Effectiveness of Network Interdiction Strategies to Limit Contagion During a Pandemic

SATYAKI ROY
1, (Member, IEEE), PREETOM BISWAS
2, AND PREETAM GHOSH
3

1Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27515, USA
2School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85281, USA
3Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA

Corresponding author: Satyaki Roy (satyakir@unc.edu)

This work was supported in part by the National Science Foundation under Grant CBET-1802588.

ABSTRACT COVID-19 is an infectious disease that has been declared a global public health emergency by the World Health Organization. Besides claiming over 3 million lives worldwide, COVID-19 led to unprecedented disruption in industrial productivity, trading, and global food supply, resulting in loss of livelihood. Despite initial success in curbing the spread of diseases through a lockdown and rapid vaccine development, human lives are threatened by sudden outbreaks from new strains of the virus. This motivates the conceptualization of effective interdiction rules to inform human mobility in a manner that the damage to lives as well as the economy could be minimized. In this work, we present three interdiction rules that employ machine learning-based network inference on daily infected cases to infer the influence of contagion between neighboring zones. The proposed rules leverage network science concepts such as coloring and clustering to attain time-varying partial or complete travel restrictions. Through extensive simulation experiments, we show that these strategies yield lower infection spread than greedy and random migration-based tie elimination approaches as well as a balance between contagion mitigation and economic gain.

INDEX TERMS COVID-19, travel rule, machine learning, network science, graph coloring.

I. INTRODUCTION Coronavirus disease (COVID-19) is an infectious disease that causes mild to severe respiratory illness. It has caused 3.2 million deaths worldwide and led to a prolonged period of uncertainty in public health, livelihood and well-being [1]. In the absence of vaccination and drugs, the nations were exercising social distancing and lockdown to curb the transmission of the virus. The ongoing restrictions has caused a decline in industrial productivity, stock exchange percentage, hike in goods prices and contraction of global GDP. Border and trade restrictions prevent farmers and markets from trading, disrupting global food supply chains [2]. Recent successes in expedient vaccine developmental efforts have allowed nations to relax these distancing regulations [3].

Despite successes in vaccine development, clinicians, immunologists and epidemiologists have established that the war against COVID-19 is far from over. COVID-19 will continue to exist as a seasonal fixture and continue to claim lives much like the influenza and other coronavirus strains.

Our future depends on how quickly and effectively we immunize ourselves against the evolving strains of the virus [4]. This necessitates continuous vaccine research and rollout endeavors to protect the population from severe symptoms in the event of sudden outbreaks from new strains. It is also imperative to design effective interdiction rule strategies that can leverage statistics of infection to guide human mobility in order to curb contagion with minimum economic ramifications.

A. RELATED WORKS Computational methods, such as artificial intelligence, machine learning, computer vision, signal processing, pattern recognition, and robotics, are being devised to detect, predict and mitigate COVID-19 cases through analytics [5]. We briefly survey the literature on the applications of epidemic models, machine learning, vaccine production and distribution and social distancing.

1) EPIDEMIC MODELS Existing literature models the impact of isolation, social distancing, vaccines, infected population and age on contagion...
spread. Pandey et al. applied the susceptible-exposed-infected-recovered (SEIR) and regression models to study the epidemiological data of COVID-19 in India between a period of January to March, 2020 to infer the reproduction number [6]. Annas et al. incorporated vaccination and social distancing into the calculation of the SEIR parameters for Indonesia [7]. Radulaescu et al. performed a case study on communities in New York to analyse spread dynamics in a age-heterogeneous scenario [8]. Lopez et al. showed through the SEIR epidemic model that social distancing and isolation can cause a manifold decrease in infection spread in the context of Spain and Italy [9]. Tang et al. used SEIR to predict contagion trends in scenarios where the infected individual may cause spread during the incubation period [10].

2) IDENTIFICATION OF FACTORS THROUGH MACHINE LEARNING

Epidemiologists, health workers, clinicians, and social scientists relied on machine learning to identify the different socioeconomic and demographic factors contributing towards pandemic spread [11]. Prediction models are built around the epidemiological and clinical data to pinpoint vulnerable age groups [12]–[14]. Khan et al. applied regression tree analysis, cluster analysis and principal component analysis to study the variations of testing in the identification of confirmed cases [3]. Roy et al. explored the various socioeconomic and demographic factors in the US states affecting the post-lockdown infection spread [15], most-affected job sectors [16], and the social acceptance of mask-wearing, vaccination and distancing in the US [17]. Liu et al. used neural networks to design a nonlinear, self-adaptive dynamic co-fication and distancing in the US [17]. Liu et al. showed through the SEIR epidemic model that social distancing and isolation can cause a manifold decrease in infection spread in the context of Spain and Italy [9]. Tang et al. used SEIR to predict contagion trends in scenarios where the infected individual may cause spread during the incubation period [10].

3) NEW INTERDICTION AND VACCINE DISTRIBUTION STRATEGIES

Sarlas et al. presented the notion of betweenness-accessibility, a network centrality measure that combines the well-studied concepts of network science and transport geography, namely, betweenness with accessibility. They demonstrated through a case study on Zurich map that the proposed metric identifies critical nodes and links whose interdiction hampers a network’s serviceability [26]. Scott et al. presented a metric for identification of critical transport network links, called Network Robustness Index (NRI), measured as the change in travel-time cost associated with rerouting network traffic should that link become unusable. They demonstrate that NRI offers better high planning solutions resulting in improved travel-time savings [27].

The non-pharmaceutical intervention of social distancing has been a primary rule to reduce contagion for a big part of the pandemic. Block et al. [28] and Roy et al. [29] proposed social distancing measures using network science concepts of clustering and homophily to enable individuals in making local decisions to curb spread. With regard to vaccines, there has been scepticism on the part of the public to accept the emerging vaccinations [30]. The national governments are challenged by the task of rapid clinical trials, industrial production of vaccines as well as their distribution at an affordable price [31]. Shim presented a dynamic model to distribute limited vaccines in South Korea [32]. Matrajt et al. presented an age-stratified optimization approach based on deaths, symptomatic infections, and ICU/non-ICU hospitalizations [33]. Roy et al. presented a time-varying vaccination distribution strategy that incorporates socioeconomic and demographic factors to disseminate vaccines among zones [34].

B. CONTRIBUTIONS

It is evident from our discussion in Secs. I-A2 and I-A3 that human interaction via transportation networks contributes greatly to epidemic spread. It is essential to design strategic interdiction measures to limit human interaction by managing mobility across zones. In this work, we study strategies for network interdiction using social network and machine learning techniques. We first employ a well studied random forest-based network inference model, called GENIE3, to derive the strength of influence of contagion between any two zones. We then apply the three proposed interdiction rules, based on network science and optimization principles such as graph coloring, community detection and knapsack problem, to achieve time-varying travel restrictions.

We carry out extensive simulation experiments on the counties of New York City to demonstrate the efficacy of the rules. We show that these rules differ from existing social distancing algorithms because it operates on the daily infected
Let us review the network science concepts used in this work, remarks in Secs. 4 and 5 of the manuscript. We present the limitations of the study and concluding remarks in Secs. 4 and 5 of the manuscript.

II. APPROACH

Let us review the network science concepts used in this work, followed by the three proposed travel rules.

A. PRELIMINARIES

1) GRAPH THEORY

A graph is an ordered pair $G = (V, E)$ where $V$ is a finite, non-empty set of objects called vertices (or nodes); and $E$ is a (possibly empty) set of 2-subsets of $V$, called edges [35]. A directed graph has directed edges $(u, v) \in E$, allowing unidirectional information flow from vertex $u$ to $v$ and not necessarily from $v$ to $u$. In a directed graph, each node $v \in V$ has in-neighbors defined as a set of nodes $v$ such that there exists an edge from $v$ into $u$, i.e., $e(v, u) \in E$. The out-neighborhood of $u \in V$ consists of nodes $v$ such that there exists an edge from $u$ into $v$, i.e., $e(u, v) \in E$. In a weighted directed graph, $(u, v) \in E$ has weight $w_{uv} \in [0, 1]$, measuring the strength of influence of $u$ on $v$.

Algorithm 1: Inter-Zone Mobility Model

Input. Transition matrix $A$, Migration rate $\xi$, $b \in B$, $N_b = \sum_{st} n_b(st)$ $h = [N_b \times \xi]$ $i = 0$

while $i < h$

$p_{st} \sim \text{Multinomial}(A_b)$ $p(st) \sim \text{Multinomial}(\frac{n_b(st)}{n_b(dead)})$ $\forall st \neq \text{dead}$

$i++$

$n_{dest}(st) = n_{dest}(st) + 1$

$n_b(st) = n_b(st) - 1$

end

2) INTER-ZONE MOBILITY MODEL

Consider a set of geographical sub-regions (termed zones) $B$. We define a frequency matrix $F \in \mathbb{M}_{|B| \times |B|}(\mathbb{R})$, such that $f_{ij} \in F$ denotes the number of trips made from zone $b_i \in B$ to $b_j \in B$. We also define a transition matrix $A \in \mathbb{M}_{|B| \times |B|}$ derived from $F$ as follows: each element of the matrix $a_{ij} \in A$ is the probability of making a trip from $b_i$ to $b_j$, calculated as $\frac{f_{ij}}{h_{ij}}$. These matrices ($F$ and $A$) determine the number of people migrating from one region to another based on the inter-zone mobility procedure.

a: DESCRIPTION

Procedure 1 is invoked periodically by a zone $b$ with population $N_b$, where $h = \lceil \xi \times N_b \rceil$ people move, with migration rate $\xi$ ranging between 0 and 1. Each individual within a zone chooses his destination $dest \in B$ using the multinomial distribution on the $b$-th row of transition matrix $A$, i.e., $A_{k \times b}$. Once the destination is finalized, the epidemic state of a person (i.e., S, E, I, R) is proportional to the fraction of people within that state, i.e., $\frac{n_{dest}(st)}{N_b}$. The procedure iteratively decrements the number of people in $b$ with state $st$, $n_b(st)$, and increments $n_{dest}(st)$ to reflect migration from borough $b$ to $dest$.

3) EPIDEMIC MODEL

We implement the susceptible-exposed-infected-recovered-death (SEIRD) model (see Fig. 1) [36], where susceptible (S) class comprises individuals who are not exposed to the infection; upon exposure, they may transfer to the exposed (E) category of asymptomatic or untested individuals, and this transition is controlled by $\beta$ (that is not a probability). The E class individuals transition to the (tested) infected (I) with probability $\sigma$. The individuals in I transition to another state with a probability $\gamma$, either recovering (R) or dying (D) with probabilities $1 - \alpha$ and $\alpha$, respectively, as shown below.

\[\frac{dS(t)}{dt} = -\beta S(t)I(t)\]

\[\frac{dE(t)}{dt} = \beta S(t)I(t) - \sigma E(t)\]

\[\frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t)\]

\[\frac{dR(t)}{dt} = (1 - \alpha)\gamma I(t)\]

\[\frac{dD(t)}{dt} = \alpha \gamma I(t)\]

In the above differential equations, $\beta = p \times C$, where $p$ and $C$ are the infection probability and contact rates, respectively. Given constants $c_1, c_2$, collision diameter $d$ and population density $\rho$, $C = \sqrt{\frac{2\pi}{\rho}} C_1\sqrt{\rho(1 + c_2\rho)}$. (In our simulation experiments, we represent $\beta = p \times C = p \times f(p) = \kappa \times \rho$.)

4) NETWORK INFERENCE

Let us consider a timeline of $T$ days and infection matrix $X = \{x_1, x_2, \ldots, x_n\}$, where $n$ is the number of zones and
subset of features influencing $x_2$ employing machine learning on features $X$.

FIGURE 2. GENIE3 employing machine learning on features $X^t \subseteq X$ to select subset of features influencing $x_2$.

$x_i \in \mathbb{R}^T$ is the data corresponding to the $i$-th zone (i.e., $x_i^t$ is the measurement of zone $i$ at time $t$). For any pair of zones $u$ and $v$, we employ GENIE3 [37] inference model to leverage $X$ in order to gauge how influences $u$ on $v$ ($u \neq v$), and vice versa. GENIE3 operates under the assumption that any two nodes sharing a GENIE3 link are assigned different colors [38]. The minimum number of the graph is called the chromatic number of colors needed to color the nodes is called the chromatic coloring approach [39] on the (weighted) conflict network $G$.

In this equation, $X_{-u}^t$ is the matrix containing the $t$-th measurement of all vectors except $x_u$, i.e.,

$$x'_u = f(X'_{-u}) + e \forall t \in \{1, 2, \ldots, T\}$$

In this equation, $X'_{-u}$ is the matrix containing the $t$-th measurement of all vectors except $x_u$, i.e.,

$$X'_{-u} = \{x_1, x_2, \ldots, x_{u-1}, x_{u+1}, \ldots, x_n\}.$$

As we illustrate in Fig. 2, GENIE3 applies machine learning on features $X'_{-u}$ to select a subset of features influencing $x_u$. The interaction between zone pairs $u$ and $v$ at any time-point, $w(u, v)$, is represented as a fully-connected, directed weighted network of $V$ nodes (s.t., $|V| = n$). In the rest of the paper, we will refer to this directed weighted network as the GENIE3 network or graph.

5) GRAPH COLORING

It is a graph labeling approach in which two nodes sharing a link are assigned different colors [38]. The minimum number of colors needed to color the nodes is called the chromatic number of the graph $\kappa$. It is possible to split the graph into $\kappa$ partitions, where links can only exist between nodes of different partitions, called color classes. We employ the greedy coloring approach [39] where the first node is assigned color 0. The subsequent nodes are assigned an existing color that is not taken up by their neighbors, while a new color is assigned if all the colors are taken by neighbor nodes.

B. PROPOSED APPROACH

We utilize the GENIE3 network inference approach (see Sec. II-A4) on the daily infection numbers of zones in a geographical region to create networks capturing the epidemiological influence of each zone on the other. We present three travel restriction rules that employ network science and optimization principles on the GENIE3 networks to identify GENIE3 links, which, when removed, can curb contagion. It is noteworthy, the travel restrictions are imposed by eliminating the removable links from the inter-zone migration matrix $A$ that we introduced in Sec. II-A2.

1) COLORING BASED TRAVEL RULE

We propose a simple interdiction rule strategy where at each time $t$, we calculate the GENIE3 network on the daily infection count of the zones between $[t - W, t]$, $G_{t-W \rightarrow t}$. We find the weighted GENIE3 network as follows: Each node $u$ is assigned a weight, $z(u)$, equal to the mean infection numbers over the last $d$ days and replacing the weight of each GENIE3 link $(u, v)$ by $w(u, v) \times z(u)$ ($v \in V$). The intuition behind this step is to attach higher importance to strong GENIE3 influence weight $w(u, v)$, high infection numbers of the source node $u$, or both.

Given an integral threshold $\zeta$, we temporarily eliminate the directed links $e \in E(G_{t-W \rightarrow t})$ if weight $w(e) \leq \zeta$. These remnant links, called conflict links, represent the strong inter-zone interaction leading to contagion. Consider the simple and weighted GENIE3 graph shown in Figs. 3a and 3b (where the diagonal links, although existent, are not shown in the interest of readability). Fig. 3c shows the conflict network $G'_{\zeta}$ for $\zeta = 7.5$. Next, since each link $(u, v)$ in the conflict network represents high likelihood of spread from zone $u$ to $v$, we create two networks, only preserving the non-conflict links, i.e., $(u, v) \in E(G'_{\zeta})$. We apply greedy coloring (see Sec. II-B1) on the (weighted) conflict network and identify the color classes for both networks ($\kappa = 2$ color classes as shown in Figs. 4a, 4b). Each link $(u, v)$ in the two networks follow the two ordering rules: (1) $u < v$ and (2) $u > v$, respectively. Finally, we coalesce the two ordered

FIGURE 3. Identification of conflict links. (a) original and (b) weighted GENIE3 networks; (c) weighted conflict network after eliminating links with weights $w \leq \zeta = 7.5$.

FIGURE 4. Conflict resolution. (a, b) two ordered networks with colored nodes comprising non-conflict links; (c) Coalesced GENIE3 network after conflict resolution.
networks (Figs. 4c) and preserve their links in the migration matrix of the next time window \([t, t + W]\).

2) KNAPSACK-BASED INTERDICTION RULE

We create a time-varying interdiction rule based on the GENIE3 network \(G(V, E)\). Given directed link \(e = (u, v) \in E\), we define decision variable \(X_e\) which denotes the fraction of the time window for which individuals are allowed to move from zone \(u\) to zone \(v\). Each link \(e\) has an economic gain \(\alpha(e)\).

\[
\min \sum_{e \in E} X_e \times w(e) \\
\text{s.t. } X_e \times \alpha(e) = B, \quad X_e \in [0, 1]
\]

We apply linear optimization modeled on fractional knapsack [40], where we drop links \(e \in E\) with the highest weight \(w(e)\), such that economic loss is restricted by a budget \(B\). In Constraint 9, the economic cost associated with a link \(e = (u, v)\), is calculated on the basis of GDP of zone \(u\) as follows:

\[
\alpha_{e=(u,v)} = \frac{A_{u,v}}{\sum_{v} A_{u,v}} \times \text{GDP}(u) \tag{10}
\]

3) QUARANTINE ZONES

Let us consider a set of nodes \(C\) that experience an outbreak at any given time \(t\) with social network \(G_t\). We define an outbreak as a sustained rise in daily infection numbers such that \(x\%\) of the overall population is infected. The most intuitive approach to prevent the spread to non-outbreak zones is to mitigate spread of infection from \(c \in C\) to the rest of the social network by isolating \(c\); this can be achieved by cutting communication links \(E'\) from and to \(c\) (i.e., \(E' = \{(u, v) : (u, v) \in E, u = c|v = c\}\)). However, as we discuss in Sec. I, there are economic implications to a lockdown. Moreover, it may be logistically difficult from a policymaking standpoint to isolate the strongly connected links on a short notice. Instead, we propose a community detection approach to identify and isolate the strongly connected communities around each \(c \in C\), where the strength of connectivity is measured in terms of the GENIE3 link weights. Unlike the isolation approach, the community based quarantine attempts to eliminate links with low GENIE3 weights (i.e., low interaction).

\(\alpha\) APPROACH

For each \(c \in C\), we initialize \(H_c\) as a graph with a single node \(c\). We iteratively add the node \(u\) s.t. \((c, u)\) is the outgoing link with the highest weight. We choose the final configuration \(H_c\) with the highest cost \(C\), given by:

\[
C = d(H_c) + d(I_c) - d(G(H_c \cup I_c)) \tag{11}
\]

Here, \(H_c\) is the community of \(c\), while \(I_c\) is a complement community of \(c\) represented by a graph \(H_c\) with node set given by \(V(G), V(H_c)\). Like all data clustering techniques [41], Eq. 12 maximizes the mean intra-cluster density and minimizes the intra-cluster density between \(I_c\) and \(H_c\). The density of a graph is calculated as:

\[
d(G) = \frac{1}{|V(G)|} \sum_{e \in E(G)} w_e \tag{12}
\]

Fig. 5 shows the GENIE3 network where the link thickness is proportional to the GENIE3 weight. Consider an outbreak zone \(A\) marked in red. The community approach iteratively includes zone \(D\) to the outbreak community and removes the links across communities.

\(b\) TIME COMPLEXITY

We employ the greedy strategies to solve the three interdiction rules proposed in this work. Given a graph \(G(V, E)\), the greedy coloring based rule (see Sec. II-B1) has a time complexity of \(O(|V|^2)\) since a node chooses the least color that has not already been used by its neighbor. Greedy fractional knapsack rule (see II-B2) has complexity of \(O(|E| \log |E|)\) since it ranks the links \(e \in E\) in the increasing order of value:cost ratio, i.e., \(\frac{w(e)}{\alpha(e)}\) and selects the links with the least ratio (s.t. \(X_e \times \alpha(e) = \text{budget } B\)). Finally, given \(c \in C\) outbreak zones, the quarantine strategy (see Sec. II-B3) explores \(|C|\) neighborhoods incurring a complexity of \(O(|C| \times |V|) = O(|V|^2)\). Note that the proposed approaches are invoked in intervals of several days, allowing the organization deciding the interdiction rule to have enough time and computational resource to execute them.

III. EXPERIMENTAL RESULTS

The results are classified into five subsections: (A) baseline approaches, (B) data collection, (C) Mobility and GDP of New York City, (D) coloring-based interdiction rule, (E) Knapsack and community interdiction rules.

The proposed systems are implemented in Python. We employ the GENIE3 [37] and PuLP [42] libraries for network inference and optimizations operations respectively. We carry out experiments on the 62 counties of New York City, where we consider a high yet heterogeneous initial infected population (ranging between approximately 200 individuals in small counties like Hamilton to 20,000 individuals in large counties like Queens). Unless
otherwise stated, we record the mean results for simulation experiments for 50 days over 10 runs. Default simulation parameters are summarized in Table 1.

### TABLE 1. List of parameters and their values.

| Parameter                     | Notation | Value          |
|-------------------------------|----------|----------------|
| Number of iterations          | -        | 10, 50         |
| Simulation duration           | $T$      | 50 days        |
| Number of zones (counties)    | -        | 62             |
| Contact parameters            | $c_1, c_2, d$ | 1.30, $-0.137, 1$m |
| $E\rightarrow I$ transition   | $\sigma$ | 0.25           |
| $I\rightarrow R, D$ transition | $\gamma, \alpha$ | 0.25, 0.05 |
| Outbreak infection percentage | $x$      | 10, 20         |

### A. BASELINES APPROACHES

The interdiction rules are implemented by partial and complete removal of migration ties between zone pairs. We compare the proposed interdiction rules against two baselines.

- Arbitrary (or random) removal of migration ties.
- Isolation of outbreak zones (see Sec. II-B3).

### B. DATA COLLECTION

We discuss the NYC map and mobility traces and the NYC (and global) infected and death numbers.

1) **MAP GENERATION AND LOCATION IDENTIFICATION**

The list of NYC boroughs and districts is acquired from [43], and the latitude and longitude of the 5 boroughs and 62 districts are taken from the Python library for geocoding services, called GeoPy [44]. The distance between any pair of points (i.e., boroughs or districts) on the NYC map is calculated using the geodesic distance function of GeoPy.

2) **NYC MOBILITY DATA**

We source the mobility data of NYC traffic from NYCOpenData [45] – a data repository for fields ranging from city government, education, environment, health to public safety, recreation, social services and transportation. The stated data (spanning a period from 2014 to 2019), collected by the Department of Transportation of New York Metropolitan Transportation Council (NYMTC), has the following fields: ID, road name, source and destination intersecting street name, compass direction, date and time. We use this data to calculate the transition matrix (see Sec. II-A2) that captures the probability of travelling within and across boroughs.

### C. MOBILITY AND GDP OF NEW YORK CITY

We discussed in Sec. II-A2 that inter-zone mobility mimics the migration matrix. Although the experiments discussed hereafter are designed around the NYC (districts or) counties, Fig. 6a presents a high-level generalized view of the migration and mobility matrices that are highly similar for the 5 NYC boroughs, (where a borough is a collection of several counties (refer Sec.III-B1 for details). Fig. 6b shows the latitude-longitude coordinates of the NYC counties (shown as circles proportional to their GDP). Fig. 6c shows the 10 counties with highest and least GDP, where the prosperous counties such as, New York, Bronx, Suffolk, Kings, Queens, etc., are fairly clustered together (in Fig. 6b). Note that the migration matrix is not a part of the proposed interdiction strategies and merely employed here to validate them. Therefore, the three approaches can be used on any mobility model.

### D. COLORING-BASED INTERDICTION RULE

Before we demonstrate the working of the coloring interdiction rule (introduced in Sec. II-B1), we show the effect of an outbreak in a single zone on the contagion of other zones. Fig. 7a shows the daily infected count when one borough experiences an outbreak (where $x = 10\%$ of total population is infected). Note that outbreaks at Manhattan and Brooklyn cause the highest contagion because the two NYC boroughs have the highest population density and population, respectively (see Table 2 for details).

To show the effect of the coloring-based interdiction rule, we first consider different thresholds for link elimination ($\zeta = 0.4, 0.5, 0.6, 0.7$) in the migration matrix. Recall from our discussions in Sec. II-B1, the higher the $\zeta$, the lower will be the degree of travel restriction. Thus, as shown in Fig. 7b,
Effectiveness of Network Interdiction Strategies to Limit Contagion During Pandemic

S. Roy et al.: Effectiveness of Network Interdiction Strategies to Limit Contagion During Pandemic

FIGURE 7. Coloring-based interdiction rule. Comparison of (a) Effect of outbreak in one borough on the contagion in other boroughs; Daily infected cases for (b) different GENIE3 weight elimination threshold \( \zeta = 0.4, 0.5, 0.6, 0.7 \) and (c) coloring versus random link elimination.

Increasing \( \zeta \) results in higher mean infection numbers on the NYC counties over 10 iterations. We compare the daily infected cases against a scenario where we randomly remove equal number of links in the migration matrix to show that the coloring-based interdiction rule is able to show increasing improvement in curbing contagion over time.

E. KNAPSACK AND COMMUNITY INTERDICTION RULES

Similar to the coloring-based rule, we compare the performance of the knapsack interdiction rule (discussed in Sec. II-B2) against that of random migration based link elimination. Fig. 8 shows that as the permissible economic budget \( B \) increases, more migration links are retained resulting in higher total contagion over a period of 50 days. Moreover, the daily infected cases of the proposed knapsack is notably lower than that of random removal.

We discuss in Sec. II-B3 that the community-based rule eliminates links with low inter-zone interaction, offering an improved contagion mitigation versus economic gain trade-off. Fig. 9a shows that the community approach exhibits lower economic loss due to elimination of low weighing GENIE3 links (as per Eq. 12) across communities, over 50 iterations, while the isolation approach has lower spread. We carry out a simulation, where we create new outbreak zones with \( x = 10, 20\% \) every 6 days. We compare the daily infection spread for community approach against no travel restriction scenario to show that the proposed approach yields visibly lower contagion (Fig. 9c).

Effect of delay in interdiction rule imposition. The delay due to administrative, political and logistical, etc., reasons also contribute towards contagion. In the community-based interdiction rule, the migration links are eliminated and reinstated by the next day before estimating the new GENIE3 network. In this experiment, we investigate the effect of delay in lifting the travel restrictions (achieved by reinstating the migration links) by a delay of 1, 2, 3, 4 days. Figs. 10a and 10b show that the increase in delay results in

FIGURE 8. Contagion for the knapsack-based interdiction rule against random migration based link elimination for varying budgets \( B \).

FIGURE 9. Community-based interdiction rule. Comparison against zone isolation approach in terms of (a) economic overhead and (b) daily infected cases; (c) no travel restrictions against daily infected cases.

VOLUME 9, 2021
FIGURE 10. Effect of delay in interdiction rule imposition. Cumulative (a) GENIE3 network weights and (b) infected cases for delay of 1, 2, 3, 4 days in reinstating GENIE3 weights.

TABLE 2. Demographic and economic characteristics of New York City boroughs.

| Zone     | Population | GDP (billions) | Density ($\rho$) |
|----------|------------|----------------|------------------|
| Manhattan| 1,628,706  | 690.244        | 71,341           |
| Bronx    | 1,418,207  | 42.695         | 32,867           |
| Brooklyn | 2,559,903  | 91.509         | 36,147           |
| Queens   | 2,253,858  | 93.310         | 20,767           |
| Staten Island | 476,143 | 14.514         | 8,157            |

the preservation of more inter-zone migration links, resulting in lower cumulative infection spread.

F. CHOICE OF INTERDICTION STRATEGY

The interdiction strategies applied on the daily infection count informs the level of inter-zone mobility resulting in contagion. This allows the policymakers to modulate the migration matrix in a manner that the interactions resulting in potential contagion can be avoided. To illustrate this point, we show the relationship between the GENIE link and the inter-zone migration weights for the knapsack interdiction strategy for four zone pairs, shown in different panels in Fig. 11.

We observe that the migration edge weight (annotated as activity fraction and marked green) is inversely correlated with the GENIE link weights (colored yellow) for all cases.

Comparative study. We compare the performance of coloring- and community-based approaches for varying number of outbreak zones. We consider the following two scenarios: (1) 2 outbreak zones with a high (20%) initial infected fraction and (2) 8 outbreak zones with a moderate (10%) initial infected fraction. For both scenarios, the non-outbreak zones maintain a 2%–5% initial infection. Fig. 12 shows that the community approach that relies on isolation of outbreak zones is more effective when there are fewer (say, 2) outbreak zones experiencing high infection rates. Conversely, the coloring approach is more apt for scenarios where there are several (say, 8) moderate outbreaks. Note that the knapsack interdiction strategy, unlike the other two, allows for fractional inter-zone migration and is particularly useful in scenarios where partial lockdown needs to be enforced.

IV. LIMITATIONS OF THE STUDY

The interdiction rule approaches proposed in this study take the first steps towards curbing contagion, while balancing economic, epidemiological and demographic considerations. There are several limitations in the study. First, in our simulation experiments, we consider outbreak zones to be places with a high daily infection rate. However, incorporating outbreak prediction ahead of time will make the interdiction rules more effective. This motivates the design of machine learning-based preemptive approaches to identify potential outbreak zones. Second, it is imperative to incorporate important factors such as population density and healthcare budget into the interdiction rule approaches. Analogously, we study the effect of delay in lifting travel restrictions $i$ (see Sec. III-E). Our future work will incorporate the effect of delays in the policymaking approaches. Finally, our present approach assumes that the cost of eliminating communication ties between zones is proportional to the inter-zone mobility.
associated with that tie. It is important to incorporate more elaborate economic metrics into the policymaking.

**V. CONCLUSION**

We presented three interdiction rule strategies that employ a machine learning-based network inference on daily infected data to mitigate infection spread. While two of the strategies employed complex network concepts such as graph coloring and network clustering, the third is a linear optimization based on the knapsack problem. We carry out simulations on the 62 counties of New York City to demonstrate the efficacy of the proposed schemes. Our analysis shows that the proposed schemes outperform greedy and random baselines for travel tie elimination, and provide a good trade-off between mitigation of contagion and economic losses. This work serves as the basis for elaborate computational models for public interdiction rules that incorporate socioeconomic and demographic factors of geographical regions.

**COMPETING INTERESTS**

The authors have declared that no competing interests exist.

**REFERENCES**

[1] (2020). Coronavirus World Map: Which Countries Have the Most Cases and Deaths?. [Online]. Available: https://www.theguardian.com/world/2020/jan/28/covid-world-map-which-countries-have-the-most-coronavirus-vaccinations-cases-and-deaths

[2] (2020). Impact of COVID-19 on People’s Livelihoods, Their Health and Our Food Systems. [Online]. Available: https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems

[3] N. Khan, M. Naushad, S. Faisal, and A. Muhammad, “COVID-2019 and world economy,” *Electron. J.* 2020, doi: 10.2139/ssrn.3566632.

[4] A. Joshi, N. Dey, and K. Santosh, Intelligent Systems and Methods to Combat Covid-19. Springer, 2020.

[5] G. Pandey, P. Chaudhary, R. Gupta, and S. Pal, “SEIR and regression model based COVID-19 outbreak predictions in India,” 2020, arXiv:2004.00958, [Online]. Available: http://arxiv.org/abs/2004.00958

[6] S. Roy and P. Ghosh, “Factors affecting COVID-19 infected and death rates inform lockdown-related policymaking,” *PLoS ONE*, vol. 15, no. 10, Oct. 2020, Art. no. e0241165, doi: 10.1371/journal.pone.0241165.

[7] S. Roy, R. Dutta, and P. Ghosh, “Recreational and philanthropic sectors are the worst hit US industries in the COVID-19 aftermath,” *Social Sci. Humanities Open*, vol. 3, no. 1, 2021, Art. no. 100098, doi: 10.1016/j.ssaho.2020.100098.

[8] S. Roy and P. Ghosh, “A comparative study on masking, and vaccine adoption rates from global Twitter trends,” *Healthcare*, vol. 9, no. 5, pp. 488, Apr. 2021, doi: 10.3390/healthcare9050488.

[9] X. Liu, S. Feng, N. Dey, R. Crespo, and E. Herrera-Viedma, “A new SEIRD pandemic prediction model with clinical and epidemiological data analysis on COVID-19 outbreak,” *Appl. Intell.*, vol. 51, no. 7, p. 1–37, Jan. 2021.

[10] A. Ascani, A. Faggian, and S. Montresor, “The geography of COVID-19 and the structure of local economies: The case of Italy,” *J. Regional Sci.*, vol. 61, no. 2, pp. 407–441, Mar. 2021, doi: 10.1111/jrs.12510.

[11] G. R. Shinde, A. B. Kalamkar, P. N. Mahalle, N. Dey, J. Chaki, and A. A. Hassaini, “Forecasting models for coronavirus disease (COVID-19): A survey of the state-of-the-art,” *Soc. Netw. Comput. Sci.*, vol. 1, no. 4, pp. 1–15, Jul. 2020.

[12] A. Paez, F. Lopez, T. Menezes, R. Cavalcanti, and M. Pitta, “A spatio-temporal analysis of the environmental correlates of COVID-19 incidence in Spain,” *Geographical Anal.*, 2020, doi: 10.1111/gean.12241.

[13] A. Paez, “Using Google community mobility reports to investigate the incidence of COVID-19 in the United States,” *Findings*, May 2020, Art. no. 12976, doi: 10.1002/fnd.12976.

[14] R. B. Noland, “Mobility and the effective reproduction rate of COVID-19,” *J. Transp. Health*, vol. 20, Mar. 2021, Art. no. 101016, doi: 10.1016/j.jth.2021.101016.

[15] C. Musselwhite, E. Avineri, and Y. Susilo, “Restrictions on mobility due to the coronavirus Covid19: Threats and opportunities for transport and health,” *J. Transp. Health*, vol. 20, Mar. 2021, Art. no. 101042, doi: 10.1016/j.jth.2021.101042.

[16] D. Höchler, R. Singh, and D. J. Graham, “Social distancing in public transport: Mobilising new technologies for demand management under the Covid-19 crisis,” *Transportation*, pp. 1–30, Apr. 2021, doi: 10.1007/s11116-021-10192-6.

[17] G. Sarlas, A. Paez, and K. W. Axhausen, “Betweenness-accessibility: Estimating impacts of accessibility on networks,” *J. Transp. Geography*, vol. 22, Apr. 2020, Art. no. 102680, doi: 10.1016/j.jtrangeo.2020.102680.

[18] D. M. Scott, D. C. Novak, L. Aultman-Hall, and F. Guo, “Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks,” *J. Transp. Geography*, vol. 14, no. 3, pp. 215–227, May 2006, doi: 10.1016/j.trangeo.2005.10.003.

[19] P. Block, M. Hoffman, I. J. Raabe, J. B. Dowd, C. Rahal, R. Kashyap, and M. C. Mills, “Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world,” *Nature Human Behaviour*, vol. 4, no. 6, pp. 580–586, Jun. 2020, doi: 10.1038/s41562-020-0898-6.

[20] S. Roy, A. Cherevko, S. Chakraborty, N. Ghosh, and P. Ghosh, “Leveraging network science for social distancing to curb pandemic spread,” *IEEE Access*, vol. 9, pp. 26196–26207, 2021, doi: 10.1109/ACCESS.2021.3058206.

[21] (2020). Mistrust of a Coronavirus Vaccine Could Imperil Widespread Immunity. [Online]. Available: https://www.nytimes.com/2020/07/18/health/coronavirus-anti-vaccine.html

[22] (2020). Testimony on Operation Warp Speed: Researching, Manufacturing, & Distributing a Safe & Effective Coronavirus Vaccine. [Online]. Available: https://www.nih.gov/about-nih/who-we-are/nih-director/testimony-operation-warp-speed-researching-manufacturing-distributing-safe-effective-coronavirus-vaccine

[23] E. Shim, “Optimal allocation of the limited COVID-19 vaccine supply in South Korea,” *J. Clin. Med.*, vol. 10, no. 4, p. 591, Feb. 2021, doi: 10.3390/jcm10040591.
[33] L. Matrajt, J. Eaton, T. Leung, and E. R. Brown, “Vaccine optimization for COVID-19: Who to vaccinate first?” *Sci. Adv.*, vol. 7, no. 6, Feb. 2020, Art. no. eabf1374, doi: 10.1126/sciadv.abf1374.

[34] S. Roy, R. Dutta, and P. Ghosh, “Optimal time-varying vaccine allocation amid pandemics with uncertain immunity ratios,” *IEEE Access*, vol. 9, pp. 15110–15121, 2021, doi: 10.1109/ACCESS.2021.3053268.

[35] M. E. J. Newman, “The structure and function of complex networks,” *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, Jan. 2003, doi: 10.1137/S003614450342480.

[36] H. W. Hethcote, “The mathematics of infectious diseases,” *SIAM Rev.*, vol. 42, no. 4, pp. 599–653, Jan. 2000, doi: 10.1137/S0036144500371907.

[37] V. A. Huynh-Thu, A. Irrthum, L. Wehenkel, and P. Geurts, “Inferring regulatory networks from expression data using tree-based methods,” *PLoS ONE*, vol. 5, no. 9, Sep. 2010, Art. no. e12776, doi: 10.1371/journal.pone.0012776.

[38] T. Jensen and B. Toft, *Graph Coloring Problems*, vol. 39. Hoboken, NJ, USA: Wiley, 2011.

[39] L. Kučera, “The greedy coloring is a bad probabilistic algorithm,” *J. Algorithms*, vol. 12, no. 4, pp. 674–684, 1991, doi: 10.1016/0196-6774(91)90040-6.

[40] H. Ishii, T. Ibaraki, and H. Mine, “Fractional knapsack problems,” *Math. Program.*, vol. 13, no. 1, pp. 255–271, Dec. 1977, doi: 10.1007/BF01584342.

[41] G. Gan, C. Ma, and J. Wu, *Data Clustering: Theory, Algorithms, and Applications*. Philadelphia, PA, USA: SIAM, 2020.

[42] S. Mitchell, M. O'Sullivan, and I. Dunning, “PuLP: A linear programming toolkit for Python,” Dept. Eng. Sci., Univ. Auckland, Auckland, New Zealand, Tech. Rep., 2011, p. 65.

[43] (2020). *Neighborhoods in New York City*. [Online]. Available: https://www.ny.gov/counties

[44] (2020). *Geopy: Geocoding Library for Python*. [Online]. Available: https://github.com/geopy/geopy

[45] (2020). *Nycopendata*. [Online]. Available: https://data.cityofnewyork.us/Transportation/Traffic-Volume-Counts-2012-2013-/p424-amsu

**SATYAKI ROY** (Member, IEEE) received the Ph.D. degree in computer science from the Missouri University of Science and Technology, Rolla, MO, USA, in 2019. He is currently a Postdoctoral Research Associate with the Department of Genetics, University of North Carolina, Chapel Hill, NC, USA. His research interests include computational biology, network science and optimization, wireless sensor networks, epidemiology, machine learning, and parallel computing.

**PREETOM BISWAS** is currently pursuing the bachelor’s degree with the School of Computing, Informatics, and Decision Systems Engineering, Arizona State University (ASU), Tempe, AZ, USA. He is currently a Math and Programming Enthusiast. His research interests include application of machine learning and graph theoretic approaches to solving real-world problems.

**PREETAM GHOSH** received the B.S. degree in computer science from Jadavpur University, Kolkata, India, and the M.S. and Ph.D. degrees in computer science and engineering from the University of Texas at Arlington. He is currently a Professor with the Department of Computer Science and directs the Biological Networks Laboratory, Virginia Commonwealth University. He has several federally funded research projects from NSF, NIH, DoD, and US-VHA. His research interests include algorithms, stochastic modeling and simulation, network science, machine learning related approaches in systems biology and computational epidemiology, and mobile computing related issues in pervasive grids that have resulted in more than 170 conference and journal articles. He is currently serves as the Secretary/Treasurer for ACM SIGBio.

***