mobility restrictions.

Personalized trip recommendation models. Providing personalized trip recommendations is an optimization problem that aims to maximize a trip-related objective, while satisfying various constraints. Closer to our work is a popular variation that seeks to provide POI recommendations. For instance, Debnath et al. [14] considered preference-aware context with temporal influence in travel route recommendation and location-based social networking problems. Qian et al. [15] utilized spatiotemporal context-aware recommendation frameworks to model third-order relations among users for large-scale POI recommendation.

Trajectory data mining. With a substantial amount of work done in the area, surveys such as [16], [17] compiled and summarized the significant contributions.

V. Conclusions & Discussion

Motivated by the global pandemic and the availability of mobility data becoming available through digital contact tracing devices and mobile applications, we have developed a mobility platform that encapsulates advanced computational methods to evaluate the risk of infection associated with (*i*) specific trips, and (*ii*) visiting specific POIs of a geographic area. Based on this mobility data analysis, we designed and developed an interactive web-based recommender service that allows users to submit queries and obtain vital information regarding the associated risk of specified origin-destination trips, or risk of visiting specific POIs in their vicinity. The broader impact of our work is that the system helps individuals to make informed decisions and promote responsible behavior in the community that can effectively control the epidemic.

Inherent system limitations. An inherent limitation of the system is that it assumes availability of mobility data. Although these data are available and can be obtained by third-parties, in this work we relied on realistic synthetic mobility data (trajectories). Another assumption is that information about POIs' occupancy levels over time is available. Again, these data can be easily obtained by third-party services, such as SafeGraph[6], but in this work we relied on simulated data to represent POIs' different levels of occupancy (as %) to ensure actual risk of visiting a specific POI is kept private.

[6]Places Data & Foot Traffic Insights – https://www.safegraph.com

Mobility platform extensibility. While our platform was designed with epidemic-specific services in mind, it is general and can thus be easily adopted to serve other domain-specific applications. For instance, the risk map could be informed by public data related to traffic incidents to inform about safe trips in road networks. Similarly, we can extend the safe POI recommendation model to other risk models, such as building code violations or neighborhood crime rates.

References

[1] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser, "Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing," *Science*, vol. 368, no. 6491, 2020.

[2] G. Cencetti, G. Santin, A. Longa, E. Pigani, A. Barrat, C. Cattuto, S. Lehmann, M. Salathé, and B. Lepri, "Digital proximity tracing on empirical contact networks for pandemic control," *Nature Communications*, vol. 12, no. 1, p. 1655, Dec 2021.

[3] A. Aleta, D. Martín-Corral, A. Pastore y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini Jr, S. Merler, A. Pentland, A. Vespignani, E. Moro, and Y. Moreno, "Modelling the impact of testing, contact tracing and household quarantine on second waves of covid-19," *Nature Human Behaviour*, vol. 4, no. 9, pp. 964–971, Sep 2020.

[4] S. Chang, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, and J. Leskovec, "Mobility network models of COVID-19 explain inequities and inform reopening," *Nature*, vol. 589, no. 7840, pp. 82–87, 2020.

[5] M. U. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, O. C.-. D. W. Group†, L. du Plessis, N. R. Faria, R. Li *et al.*, "The effect of human mobility and control measures on the covid-19 epidemic in china," *Science*, vol. 368, no. 6490, pp. 493–497, 2020.

[6] C. P. Birch, S. P. Oom, and J. A. Beecham, "Rectangular and hexagonal grids used for observation, experiment and simulation in ecology," *Ecological modelling*, vol. 206, no. 3-4, pp. 347–359, 2007.

[7] S. Manikandan, "Measures of dispersion," *Journal of Pharmacology & Pharmacotherapeautics*, vol. 2, no. 4, p. 315, 2011.

[8] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic traffic simulation using sumo," in *21st Int. Conf. on intelligent transportation systems (ITSC)*. IEEE, 2018, pp. 2575–2582.

[9] A. Sawas, A. Abuolaim, M. Afifi, and M. Papagelis, "Tensor methods for group pattern discovery of pedestrian trajectories," in *2018 19th IEEE International Conference on Mobile Data Management (MDM)*, 2018, pp. 76–85.

[10] T. Pechlivanoglou and M. Papagelis, "Fast and accurate mining of node importance in trajectory networks," in *2018 IEEE International Conference on Big Data (Big Data)*. Los Alamitos, CA, USA: IEEE Computer Society, Dec 2018, pp. 781–790.

[11] S. Mehmood and M. Papagelis, "Learning semantic relationships of geographical areas based on trajectories," in *2020 21st IEEE International Conference on Mobile Data Management (MDM)*, 2020, pp. 109–118.

[12] S. Hazarie, D. Soriano-Paños, A. Arenas, J. Gómez-Gardeñes, and G. Ghoshal, "Interplay between population density and mobility in determining the spread of epidemics in cities," *Communications Physics*, vol. 4, no. 1, pp. 1–10, 2021.

[13] T. Pechlivanoglou, J. Li, J. Sun, F. Heidari, and M. Papagelis, "Epidemic spreading in trajectory networks," *Big Data Research*, vol. 27, p. 100275, 2022.

[14] M. Debnath, P. K. Tripathi, A. K. Biswas, and R. Elmasri, "Preference aware travel route recommendation with temporal influence," in *Proc. of the 2nd ACM SIGSPATIAL/LBSN Workshop*, 2018, pp. 1–9.

[15] T. Qian, B. Liu, Q. V. H. Nguyen, and H. Yin, "Spatiotemporal representation learning for translation-based poi recommendation," *ACM Trans. Inf. Syst.*, vol. 37, no. 2, Jan 2019.

[16] G. Atluri, A. Karpatne, and V. Kumar, "Spatio-temporal data mining: A survey of problems and methods," *ACM Comput. Surv.*, vol. 51, no. 4, Aug 2018.

[17] A. Hamdi, K. Shaban, A. Erradi, A. Mohamed, S. K. Rumi, and F. D. Salim, "Spatiotemporal data mining: a survey on challenges and open problems," *Artificial Intelligence Review*, vol. 55, no. 2, p. 1441–1488, Feb 2022.