IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 20, 2021, accepted October 8, 2021, date of publication October 11, 2021, date of current version October 21, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3119459*

# Modeling and Evaluating Epidemic Control Strategies With High-Order Temporal Networks

**ALESSIA ANTELMI[1], GENNARO CORDASCO[2], VITTORIO SCARANO[1], AND CARMINE SPAGNUOLO[1]**

[1]Dipartimento di Informatica, Università degli Studi di Salerno, 84084 Fisciano, Italy
[2]Dipartimento di Psicologia, Università degli Studi della Campania Luigi Vanvitelli, 81100 Caserta, Italy

Corresponding authors: Alessia Antelmi (aantelmi@unisa.it) and Carmine Spagnuolo (cspagnuolo@unisa.it)

This work was supported in part by the European Union Horizon 2020 Research and Innovation Programme under Agreement 769872 (EMPATHIC) and 823907 (MENHIR).

**ABSTRACT** Non-Pharmaceutical Interventions (NPIs) are essential measures that reduce and control a severe outbreak or a pandemic, especially in the absence of drug treatments. However, estimating and evaluating their impact on society remains challenging, considering the numerous and closely tied aspects to examine. This article proposes a fine-grain modeling methodology for NPIs, based on high-order relationships between people and environments, mimicking direct and indirect contagion pathways over time. After assessing the ability of each intervention in controlling an epidemic propagation, we devise a multi-objective optimization framework, which, based on the epidemiological data, calculates the NPI combination that should be implemented to minimize the spread of an epidemic as well as the damage due to the intervention. Each intervention is thus evaluated through an agent-based simulation, considering not only the reduction in the fraction of infected but also to what extent its application damages the daily life of the population. We run experiments on three data sets, and the results illustrate how the application of NPIs should be tailored to the specific epidemic situation. They further highlight the critical importance of correctly implementing personal protective (e.g., using face masks) and sanitization measures to slow down a pathogen spreading, especially in crowded places.

**INDEX TERMS** Agent-based modeling, complex networks, epidemic, high-order relationships, hypergraphs, non-pharmaceutical interventions.

## I. INTRODUCTION

Since ancient times, different populations have adopted varying strategies to prevent and contain diseases, from the isolation of sick individuals to establishing a time limit to the manifestation of symptoms to magical practices. [29]. The concept of modern and preventive quarantine dates back only to 1377; still today, it represents a general preventive intervention in the absence of a targeted vaccine, along with high healthcare surveillance and public information [29], [44], [51]. Generally, all healthcare policies intended to mitigate the effects of the spread of a new virus or pathogens when no vaccines or medicine are available yet are commonly referred to as Non-Pharmaceutical Interventions (NPIs) [51]. Current development of the recent pandemic highlighted to what extent the increase of human mobility and goods exchange made NPIs, such as lockdown and border closure, more challenging to apply to past cases in history, mainly because of their negative impact on the worldwide economy [9], [24] as well as on the psychological wellness of society [33], [56], [68], [69]. Choosing the right control policy to adopt is a burden that governments bear as their decisions have repercussions also when the epidemic is under control. Thus, each country is responsible for adopting NPIs according to its territory's specific needs. As examined in a document published by the World Health Organization (WHO) in 2019 [51], each NPI has different effects, resource implications, and ethical considerations. To precisely study the consequences of each intervention on the region it has to be applied, the models used to simulate an epidemic propagation need to guarantee a high level of accuracy to ensure both efficacy and efficiency in evaluating the specific control policy.

Most of the well-known models adopted to evaluate the outcomes of control policies are based on math equations

The associate editor coordinating the review of this manuscript and approving it for publication was Azwirman Gusrialdi.

(Equation-Based Models, EBM) that have proved their ability to mimic the epidemic spreading in individuals [23], [34], [36]. However, these models assume that the population behavior and individual contact types are homogeneous [10]. The aforementioned is a severe limitation for real-world scenarios. It reduces the modeling effectiveness in describing different individuals and social behaviors, urban mobility patterns, geographic information systems (GIS), and so on [41]. Further, equation-based models do not provide an easy way to model different contacts [28]. In particular, many epidemic contagions operate in two ways: *direct* contagion (person-to-person infection) and *indirect* contagion (infection via an intermediary, such as furnishings or clothing). In other terms, EBMs provide us quantitative information about the number of infected individuals in the worst-case scenario. Yet, to reduce the negative impact of control interventions, policymakers need to introduce targeted actions and obligations, such as closing a specific area or building (school, workplace, metro), or wearing a face mask in crowded places.

In this sense, designing, implementing, and evaluating epidemic control interventions has become a challenging task. Agent-Based Models (ABMs) support researchers in this direction. ABMs are a modeling tool able to easily incorporate features related to population and society and are widely adopted to simulate human behaviors under specific conditions [48], [62], in particular in epidemiological studies [45]. They allow researchers to naturally include human mobility data to model humans interactions between the environment or other individuals. Typically, many epidemic ABMs also exploit networks to define possible agent interactions. Such data can be retrieved from online social networks (OSNs), where users share their real-time location (Foursquare), geotag media posts (Facebook and Instagram), or review businesses (Yelp). The growing popularity of these online platforms and the ubiquitous online access provide gold data for studying users' habits, lifestyles, and mobility patterns to include in ABMs.

The above considerations have been the seed of our previous work [6]: high-order interactions cannot be ignored when dealing with models of diseases spreading through the air or any other infected object or environment. Here hypergraphs come into play, as a hyperedge can naturally model a group of people being in the same location (or environment) in a given time, even though they did not have any direct contact. Specifically, we formally defined time-varying hypergraphs (TVH), where a node's weight within a hyperedge represents the last time a user has visited that specific location. We further developed the SIS compartmental model into an ABM, exploiting our model to simulate interactions between agents and locations, which interact according to the human mobility pattern computed from a real-world data set of the social network Foursquare. We evaluated and compared the effect of direct and indirect contacts on disease propagation, studying the impact of accurately modeling the time interval within which the two contagion pathways may happen and hasten the epidemic diffusion.

This contribution is an extension of our previous work just mentioned. In this study, we specifically focus on formally introducing NPIs within our TVH framework, assessing the effectiveness of each control policy in reducing direct and indirect contagion pathways. Each intervention is evaluated via an ABM simulation, considering the reduction in the number of infected and to what extent its application damages the daily life of the population. We then embed these two contrasting goals within a multi-objective optimization framework to guide the choice of which NPI combination should be implemented. The major contributions of this paper can be summarized as follows:

- A discussion about NPIs and how they can be embedded within an epidemiological model based on high-order networks and ABMs, emphasizing the importance of modeling individual and global population behaviors;
- The formal definition of NPIs, described by the WHO in [51], for our epidemiological framework based on TVHs, proposed in [6];
- Evaluation of each NPI applying the SIS compartmental equation-model into an ABM that exploits our diffusion algorithm to simulate interactions between agents and locations;
- The design and implementation of a genetic algorithm-based methodology to optimize the choice of which NPI combination has to be adopted when contrasting objectives are considered.

The remainder of this paper is organized as follows. Section II reviews some relevant literature about epidemic models on hypergraphs and NPIs. Temporal hypergraphs, our ABM design-methodology, and the epidemic diffusion algorithm are summarized in Section III. In Section IV, we discuss NPIs and how we embedded them within the epidemiological framework based on THVs. Section V describes the experimental setting and the data sets used in our analysis and Section VI presents the sensitivity analysis of the model. In Section VII, we examine and discuss how evaluating the choice of an NPI implementation within a multi-objective optimization framework. Finally, Section VIII details the conclusion and future works.

## II. RELATED WORK

### A. EPIDEMIC MODELS ON HYPERGRAPHS

A complete overview of mathematical frameworks capably of explicitly and naturally describing group interactions is given by Battiston *et al.* [11]. Specifically, the authors outline the dynamics of structures with many-to-many interactions and discuss higher-order diffusion models, including spreading dynamics on hypergraphs.

Bodò *et al.* [12] first proposed modeling communities as hyperedges, based on the concept that an actual model of an epidemic outbreak has to take into account two factors: community structure and infection pressure. They translated this approach into practice using different contagion probabilities according to the place. In addition, they bounded

the likelihood that a susceptible individual becomes infected in a unit to be not proportional to the number of infected individuals within that unit. The authors show that using a non-linear function to model the infection pressure is crucial not to overestimate the epidemic propagation. In that way, they demonstrated that graphs are not a well-suited structure to capture the many-to-many relationships that come into play during epidemic propagation processes. In this model, Poisson processes govern both infection and recovery, where the infection rate $r$ takes into account connectivity patterns. In contrast, recovery is a spontaneous process controlled by a fixed recovery rate $\gamma$.

Suo *et al.* [63] investigated a similar SIS model on hypergraphs in the context of rumor spreading on social media. They proposed two information diffusion models by considering how an individual might decide to share content on a social media platform, either to all the contacts or targeting a particular group. In the global strategy, at each time step, an infected node $i$ can infect with a probability $\beta$ all the susceptible neighboring nodes connected to $i$ via a hyperedge. With the local approach, an infected node $i$ randomly chooses one of its hyperedges $e$ and then tries to infect all the nodes in $e$ with a probability of success $\beta$.

Another version of a high-order contagion model for spreading dynamics occurring at group level was proposed by Jhun *et al.* [40]. Specifically, they studied the spreading process on scale-free $d$-uniform hypergraphs. In this model, a susceptible node in a hyperedge $e$ of size $d$ may be infected from $e$, with rate $\beta_d$, only if the remaining $d-1$ nodes composing $e$ are infectious. A standard recovery probability $\mu$ is used for recovery.

Recently, de Arruda *et al.* [8] presented an SIS framework that explicitly includes critical-mass dynamics into the contagion model. The authors generalize the simplicial contagion model proposed by Iacopini *et al.* [37] both structurally and dynamically. They (i) moved from simplicial complexes to hypergraphs and (ii) allowed a hyperedge $e$ to be potentially infectious for a node $i \in e$ if the number of infected nodes composing $e$ is greater or equal to a given threshold $t_e$.

The epidemic propagation model, proposed in our previous work [6] and enriched in this paper, extends the work of Bodò *et al.* [12] by accounting for both direct and indirect contacts in the spreading process. As for Bodò *et al.*, infection and recovery are Poisson processes. Further, we used the same non-linear function to model the infection pressure. In contrast, in our model, a location may become contaminated (and thus spreading the epidemic) according to Poisson processes proportional to the number of infected nodes within that hyperedge. This concept also differs from the models of Jhun *et al.* [40], and de Arruda *et al.* [8], where a hyperedge of degree $d$ becomes infectious if $d-1$ nodes within it are infected [40] or if they are higher than a given threshold [8].

### B. NON-PHARMACEUTICAL INTERVENTIONS

Non-pharmaceutical interventions (NPIs) have been the subject of vast literature even before the COVID-19 pandemic [27], [67]. Still, until 2007, Aledort *et al.* [2] report a generally poor quality of evidence on which to base non-pharmaceutical pandemic planning decisions, mainly due to the lack of representative data and a validation process [67]. Unfortunately, when the COVID-19 started spreading worldwide, NPIs were the only possible measures to stop its diffusion. This event gave birth to an unseen joint effort of the academic community and tech giants in understand, model and assess the connection between human behaviors and disease diffusion and the effects of applying NPIs [54].

In a recent survey [54], Nicola Perra thoroughly describes current literature about NPIs during the COVID-19 pandemic. The author classifies the models adopted into four categories: i) compartmental models, ii) metapopulation models, iii) statistical models, and iv) agent-based models. In this section, we will focus on agent-based models as, in this work, we exploit the ABM paradigm to simulate the epidemic spreading and the application of NPIs. For a detailed discussion about ABMs and a comparison with EBMs, we refer the reader to the following works [7], [65], [66]. Another recent survey about the use of ABMs to simulate the Covid-19 pandemic can be found in [47].

Hoertel *et al.* [35] develop a detailed ABM for France to evaluate the effectiveness of different NPIs in the reopening phases after the first wave. The proposed model has 194 parameters, describing the socio-demographic features of the French population (140), the contact networks (33), and the features of the virus (21). The authors study lockdown and post-lockdown measures, including physical distancing, mask-wearing, and isolation approaches, highlighting how the interventions after lifting the first lockdown were not enough to overburden the healthcare system. With a similar approach, Aleta *et al.* [3] build a multilayer synthetic population that models the socio-demographic features of the Boston metropolitan area using high-resolution data describing the movements and potential interactions of people in the city to investigate the impact of different reopening scenarios. Their results suggest how a proactive policy of testing, contact tracing, and household quarantine could gradually reopen economic activities and workplaces with a low impact on the healthcare system. Analogous to this work, the same authors also propose an ABM for the metropolitan areas of New York and Seattle [4], informing the model with mobile phone data and Foursquare data to identify POIs.

Wilder *et al.* [70] propose an ABM to study the spreading of COVID-19 in Hubei, Lombardy (Italy), and New York City. The synthetic population is formed by individuals stratified for age, comorbidities, and assigned to a household, while contacts among agents in different contexts than home are modeled via contact matrices. Their results suggest that measures should be tailored to the specific socio-demographic features of each population as the efficacy of NPIs varied across the analyzed location. Contact matrices are also used by Ogden *et al.* [50] to model a synthetic sample of the Canadian population to evaluate social distancing and isolation measures to control the disease spread.

Their outcomes indicate that lifting disruptive NPIs such as shut-downs must be accompanied by enhancements to other NPIs to prevent new introductions and to identify and control new transmission chains.

Yang *et al.* [73] use a network-based model to represent contact happening inter- and intra- different cities in the Hubei province. They specifically examined the use of personal protective, social distancing, and a combination of those measures to decrease the infection rate.

Bouchnita *et al.* [13] design a multi-scale ABM in which agents move according to a social force model and which considers both direct and indirect transmission mechanisms. The authors do not explicitly consider the notion of location as agents move on a grid. The model simulates indirect contagion based on the normalized concentration of deposited SARS-CoV-2 on hard surfaces, the averaged rate of SARS-CoV-2 secretion by contagious agents, and the decay rate of the virus. For each agent, indirect transmission can occur only once every day at a random moment. As discussed in Section II-A, the indirect diffusion mechanism described by Bouchnita *et al.* is profoundly different from the process described in our previous work [6] as hyperedges encode places where agents can meet and may become contaminated according to Poisson processes proportional to the number of infected nodes within that hyperedge.

Silva *et al.* [59] propose an agent-based framework (COVID-ABS) to simulate people, business, government, and the health care system. The framework allows for the implementation of several NPIs and measures the impact in terms of disease and economic burden. Other few works also use ABMs to simulate the economic consequences of the COVID-19. For instance, Inoue and Todo [38] quantify the economic effect of a possible lockdown of Tokyo, estimating that the lockdown would result in an 86% reduction of the daily production in Japan after one month. Similar to COVID-ABS, Dignum *et al.* [18] present an ABM simulation tool to analyze possible repercussions of policy interventions, combining social, economic, and health aspects. Generally, all works agree that continued intervention should be considered to keep the transmission of an epidemic under control, as well as mixing NPIs to best regulate contagion dynamics.

In this work, we exploit the previously proposed TVH framework [6] to model direct and indirect transmission dynamics of disease spreading. Alike (temporal) graphs, TVHs abstract and formalize contact among agents simulated with an ABM, but adding information about where the contact is happening. Hence, such structures allow to formally analyze diffusion mechanisms while accounting for group interactions and indirect contagion processes via contaminated locations.

## III. HIGH-ORDER TEMPORAL EPIDEMIC DYNAMICS

This section describes some useful concepts related to the framework presented in [6], formally introducing TVHs and the high-order diffusion process to simulate human-to-human (direct) and human-to-environment (indirect) infection propagation.

### A. A HYPERGRAPH MODEL

Propagation of contagious diseases is a complex dynamic process that holds abounding human behavior aspects. Graphs are not expressive enough to easily simulate direct and indirect contact among individuals; furthermore, such structures do not consider when a connection happens, which is crucial in epidemic dynamics [53]. A contact network can be easily extended to include the time dimension by using *Time-Varying Graphs* (TVGs) [16], [22], a variant of the graph model, where a link between two nodes is valid only for a given time interval. Nonetheless, even using TVGs, the information tying together a group of individuals simultaneously in a particular geo-location is lost.

To correctly model an epidemic propagation in a many-to-many fashion and capture that people, moving through different locations, form a community in a given time and space, we adopt *hypergraphs*. A hypergraph is an ordered pair $\mathcal{H} = (V, E)$, where $V$ is the set of nodes, and $E$ is the set of hyperedges. Each hyperedge is a non-empty subset of nodes. Hypergraphs generalize the well-known graph model as graphs represent only binary relationships, while hypergraphs relationships of any arity. We will refer to $n = |V|$ and $m = |E|$ as the size of the vertex set and the edge set, respectively.

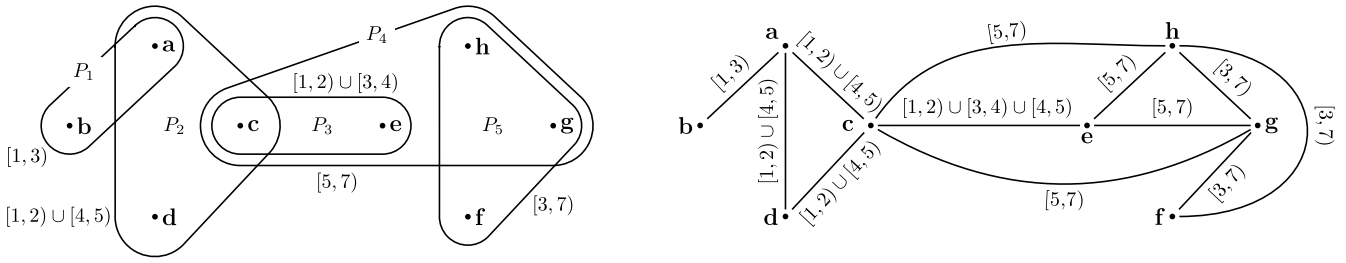
#### 1) TIME-VARYING HYPERGRAPHS

To better mimic an epidemic spreading, we extend the definition of TVGs, presented by Casteigts *et al.* in [16], to hypergraphs. Employing a TVH to describe a contact network enables us to minimize the effect of time and the presence of only direct contacts. Formally, a TVH is defined as follows.

*Definition 1:* A **TVH** is a hypergraph $\mathcal{H} = (V, E, \mathcal{T}, \rho)$, where $\mathcal{T}$ is the lifetime of the system and $\rho: E \times \mathcal{T} \rightarrow \{0, 1\}$ is an existing function, indicating whether a hyperedge exists in a given timeframe.

For each $t \in \mathcal{T}$, we refer to the hypergraph $\mathcal{H}_t = (V, E_t)$ as the hypergraph corresponding to a particular time $t$, i.e., $E_t = \{e \in E : \rho(e, t) = 1\}$, where $E_t$ denotes the set of existing hyperedges at time $t$.

As described by Bretto in [14], the two-section (or clique) representation of $\mathcal{H}$, denoted with $[\mathcal{H}]_2$, is a graph whose vertices are the vertices of $\mathcal{H}$, and where two vertices form an edge if they are in the same hyperedge. Figure 1 presents an instance of a TVH (see Figure 1a) compared to its corresponding two-section graph (see Figure 1b). It illustrates a trivial TVH made up by 8 individuals (nodes), $V = \{a, b, c, d, e, f, g, h\}$, and 5 geographical locations (hyperedges), $E = \{P_1, P_2, P_3, P_4, P_5\}$. Each hyperedge is labeled with its corresponding availability time interval $t = [t_s, t_e)$, $\mathcal{T} = \{[1, 2), [1, 3), [3, 4), [4, 5), [5, 7), [3, 7)\}$. It is worth noting that the $[\mathcal{H}]_2$ representation introduces a loss of information in the contact network. For instance, it is not possible




(a) TVH $\mathcal{H}$. (b) $\mathcal{G} = [\mathcal{H}]_2$.

**FIGURE 1.** A simple TVH $\mathcal{H}$ (left) and its clique representation $\mathcal{G} = [\mathcal{H}]_2$ (right). Each hyperedge/edge is labeled with the corresponding availability time span.

to recognize which is the time interval when the individuals $c$ and $e$ were both in the venue $P_3$ or $P_4$.

In our epidemiological modeling framework, we used the following definition of TVH and check-in function $\omega$.

*Definition 2:* A **TVH for an epidemic diffusion** is a hypergraph $\mathcal{H} = (V, E, \mathcal{T}, \omega)$, where

- $V$ is the set of $n$ vertices (users/agents);
- $E$ is the set of $m$ hyperedges (locations);
- $\mathcal{T}$ is the lifetime of the system;
- $\omega$: $V \times E \times [t_s, t_e] \rightarrow \{0, ct_{v,\ell}\}$, where $ct_{v,\ell} \in \mathcal{T}$ is the last check-in time of an agent $v \in V$ in a location $\ell \in E$ during the time interval $[t_s, t_e]$ (0 means that $v$ has never checked-in $\ell$ during $[t_s, t_e]$).

The function $\omega$ only keeps track of the last check-in time $ct_{v,\ell}$ for a vertex $v$ in a given location $\ell$ during a time interval. If $v$ did not checked-in in the same location $\ell$ or in another location $\ell'$ in the next time interval $t'$, the value of $ct_{v,\ell}$ is still considered a valid check-in time in the current timestamp $t$. We store the last check-in time $ct_{v,\ell}$ of a vertex $v$ in a location $\ell$ as the weight of $v$ in the hyperedge representing $\ell$. This modeling strategy allows us to simulate direct and indirect contagion processes easily: for each time interval, we can effortlessly know where the user is. We include this approach in an ABM simulation to enrich the TVH model with crucial social dynamics. The simulation time is split into fixed-width intervals of length $\Delta$, corresponding to the time interval when indirect contagion may happen. Given a sampling time $\Phi$ of a check-in data set, the total number of time intervals considered is $|\mathcal{T}| = \lceil \frac{\Phi}{\Delta} \rceil$.

### B. THE DESIGN METHODOLOGY

We based the idea of our ABM design methodology on the assumption that two spreading policies regularize an epidemic process: *direct* and *indirect* contagions between individuals and environments. Direct contact implies a pairwise interaction between two individuals in the same place. In contrast, an indirect contagion embodies the interchanges that may happen between agents and locations. These two types of contacts are a natural consequence of each person's daily activities and commuting routes. For instance, when an agent moves from its home to its workplace, it may be either infected by touching some furniture or simply breathing contaminated air (indirect interaction) or by a face-to-face talk with another agent (direct exchange).

In [6], we proposed a diffusion algorithm whose spreading process is designed in a discrete-event fashion and exploits the TVH structure to discover whether direct or indirect interactions may happen. During each time interval, agents are simulated according to their scheduling policy. Then, our diffusion algorithm is performed. As agents are free to move, the epidemic has the chance of spreading from one location to another. Simultaneously, the outbreak may still spread across agents located in the same place at a particular interval of time. While direct contaminations require agents' co-presence, indirect connections happen between agents and the environment, and co-presence is not needed. In the following, we provide definitions of the concepts adopted in the next sections.

- $\Phi$ is the time span of the data.
- $\mathcal{T}$ is the set of time intervals describing the evolution of the relationships between agents and geo-locations.
- $\Delta$ is a real value (minutes, hours or days) corresponding to the time discretization parameter. It further refers to the time span when *indirect* contagions may happen.
- $\delta$ is a real - small - value (milliseconds, seconds or minutes) defining when direct infections may take place. A direct contagion is established if two agents stay in the same location within a time difference less than $\delta$.
- $t$ is the current simulation interval ($t \in \mathcal{T}$).
- $\mathcal{H} = (V, E, \mathcal{T}, \omega)$ is the TVH representing the data (see Definition 2).
- $\Gamma_t$ and $\mathrm{N}_t$ define the neighborhood functions of an agent $a \in V$ in a given simulation time $t$. Specifically, $\Gamma_t(a) = \{\ell \in E : \omega(a, \ell, t) \neq 0\}$ is the set of locations visited by $a$ during the interval $t$; $\mathrm{N}_t(a) = \bigcup_{\ell \in \Gamma_t(a)} V_t(\ell)$ is the set of neighbors of $a$ during the simulation time $t$, where $V_t(\ell)$ denotes the set of agents that visited the location $\ell$ during the interval $t$.
- $\Upsilon(a, \ell)$ is a time function providing the last check-in time of the agent $a$ in the location $\ell$.
- $\mathrm{T}_t(a)$ and $\mathrm{T}_t(\ell)$ denote the infection state (1 infected, 0 not infected) of an agent or a location in a given simulation time $t$, respectively.
- $\mathrm{X}_t(a, b)$ is a direct contact function. Given two agents $a$ and $b$, it returns 1 if they have direct contact in the time

interval $t$; 0 otherwise. Formally,

$$\mathrm{X}_t(a,b) = \begin{cases} 1, \text{if } \exists\ \ell \in \Gamma_t(a) \cap \Gamma_t(b) \wedge |\Upsilon(a,\ell) - \Upsilon(b,\ell)| < \delta \\ 0, \text{otherwise.} \end{cases}$$

In the following, we will remove the subscript $t$ when the simulation time is clear from the context.

The scheduling routine of each agent is defined by an input check-ins data set, describing where each agent is and when the agent entered that specific location. As described in Section III-A, the overall check-in data set is split into fixed-width intervals of length $\Delta$, corresponding to the time-span when an indirect contagion may happen. In practice, all check-ins happening within the same location in a $\Delta$ time window are grouped together, and a hyperedge is added to the TVH to represent the location where a group interaction (or indirect contact) is happening. For each time window, if the time difference between two check-ins in the same location is lower than a given small value $\delta$, we consider direct contact happening between those two agents. This methodology approximates the notion of direct contact when high-granular data on the individual routine is not available.

### C. THE DIFFUSION PROCESS

In a typical SIS model, the infection and recovery states are ruled by a Poisson process. Thus, either a susceptible individual or location becomes (directly or indirectly) infected with a probability $1 - e^{-\beta_x f(n)}$. Here, $\beta_x$ denotes the infection rate per-contact (considering either direct or indirect contacts), $n$ is the number of infected entities (individuals or locations), and $f^e(\cdot)$ is a non-linear function used to bound the infection pressure for large value of $n$, as described in Section II. Similarly, an infected agent or location recovers with probability $1 - e^{-\gamma_x}$, where $\gamma_x$ denotes the recovery rate for agents and locations. In our previous work [6], we devised an SIS variant on hypergraphs by means of a diffusion algorithm, whose pseudo-code is described in Algorithm 1, Section 5.4, of [6]. In the following, we detail its main components.

During each time step, our diffusion algorithm proceeds in three contagious phases.

1. *Agent-to-Environment*. The first phase simulates the *environment infectiveness*. For all non contaminated locations, (i.e., $\ell \in E : T(\ell) = 0$), we compute the number of infected agents that have visited that location:

$$I^e(\ell) = \sum_{a \in V(\ell)} \mathrm{T}(a).$$

This value is then used to update the infection state of a susceptible location $\ell$, as expressed by the following rule:

$$\mathrm{T}(\ell) = \begin{cases} 1, & \text{infected with probability } 1 - e^{-f^e(I^e(\ell))} \\ 0, & \text{not infected otherwise,} \end{cases}$$

where $f^e(\cdot)$ is a non-linear function, typically adopted to govern the behavior of the epidemic outbreak over hyperedges [12]. In our experiments (see Sections VI and VII), we considered the following regularization function:

$$f^e(x;c) = \begin{cases} x, & \text{if } 0 \le x \le c \\ c, & \text{if } x > c, \end{cases}$$

where $c$ is a constant given as parameter.

2. *Agent-to-Agent*. The second phase simulates the *direct* propagation process. For all susceptible agents (i.e., $a \in V : T(a) = 0$), we compute the total number of infected neighbors. Formally,

$$I^d(a) = \sum_{b \in \mathrm{N}(a)} \mathrm{T}(b)\mathrm{X}(a,b).$$

This value is then used to update the infection state of a susceptible agent $a$, as expressed by the following rule:

$$\mathrm{T}(a) = \begin{cases} 1, & \text{infected with probability } 1 - e^{-I^d(a)} \\ 0, & \text{not infected otherwise.} \end{cases}$$

3. *Environment-to-Agent.* The third and last phase simulates the *indirect* propagation process. For all susceptible agents, (i.e., $a \in V : T(a) = 0$), we compute the number of infected locations visited. Formally,

$$I^i(a) = \sum_{\ell \in \Gamma(a)} \mathrm{T}(\ell).$$

This value is then used to update the infection state of a susceptible agent $a$, as expressed by the following rule:

$$\mathrm{T}(a) = \begin{cases} 1, & \text{infected with probability } 1 - e^{-I^i(a)} \\ 0, & \text{not infected otherwise.} \end{cases}$$

We consider the simulation proceeding in $|\mathcal{T}|$ discrete steps. At each simulation step $t$, every agent independently runs its *step function* and updates its internal state, which will be effective in the next simulation phase $t + 1$.

To simulate the introduction of the NPIs described in Section IV, we extended the diffusion algorithm to accommodate the specific details of each intervention and modeling supplementary agent behaviors. The description of each NPI, how it is formally embedded within a high-order diffusion process based on hypergraphs, and how it is implemented within the simulation are discussed in the following section. All the parameters in input to our model are specified in Section V-C.

## IV. MODELING NPIs

NPIs are healthcare policies readily available at all times and in all countries, intended to mitigate the effects of the spread of a new virus or pathogens when no vaccines or medicine are available yet [55]. The potential impacts of NPIs on an influenza epidemic are to delay the introduction of the infection into a population, delay the height and peak of the outbreak if it has started, reduce transmission by personal protective or environmental measures, and reduce the number of infections and hence the number of severe cases [51]. In 2019,

the WHO Global Influenza Programme and the WHO Collaborating Centre for Infectious Disease Epidemiology and Control (School of Public Health, University of Hong Kong), published a report providing recommendations for the use of NPIs based on a systematic review of the evidence on their effectiveness, including personal protective measures, environmental measures, social distancing measures, and travel-related measures [51].

In the remainder of this section, we will review personal protective, environmental, and social distancing measures and how they can be embedded into an epidemiological model based on high-order networks, ABMs, and the SIS equation-based model. Moreover, we will describe how we formally enriched our modeling framework to support the evaluation of NPIs. In this work, we do not consider travel-related measures as they require specific mobility data. Those will be the focus of future investigations.

### A. PERSONAL PROTECTIVE MEASURES

Examples of personal protective measures (PPMs) are hand hygiene, respiratory etiquette, and face masks. While the first two actions are a well-established and straightforward practice concerning personal daily hygiene, face masks are only conditionally recommended during a severe epidemic or a pandemic or for symptomatic individuals. PPMs can be implemented and simulated by decreasing the transmission probability of a pathogen from one individual to another when they come in contact and changing the infectiousness probability caused by interacting with environments. PPMs are an instance of individual behavior interventions as their efficiency is defined by how many people respect them.

We embedded the adoption of PPMs into our model by reducing the epidemic spreading opportunity (transmission probabilities) for both direct and indirect contagion pathways. Specifically, we decreased the values of the parameters $\beta_d$ (*agent-to-agent*), $\beta_i$ (*environment-to-agent*), and $\beta_e$ (*agent-to-environment*) to simulate the introduction of PPMs into the agent population; hence, leaving unchanged the three phases of the diffusion algorithm. It is worth noting that we lower the value of those parameters only for the agents adopting PPMs.

### B. ENVIRONMENTAL MEASURES

Surface and object cleaning actions are environmental measures (EMs) recommended as a public health intervention in all settings to reduce influenza transmission. As for PPMs, environmental measures can be implemented by decreasing a pathogen's transmission probability from a contaminated object to an individual. For instance, if a person sneezes on a table and, soon after, another person touches that surface, a contagion may indirectly occur from one person to another. To simulate EMs, we need to embed within the network of contacts the notion of *environment* or *location* where people can interact either with it or with other individuals. In other words, the network model used to study epidemic propagation has to be location-aware. Here, hypergraphs come into play as a hyperedge can naturally model a group of people being in the same location in a given time, even though they did not have any direct contact (see Section III).

Similar to PPMs, the adoption of EMs can be simulated by modifying the contagiousness probabilities $\beta_i$ and $\beta_e$. Specifically, we reduced the parameter regulating the contagiousness of an indirect contact $\beta_i$ by a factor $\beta_\ell$ for all locations $\ell \in \mathcal{L}$ sanitized at given time intervals. Formally, $\beta_i \leftarrow \beta_i - \mathrm{X}_{\mathcal{L}}(\ell)\beta_\ell$, where $\mathrm{X}_{\mathcal{L}}(l)$ is an indicator function equal to 1 for all the locations that are continuously sanitized. If $\beta_\ell = \beta_i$, the transmission probability becomes $\beta_i = \beta_i - \beta_\ell = 0$. Thus, it causes the place to be no more infectious. As detailed in Section VI-B, we simulate the cleaning of all locations restoring their status to susceptible and not reducing the parameter $\beta_i$ to reproduce the fact that a location may become infected again if it is not sanitized after each use. The implementation of EMs required the addition of a fourth phase after the three described in Section III-C.

### C. SOCIAL DISTANCING MEASURES

Social distancing measures (SDMs) represent interventions on individuals' sociality and involve the population or sub-population behaviors. Examples of these measures are i) isolation of sick individuals, ii) quarantine of exposed individuals, ii) contact tracing, iv) avoiding crowding, and v) school, workplace, and, in general, public or private structure closures. In the following, we describe how we embedded these measures within our modeling approach.

#### 1) ISOLATION

The word isolation indicates the separation or restriction of movement of ill individuals with an infectious disease to prevent transmission to others [19]. We embedded this policy within the behavior of each agent, able to recognize whether it is sick. At each time interval and before the execution of the three phases of the diffusion algorithm, every agent $a$ may go in self-isolation according to a probability $\beta_{isolation}$, regulated by the following Poisson process $1 - e^{-T(a)\cdot\beta_{isolation}}$, where $T(a)$ is the infection state of the agent $a$. When an agent is isolated, it does not contribute to the epidemic spreading; in other words, other agents or environments cannot interact with an isolated agent. Once isolated, the agent does not exit this status until it recovers with a probability $\gamma_a$.

#### 2) QUARANTINE

The word quarantine indicates an imposed separation or restriction of movement of individuals exposed, who may or may not be infected but are not ill, and who may become infectious to others [19]. As for the isolation measure, we embedded this policy within the behavior of the agents. At each time interval $t$ and, specifically, during the *Agent-to-Agent* contagion phase, an agent $a$ enters the quarantine state according to an overall probability $\beta_{quarantine}$ proportional to the number of infected agents $I_t^d(a)$ it has met during the previous time step. This scenario is regulated by the Poisson process $1 - e^{-I_t^d(a)\beta_{quarantine}}$, where the number of ill

agents $I_t^d(a)$ is computed based on the simulation status of the previous time step $(t-1)$. A quarantined agent no longer contributes to the epidemic until it exits the confinement state or recovers with a probability $\gamma_a$.

#### 3) CONTACT TRACING

The idea behind contact tracing applications is to rapidly identify at-risk individuals once a case has been detected [21]. However, even though such systems can substantially increase the proportion of people quarantined, it can lead to ethical issues such as leakage of information [51]. Furthermore, successful implementation relies on the availability of resources and technology [25], [43]. Generally, a design tool has to trace all contacts that agents have during the simulation and identify infected agents. To simplify the evaluation of a contact tracing intervention, we introduced several abstractions in our model. Specifically, we did not consider any error in the retracement process, assuming a 100% accuracy of location data, and we did not implement any particular contact tracing protocol [1]. Further, we allowed all agents access to global knowledge on the status of the epidemic. For each agent $a$ using the tracing application, we introduced a list $\mathcal{N}(a)$ comprising all the agents (using the tracing application themselves) that interacted with $a$ during the simulation.

Contact tracing is modeled similarly to the quarantine measure, but in this case, the tracing system is more powerful and enables tracking the number of infected individuals the agent has met during all previous steps instead of only the previous one. In other words, we can consider this measure like an informed quarantine as the main difference between these two interventions lies in the information used by an agent to decide whether to enter the quarantine state. In this case, the agent *remembers* the contacts had in all previous steps since the tracing measure was adopted. The usefulness of distinguishing these two scenarios is also evident in a more complex compartmental model with an Exposed state. In this case, adopting a tracing technology allows the agent to keep track of all its contacts and know whether they have manifested symptoms or tested positive on the pathogen even though they have come across many simulation steps before. When only quarantine measures are implemented, the agent only uses the contact information available in the previous simulation step. As a consequence, the agent may self-quarantine with a lower probability.

Formally, at each simulation step and before the execution of the three phases of the diffusion algorithm, every agent $a$, either infected or not, may decide to self-quarantine according to the Poisson process $1 - e^{-I^d(a)\beta_{tracing}}$, where $I^d(a)$ is the number of infected individuals the agent has met during all previous steps. If an agent set its state to quarantine, then the simulation proceeds as for the implementation of the quarantine measure.

#### 4) AVOIDING CROWDING

Avoiding crowding is another example of SDM, often used in combination with other policies to reduce influenza transmission. Avoiding crowding may have cultural or religious implications; for instance, gatherings are places to share information during influenza, comforting people, and reducing fear. During the SARS-CoV-2 pandemic, many countries adopted several avoiding crowding policies - some stricter than others - to contain its spread. For instance, the English government permitted only up to 6 people to meet [49], while the Italian government banned all sorts of gatherings [57].

In our framework, we implemented the avoiding crowding measure as a global policy by deterministically reducing the number of agents allowed within each place. We defined a threshold $\alpha_\ell$ representing the maximum number of people allowed in a location $\ell$ per time interval. In other words, for each location $\ell$, we only simulate the first $\alpha_\ell$ agents (sorted by their check-in time) at each time step. As the scheduling policy of each agent is determined by the input data, the newly reduced data set can be easily pre-computed without requiring modifications to the diffusion algorithm.

#### 5) LOCATION CLOSURE

Closing public or private places is a standard measure to control an epidemic spreading by reducing possible contact between individuals and environments in a severe outbreak. This extreme measure may have substantial economic consequences and result in significant societal problems.

Typically, this intervention is implemented by selectively closing places according to a specific classification, such as schools, transportation, workplace, restaurants. We implemented this policy by stopping the simulation of all contacts, both direct and indirect, happening within a location $\ell$ belonging to the set of closed places $\mathcal{L}$, starting from a time interval $t$. Hence, implementing such a measure required modifying all three phases of the diffusion algorithm to prevent the simulation of all contacts within the closed locations. This measure can be formally illustrated by reducing to 0 all parameters regulating either an agent or an environment's infectiousness $\beta_{(\cdot)} \leftarrow \beta_{(\cdot)} - \mathrm{X}_{\mathcal{L}}(\ell)\beta_{(\cdot)} \quad \forall \beta_{(\cdot)} \in \{\beta_i, \beta_d, \beta_e\}$, where $\mathrm{X}_{\mathcal{L}}(\ell)$ is an indicator function equal to 1 for all closed locations (i.e., $\ell \in \mathcal{L}$). The set $\mathcal{L}$ can be pre-computed according to the policy to be investigated (e.g., closing transportation or workplaces).

## V. EXPERIMENT SETTING

This section describes the experimental setup of the simulation, introducing the underlying epidemiological assumptions and detailing the simulation parameters. It further defines how each NPI is evaluated and finally describes the data sets used. The experiment is designed to evaluate the impact of NPIs not only in terms of reducing the fraction of infected but also considering the cost of applying each measure. The implementation of the model and the experiments are available on an open-source GitHub repository.[1]

[1] https://github.com/alessant/HGEpidemics

**TABLE 1. Simulation parameters. For each parameter, the table reports: which aspect of the simulation the parameter regulates, a short description, its domain, and its value in the simulation (*variable* indicates that the parameter changes according to the specific experiment).**

| Type | Parameter | Description | Domain | Value |
|---|---|---|---|---|
| High-order epidemic spreading | $\beta_d$ | Probability that an agent is infected by another agent via direct contact. | $[0, 1]$ | 0.57 |
| | $\beta_i$ | Probability that an agent is infected via indirect contact due to a location. | $[0, 1]$ | 0.29 |
| | $\beta_e$ | Probability that a location is infected by an agent. | $[0, 1]$ | 0.29 |
| | $\gamma_a$ | Probability that an agent spontaneously recovers. | $[0, 1]$ | 0.024 |
| | $\gamma_e$ | Probability that a location is sanitized. | $[0, 1]$ | 0.017 |
| | $c$ | Number of contacts in *Agent-to-Environment*. | $\mathbb{N}$ | 5 |
| PPMs | $\alpha_p$ | Fraction of agents using PPMs. | $[0, 1]$ | Variable |
| | $ppm\_\beta_d$ | Probability that an agent is infected by another agent via direct contact when PPMs are in use. | $[0, 1]$ | 0.1 |
| | $ppm\_\beta_i$ | Probability that an agent is infected via indirect contact due to a location when PPMs are in use. | $[0, 1]$ | 0.05 |
| | $ppm\_\beta_e$ | Probability that a location is infected by an agent when PPMs are in use. | $[0, 1]$ | 0.05 |
| EMs | `sanitize` | Whether locations are regularly sanitized. | {`true`, `false`} | Variable |
| SDMs | $\beta_{isolation}$ | Probability that an agent enters the isolation state. | $[0, 1]$ | Variable |
| | $\beta_{quarantine}$ | Probability that an agent enters the quarantine state. | $[0, 1]$ | Variable |
| | $\alpha_e$ | Fraction of location closed during the lockdown. | $[0, 1]$ | Variable |
| | $\alpha_i$ | Fraction of agent using a tracing application. | $[0, 1]$ | Variable |
| | $\beta_{tracing}$ | Probability that an agent enters the quarantine state if it adopts tracing technologies. | $[0, 1]$ | 0.6 |
| | `avoiding crowding` | Whether avoiding crowding measures are applied. | {`true`,`false`} | Variable |

### A. ASSUMPTIONS

As mentioned in Section III, the proposed epidemic design methodology is based on high-order relationships that may happen during an epidemic outbreak between humans and environments, via direct and indirect contagious pathways. From the epidemiological point-of-view, it is crucial to clarify the assumptions lying underneath our model. First, the diffusion procedure assumes that all infected individuals are asymptomatic. This choice advantages the epidemic spreading as all infected continue to propagate the infection, representing an epidemic outbreak's optimal case. Second, we do not consider the notion of incubation within our model, intended as the time elapsed between exposure to a pathogenic organism, and when symptoms and signs are first apparent. In our model, each agent contracts the infection according to a transmission probability and proportionally to the number of infected individuals it met, or to the number of locations it visited. As for the previous point, this choice helps the epidemic spread out as the person becomes sick immediately. Third, we do not study how the time-length individuals spend together plays a role in the infection dynamics as we thoroughly analyzed the sensitivity of the TVH model to the simulation's parameters in [6]. Lastly, we fixed the value of each epidemic parameter during the whole simulation steps and we assume that each NPI is perfectly applied.

### B. NPIs EVALUATION

To implement each NPI, we modified the diffusion algorithm described in Section III-C according to the policy to simulate. The parameters and framework enrichment for each technique are detailed in Section IV. The impact of an NPI or a combination of NPIs is compared against an unmitigated scenario in which no interventions are implemented. We evaluated the effectiveness of each NPI (or a combination of them) according to two parameters: the reduction of the final fraction of infected agents and how much the intervention affects the population. To quantify each action's impact and compare the results across all data sets, we operatively defined two domain-agnostic notions of *damage*. Specifically, for each agent, we determined the fraction of agents it was unable to meet because of the adopted interventions. We then defined the overall social damage of the intervention $\mathcal{D}_a$ as the average over the whole population. Similarly, for each agent, we computed the fraction of the locations it was not able visiting, due to the adopted interventions, and then evaluated the overall commuting damage $\mathcal{D}_l$ as the average over the population.

### C. SIMULATION PARAMETERS

In the unmitigated (baseline) scenario or when no specified otherwise, we set the epidemic parameters as follows: $\beta_d = 0.57$, $\beta_i = 0.29$, $\beta_e = 0.29$, $\gamma_e = 0.017$, $\gamma_a = 0.024$, and $c = 5$. To have reasonable parameter values, we based their choice on the mathematical model *SIDHARTE* proposed by Giordano *et al.* [30] for the SARS-CoV-2 pandemic. We fixed the values of $\Delta = 4$ (4 hours) for an indirect contact to happen, to mimic the resistance of the COVID-19 on surfaces [52], and $\delta = 15$ (15 minutes), based on the Immuni mobile tracing app that considers a direct

contact happening in a time window of 15 minutes [32]. Thus, in our model, an indirect contact may occur if two people have been in the same place within 4 hours of difference, while a direct contact may happen if two people have been in the same place within 15 minutes. It is worth underlining that our goal is not to simulate the diffusion of the SARS-CoV-2 virus, but more generally, the diffusion of any epidemic that can spread even indirectly. In this setting, the average time an agent is infected is $42 \cdot \Delta$ intervals (7 days), while for a location is $6 \cdot \Delta$ intervals (1 day). Table 1 lists all the simulation parameters used to control the epidemic spreading and the introduction of NPIs.

We generated the population mobility pattern according to the three data sets that will be presented in Section V-D. The simulations last 185 steps while the interventions (if any) are introduced at the 100*th* step. As the fraction of infected results from probabilistic processes, we ran each simulation scenario 80 times, and we consider the averaged value as a result. The simulation is initialized with a single infected agent, the patient zero.

It should be emphasized that, in an SIS model, recovered individuals become susceptible again without gaining any immunity against the pathogen. In other words, there is no memory of past infections. This is the typical spreading model of infections like the common cold and influenza, which do not confer any long-lasting immunity. In contrast, in the SIR model, all agents will eventually recover and not contract the infection anymore. In other terms, in the SIR model, the outbreak will ultimately drop out. Based on this consideration, and as we are interested in evaluating control actions in the worst scenario, we analyze each intervention under the SIS model.

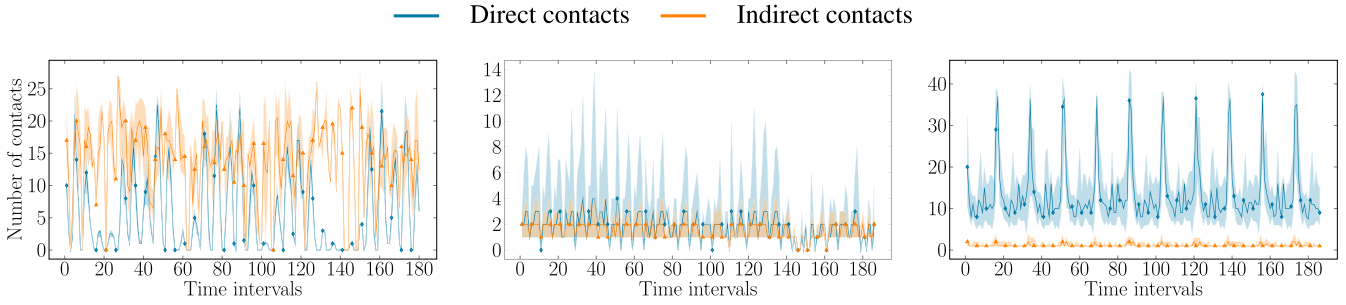
### D. DATA SETS

In the ABM model we developed, each agent moves between geo-locations over time and comes in contact, via direct or indirect pathways, with other agents and different environments (geo-locations). To model individuals (agents) mobility patterns, we adopted three data sets describing human interactions at different scale: from i) a location-aware sensing infrastructure (BLEBeacon data set [58]), ii) a metropolitan area scenario (Foursquare data set [72]), and iii) a virtual society scenario (Game of Thrones data set [39]).

Figure 2 shows the median number of direct and indirect contacts across all data sets, considering $\Delta = 4$ hours and $\delta = 15$ minutes. For each value, we also reported the 25% and 75% quantiles. In more detail, as direct contacts, we evaluated for each agent the number of different other agents it met; while, as indirect contacts, the number of different locations the agent has been. These plots depict an *agent-centered* vision of the data, in the sense that we can grasp, on average, how many different ways an agent may be infected by counting the number of direct contacts and location visited (indirect contacts). Figure 2 reveals the diverse nature of the data sets, which results in distinct contact patterns. The BLEBeacon data set refers to high-granular small-scale check-ins happening within a building. Based on that, it is reasonable to think that an individual tend to meet always the same people, but - at the same time - they are free to walk in the structure. Figure 2a details this pattern where the number of indirect contacts is generally higher than the number of direct contacts. We can observe a completely different picture in Figure 2c, related to the Game of Thrones (GoT) data set, where the number of direct contacts is significantly higher than the number of indirect contacts. Once again, this pattern is due to the constitution of the data as the majority of the GoT characters tend not to move across many different locations, but they still have many direct contacts (e.g., council or battle scenes). Figure 2b refers to the real-world large-scale check-in data set from the social platform Foursquare. In this case, no class of contacts strongly prevails on the other. In contrast with the other two data sets, we can notice a consistent variance in the number of locations visited by the users. As already discussed in our previous work [6] and other several studies [5], [31], [71], this behavior is typical of online social network, where the minority of the users accounts for the most content. A detailed description of each data set follows.

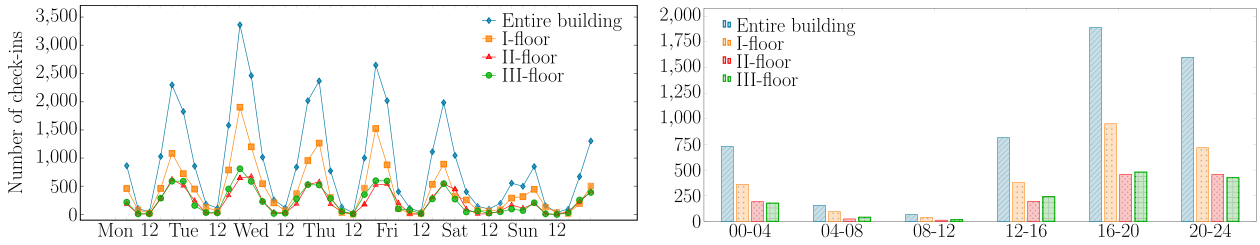
#### 1) THE BLEBeacon DATA SET

The BLEBeacon data set [58] is a collection of Bluetooth Low Energy (BLE) advertisement packets/traces generated from BLE beacons carried by people following their daily routine inside a university building for a whole month. The main objective of this data set was to present a real-life realization of a location-aware sensing infrastructure that




(a) BLEBeacon. (b) Foursquare. (c) GoT.

**FIGURE 2.** Distribution of direct and indirect contacts, fixing $\Delta = 4$ hours and $\delta = 15$ minutes.




(a) Aggregated weekly check-ins. (b) Aggregated daily check-ins.

**FIGURE 3. Aggregated weekly (on the left) and daily (on the right) check-ins distribution of the overall BLEBeacon data set referring to 30 days of people daily routine inside a university building.**

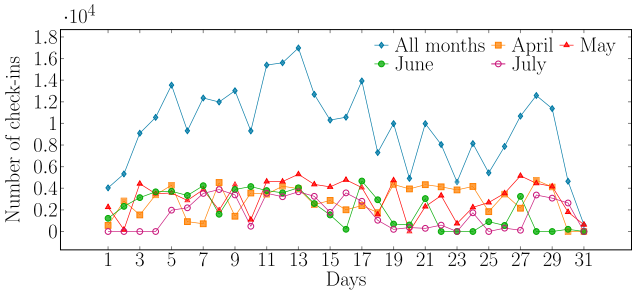



**FIGURE 4. Daily check-in distribution for the Foursquare data set.**

can provide insights for smart sensing platforms, building management, and user-localization frameworks. The data set contains 153, 868 check-ins of 62 users and 31 locations. Figure 3 provides the check-ins distribution aggregated over a week (see Figure 3a) and over a day (see Figure 3b).

### 2) THE FOURSQUARE DATA SET

The Foursquare social network data set, introduced by Yang *et al.* in [72], is a collection of check-ins originated from the city of Tokyo and crawled from 12 April 2012 to 16 February 2013. The data set contains 573, 703 check-ins of 2, 293 users and 61, 858 locations (such as restaurants, cinema, sports and so on). In our previous work [6], we analyzed the sensitivity of the proposed TVH model to the simulation's parameters over the most crowded month (May 2012) of this data. Having fixed $\Delta = 4$ and $\delta = 15$, our analysis revealed a peak in the number of infections near the 20% of the overall population. Towards this, it is worth stressing that the application of a single or a combination of non-pharmaceutical measures only makes sense in a dangerous scenario, in which the epidemic spreading is hard to control. Thus, to increase the probability of having a more virulent pathogen spreading and overcome the sparsity nature of the data set, we merged the check-ins happening from April 1*st*, 2012 up to August 1*st*, 2012 into a single month, obtaining 2, 147 users and 41, 519 locations. Figure 4 shows the number of daily check-ins over the four months considered (April-July) and the final amount used for the simulation.

### 3) THE GAME OF THRONES DATA SET

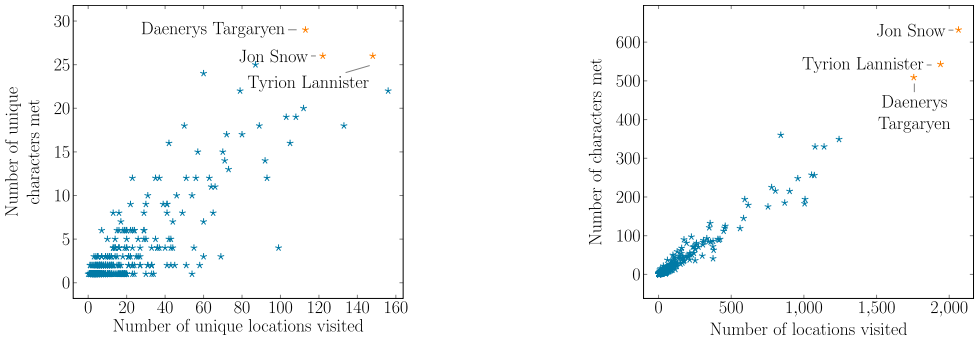
Starting from scripting data [39] of the whole 8 seasons of the *Game of Thrones* (GoT) HBO TV series, we developed a check-in data set based on the mobility patterns of the characters' series. As episodes are chronologically ordered, but no real date is available, we set a virtual clock to January 1st, updating it according to each scene's duration. As each episode has an average duration of around 1 hour, we obtained a data set spanning over 70 hours. As for the Foursquare data set, to better estimate the effect of NPIs in a virulent scenario on a longer temporal interval, we finally concatenated the obtained check-ins until having data covering a total of one month, with 577 characters and 111 locations. Figure 5 depicts an analysis of the characters' mobility patterns over the first 70 hours. Specifically, we examined a possible correlation between the number of unique (without repetition) people met and unique locations visited by each GoT character (see Figure 5a). As shown, the characters that had contacts with a higher number of people tended to have traveled across many places. This behavior becomes even more evident in Figure 5b, where we considered all characters met and locations visited.

## VI. SENSITIVITY ANALYSIS

This section describes the experiments we carried out to validate the proposed model and verify the correctness of its implementation. In our previous work [6], we extensively analyzed the sensitivity of the TVH model to the epidemic parameters and different discretization of the time intervals when either direct or indirect contacts may happen. In this study, we focus on evaluating the model's sensitivity regarding the parameters regulating each NPI; thus, studying how perturbations in the input modify the model output.

To analyze the impact in reducing the fraction of infected and the damage brought by each NPI, we defined several scenarios in which we applied just a single intervention. In this way, we not only evaluate the sensitivity of the model to the specific NPI, but we also verify the implementation of each intervention and ensure that each logical component of the model behaves as intended. In the following sections, we describe each validation scenario, comparing the fraction of infected agents at i) the peak, ii) the lowest value reached by the infection and iii) the end of each simulation against the unmitigated scenario. We further report the impact of the intervention in terms of damage as defined in Section V-B. For each scenario, we also discuss some implications of using the given measure and its efficacy.

Figure 6 shows the epidemic diffusion patterns in an unmitigated scenario, considering contagions due to i) direct and




(a) *Unique*. (b) *Complete*.

**FIGURE 5.** GoT characters (nodes) and environments (hyperedges) distribution, considering an agent's contact history without (5a) and with (5b) repetitions.

indirect contacts, ii) direct contacts, and iii) indirect contacts. In this case, the three plots report an *epidemic-centered* vision, in which we observe the fraction of infected due to either direct, indirect (infected locations), or both contact types. As expected, these patterns are a consequence of the contacts happening in the data. In fact, if we compare Figure 2 with Figure 6, we can note that indirect contacts cause higher peaks in the fraction of infected for the BLE data set since agents visits at least 10 locations on average while having a lower mean of direct contacts (see Figure 6a). On the contrary, direct contacts trigger more contagions than indirect contacts for the GoT data set, as suggested by the respective number of the two types of interactions (see Figure 6c). On the other hand, the contagion patterns in the Foursquare data set never reach the peak of 80% of infected (see Figure 6b). Once again, this outcome was expected given the high sparsity of the data. Here, indirect prevail direct contagions as we have more information about the places each user visit rather than the other people they meet. Further, it is worth noting that even though each agent visits only a few locations on average (see Figure 2b), their check-ins tend to refer to common place like transportation or general entertainment (for instance, the 41 most crowded locations in the data set refer to the transportation system). Hence, indirect contacts drive the epidemic diffusion as one location can potentially infect many agents. Consequently, these locations have the potential to spread the epidemic across a considerable number of agents.

Figure 7 presents the averaged curves of the fraction of infected and susceptible agents obtained from the simulation replications. They behave similarly to the standard SIS model, dominated by the decline of the Susceptible population and the increase of the Infected population. Dark lines in the figure represent the median of 80 simulation replications, and the shaded areas represent 25 and 75 quantiles. These curves appear more rugged than those resulting from an EBM due to the heterogeneous mixing facilitated by the contact network.

### A. SCENARIO 1: USING PPMs

#### 1) EXPERIMENTAL SETTING

As described in Section IV-A, we modeled the introduction of PPMs by reducing the infection probability of each agent adopting these measures. Specifically, we reduced the direct and indirect contagious transmission probability by about 80%. We set the new direct and indirect contagion probability as follows: $ppm\_\beta_d = 0.1$, $ppm\_\beta_e = 0.05$ and $ppm\_\beta_i = 0.05$ (see Table 1). To analyze the impact of applying PPMs, we varied the fraction $\alpha_p$ of agents using the measures, testing five values 0.0, .25, .50, .75, and 1.0. Clearly the scenario with $\alpha_p = 0.0$ corresponds to the unmitigated scenario.

#### 2) DISCUSSION

Table 2 reports the results for this scenario. First, we need to note that the application of such intervention does not cause any social damage ($\mathcal{D}_a$) nor commuting damage ($\mathcal{D}_l$). In fact, adopting PPMs like face masks and hand hygiene does not prevent an agent from meeting other agents or freely move across locations. Second, increasing the fraction of agents adopting PPMs causes a decrease in the fraction of infected, suggesting how the infection propagation is susceptible to PPMs usage by a growing number of individuals. However, at least 75% of the population have to use PPMs to visibly decrease the fraction of infected both at the end of the simulation and in the lowest peak of the infection in all data sets. As expected from real-world events and the vast amount of current literature [54], only applying PPMs cannot notably reduce the spreading of a pathogen, but - even in small percentages - can help lower the number of infections.

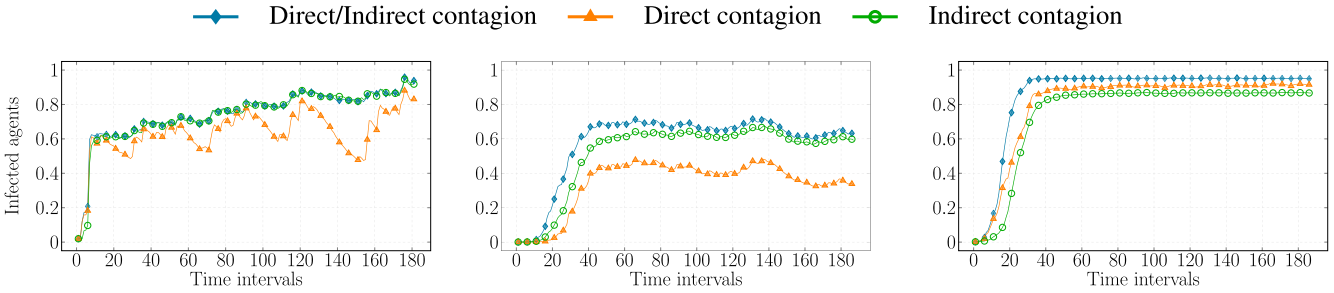
### B. SCENARIO 2: USING EMs

#### 1) EXPERIMENTAL SETTING

As introduced in Section IV-B, our modeling framework allows by design adopting EMs by manipulating the infection probability due to indirect contacts. In this validation scenario, we simulate the cleaning of all locations restoring their status to *healthy* at the end of the most crowded time intervals, i.e., 12:00-16:00, 16:00-20:00 and 20:00-24:00. We did not reduce the parameter $\beta_i$ to reproduce the fact that a location may become infected again if it is not sanitized after each use.

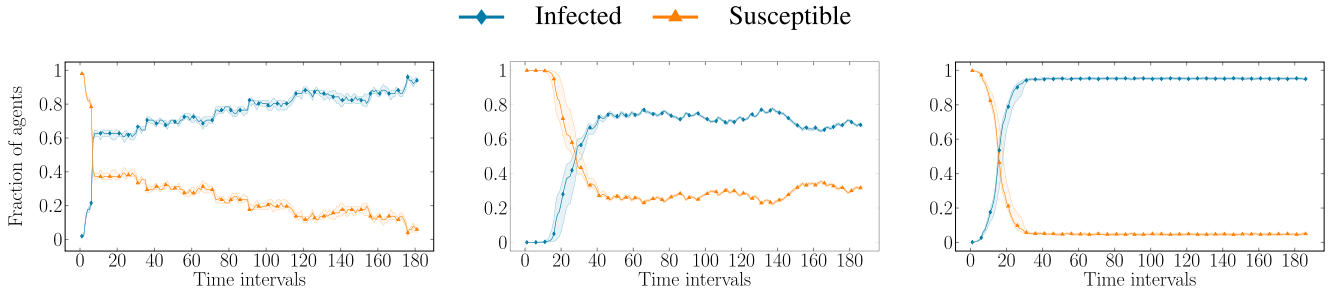
#### 2) DISCUSSION

Table 3 presents the results of this experiment. In this case, we analyze the epidemic spreading under two different




(a) BLEBeacon. (b) Foursquare. (c) GoT.

**FIGURE 6.** Pathogen diffusion in an unmitigated scenario, considering contagions due to i) both direct and indirect contacts, ii) direct contacts, and iii) indirect contacts.




(a) BLEBeacon. (b) Foursquare. (c) GoT.

**FIGURE 7.** Epidemic propagation behavior under the SIS model.

perspectives. We report the fraction of infected i) considering both direct and indirect contagious pathways and ii) focusing only on the contribution made by indirect contacts. Even though we can observe poor results in the application of cleaning procedures in the first case, we can note that the introduction of EMs has some effect in reducing the fraction of infected due to indirect contacts. This result is not surprising as agents are still free to move and propagate the epidemic making the locations infected again. As in the case of the adoption of PPMs, we have neither social nor commuting damage as there are no constraints on the mobility of the agents. Given the limited cost of EMs, these interventions can be considered good practice in real scenarios as they effectively lower the number of infected cases.

### C. SCENARIO 3: USING SDMs - ISOLATION

#### 1) EXPERIMENTAL SETTING

In this validation scenario, we analyze the sensitivity of the model to the implementation of isolation measures (see Section IV-C1). Specifically, we studied how changes in the parameter $\beta_{isolation}$, representing the willingness of an agent to enter the isolation state if infected, reflect on the number of spreaders. We ranged the parameter in the interval [0, 1]. Note that the scenario with $\beta_{isolation} = 0.0$ corresponds to the unmitigated scenario when no measures are being applied. Modifying the parameter $\beta_{isolation}$ means regulating when an infected agent enters the isolation state (the higher, the sooner). Further, the probability of entering the isolation state is also proportional to the number of infected agents met (the higher the number, the higher the probability). The time spent by each agent in this state is strictly dependent on the recovery probability as the agent may exit the isolation only if it becomes susceptible again. Having set the simulation parameters as described in Table 1, the upper bound to the time an agent remains in isolation is 7 days on average (see Section V-C).

#### 2) DISCUSSION

Table 4 shows the experiment results up to $\beta_{isolation} = 0.5$ as the fraction of infected drastically drops out even for small values of the parameter. In this case, $inf_{last}$, $inf_{peak}$, and $inf_{lower}$ do not report the total fraction of infected, but the fraction of agents that can still spread the infection, i.e., the agents that are sick but not isolated. Increasing the value of $\beta_{isolation}$ has a sensible impact on the overall spreading since - as we can see from column $isolated_{last}$ - the majority of the agents will eventually enter the isolation state, and as a such, they will no more spreading the pathogen. That results in a minimum number of effective spreaders. In a real scenario, the effectiveness of such intervention is strictly dependent on the easiness and the availability of tools to test whether a person is infected. Nevertheless, this simple experiment emphasizes the critical importance of effectively identifying contagious individuals to decrease the epidemic curve.

As expected, both damages $\mathcal{D}_a$ and $\mathcal{D}_l$ reflect that at least 50% up to 87% of the agents are isolated, preventing them from seeing other agents or visiting locations. Fixing $\beta_{isolation} = 0.5$, agents lose on average the 70% of their contacts and around the 40% of the locations. The social

**TABLE 2.** Scenario 1: Using PPMs. Fraction of infected agents (averaged over all simulation runs) at the peak ($inf_{peak}$) and the lowest value ($inf_{lower}$) of the infection, and the end of the simulation ($inf_{last}$) when $\alpha_p$ agents use PPMs. Each value is followed by the standard deviation. In all scenarios, the damage of the intervention is 0.0.

| Data set | $\alpha_p$ | $\text{inf}_{\text{last}}$ | $\text{inf}_{\text{peak}}$ | $\text{inf}_{\text{lower}}$ |
|---|---|---|---|---|
| BLEBeacon | 0.00 | $0.91_{\pm 0.04}$ | $0.96_{\pm 0.02}$ | $0.77_{\pm 0.04}$ |
| | 0.25 | $0.89_{\pm 0.04}$ | $0.95_{\pm 0.03}$ | $0.74_{\pm 0.04}$ |
| | 0.50 | $0.85_{\pm 0.05}$ | $0.93_{\pm 0.04}$ | $0.73_{\pm 0.05}$ |
| | 0.75 | $0.83_{\pm 0.05}$ | $0.90_{\pm 0.03}$ | $0.70_{\pm 0.05}$ |
| | 1.00 | $0.84_{\pm 0.06}$ | $0.90_{\pm 0.04}$ | $0.62_{\pm 0.06}$ |
| Foursquare | 0.00 | $0.60_{\pm 0.16}$ | $0.69_{\pm 0.19}$ | $0.57_{\pm 0.16}$ |
| | 0.25 | $0.60_{\pm 0.12}$ | $0.71_{\pm 0.14}$ | $0.58_{\pm 0.11}$ |
| | 0.50 | $0.54_{\pm 0.11}$ | $0.70_{\pm 0.14}$ | $0.53_{\pm 0.10}$ |
| | 0.75 | $0.48_{\pm 0.14}$ | $0.70_{\pm 0.21}$ | $0.48_{\pm 0.14}$ |
| | 1.00 | $0.39_{\pm 0.07}$ | $0.67_{\pm 0.11}$ | $0.39_{\pm 0.07}$ |
| GoT | 0.00 | $0.94_{\pm 0.01}$ | $0.94_{\pm 0.01}$ | $0.94_{\pm 0.01}$ |
| | 0.25 | $0.93_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.93_{\pm 0.01}$ |
| | 0.50 | $0.92_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.91_{\pm 0.01}$ |
| | 0.75 | $0.89_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.89_{\pm 0.01}$ |
| | 1.00 | $0.87_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.87_{\pm 0.01}$ |

**TABLE 3.** Scenario 2: Using EMs. Fraction of infected (averaged over all simulation runs) at the peak ($inf_{peak}$) and the lowest value ($inf_{lower}$) of the infection, and the end of the simulation ($inf_{last}$). Each value is followed by the standard deviation. The table also reports whether contagions are only due to indirect contacts or to both direct and indirect contacts (*Type of contagion*) and whether sanitization procedures are in place (*sanitize*). In all scenarios, the damage of the intervention is 0.0.

| Data set | Type of contagion | sanitize | $\text{inf}_{\text{last}}$ | $\text{inf}_{\text{peak}}$ | $\text{inf}_{\text{lower}}$ |
|---|---|---|---|---|---|
| BLEBeacon | Direct/Indirect | false | $0.90_{\pm 0.04}$ | $0.97_{\pm 0.03}$ | $0.78_{\pm 0.04}$ |
| | | true | $0.87_{\pm 0.05}$ | $0.93_{\pm 0.03}$ | $0.75_{\pm 0.05}$ |
| | Indirect | false | $0.92_{\pm 0.03}$ | $0.95_{\pm 0.03}$ | $0.78_{\pm 0.04}$ |
| | | true | $0.83_{\pm 0.05}$ | $0.87_{\pm 0.04}$ | $0.74_{\pm 0.05}$ |
| Foursquare | Direct/Indirect | false | $0.66_{\pm 0.17}$ | $0.76_{\pm 0.19}$ | $0.63_{\pm 0.16}$ |
| | | true | $0.62_{\pm 0.15}$ | $0.74_{\pm 0.19}$ | $0.58_{\pm 0.15}$ |
| | Indirect | false | $0.62_{\pm 0.14}$ | $0.70_{\pm 0.16}$ | $0.59_{\pm 0.14}$ |
| | | true | $0.55_{\pm 0.13}$ | $0.67_{\pm 0.15}$ | $0.53_{\pm 0.12}$ |
| GoT | Direct/Indirect | false | $0.95_{\pm 0.01}$ | $0.96_{\pm 0.01}$ | $0.95_{\pm 0.01}$ |
| | | true | $0.94_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.93_{\pm 0.01}$ |
| | Indirect | false | $0.85_{\pm 0.20}$ | $0.86_{\pm 0.20}$ | $0.85_{\pm 0.20}$ |
| | | true | $0.78_{\pm 0.20}$ | $0.85_{\pm 0.21}$ | $0.76_{\pm 0.20}$ |

damage $\mathcal{D}_a$ assumes lower values for the GoT data set. Once again, we have to recall the nature of the data set: most of the GoT characters meet a reduced number of other characters, thus explaining the lower values. A similar comment holds for the commuting damage $\mathcal{D}_l$.

### D. SCENARIO 4: USING SDMs - QUARANTINE

#### 1) EXPERIMENTAL SETTING

Similar to the previous scenario, in this experiment, we study the sensitivity of the model to the implementation of quarantine measures, regulated by the parameter $\beta_{quarantine}$. This parameter represents the willingness of an agent to enter the quarantine state, and it is proportional to the number of infected agents met. As for isolation measures, modifying the parameter $\beta_{quarantine}$ means regulating when an infected agent enters the quarantine state (the higher, the sooner). The time spent by each agent in this state is strictly dependent on i) whether the agent is infected and ii) the recovery probability of the agents. If an agent has been quarantined even though being susceptible, then it will exit the quarantine status immediately in the next simulation step. Otherwise, the same rules applied for exiting the isolation state hold. It is worth recalling that our experiments simulate an SIS compartmental epidemic model, where there is no latency between the infection and the actual manifest of symptoms. The absence of an Exposed state translates into safely assuming that if an agent gets infected in a time step $t$, it will enter the Infected state in the following time step $(t + 1)$. Hence, an agent may leave the quarantine state after a time window $\Delta$ (4 hours) only if it is susceptible; otherwise, it will remain quarantined until it heals. Clearly, in a compartmental model including the Exposed state, the quarantine should last more than the time required by the symptoms to become evident or to have a test result.

**TABLE 4.** Scenario 3: Using SDMs - Isolation. Fraction of infected but not isolated agents (averaged over all simulation runs) at the peak ($inf_{peak}$) and the lowest value ($inf_{lower}$) of the infection, and the end of the simulation ($inf_{last}$). The column $isolated_{last}$ reports the average fraction of isolated agents at the end of the simulation. $\mathcal{D}_a$ and $\mathcal{D}_l$ represent the social and commuting damage of the intervention. Each value is followed by the standard deviation.

| Data set | $\beta_{isolation}$ | $\text{inf}_{\text{last}}$ | $\text{inf}_{\text{peak}}$ | $\text{inf}_{\text{lower}}$ | $\text{isolated}_{\text{last}}$ | $\mathcal{D}_a$ | $\mathcal{D}_l$ |
|---|---|---|---|---|---|---|---|
| BLEBeacon | 0.00 | $0.91_{\pm 0.04}$ | $0.96_{\pm 0.03}$ | $0.79_{\pm 0.04}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | 0.05 | $0.28_{\pm 0.06}$ | $0.79_{\pm 0.03}$ | $0.25_{\pm 0.07}$ | $0.62_{\pm 0.06}$ | $0.37_{\pm 0.07}$ | $0.15_{\pm 0.04}$ |
| | 0.10 | $0.19_{\pm 0.05}$ | $0.80_{\pm 0.04}$ | $0.12_{\pm 0.05}$ | $0.73_{\pm 0.06}$ | $0.51_{\pm 0.05}$ | $0.26_{\pm 0.04}$ |
| | 0.20 | $0.12_{\pm 0.04}$ | $0.80_{\pm 0.03}$ | $0.06_{\pm 0.04}$ | $0.78_{\pm 0.05}$ | $0.59_{\pm 0.05}$ | $0.32_{\pm 0.05}$ |
| | 0.30 | $0.07_{\pm 0.04}$ | $0.82_{\pm 0.03}$ | $0.04_{\pm 0.04}$ | $0.85_{\pm 0.06}$ | $0.66_{\pm 0.04}$ | $0.37_{\pm 0.04}$ |
| | 0.40 | $0.06_{\pm 0.03}$ | $0.81_{\pm 0.03}$ | $0.03_{\pm 0.02}$ | $0.84_{\pm 0.05}$ | $0.67_{\pm 0.04}$ | $0.37_{\pm 0.05}$ |
| | 0.50 | $0.05_{\pm 0.03}$ | $0.80_{\pm 0.04}$ | $0.03_{\pm 0.02}$ | $0.87_{\pm 0.05}$ | $0.68_{\pm 0.04}$ | $0.38_{\pm 0.05}$ |
| Foursquare | 0.00 | $0.62_{\pm 0.07}$ | $0.72_{\pm 0.09}$ | $0.60_{\pm 0.08}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | 0.05 | $0.19_{\pm 0.05}$ | $0.69_{\pm 0.18}$ | $0.17_{\pm 0.16}$ | $0.41_{\pm 0.10}$ | $0.42_{\pm 0.11}$ | $0.28_{\pm 0.07}$ |
| | 0.10 | $0.11_{\pm 0.01}$ | $0.69_{\pm 0.08}$ | $0.08_{\pm 0.01}$ | $0.47_{\pm 0.06}$ | $0.55_{\pm 0.06}$ | $0.36_{\pm 0.04}$ |
| | 0.20 | $0.05_{\pm 0.02}$ | $0.68_{\pm 0.19}$ | $0.03_{\pm 0.01}$ | $0.49_{\pm 0.14}$ | $0.62_{\pm 0.18}$ | $0.41_{\pm 0.12}$ |
| | 0.30 | $0.03_{\pm 0.01}$ | $0.68_{\pm 0.16}$ | $0.02_{\pm 0.01}$ | $0.49_{\pm 0.10}$ | $0.66_{\pm 0.15}$ | $0.43_{\pm 0.10}$ |
| | 0.40 | $0.03_{\pm 0.01}$ | $0.69_{\pm 0.14}$ | $0.02_{\pm 0.00}$ | $0.50_{\pm 0.10}$ | $0.68_{\pm 0.14}$ | $0.45_{\pm 0.10}$ |
| | 0.50 | $0.02_{\pm 0.00}$ | $0.71_{\pm 0.01}$ | $0.01_{\pm 0.00}$ | $0.50_{\pm 0.06}$ | $0.70_{\pm 0.08}$ | $0.47_{\pm 0.05}$ |
| GoT | 0.00 | $0.95_{\pm 0.01}$ | $0.96_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | 0.05 | $0.31_{\pm 0.02}$ | $0.95_{\pm 0.01}$ | $0.31_{\pm 0.02}$ | $0.62_{\pm 0.02}$ | $0.18_{\pm 0.02}$ | $0.02_{\pm 0.00}$ |
| | 0.10 | $0.19_{\pm 0.02}$ | $0.95_{\pm 0.01}$ | $0.18_{\pm 0.02}$ | $0.73_{\pm 0.02}$ | $0.28_{\pm 0.02}$ | $0.04_{\pm 0.01}$ |
| | 0.20 | $0.10_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.10_{\pm 0.01}$ | $0.79_{\pm 0.02}$ | $0.38_{\pm 0.08}$ | $0.07_{\pm 0.01}$ |
| | 0.30 | $0.07_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.07_{\pm 0.01}$ | $0.80_{\pm 0.02}$ | $0.43_{\pm 0.02}$ | $0.09_{\pm 0.01}$ |
| | 0.40 | $0.06_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.06_{\pm 0.01}$ | $0.80_{\pm 0.02}$ | $0.46_{\pm 0.02}$ | $0.11_{\pm 0.01}$ |
| | 0.50 | $0.05_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.05_{\pm 0.01}$ | $0.80_{\pm 0.02}$ | $0.47_{\pm 0.01}$ | $0.12_{\pm 0.01}$ |

We ranged the parameter in the interval [0, 1]. Note that the scenario with $\beta_{quarantine} = 0.0$ corresponds to the unmitigated scenario when no measures are being applied.

#### 2) DISCUSSION

Table 5 reports the results of this scenario ranging the parameter $\beta_{quarantine} \in \{0.00, 0.05, 0.10, 0.20, 0.60, 0.90\}$. Also in this case, $inf_{last}$, $inf_{peak}$, and $inf_{lower}$ do not report the total fraction of infected, but the fraction of agents that can still spread the infection, i.e., the agents that are sick but not quarantined. As before, increasing the value of $\beta_{quarantine}$ helps reducing the fraction of infected (see the columns $inf_{lower}$ and $inf_{last}$). However, even though we can observe a drastic drop of this value, the epidemic still has a consistent probability of spreading. This behavior is due to the nature of the intervention itself: quarantining individuals is a preventive measure, and, as such, even susceptible agents may enter this state. On the contrary, isolation measures directly target and isolate infected individuals. For this reason, in the first case, we have a lower, even still considerable, efficacy in reducing the fraction of infected. Once again, in a real-world scenario, the effectiveness of such intervention closely relies on the time window during which a person may infect other people but still does manifest any symptoms and the possibility of observing asymptomatic spreaders.

In this scenario, we can note very low values for $\mathcal{D}_a$ and $\mathcal{D}_l$ even when $\beta_{quarantine} = 0.9$. This outcome is well explained by how the intervention works. As already discussed, both susceptible and infected agents may be quarantined. When a susceptible enters the quarantine, it will exit the state in the following iteration; thus, the intervention causes negligible damage to those agents, still protecting them from getting the infection - even if for a small time window. The combination of these two elements (quarantining infected and protecting susceptible agents) ensures a reasonable trade-off between the efficacy of the intervention in reducing the number of infected and the damage brought to the population.

### E. SCENARIO 5: USING SDMs - TRACING CONTACTS

#### 1) EXPERIMENTAL SETTING

In this scenario, we evaluate the sensitivity of the model to the introduction of a tracing application as a control strategy to inform each agent whether and how many infected it has met in the previous simulation steps. The parameter $\alpha_i$, varying in the range [0, 1], regulates the fraction of the population adopting the tracing technology. The scenario with $\alpha_i = 0.0$ corresponds to the unmitigated scenario when no measures are being applied. The parameter $\beta_{tracing}$ controls the probability that an agent enters the quarantine state, based on the number of infected the agent knows it has been in contact. We fixed $\beta_{tracing} = 0.6$ to model that if an agent uses a tracing application, it will be more likely to enter the quarantine state if needed. As expected, we traced the contacts between individuals only if both agents are using the tracing measure.

#### 2) DISCUSSION

Table 6 shows the results for $\alpha_i \in \{0.00, 0.25, 0.50, 0.75, 1.00\}$. In line with the previously simulated interventions, the more agents adopt the measure, the more the intervention effectively reduces the fraction of infected. When the whole population adopts the tracing technology, the epidemic trend drastically drops in all scenarios. Nevertheless, at least 50% of the agents must use the application to observe a significant reduction. Comparing this experiment with the

**TABLE 5.** Scenario 4: Using SDMs - Quarantine. Fraction of infected but not quarantined agents (averaged over all simulation runs) at the peak ($inf_{peak}$) and the lowest value ($inf_{lower}$) of the infection, and the end of the simulation ($inf_{last}$). The column $quarantined_{last}$ reports the average fraction of quarantined agents at the end of the simulation. $\mathcal{D}_a$ and $\mathcal{D}_l$ represent the social and commuting damage of the intervention. Each value is followed by the standard deviation.

| Data set | $\beta quarantine$ | $\text{inf}_{\text{last}}$ | $\text{inf}_{\text{peak}}$ | $\text{inf}_{\text{lower}}$ | $\text{quarantined}_{\text{last}}$ | $\mathcal{D}_a$ | $\mathcal{D}_l$ |
|---|---|---|---|---|---|---|---|
| BLEBeacon | 0.00 | 0.92±0.04 | 0.95±0.03 | 0.78±0.03 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| | 0.05 | 0.75±0.05 | 0.81±0.03 | 0.72±0.05 | 0.15±0.05 | 0.07±0.02 | 0.02±0.01 |
| | 0.10 | 0.68±0.05 | 0.79±0.03 | 0.68±0.05 | 0.23±0.05 | 0.09±0.03 | 0.02±0.01 |
| | 0.20 | 0.67±0.06 | 0.80±0.03 | 0.62±0.05 | 0.26±0.05 | 0.11±0.03 | 0.02±0.01 |
| | 0.60 | 0.59±0.06 | 0.81±0.04 | 0.58±0.06 | 0.32±0.06 | 0.14±0.03 | 0.03±0.01 |
| | 0.90 | 0.57±0.06 | 0.81±0.03 | 0.57±0.07 | 0.33±0.05 | 0.13±0.03 | 0.03±0.01 |
| Foursquare | 0.00 | 0.64±0.13 | 0.74±0.15 | 0.62±0.12 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| | 0.05 | 0.58±0.16 | 0.67±0.18 | 0.56±0.15 | 0.03±0.01 | 0.04±0.01 | 0.02±0.00 |
| | 0.10 | 0.57±0.07 | 0.68±0.08 | 0.55±0.07 | 0.05±0.01 | 0.06±0.01 | 0.03±0.01 |
| | 0.20 | 0.56±0.06 | 0.70±0.01 | 0.54±0.06 | 0.08±0.01 | 0.10±0.01 | 0.05±0.01 |
| | 0.60 | 0.47±0.09 | 0.69±0.14 | 0.47±0.09 | 0.14±0.03 | 0.17±0.03 | 0.08±0.02 |
| | 0.90 | 0.45±0.07 | 0.70±0.11 | 0.44±0.07 | 0.16±0.03 | 0.19±0.03 | 0.10±0.02 |
| GoT | 0.00 | 0.95±0.01 | 0.96±0.01 | 0.95±0.01 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| | 0.05 | 0.80±0.02 | 0.95±0.01 | 0.80±0.02 | 0.14±0.01 | 0.05±0.01 | 0.00±0.00 |
| | 0.10 | 0.70±0.02 | 0.95±0.01 | 0.70±0.02 | 0.22±0.01 | 0.07±0.01 | 0.01±0.00 |
| | 0.20 | 0.57±0.02 | 0.94±0.01 | 0.57±0.02 | 0.32±0.01 | 0.10±0.02 | 0.01±0.00 |
| | 0.60 | 0.36±0.02 | 0.95±0.01 | 0.36±0.02 | 0.47±0.01 | 0.15±0.02 | 0.07±0.00 |
| | 0.90 | 0.30±0.02 | 0.95±0.01 | 0.30±0.02 | 0.52±0.01 | 0.16±0.02 | 0.02±0.00 |

quarantine scenario when $\beta_{quarantine} = 0.6$ (see Table 5), we can note a higher fraction of quarantined agents on average at the end of the simulation. This outcome is explained by how the different interventions work in the simulation. An agent may decide to enter the quarantine state according to both the probability $\beta_{quarantine}$ and the number of infected agents met during each interval $\Delta$. When the agent uses the tracing application, it may enter the quarantine still based on the probability $\beta_{quarantine}$ but, this time, the number of infected agents is evaluated over the previous intervals (starting from the interval when the intervention is applied). Hence, the number of infected agents met is generally higher in this second case and, consequently, the overall probability of entering the quarantine state.

Consistently with the previous experiment on quarantining individuals (see Section VI-D), the social damage $\mathcal{D}_a$ and especially the commuting damage $\mathcal{D}_l$ assume low if not negligible values. The explanation for that is the same as discussed in the quarantine scenario. These results suggest the potential impact such measure can have, even though, in the real world, policymakers have to consider a plethora of constraints when implementing similar interventions, among all privacy concerns.

### F. SCENARIO 6: USING SDMs - AVOIDING CROWDING

#### 1) EXPERIMENTAL SETTING

In this validation scenario, we examine the sensitivity of the model when avoiding crowding measures are applied. As described in Section IV-C4, this intervention consists in increasing the social distancing between individuals, which implicitly means reducing the number of possible direct interactions. To implement such a policy for the BLEBeacon data set, we reduced the number of agents who could access the building during a day by simulating only half of the entire population's movements. We applied a similar approach for the Foursquare and the GoT data sets, halving the number of agents that could access a given location.

#### 2) DISCUSSION

Table 7 reports the results of this experiment. As for the EMs (see Section VI-B), we analyze the epidemic spreading reporting the number of infected i) considering both direct and indirect contagion pathways and ii) focusing only on the contribution made by direct contacts. Also in this scenario, when we look at the final fraction of infected due to both types of interactions, we can observe little or no impact on the epidemic spreading. On the other hand, we can note that the introduction of such a measure effectively lowers the fraction of infected due to direct contacts and the lowest peak of the infection. This outcome is somewhat expected given the nature of the intervention itself. This measure does not affect interactions between people and environments; hence the pathogen is free to propagate via indirect contacts.

The intervention cost in terms of damage is tendentially higher than the other scenarios described except for the isolation measure. These values are clearly explained by the fact that some agents cannot enter a location if this already contains a number of agents equal to half of its total capacity (evaluated over the unmitigated scenario).

### G. SCENARIO 7: USING SDMs - LOCATION CLOSURE

#### 1) EXPERIMENTAL SETTING

In this last validation scenario, we assess the model's sensitivity to the implementation of lockdown measures (see Section IV-C5). We selected the locations to close according to two possible scenarios. In the first scenario, we randomly chose the locations; in the second configuration, we picked the most crowded places first. The parameter $\alpha_e$ regulates the number of locations to close. We ranged it in the interval [0, 1] to simulate partial and complete closure policies.

**TABLE 6.** Scenario 5: Tracing. Fraction of infected but not quarantined agents (averaged over all simulation runs) at the peak ($inf_{peak}$) and the lowest value ($inf_{lower}$) of the infection, and the end of the simulation ($inf_{last}$) when $\alpha_{tracing}$ agents use a tracing application. The column $quarantined_{last}$ reports the average fraction of quarantined agents at the end of the simulation. $\mathcal{D}_a$ and $\mathcal{D}_l$ represent the social and commuting damage of the intervention. Each value is followed by the standard deviation.

| Data set | $\alpha_{tracing}$ | $\text{inf}_{\text{last}}$ | $\text{inf}_{\text{peak}}$ | $\text{inf}_{\text{lower}}$ | $\text{quarantined}_{\text{last}}$ | $\mathcal{D}_a$ | $\mathcal{D}_l$ |
|---|---|---|---|---|---|---|---|
| BLEBeacon | 0.00 | 0.94±0.03 | 0.96±0.03 | 0.77±0.04 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| | 0.25 | 0.85±0.05 | 0.88±0.05 | 0.78±0.04 | 0.06±0.05 | 0.03±0.02 | 0.00±0.00 |
| | 0.50 | 0.57±0.07 | 0.81±0.03 | 0.57±0.07 | 0.27±0.06 | 0.10±0.03 | 0.01±0.01 |
| | 0.75 | 0.34±0.07 | 0.81±0.03 | 0.34±0.07 | 0.49±0.07 | 0.20±0.03 | 0.03±0.01 |
| | 1.00 | 0.06±0.06 | 0.81±0.03 | 0.06±0.06 | 0.68±0.07 | 0.32±0.04 | 0.06±0.01 |
| Foursquare | 0.00 | 0.66±0.19 | 0.76±0.22 | 0.63±0.18 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| | 0.25 | 0.61±0.12 | 0.72±0.14 | 0.59±0.12 | 0.04±0.01 | 0.04±0.01 | 0.01±0.00 |
| | 0.50 | 0.49±0.10 | 0.68±0.14 | 0.49±0.10 | 0.12±0.03 | 0.12±0.02 | 0.04±0.01 |
| | 0.75 | 0.37±0.09 | 0.70±0.18 | 0.37±0.09 | 0.22±0.06 | 0.23±0.06 | 0.08±0.02 |
| | 1.00 | 0.22±0.01 | 0.71±0.01 | 0.22±0.01 | 0.34±0.01 | 0.35±0.01 | 0.13±0.00 |
| GoT | 0.00 | 0.95±0.01 | 0.95±0.01 | 0.95±0.01 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| | 0.25 | 0.85±0.01 | 0.95±0.01 | 0.84±0.01 | 0.10±0.01 | 0.04±0.01 | 0.00±0.00 |
| | 0.50 | 0.64±0.01 | 0.95±0.01 | 0.64±0.01 | 0.27±0.01 | 0.11±0.01 | 0.01±0.00 |
| | 0.75 | 0.38±0.02 | 0.95±0.01 | 0.38±0.02 | 0.47±0.02 | 0.18±0.01 | 0.01±0.00 |
| | 1.00 | 0.10±0.02 | 0.95±0.01 | 0.10±0.02 | 0.66±0.02 | 0.28±0.02 | 0.02±0.00 |

The scenario with $\alpha_e = 0.0$ corresponds to the baseline scenario when no measures are being applied.

### 2) DISCUSSION

Table 8 reports the results of the application of the two tested lockdown scenarios, simulating an epidemic spreading varying the parameter $\alpha_e \in \{0.00, 0.30, 0.60, 0.90, 1.00\}$.

If we look at the final fraction of infected at the end of the simulation (column $inf_{last}$), we can note how such measure has a different impact according to the nature of the data set. For instance, we need to deny access to all rooms within the building described by the BLEBeacon data if we want to reduce the contagions drastically. In addition, closing the locations giving priority to crowded places does not seem to improve the performance over closing random rooms when $\alpha_e \leq 0.90$. Although surprising at first glance, the explanation for this outcome is due to the fact that the number of check-ins per location is evaluated over the whole data set and not only over the portion when the intervention is applied. Thus, the most crowded locations may contain the most check-ins at the start of the simulation rather than the end, leading to the closure of less crowded rooms when the intervention is applied. On the other hand, we can observe a completely different picture for the Foursquare data set, in which closing 30% of the most crowded locations leads to a notably decreasing in the fraction of infected. Once again, this outcome is due to the characteristic of the data. The majority of the check-ins of the original data set happens in a limited amount of places, like transportation or general entertainment. Hence, even closing a limited number of places brings down the epidemic. We have a similar outcome for the GoT scenario in which closing the most crowded locations generally achieves better results in lowering the fraction of infected than closing random locations. The results of these two scenarios are aligned with the measures we would expect in a real-life scenario. In fact, if we consider lockdown policies issued for the COVID-19 pandemic, we can note that they usually tended to penalize aggregation and leisure places. Clearly, we can note very high values for the commuting damage $\mathcal{D}_l$ due to the locations closed by the intervention. As a consequence, also the social damage $\mathcal{D}_a$ assumes high values when $\alpha_e \geq 0.50$ as agents cannot meet other agents in a closed place. When all locations are closed, we have the maximum damage.

## VII. THE EFFECT OF COMBINING NPIs ON THE EPIDEMIC DYNAMICS

In Section VI, we investigated the model output when a single intervention is applied. Other than evaluating the model sensitivity, we thus explored, at the same time, the effects of applying a single NPI in terms of effectively reducing the number of infected and costs required to implement the given measure. However, we expect to see the application of different interventions combined to contain the epidemic spreading in a real-world scenario. Which NPIs should be applied or how strict the measures should be implemented closely depends upon the gravity of the current situation, like the pressure on the hospitalization system. For instance, we experienced rigid lockdown policies during the first and second waves of the COVID-19 pandemic, while even the use of face masks was lifted during summer 2020.

Current literature on the topic investigates the application of NPIs based on the complexity and heterogeneity of the data fed into the ABM. When the model is highly detailed, like in the case of France [35], Boston [3], Seattle, and New York [4], the simulation usually focuses on reproducing and improving the measures really implemented by the government. In the case of simpler models, other works manually combine several NPIs and examine their outcomes [13], [59], [73]. In this work, we tackle the issue of identifying feasible combinations of NPIs by approaching the problem under an optimization framework with two contrasting objectives - the fraction of infected and the damage brought by the intervention. The optimal solutions (in terms of NPI combinations)

**TABLE 7.** Scenario 6: Using SDMs - Avoiding Crowding. Fraction of infected (averaged over all simulation runs) at the peak ($inf_{peak}$) and the lowest value ($inf_{lower}$) of the infection, and the end of the simulation ($inf_{last}$). $\mathcal{D}_a$ and $\mathcal{D}_l$ represent the social and commuting damage of the intervention. Each value is followed by the standard deviation. The table also reports whether avoiding crowding measures are applied (*avoiding crowding*).

| Data set | Type of contagion | Avoiding crowding | $\text{inf}_{\text{last}}$ | $\text{inf}_{\text{peak}}$ | $\text{inf}_{\text{lower}}$ | $\mathcal{D}_a$ | $\mathcal{D}_l$ |
|---|---|---|---|---|---|---|---|
| BLEBeacon | Direct/Indirect | false | $0.93_{\pm 0.04}$ | $0.96_{\pm 0.03}$ | $0.77_{\pm 0.04}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | | true | $0.89_{\pm 0.05}$ | $0.90_{\pm 0.04}$ | $0.76_{\pm 0.04}$ | $0.65_{\pm 0.00}$ | $0.13_{\pm 0.00}$ |
| | Direct | false | $0.75_{\pm 0.05}$ | $0.86_{\pm 0.04}$ | $0.46_{\pm 0.06}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | | true | $0.65_{\pm 0.04}$ | $0.72_{\pm 0.03}$ | $0.31_{\pm 0.04}$ | $0.55_{\pm 0.00}$ | $0.13_{\pm 0.00}$ |
| Foursquare | Direct/Indirect | false | $0.68_{\pm 0.20}$ | $0.78_{\pm 0.24}$ | $0.65_{\pm 0.20}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | | true | $0.56_{\pm 0.14}$ | $0.69_{\pm 0.18}$ | $0.54_{\pm 0.14}$ | $0.84_{\pm 0.00}$ | $0.12_{\pm 0.00}$ |
| | Direct | false | $0.33_{\pm 0.14}$ | $0.48_{\pm 0.20}$ | $0.33_{\pm 0.14}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | | true | $0.14_{\pm 0.05}$ | $0.41_{\pm 0.15}$ | $0.14_{\pm 0.05}$ | $0.84_{\pm 0.00}$ | $0.12_{\pm 0.00}$ |
| GoT | Direct/Indirect | false | $0.95_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | | true | $0.93_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.93_{\pm 0.01}$ | $0.40_{\pm 0.00}$ | $0.01_{\pm 0.00}$ |
| | Direct | false | $0.92_{\pm 0.01}$ | $0.92_{\pm 0.01}$ | $0.90_{\pm 0.01}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | | true | $0.73_{\pm 0.01}$ | $0.91_{\pm 0.01}$ | $0.73_{\pm 0.01}$ | $0.41_{\pm 0.00}$ | $0.01_{\pm 0.00}$ |

**TABLE 8.** Scenario 7: Using SDMs - Location closure. Fraction of infected (averaged over all simulation runs) at the peak ($inf_{peak}$) and the lowest value ($inf_{lower}$) of the infection, and the end of the simulation ($inf_{last}$). $\mathcal{D}_a$ and $\mathcal{D}_l$ represent the social and commuting damage of the intervention. Each value is followed by the standard deviation. The table also reports the policy according to which locations are closed (*Sorting*).

| Data set | $\alpha_e$ | Sorting | $\text{inf}_{\text{last}}$ | $\text{inf}_{\text{peak}}$ | $\text{inf}_{\text{lower}}$ | $\mathcal{D}_a$ | $\mathcal{D}_l$ |
|---|---|---|---|---|---|---|---|
| BLEBeacon | 0.00 | - | $0.93_{\pm 0.03}$ | $0.95_{\pm 0.03}$ | $0.78_{\pm 0.03}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | 0.30 | random | $0.82_{\pm 0.04}$ | $0.91_{\pm 0.03}$ | $0.66_{\pm 0.06}$ | $0.10_{\pm 0.00}$ | $0.38_{\pm 0.00}$ |
| | 0.60 | random | $0.79_{\pm 0.04}$ | $0.88_{\pm 0.04}$ | $0.60_{\pm 0.06}$ | $0.25_{\pm 0.00}$ | $0.67_{\pm 0.00}$ |
| | 0.90 | random | $0.64_{\pm 0.05}$ | $0.80_{\pm 0.03}$ | $0.45_{\pm 0.06}$ | $0.69_{\pm 0.00}$ | $0.91_{\pm 0.00}$ |
| | 0.30 | most crowded | $0.90_{\pm 0.04}$ | $0.96_{\pm 0.03}$ | $0.73_{\pm 0.05}$ | $0.01_{\pm 0.00}$ | $0.32_{\pm 0.00}$ |
| | 0.60 | most crowded | $0.81_{\pm 0.05}$ | $0.87_{\pm 0.03}$ | $0.62_{\pm 0.06}$ | $0.28_{\pm 0.00}$ | $0.67_{\pm 0.00}$ |
| | 0.90 | most crowded | $0.50_{\pm 0.05}$ | $0.82_{\pm 0.03}$ | $0.35_{\pm 0.07}$ | $0.86_{\pm 0.00}$ | $0.96_{\pm 0.00}$ |
| | 1.00 | - | $0.10_{\pm 0.04}$ | $0.80_{\pm 0.03}$ | $0.10_{\pm 0.04}$ | $1.00_{\pm 0.00}$ | $1.00_{\pm 0.00}$ |
| Foursquare | 0.00 | - | $0.66_{\pm 0.11}$ | $0.76_{\pm 0.12}$ | $0.63_{\pm 0.10}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | 0.30 | random | $0.54_{\pm 0.13}$ | $0.68_{\pm 0.16}$ | $0.53_{\pm 0.12}$ | $0.37_{\pm 0.00}$ | $0.32_{\pm 0.00}$ |
| | 0.60 | random | $0.46_{\pm 0.05}$ | $0.71_{\pm 0.08}$ | $0.45_{\pm 0.05}$ | $0.57_{\pm 0.00}$ | $0.60_{\pm 0.00}$ |
| | 0.90 | random | $0.23_{\pm 0.04}$ | $0.71_{\pm 0.11}$ | $0.23_{\pm 0.04}$ | $0.89_{\pm 0.00}$ | $0.90_{\pm 0.00}$ |
| | 0.30 | most crowded | $0.10_{\pm 0.03}$ | $0.67_{\pm 0.20}$ | $0.10_{\pm 0.03}$ | $0.99_{\pm 0.00}$ | $0.76_{\pm 0.00}$ |
| | 0.60 | most crowded | $0.10_{\pm 0.02}$ | $0.69_{\pm 0.11}$ | $0.10_{\pm 0.02}$ | $1.00_{\pm 0.00}$ | $0.89_{\pm 0.00}$ |
| | 0.90 | most crowded | $0.10_{\pm 0.01}$ | $0.70_{\pm 0.08}$ | $0.101_{\pm 0.0}$ | $1.00_{\pm 0.00}$ | $0.97_{\pm 0.00}$ |
| | 1.00 | - | $0.10_{\pm 0.02}$ | $0.71_{\pm 0.14}$ | $0.10_{\pm 0.02}$ | $1.00_{\pm 0.00}$ | $1.00_{\pm 0.00}$ |
| GoT | 0.00 | - | $0.95_{\pm 0.01}$ | $0.96_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.00_{\pm 0.00}$ | $0.00_{\pm 0.00}$ |
| | 0.30 | random | $0.64_{\pm 0.00}$ | $0.94_{\pm 0.00}$ | $0.64_{\pm 0.00}$ | $0.43_{\pm 0.00}$ | $0.44_{\pm 0.00}$ |
| | 0.60 | random | $0.51_{\pm 0.00}$ | $0.95_{\pm 0.00}$ | $0.51_{\pm 0.00}$ | $0.62_{\pm 0.00}$ | $0.66_{\pm 0.00}$ |
| | 0.90 | random | $0.27_{\pm 0.00}$ | $0.95_{\pm 0.00}$ | $0.27_{\pm 0.00}$ | $0.87_{\pm 0.00}$ | $0.91_{\pm 0.00}$ |
| | 0.30 | most crowded | $0.41_{\pm 0.05}$ | $0.95_{\pm 0.11}$ | $0.41_{\pm 0.05}$ | $0.77_{\pm 0.00}$ | $0.75_{\pm 0.00}$ |
| | 0.60 | most crowded | $0.14_{\pm 0.02}$ | $0.95_{\pm 0.01}$ | $0.14_{\pm 0.02}$ | $0.95_{\pm 0.00}$ | $0.94_{\pm 0.00}$ |
| | 0.90 | most crowded | $0.12_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.12_{\pm 0.01}$ | $1.00_{\pm 0.00}$ | $1.00_{\pm 0.00}$ |
| | 1.00 | - | $0.13_{\pm 0.01}$ | $0.95_{\pm 0.01}$ | $0.13_{\pm 0.01}$ | $1.00_{\pm 0.00}$ | $1.00_{\pm 0.00}$ |

are thus found by exploring the model's parameter space via a multi-objective genetic algorithm. Hence, no handcrafted configuration is required. In this manner, we are able to unbiasedly explore the model's behavior without any assumption about the interaction of model parameters and their effects on the overall simulation. In the following, we will briefly describe the problem of exploring the parameter space of ABMs, the method used for our experiments, and the results obtained. A discussion section in which we review possible real-world implications follows.

### A. MODEL EXPLORATION

ABMs are usually characterized by a large number of controlling parameters and range of parameter values,

a significant amount of computation required to run a model, and a stochastic nature which requires multiple trials to assess the model's behavior [61]. As the number of parameters increases, it becomes unfeasible to manually handle the exploration of the parameter space, which comprises the selection of parameter selection, the simulation run, and the evaluation of the output [15]. The process of i) choosing a starting configurations, ii) running the simulation, iii) evaluating the outcomes, and iv) selecting new candidates is known as Simulation Optimization (SO) process [26], [64]. A SO framework can be formally described as a general optimization problem whose goal is to find a setting of controllable parameters that minimizes a given objective function, i.e.,

$$\min_{\theta \in \Theta} J(\theta), \tag{1}$$

where $\Theta$ is the admissible decision space, $\theta \in \Theta$ is the vector of input variables representing a single configuration, and $J(\theta)$ is the scalar objective function estimated via the simulation. Because of their nature, simulations provide a noisy estimate of $J(\theta)$; for this reason, the most common form for $J$ is an expectation

$$J(\theta) = \mathbb{E}[L(\theta, \epsilon)],$$

where and $L(\cdot)$ is the sample performance measure and $\epsilon$ represents the stochastic effect in the system.

Having formalized the SO process as a general optimization problem, the parameter space search can thus be done via any optimization algorithms [64]. The literature widely explored genetic algorithms (GAs) in this direction, working on parameter-search and exploration in ABMs, as well as on the problem of parameter-search in general [61]. When the optimization problem is multi-objective ($J(\theta) \in \mathcal{R}^n$), multi-objective GAs (MOGAs) come into play, whose central idea lies in generating the optimal Pareto Front in the objective space so that it is not possible to further enhance any fitness function without disturbing the other fitness functions [46]. In this work, we use the algorithm NSGA-II (fast elitist non-dominated sorting genetic algorithm), a Pareto-based MOGA proposed by Deb *et al.* [17]. This algorithm alleviates the problems of lack of elitism, the need of sharing parameters, and high computation complexity characterizing its predecessor NSGA [60], still able to find a diverse set of solutions and converge near the actual Pareto-optimal set.

For an exhaustive description of the SO process and a review on genetic algorithms, we refer the reader to [26], [64] and [42], respectively.

### B. EXPERIMENTAL SETTING

As introduced in the previous paragraphs, we formalize the parameter space exploration of our model under the MOGA framework. Specifically, we used the NSGA-II algorithm and its Julia implementation, available at the following GitHub repository.[2] We examine the multi-objective problem of balancing the use of NPIs to control an epidemic spreading and the negative impact on the overall functioning of society considering as contrasting objectives i) reducing the final fraction of infected at the end of the simulation, ii) the social damage $\mathcal{D}_a$, and iii) the commuting damage $\mathcal{D}_l$.

Each individual (i.e., a setting) is described by a 7-element vector, where each item, ranging in the interval [0, 1], represents whether or not a given NPI is applied and to what extent. We considered the same measures described in Section IV and analyzed in Section VI, namely:

- *PPMs*. Measure regulated by the parameter $\alpha_p$, describing the fraction of agents using PPMs.
- *EMs*. Measure regulated by the boolean parameter *sanitize*, representing whether locations are regularly sanitized.
- *SDMs*
  - *Isolation*. Intervention regulated by the parameter $\beta_{isolation}$, representing the willingness of an agent to enter the isolation state if infected.
  - *Quarantine*. Intervention regulated by the parameter $\beta_{quarantine}$, describing the willingness of an agent to enter the quarantine state.
  - *Tracing*. Measure regulated by the parameter $\alpha_i$, indicating the fraction of the population adopting tracing technologies.
  - *Avoiding crowding*. Measure regulated by the boolean parameter *avoiding crowding*, representing whether avoiding crowding measures are applied.
  - *Lockdown*. Intervention regulated by the parameter $\alpha_e$, describing the fraction of locations to close. We picked the most crowded places first, given the best outcome in terms of reduction on the fraction of infected evaluated in Section VI-G.

The boolean parameters are considered `true` in the simulation if the value of the corresponding item in the individual is at least equal to 0.5.

We run NSGA-II for 100 generations, using a population of 100 individuals. The remaining parameters of the simulation are described in Section V.

### C. RESULTS

In this section, we go through the outcomes of our analysis, first observing the Pareto front output of the algorithm NSGA-II. We then zoom into the characteristics of the solution nearest to the ideal point, trivially identified with the origin (0, 0, 0), indicating that there are no infected and no damage due to the interventions. This solution is further compared with the two NPI configurations reaching the lowest number of infected and the lowest damage. We show the characteristics of each configuration in a radar chart, in which we report all interventions counterclockwise according to their damage. The color of the external circular shape (yellow-orange-red) encodes the damage of each intervention: the darker, the higher. Visually, low-damage solutions tend to pick interventions in the upper half of the circle. On the contrary, high-damage configurations tend to

[2]https://github.com/gsoleilhac/NSGAII.jl

select higher values for interventions located in the lower half of the chart. We finally discuss how these configurations impact the epidemic spreading and the damage paid for their application.

In more detail, every plot highlights three NPI configurations: i) the configuration reaching the lowest number of infected at the end of the simulation (*Lowest-Infected*, depicted in blue), the configuration obtaining the lowest damage (*Lowest-Damage*, depicted in orange), and iii) the configuration nearest to the optimal point $(0, 0, 0)$ according to the euclidean distance (*Nearest-to-Ideal*, depicted in green). We also added to each plot the two settings when no interventions are in place (*No-NPIs*) and when all NPIs are used to the fullest (*Full-NPIs*). All configurations come from the Pareto optimization and lies on the Pareto front analyzed. It is worth noting that the optimization algorithm explores the overall Pareto front; however, the NPI combinations that decision-makers should investigate are the solutions that guarantee the best balance between the three contrasting objectives. As a matter of fact, in this specific context is straightforward to handcraft an intervention that brings no damage (trivially selecting no NPIs or only PPMs or EMs as they provide no damage) or aims to minimize the number of infected (e.g., implementing all NPIs available). Still, none of these interventions are truly useful in a real-world situation. The optimization algorithm can hence guide policymakers towards examining the best NPI combination according to the actual data.

#### 1) THE BLEBeacon SCENARIO

This paragraph discusses the outcome of the BLEBeacon data set, resembling a social event happening in a month timeframe in a university building (see Section V-D). Figure 8 reports the results.

As we can see in Figure 8a, the Pareto front evaluated by the algorithm NSGA-II spans over a broad spectrum of solutions, but, as observed in the introduction of this section, we will focus on the characteristics of the combination nearest to the optimal point (Nearest-to-Ideal). The first element to note is that this NPI configuration only selects policies located in the yellow-orange zone of the radar chart, causing the lowest damage (see Figure 8b(ii)). Among those, we can observe that all agents use PPMs ($\alpha_p = 0.99$) and that locations are regularly cleaned. We can further notice that almost all agents exploit tracing technologies ($\alpha_i = 0.97$) and that the probability of entering the quarantine state is relatively high ($\beta_{quarantine} = 0.63$). This combination translates into an overall significant probability for an agent to self-quarantine itself. It is interesting to observe that the configuration Nearest-to-Ideal can reduce by 70% the final number of infected with the only use of preventive measures. Clearly, more restrictive measures should be used in this kind of environment to reduce the contagion further. The application of this intervention brings to a social damage $\mathcal{D}_a$ of 0.50, meaning that each agent loses, on average, half of the individuals it should have met in normal conditions. The commuting damage $\mathcal{D}_l$ is 0.26, meaning that agents lose, on average, a quarter of the locations they should have visited in normal conditions.

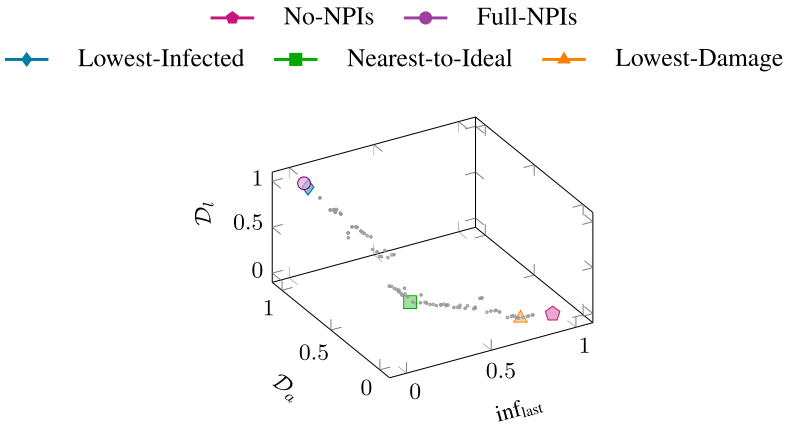
As expected, the configuration conveying to the lowest damage favors zero-damage interventions, like PPMs and EMs (see Figure 8b(iii)). Nevertheless, even though the final fraction of infected is lowered only by a small percentage, we can note that the overall trend (and the lowest peak) after introducing the intervention is generally reduced than the unmitigated scenario (see Figure 8c). Regarding the configuration bringing to the lowest fraction of infected, we can observe that this configuration selects the NPIs with the highest damage, located in the red zone of the radar chart (see Figure 8b(i)). In particular, when this combination is applied, almost all rooms of the building become inaccessible ($\alpha_e = 0.87$), thus preventing the pathogen from spreading and bringing to very serious damage. For this reason, this configuration achieves results comparable to the intervention implementing all protective measures (Full-NPIs).

#### 2) THE FOURSQUARE SCENARIO

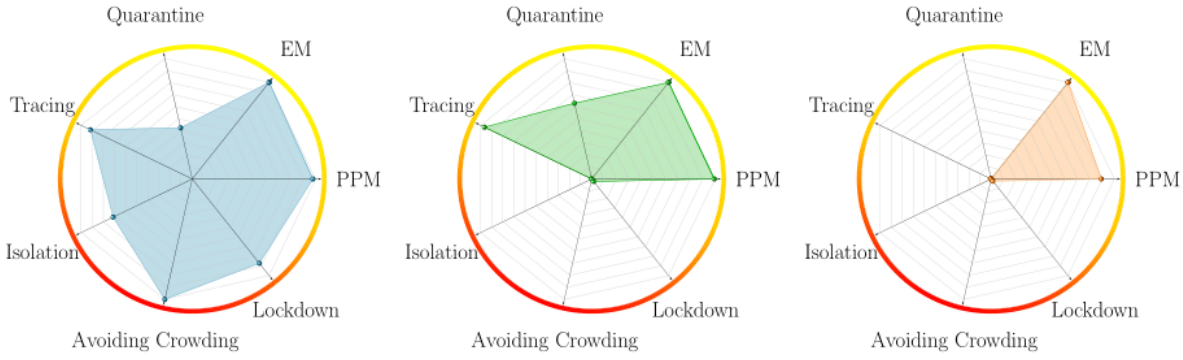
In this paragraph, we discuss the outcomes for the Foursquare data set, resembling real-life movements of users of the Foursquare social network in Tokyo (see Section V-D). Figure 9 reports the results.

As we can note from the Pareto front plotted in Figure 9a, all solutions evaluated by the algorithm encode an NPI combination able to reduce the fraction of infected in the unmitigated scenario by at least 25%, paying highly variable damage. In this case, the Nearest-to-Ideal NPI configuration (see Figure 9b(ii)) appears quite different from the corresponding solution analyzed for the BLEbeacon data set. We can still observe a high value for $\alpha_p = 0.97$, meaning that the majority of the agents adopt PPMs and the use of sanitization measures. However, contrarily to BLEBeacon, very low values are selected for the parameters $\beta_{isolation}$, $\beta_{quarantine}$ and $\alpha_i$, equal to 0.002, 0.06 and 0.013 respectively. Interestingly, this NPI combination can still halve the fraction of infected and hence decrease the epidemic trend (see Figure 9c). At the same time, the damage brought by the intervention is minimal ($\mathcal{D}_a = 0.10$, $\mathcal{D}_l = 0.05$) and in practice due to the reduced percentage of quarantined and isolated agents. The most surprising aspect of this outcome is that despite the lack of lockdown measures, the epidemic trend still constantly diminishes. This result should be contextualized to the sparse nature of the data set; however, it suggests the fundamental role of PPMs and EMs in reducing an epidemic spreading at zero-damage also in extremely crowded places.

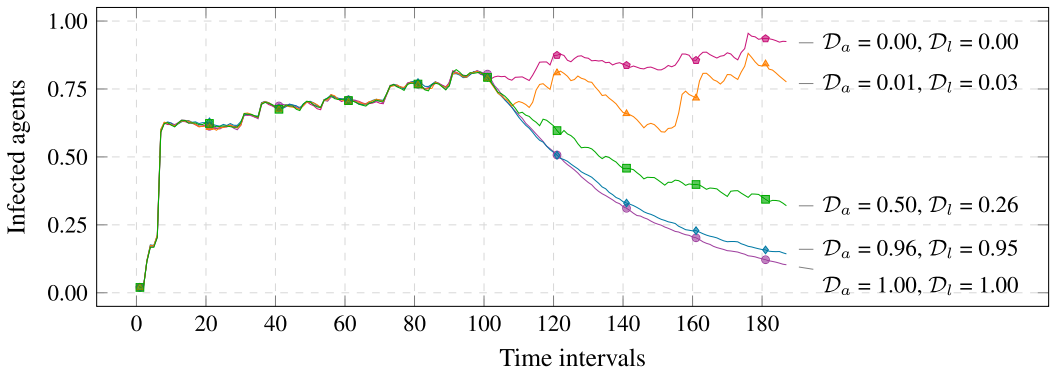
Regarding the configurations at the extremes of the Pareto front, the Lowest-Damage configuration almost corresponds to the Nearest-to-Ideal (see Figure 9b(iii)); while the Lowest-Infected combination (see Figure 9b(i)) reaches the same effectiveness of Full-NPIs, with a lower commuting damage $\mathcal{D}_l$ and a non-significant difference for the social damage $\mathcal{D}_a$ (see Figure 9c).




(a) Pareto front evaluated by the algorithm NSGA-II for the BLEBeacon data set using as objective function: ($\text{inf}_{\text{last}}$, $\mathcal{D}_a$, $\mathcal{D}_l$).




(b) Combination of NPIs, corresponding to the configurations obtaining the lowest number of infected at the end of the simulation (Lowest-Infected, left), the lowest damage (Lowest-Damage, right), and the nearest to the optimal point (Nearest-to-Ideal, center) evaluated over the BLEBeacon data set.



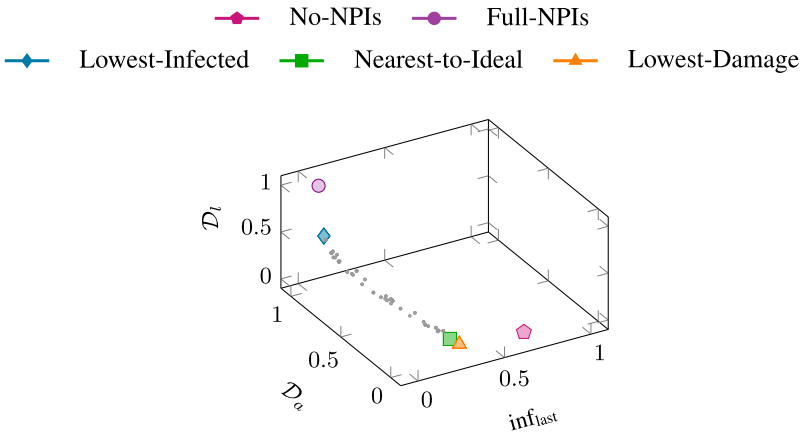

(c) Epidemic trend in the BLEBeacon data set when i) no NPIs are implemented (No-NPIs), ii) all NPIs are in place (Full-NPIs), iii) the configuration bringing to the lowest number of infected ad the end of the simulation is used (Lowest-Infected), iv) the configuration bringing to the lowest damage is applied (Lowest-Damage), and v) the configuration nearest to the optimum is used (Nearest-to-Ideal).

**FIGURE 8.** **BLEBeacon data set.**

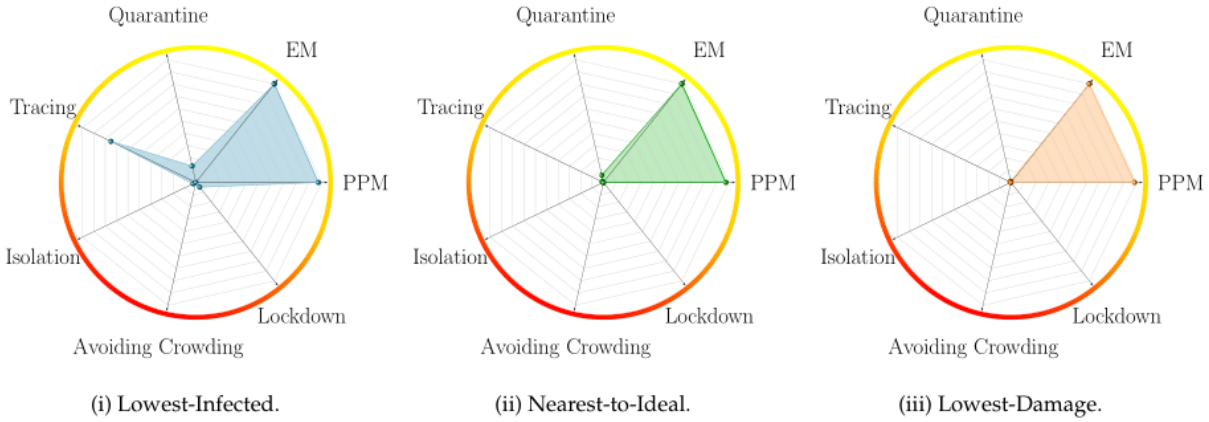
### 3) THE GoT SCENARIO

In this paragraph, we discuss the outcomes for the GoT data set, representing the characters' mobility pattern of the GoT TV series (see Section V-D). Figure 10 reports the results.

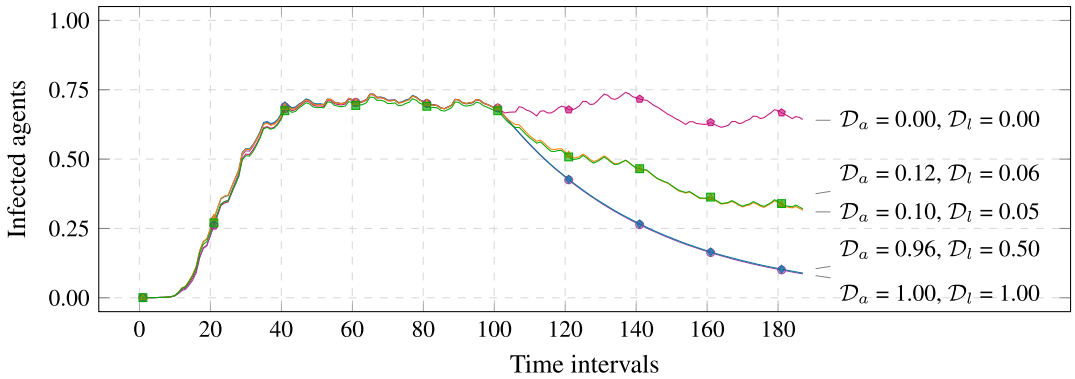
As for the Foursquare data set, all the solutions elaborated by the genetic algorithm can at least halve the fraction of infected in the unmitigated scenario with variable damages (Figure 10a). Focusing on the configuration Nearest-to-Ideal (see Figure 10b(ii)), we can notice values similar to the solution evaluated for the BLEBeacon scenario. As before, almost all agents use PPMs ($\alpha_p = 0.97$) and have a probability $\beta_{quarantine} = 0.72$ of self-quarantine themselves, while




(a) Pareto front evaluated by the algorithm NSGA-II for the Foursquare data set using the objective function: ($\text{inf}_{\text{last}}$, $\mathcal{D}_a$, $\mathcal{D}_l$).



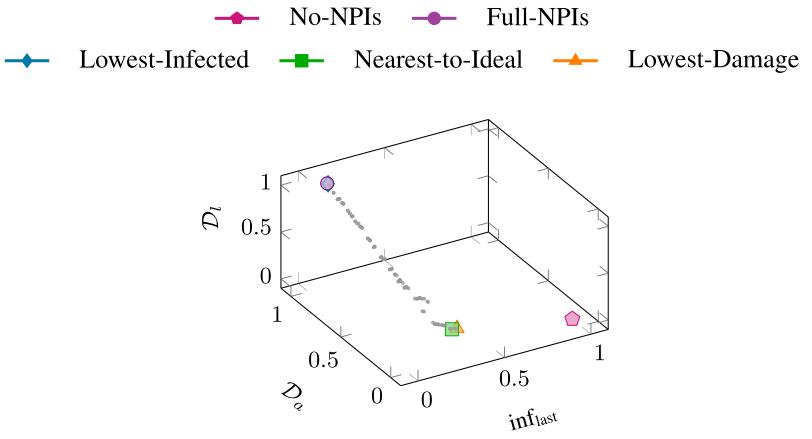

(b) Combination of NPIs, corresponding to the configurations obtaining the lowest number of infected at the end of the simulation (Lowest-Infected, left), the lowest damage (Lowest-Damage, right), and the nearest to the optimal point (Nearest-to-Ideal, center) evaluated over the Foursquare data set.



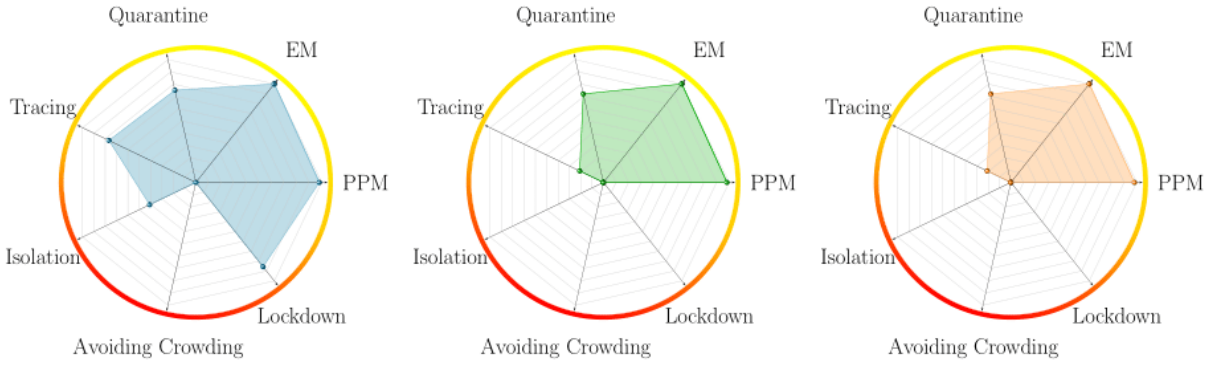

(c) Epidemic trend in the Foursquare data set when i) no NPIs are implemented (No-NPIs), ii) all NPIs are in place (Full-NPIs), iii) the configuration bringing to the lowest number of infected ad the end of the simulation is used (Lowest-Infected), iv) the configuration bringing to the lowest damage is applied (Lowest-Damage), and v) the configuration nearest to the optimum is used (Nearest-to-Ideal).

**FIGURE 9.** **Foursquare data set.**

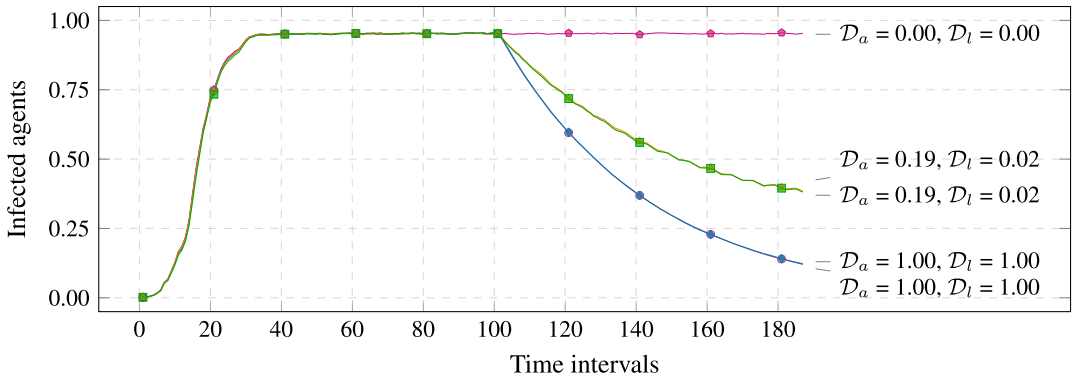
locations are regularly sanitized. On the other hand, we can note a lower use of tracing technologies, with only 20% of the agents adopting this measure. Once again, the nature of the data set may explain the massive use of preventive measures, such as quarantine policies. In this scenario, only a limited number of characters move across most of the locations, thus coming in contact with many other characters. Hence, increasing the probability of entering the quarantine state corresponds to a higher chance to quarantine agents, especially the super-spreader characters. Interestingly, this




(a) Pareto front evaluated by the algorithm NSGA-II for the GoT data set using the objective function: ($\text{inf}_{\text{last}}$, $\mathcal{D}_a$, $\mathcal{D}_l$).




(b) Combination of NPIs, corresponding to the configurations obtaining the lowest number of infected at the end of the simulation (Lowest-Infected, left), the lowest damage (Lowest-Damage, right), and the nearest to the optimal point (Nearest-to-Ideal, center) evaluated over the GoT data set.




(c) Epidemic trend in the GoT data set when i) no NPIs are implemented (No-NPIs), ii) all NPIs are in place (Full-NPIs), iii) the configuration bringing to the lowest number of infected ad the end of the simulation is used (Lowest-Infected), iv) the configuration bringing to the lowest damage is applied (Lowest-Damage), and v) the configuration nearest to the optimum is used (Nearest-to-Ideal).

**FIGURE 10.** GoT data set.

NPI combination can decrease by 75% the final fraction of infected without imposing strict restrictions that agents have to follow. Quantitatively speaking, we can observe a modest value for the social damage $\mathcal{D}_a$ (0.19) and a negligible commuting damage $\mathcal{D}_l$ (0.02).

As for the BLEBeacon scenario, the configuration Lowest-Infected perfectly overlaps with the Full-NPIs even though selecting a different NPI combination (see Figure 10b(i)). Also in this case, the Lowest-Damage solution corresponds to the Nearest-to-Ideal (see Figure 10b(iii)).

### D. DISCUSSION

In the following, we sum up the most interesting elements from our experiments and discuss some real-world implications of these findings.

#### 1) NO NPIs CAN RULE THEM ALL

If we focus our attention on the Nearest-to-Ideal NPI configuration, we can note how it largely varies across all data sets. For instance, we can observe that in the BLEBeacon data set, this NPI combination heavily counts on the probability of the agents to self-quarantine, translated into the massive use of tracing technology and a high probability of entering the quarantine state. We can observe a similar scenario in the GoT data set, in which the use of tracing measures is decreased, but the self-quarantine probability is still considerable. Entirely opposite, the Nearest-to-Ideal configuration in the Foursquare scenario exploits a reduced quarantine probability and no tracing policies. Clearly, these results reflect the nature of the data fed into the simulation model. The BLEBeacon data set represents a closed environment; hence, it is easier for a pathogen to spread. As a consequence, stricter measures should be implemented. Although we shrank four months of check-ins into a single month timeframe, the Foursquare data set remains highly sparse, with few locations containing most user check-ins. As already observed, the sparse nature of this data may partially explain why the Nearest-to-Ideal solution selects little or zero values for isolation, quarantine, and tracing measures. Similarly, the mobility patterns encoded by the GoT data set make it reasonable to use tracing and quarantine measures, which are preventive interventions, as GoT characters tend to meet in communities, and only a few of them travel across many different places (see Section V-D). In a real-world scenario, these results translate into applying different combinations of NPIs and tailor their severity according to local specific data, like the current epidemiological situation and other socio-demographic features. This outcome is further aligned with the guidelines of the European Centre for Disease Prevention and Control [20] and with previous literature [70].

#### 2) PLEASE, WEAR YOUR MASK AND WASH YOUR HANDS

Other than a widespread slogan to raise awareness in the population to fight the COVID-19, this phrase well describes another primary outcome of our experiments. Despite the different nature of the three data sets, we can easily spot that all the Nearest-to-Ideal NPI configurations share the application of PPMs and EMs, often selected in combination with other interventions (located in the yellow-orange zone of the radar chart, meaning that they generally cause low damage). Nonetheless, as discussed in the description of the results for the Foursquare experiment, the sole application of those interventions can still induce a decrease in the epidemic trend with the introduction of no damage. This result may have critical importance in a real-world situation as PPMs and EMs give a fundamental contribution in slowing the epidemic spreading at zero damage even in crowded locations and gathering places. However, the effectiveness of these policies remains strictly dependent upon the correct use made by the population.

#### 3) NOT A TREASURE HUNT, BUT A GUIDED SEARCH

As discussed in Section VII-A, GAs are widely exploited in ABMs to explore the best combination of parameters for the model. In the specific context of epidemic simulations, GAs may become a valuable tool to support policymakers as they can manually analyze some of the best solutions to control a pandemic and tune them according to specific needs. For instance, this approach can help investigate whether it is possible to implement an NPI combination that performs as well as closing gathering places. In this work, we examined GAs to optimize the implementation of NPIs while preserving two contrasting objectives: i) the fraction of infected and ii) damage brought by the intervention. We considered two types of damage to include the need of people to meet other individuals ($\mathcal{D}_a$) and the possibility of visiting a given location ($\mathcal{D}_l$), e.g., simulating a person going to jog. GAs represents a general framework as the objective functions to optimize can be defined according to a precise need and based on the specific target to study. For example, an objective can model the economic damage caused by closing given business categories or the cost of hospitalizations. Similarly, our TVH framework can be tuned according to the specific compartmental model to implement, and the high-order network can be instantiated according to specific socio-demographic data.

#### 4) COMMENT ON THE USE OF GAs

In this study, we used 100 individuals and 100 generations because of the long computational time required to run each simulation (e.g., forty minutes on average on the Foursquare data set) and because the implementation of the NSGA-II algorithm does not currently support the parallel evaluation of the individuals. However, the population size and the number of generations should be tuned to the complexity of the model when simulating real-world scenarios.

## VIII. CONCLUSION AND FUTURE WORK

NPIs gained attention during the SARS-CoV-2 pandemic in 2019 as the only force to resist an unforeseen epidemic diffusion in the absence of effective pharmaceutical interventions. Thus, deeply understanding the potential impact of introducing regulation policies aiming to control and reduce the epidemic propagation is of fundamental importance to minimize the effect on both the economy and the psychological wellness of the society.

In this work, we delve into discussing how such controlling measures can be embedded within an epidemiological model based on high-order relationships between people and environments, mimicking direct and indirect contagion pathways over time. To model and evaluate the impact of NPIs, we provide a formal definition of each intervention for the epidemiological framework based on *Time Varying Hypergraphs*. We evaluated each NPI applying the SIS com-

partmental model into an ABM that exploits our methodology to simulate interactions between agents and environments by designing the agent mobility behavior accordingly real-world data sets. The results of our experiments resemble previous literature and the guidelines of the WHO. Introduced alone, each NPI cannot extinguish an epidemic, even though some drastic measures, such as isolation and strict lockdown, have a higher impact in controlling the pathogen diffusion and considerably reduce its spread as indicated by the results in Section VI. The discussed outcomes further highlight that different combination of NPIs and their severity should be tailored according to local specific epidemiological data and that basic hygiene procedures are fundamental to reduce the spreading. Broadly translated, our findings indicate that the effects of NPIs in controlling are indisputable; however, the potential benefit of introducing NPIs in our society is functional to a massive and correct application in the population.

To enrich our TVH framework, we are currently working towards developing and analyzing immunization strategies that exploit the high-order nature of the underlying contact network. We will further develop other compartmental models to offer a complete framework within which higher-order diffusion phenomena can be examined.

## ACKNOWLEDGMENT

This work was supported in part by the European Union Horizon 2020 Research and Innovation Programme under Agreement 769872 (EMPATHIC) and 823907 (MENHIR). A preliminary research concerning of this work was presented in part at the Proceedings of 19th International Conference on Autonomous Agents and MultiAgent Systems in 2020 in "A Design-Methodology for Epidemic Dynamics via Time-Varying Hypergraphs" [6].

## REFERENCES

[1] N. Ahmed, R. A. Michelin, W. Xue, S. Ruj, R. Malaney, S. S. Kanhere, A. Seneviratne, W. Hu, H. Janicke, and S. K. Jha, "A survey of COVID-19 contact tracing apps," *IEEE Access*, vol. 8, pp. 134577–134601, 2020, doi: 10.1109/ACCESS.2020.3010226.

[2] J. E. Aledort, N. Lurie, J. Wasserman, and S. A. Bozzette, "Non-pharmaceutical public health interventions for pandemic influenza: An evaluation of the evidence base," *BMC Public Health*, vol. 7, no. 1, p. 208, Dec. 2007, doi: 10.1186/1471-2458-7-208.

[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, 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 Hum. Behav.*, vol. 4, no. 9, pp. 964–971, Sep. 2020, doi: 10.1038/s41562-020-0931-9.

[4] A. Aleta, D. Martín-Corral, M. A. Bakker, A. P. Y. Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini, A. Pentland, A. Vespignani, Y. Moreno, and E. Moro, "Quantifying the importance and location of SARS-CoV-2 transmission events in large metropolitan areas," *medRxiv*, vol. 2020, pp. 1–18, Dec. 2020, doi: 10.1101/2020.12.15.20248273.

[5] A. Antelmi, D. Malandrino, and V. Scarano, "Characterizing the behavioral evolution of Twitter users and the truth behind the 90-9-1 rule," in *Proc. Companion Proc. World Wide Web Conf.*, May 2019, pp. 1035–1038, doi: 10.1145/3308560.3316705.

[6] A. Antelmi, G. Cordasco, C. Spagnuolo, and V. Scarano, "A design-methodology for epidemic dynamics via time-varying hypergraphs," in *Proc. 19th Int. Conf. Auto. Agents MultiAgent Syst.*, 2020, pp. 61–69.

[7] H. Arora, T. S. Raghu, and A. Vinze, "Decision support for containing pandemic propagation," *ACM Trans. Manage. Inf. Syst.*, vol. 2, no. 4, pp. 1–25, Dec. 2011, doi: 10.1145/2070710.2070714.

[8] G. F. de Arruda, G. Petri, and Y. Moreno, "Social contagion models on hypergraphs," *Phys. Rev. Res.*, vol. 2, no. 2, Apr. 2020, Art. no. 023032, doi: 10.1103/PhysRevResearch.2.023032.

[9] A. Atalan, "Is the lockdown important to prevent the COVID-19 pandemic? Effects on psychology, environment and economy-perspective," *Ann. Med. Surg.*, vol. 56, pp. 38–42, Aug. 2020, doi: 10.1016/j.amsu.2020.06.010.

[10] S. Bansal, B. T. Grenfell, and L. A. Meyers, "When individual behaviour matters: Homogeneous and network models in epidemiology," *J. R. Soc. Interface*, vol. 4, no. 16, pp. 879–891, 2007, doi: 10.1098/rsif.2007.1100.

[11] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, "Networks beyond pairwise interactions: Structure and dynamics," *Phys. Rep.*, vol. 874, pp. 1–92, Aug. 2020, doi: 10.1016/j.physrep.2020.05.004.

[12] Á. Bodó, G. Y. Katona, and P. L. Simon, "SIS epidemic propagation on hypergraphs," *Bull. Math. Biol.*, vol. 78, no. 4, pp. 713–735, Apr. 2016, doi: 10.1007/s11538-016-0158-0.

[13] A. Bouchnita and A. Jebrane, "A hybrid multi-scale model of COVID-19 transmission dynamics to assess the potential of non-pharmaceutical interventions," *Chaos, Solitons Fractals*, vol. 138, Sep. 2020, Art. no. 109941, doi: 10.1016/j.chaos.2020.109941.

[14] A. Bretto, *Hypergraph Theory: An Introduction* (Mathematical Engineering). New York, NY, USA: Springer, 2013, doi: 10.1007/978-3-319-00080-0.

[15] M. Carillo, G. Cordasco, F. Serrapica, V. Scarano, C. Spagnuolo, and P. Szufel, "Distributed simulation optimization and parameter exploration framework for the cloud," *Simul. Model. Pract. Theory*, vol. 83, pp. 108–123, Apr. 2018, doi: 10.1016/j.simpat.2017.12.005.

[16] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro, "Time-varying graphs and dynamic networks," in *Ad-hoc, Mobile, Wireless Networking*, H. Frey, X. Li, and S. Ruehrup, Eds. Berlin, Germany: Springer, 2011, pp. 346–359, doi: 10.1007/978-3-642-22450-8_27.

[17] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002, doi: 10.1109/4235.996017.

[18] F. Dignum, V. Dignum, P. Davidsson, A. Ghorbani, M. van der Hurk, M. Jensen, C. Kammler, F. Lorig, L. G. Ludescher, A. Melchior, R. Mellema, C. Pastrav, L. Vanhee, and H. Verhagen, "Analysing the combined health, social and economic impacts of the corovanvirus pandemic using agent-based social simulation," *Minds Mach.*, vol. 30, no. 2, pp. 177–194, Jun. 2020, doi: 10.1007/s11023-020-09527-6.

[19] *Guide to Public Health Measures to Reduce the Impact of Influenza Pandemics in Europe: The ECDC Menu*, ECDC, Solna Municipality, Sweden, 2009.

[20] *Guidelines for Non-Pharmaceutical Interventions to Reduce the Impact of COVID-19 in the EU/EEA and the UK*, ECDC, Solna Municipality, Sweden, 2020.

[21] K. T. Eames, C. Webb, K. Thomas, J. Smith, R. Salmon, and J. M. F. Temple, "Assessing the role of contact tracing in a suspected H7N2 influenza a outbreak in humans in Wales," *BMC Infectious Diseases*, vol. 10, no. 1, p. 141, Dec. 2010, doi: 10.1186/1471-2334-10-141.

[22] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World* (Cambridge Books). Cambridge, U.K.: Cambridge Univ. Press, 2010, doi: 10.1017/CBO9780511761942.

[23] M. Ehrhardt, J. Gaáper, and S. Kilianová, "SIR-based mathematical modeling of infectious diseases with vaccination and waning immunity," *J. Comput. Sci.*, vol. 37, Oct. 2019, Art. no. 101027, doi: 10.1016/j.jocs.2019.101027.

[24] H. Fang, L. Wang, and Y. Yang, "Human mobility restrictions and the spread of the novel coronavirus (2019-ncov) in China," Nat. Bureau Econ. Res., Cambridge, MA, USA, Tech. Rep. 26906, 2020, doi: 10.3386/w26906.

[25] 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, May 2020, doi: 10.1126/science.abb6936.

[26] M. Fu, F. Glover, and J. April, "Simulation optimization: A review, new developments, and applications," in *Proc. Winter Simul. Conf.*, 2005, p. 13, doi: 10.1109/WSC.2005.1574242.

[27] S. Funk, M. Salathé, and V. A. A. Jansen, "Modelling the influence of human behaviour on the spread of infectious diseases: A review," *J. Roy. Soc. Interface*, vol. 7, no. 50, pp. 1247–1256, 2010, doi: 10.1098/rsif.2010.0142.

[28] N. Ganguly, T. Krueger, A. Mukherjee, and S. Saha, "Epidemic spreading through direct and indirect interactions," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 90, no. 3, Sep. 2014, Art. no. 032808, doi: 10.1103/PhysRevE.90.032808.

[29] G. Gensini, "The concept of quarantine in history: From plague to SARS," *J. Infection*, vol. 49, no. 4, pp. 257–261, Nov. 2004, doi: 10.1016/j.jinf.2004.03.002.

[30] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, and M. Colaneri, "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy," *Nature Med.*, vol. 26, no. 6, pp. 855–860, Jun. 2020, doi: 10.1038/s41591-020-0883-7.

[31] W. Gong, E.-P. Lim, and F. Zhu, "Characterizing silent users in social media communities," in *Proc. AAAI Press*, 2015, pp. 140–149.

[32] *Immuni*, India Government, New Delhi, India, 2020.

[33] M. R. Gualano, G. Lo Moro, G. Voglino, F. Bert, and R. Siliquini, "Effects of COVID-19 lockdown on mental health and sleep disturbances in Italy," *Int. J. Environ. Res. Public Health*, vol. 17, no. 13, p. 4779, Jul. 2020, doi: 10.3390/ijerph17134779.

[34] R. A. Hammond, "Considerations and best practices in agent-based modeling to inform policy," in *Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation* (Board on Population Health and Public Health Practice; Institute of Medicine). Washington DC, WA, US: National Academic Press, 2015, pp. 161–193.

[35] N. Hoertel, M. Blachier, C. Blanco, M. Olfson, M. Massetti, M. S. Rico, F. Limosin, and H. Leleu, "A stochastic agent-based model of the SARS-CoV-2 epidemic in France," *Nature Med.*, vol. 26, no. 9, pp. 1417–1421, Sep. 2020, doi: 10.1038/s41591-020-1001-6.

[36] E. Hunter, B. Mac Namee, and J. D. Kelleher, "A taxonomy for agent-based models in human infectious disease epidemiology," *J. Artif. Societies Social Simul.*, vol. 20, no. 3, pp. 1–17, 2017, doi: 10.18564/jasss.3414.

[37] I. Iacopini, G. Petri, A. Barrat, and V. Latora, "Simplicial models of social contagion," *Nature Commun.*, vol. 10, no. 1, p. 2485, 2019, doi: 10.1038/s41467-019-10431-6.

[38] H. Inoue and Y. Todo, "The propagation of economic impacts through supply chains: The case of a mega-city lockdown to prevent the spread of COVID-19," *PLoS ONE*, vol. 15, no. 9, Sep. 2020, Art. no. e0239251, doi: 10.1371/journal.pone.0239251.

[39] J. Lancaster. (2019). *Game of Thrones Datasets and Visualizations*. [Online]. Available: https://github.com/jeffreylancaster/game-of-thrones

[40] B. Jhun, M. Jo, and B. Kahng, "Simplicial SIS model in scale-free uniform hypergraph," *J. Stat. Mech., Theory Exp.*, vol. 2019, no. 12, Dec. 2019, Art. no. 123207, doi: 10.1088/1742-5468/ab5367.

[41] A. Jindal and S. Rao, "Agent-based modeling and simulation of mosquito-borne disease transmission," in *Proc. 16th Conf. Auto. Agents MultiAgent Syst.*, Richland, SC, USA, 2017, pp. 426–435.

[42] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: Past, present, and future," *Multimedia Tools Appl.*, vol. 80, pp. 8091–8126, Oct. 2020, doi: 10.1007/s11042-020-10139-6.

[43] D. Klinkenberg, C. Fraser, and H. Heesterbeek, "The effectiveness of contact tracing in emerging epidemics," *PLoS ONE*, vol. 1, no. 1, Dec. 2006, Art. no. e12, doi: 10.1371/journal.pone.0000012.

[44] P. Kostkova, M. Szomszor, and C. St. Louis, "#Swineflu: The use of Twitter as an early warning and risk communication tool in the 2009 swine flu pandemic," *ACM Trans. Manage. Inf. Syst.*, vol. 5, no. 2, pp. 1–25, Jul. 2014, doi: 10.1145/2597892.

[45] V. Leonenko, S. Arzamastsev, and G. Bobashev, "Contact patterns and influenza outbreaks in Russian cities: A proof-of-concept study via agent-based modeling," *J. Comput. Sci.*, vol. 44, May 2020, Art. no. 101156, doi: 10.1016/j.jocs.2020.101156.

[46] B. Li, J. Li, K. Tang, and X. Yao, "Many-objective evolutionary algorithms: A survey," *ACM Comput. Surv.*, vol. 48, no. 1, pp. 1–35, Sep. 2015, doi: 10.1145/2792984.

[47] F. Lorig, E. Johansson, and P. Davidsson, "Agent-based social simulation of the COVID-19 pandemic: A systematic review," *J. Artif. Societies Social Simul.*, vol. 24, no. 3, p. 5, 2021, doi: 10.18564/jasss.4601.

[48] S. Muñoz and C. A. Iglesias, "An agent based simulation system for analyzing stress regulation policies at the workplace," *J. Comput. Sci.*, vol. 51, p. 101326, 2021, doi: 10.1016/j.jocs.2021.101326.

[49] *Social Gatherings Over Six to be Banned in England*, B News, Mumbai, India, 2020.

[50] N. H. Ogden, A. Fazil, J. Arino, P. Berthiaume, D. N. Fisman, A. L. Greer, A. Ludwig, V. Ng, A. R. Tuite, P. Turgeon, and L. A. Waddell, "Modelling scenarios of the epidemic of COVID-19 in Canada," *Canada Commun. Disease Rep.*, vol. 4, pp. 198–204, Jun. 2020, doi: 10.14745/ccdr.v46i06a08.

[51] *Non-Pharmaceutical Public Health Measures for Mitigating the Risk and Impact of Epidemic and Pandemic Influenza*, WWH Organization, Mumbai, India, 2019.

[52] *Water, Sanitation, Hygiene and Waste Management for the COVID-19 Virus*, WWH Organization, Mumbai, India, 2020.

[53] P. E. Pare, C. L. Beck, and A. Nedic, "Epidemic processes over time-varying networks," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 3, pp. 1322–1334, Sep. 2018, doi: 10.1109/TCNS.2017.2706138.

[54] N. Perra, "Non-pharmaceutical interventions during the COVID-19 pandemic: A review," *Phys. Rep.*, vol. 913, pp. 1–52, 2021, doi: 10.1016/j.physrep.2021.02.001.

[55] S. Rojas-Galeano and L. Alvarez, "Simulation of non-pharmaceutical interventions on COVID-19 with an agent-based model of zonal restraint," *MedRxiv*, vol. 2020, pp. 1–18, Jun. 2020, doi: 10.1101/2020.06.13.20130542.

[56] R. Rossi, V. Socci, D. Talevi, S. Mensi, C. Niolu, F. Pacitti, A. Di Marco, A. Rossi, A. Siracusano, and G. Di Lorenzo, "COVID-19 pandemic and lockdown measures impact on mental health among the general population in Italy," *Frontiers Psychiatry*, vol. 11, p. 790, Aug. 2020, doi: 10.3389/fpsyt.2020.00790.

[57] M. D. Salute, "COVID-19, travellers," Ministero della Salute, Rome, Italy, Tech. Rep., 2020.

[58] D. Sikeridis, I. Papapanagiotou, and M. Devetsikiotis, "BLEBeacon: A real-subject trial dataset from mobile Bluetooth low energy beacons," *CoRR*, vol. abs/1802.08782, pp. 1–3, Feb. 2018.

[59] P. C. L. Silva, P. V. C. Batista, H. S. Lima, M. A. Alves, F. G. Guimarães, and R. C. P. Silva, "COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110088, doi: 10.1016/j.chaos.2020.110088.

[60] N. Srinivas and K. Deb, "Muiltiobjective optimization using nondominated sorting in genetic algorithms," *J. Evol. Comput.*, vol. 2, no. 3, pp. 221–248, 1994, doi: 10.1162/evco.1994.2.3.221.

[61] F. Stonedahl and U. Wilensky, "Finding forms of flocking: Evolutionary search in ABM parameter-spaces," in *Multi-Agent-Based Simulation* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2011, pp. 61–75, doi: 10.1007/978-3-642-18345-4_5.

[62] B. Su, P. Andelfinger, J. Kwak, D. Eckhoff, H. Cornet, G. Marinkovic, W. Cai, and A. Knoll, "A passenger model for simulating boarding and alighting in spatially confined transportation scenarios," *J. Comput. Sci.*, vol. 45, Sep. 2020, Art. no. 101173, doi: 10.1016/j.jocs.2020.101173.

[63] Q. Suo, J. L. Guo, and A.-Z. Shen, "Information spreading dynamics in hypernetworks," *Phys. A, Stat. Mech. Appl.*, vol. 495, pp. 475–487, Apr. 2018, doi: 10.1016/j.physa.2017.12.108.

[64] E. Tekin and I. Sabuncuoglu, "Simulation optimization: A comprehensive review on theory and applications," *IIE Trans.*, vol. 36, no. 11, pp. 1067–1081, Nov. 2004, doi: 10.1080/07408170490500654.

[65] M. Tracy, M. Cerdá, and K. M. Keyes, "Agent-based modeling in public health: Current applications and future directions," *Annu. Rev. Public Health*, vol. 39, no. 1, pp. 77–94, Apr. 2018, doi: 10.1146/annurev-publhealth-040617-014317.

[66] H. Van Dyke Parunak, R. Savit, and R. L. Riolo, "Agent-based modeling vs. equation-based modeling: A case study and users' guide," in *Multi-Agent-Based Simulation*. Berlin, Germany: Springer, 1998, pp. 10–25, doi: 10.1007/10692956_2.

[67] F. Verelst, L. Willem, and P. Beutels, "Behavioural change models for infectious disease transmission: A systematic review (2010–2015)," *J. Roy. Soc. Interface*, vol. 13, no. 125, Dec. 2016, Art. no. 20160820, doi: 10.1098/rsif.2016.0820.

[68] C. Wang, R. Pan, X. Wan, Y. Tan, L. Xu, R. S. McIntyre, F. N. Choo, B. Tran, R. Ho, V. K. Sharma, and C. Ho, "A longitudinal study on the mental health of general population during the COVID-19 epidemic in China," *Brain, Behav., Immunity*, vol. 87, pp. 40–48, Jul. 2020, doi: 10.1016/j.bbi.2020.04.028.

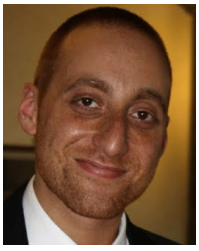
[69] L. Webb, "COVID-19 lockdown: A perfect storm for older people's mental health," *J. Psychiatric Mental Health Nursing*, vol. 28, no. 2, p. 300, Apr. 2021, doi: 10.1111/jpm.12644.

[70] B. Wilder, M. Charpignon, J. A. Killian, H.-C. Ou, A. Mate, S. Jabbari, A. Perrault, A. N. Desai, M. Tambe, and M. S. Majumder, "Modeling between-population variation in COVID-19 dynamics in hubei, lombardy, and new York city," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 41, pp. 25904–25910, Oct. 2020, doi: 10.1073/pnas.2010651117.
[71] C. Yang, X. Shi, L. Jie, and J. Han, "I know you'll be back: Interpretable new user clustering and churn prediction on a mobile social application," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2018, pp. 914–922, doi: 10.1145/3219819.3219821.
[72] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu, "Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 45, no. 1, pp. 129–142, Jan. 2015, doi: 10.1109/TSMC.2014.2327053.
[73] Q. Yang, C. Yi, A. Vajdi, L. W. Cohnstaedt, H. Wu, X. Guo, and C. M. Scoglio, "Short-term forecasts and long-term mitigation evaluations for the COVID-19 epidemic in Hubei province, China," *Infectious Disease Model.*, vol. 5, pp. 563–574, Dec. 2020, doi: 10.1016/j.idm.2020.08.001.

**ALESSIA ANTELMI** graduated in computer science from the Università degli Studi di Salerno, Italy, in 2018, under the supervision of Prof. Vittorio Scarano. Before starting her Ph.D. studies in November 2018, she joined the Unit for Social Semantics at the Data Science Institute, Galway, Ireland, led by Prof. John Breslin, under the Erasmus+ Traineeship Grant. She is currently pursuing the Ph.D. degree with the ISISLab Laboratory, Università degli Studi di Salerno, supervised by Prof. V. Scarano, Prof. G. Cordasco, and Dr. C. Spagnuolo. Her research interests include the fields related to complex networks, especially hypergraphs, and agent-based models.

**GENNARO CORDASCO** received the Laurea degree *(cum laude)* in computer science from the Università degli Studi di Salerno, Italy, in 2002, discussing a thesis in distributed algorithms titled "An optimal mapping algorithm for complete binary trees to parallel memories, and the Ph.D. degree from the aforementioned university in April 2006, supervised by Prof. Alberto Negro and Prof. Vittorio Scarano, discussing a thesis on peer to peer systems, titled "On Designing Overlay Networks for P2P Systems." He is currently an Assistant Professor with the Dipartimento di Psicologia, Facoltà di Psicologia, Università degli studi della Campania "Luigi Vanvitelli," Italy, and an Affiliate Member of the International Institute for Advanced Scientific Studies, Vietri sul Mare, Italy. From March 2005 to November 2005, he was a Research-Scholar with the Theoretical Aspect of Parallel and Distributed Systems (TAPADS) Laboratory, Department of Computer Science, University of Massachusetts at Amherst, MA, USA, doing his research activity under the supervision of Distinguished Prof. Arnold L. Rosenberg. From May 2006 to March 2011, he was a Postdoctoral Fellow (Assegnista di Ricerca) with the Dipartimento di Informatica, Università degli Studi di Salerno, cooperating in the research activity of the ISISLab Laboratory.

**VITTORIO SCARANO** received the Laurea degree in information science from the Università degli Studi di Salerno, Italy, in 1990, and the Research Doctorate degree in applied mathematics and computer science from the Università di Napoli, in 1995. In 1992, he visited Eötvös Loránd University, Budapest, Hungary; and the Department of Computer Science, University of Massachusetts at Amherst, USA, from 1992 to 1994, carrying out research in the team of Prof. Arnold Rosenberg on parallel algorithms and architectures. Since 2001, he has been an Associate Professor with the Department of Computer Science, Università degli Studi di Salerno. In 1995, he co-founded the ISISLab Laboratory, with Prof. Alberto Negro, at the Department of Computer Science, Università degli Studi di Salerno. The research activities of the ISISLab Laboratory focuses on distributed and parallel programming, social and collaborative management of Open Data, and interactive visualization. Since its foundation, the research laboratory has hosted ten Ph.D. students, more than 20 collaborators (research grants and research contracts), and supervised more than 120 theses (bachelor and master). He is the coauthor of more than 100 articles in international journals and conferences, and since 2000, he has been a member of more than 20 international conference program committees. He has been a reviewer for over 30 journals, conferences, and workshops. He has also involved in the steering committee of conference series, such as Adaptive Hypermedia and Visualization. He coordinated the European Research Project H2020 ROUTE-TO-PA "Raising open and user-friendly transparency-enabling technologies for public administrations," with 12 partners and a budget of over 3M €. He participated and coordinated local units of European, national, and regional projects.



**CARMINE SPAGNUOLO** received the master's degree *(cum laude)* in computer science from the Università degli Studi di Salerno, in 2013, and the Ph.D. degree in computer science, under the supervision of Prof. Vittorio Scarano and Prof. Gennaro Cordasco, in 2017. He is interested in parallel algorithms, distributed systems, graph theory, social networks, and agent-based simulations. In 2012, he got a grant from the Office of Naval Research (ONR) for visiting the George Mason University (GMU). In May 2017 and from October 2017 to December 2017, he was a Visiting Student with the University of Chicago and Argonne National Laboratory, under the supervision of Dott. Jonathan Ozik and exploiting a grant from ANL. In December 2019, he was a Visiting Researcher at GMU under the supervision of Prof. Sean Luke. He is currently a Postdoctoral Researcher at the Università degli Studi di Salerno, and he is a Senior Member of ISISLab Laboratory. He is the coauthor of more than 30 papers in international refereed journals and conferences.