

# An individual-based epidemic simulator

Tatsuo Unemi,<sup>1</sup> Saki Nawata, Masaaki Miyashita, and Norihiko Shinomiya

Soka University, Japan  
(Tel: 81-42-691-9429, Fax: 81-42-691-9312)

<sup>1</sup>unemi@soka.ac.jp

**Abstract:** From a view point of Artificial Life researches, a multi-agent based ecological simulation will be helpful to understand the dynamics of epidemic. The authors developed an individual-based simulator in which the user is allowed to set the parameter values and to examine what would go on in a virtual population of tens or more thousands of individuals. The parameters include the metrics for environment, mobility, pathogenesis, countermeasures, and tests. The simulator provides not only a graphical user interface and dynamic graphical monitoring functionalities, but also a batch job management and a recording capability of a number of statistical indexes. Through some cases of simulation experiments based on the parameter settings targeting SARS-CoV-2, we observed that all of implemented measures are effective to suppress the spreads. The dynamics shows almost same figures to those known from SIR model when the population size is large. However, under some conditions of strong restriction in relatively small population, the end of epidemic comes early. This fact also means it takes long period until the cessation in big cities, even when similar scale of restrictions to the country side are applied.

**Keywords:** evo-eco simulator, multi-agent system, infectious disease, epidemic dynamics

## 1 INTRODUCTION

Not only SARS-CoV-2 and COVID-19, epidemic of new infectious disease is one of the menace for human beings. From a view point of Artificial Life and its related researches, a multi-agent based ecological simulation will be a helpful tool to understand the dynamics of epidemic. The advantage comparing to another type of mathematical model, such as SIR model [1] based on differential equations, is that it allows us to introduce concrete models of people's behavior and countermeasures, even though those features bring a disadvantage for mathematical analysis. It also provides a source of visualization that helps our intuitive understanding of the complex phenomena. This paper introduces an approach to build an individual-based simulator in multiple scale, from smartphone to high-performance computer, with graphical user interface and visualization. The model and algorithm are inherited from the authors previous work [2] on evolutionary simulation for sexual dimorphism and speciation. The models of individual behavior, pathogenesis, and measures are implemented. The individuals move in a two-dimensional Euclidean space basically following a simple Newtonian mechanics with mass and friction, affected by forces for collision avoidance and gatherings. They could also jump to another position as a model of long distance travel. The measures include social distancing, mobility restriction, gathering restriction, tests, quarantine, and contact tracing.

As the real world tactics to take a balance between epidemic suppression vs. economic activity, we usually change the public restriction following the situation of both sides. To examine a strategy of such reactive countermeasures, we implemented a functionality of scenario, a set of rules to change the restriction strength under conditions on statistical indexes, though the implementation of an economy model is remaining as our future work.

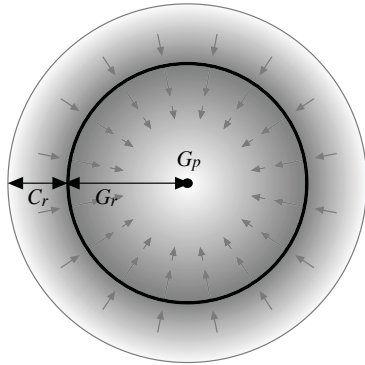
The rest of this paper describes the model overview, software architecture, and some results of simulation as examples, followed by concluding remarks and future works.

## 2 MODEL OVERVIEW

This section describes an overview of simulation model on both an individual agent and the environment.

### 2.1 Spatiotemporal model of the world

The field of the simulation is a two dimensional square plane in which a position is represented by continuous Cartesian coordinates. Each of individual agent has its own coordinate  $(x, y)$  and velocity  $(v_x, v_y)$ . The individual action is acceleration synthesized from a number of forces as its selected action following the rules described in a later subsection 2.2. The individuals in the population are placed in the field in random position and has zero velocity at the beginning. The health condition of each individual is susceptible except a number of randomly selected



**Fig. 1:** Force field of gathering. All forces are toward the center point. They are strongest on the circumference of radius  $G_r$ .

individuals as infected. The population density is adjustable by tuning the population size and the field's area size.

## 2.2 Individual behavior

Each individual moves following a force vector affected by surroundings. The force is a combination of repulsion forces proportional to the square of distance to the others and attraction force toward its favorite one. The new position is approximately calculated together with mass and friction based on Newtonian mechanics using a modified version of Euler method. The individual located near from the edge of the field is affected by a repulsion force from it to avoid travel beyond the border. The force  $F_i$  affecting the individual  $i$  is

$$F_i = \alpha \cdot \sum_j \frac{p_i - p_j}{|p_i - p_j|^3} + \beta \cdot (p_b - p_i) + \gamma \cdot \sum_{j=1}^4 \frac{p_i - w_j}{|p_i - w_j|^3} + \zeta \cdot G_i \quad (1)$$

where  $p_i$  is the position of individual  $i$ ,  $b$  is the most favorite individual around it,  $w_j$  is the nearest point on the border  $j$  from  $i$ , and  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\zeta$  are coefficients.  $G_i$  is the force toward the center of the nearest gatherings described later.

### 2.2.1 Long-distance move

An individual who doesn't show a symptom moves to another position in the field with a specified probability. The mobility restriction as a countermeasure is implemented by tuning the distance and probability of this type of move.

On the other hand besides the move in the inside of the field, an individual may move to the outside for quarantine or burial. When the individual becomes positive in test, he/she moves to the hospital area. When the individual died, he/she moves to the cemetery area. An individual who has been fully recovered come back to the original position in the field.

### 2.2.2 Gatherings

As a model of events where people gather together, *gatherings* probabilistically happen in the field. Each gathering occupies

a circular area with parameters of size, duration, and strength. The individuals without any symptom around the edge of the circle receive a force toward the center of gathering. The force strength has a maximum on the circular edge, and it reduces proportionally to the inverse of distance from there. Fig. 1 illustrates the force field applied to individuals. The value of  $G_i$  in the above equation is

$$G_i = \begin{cases} 0 & \text{if } |p_i - G_p| > G_r + C_r \\ 1 - (|p_i - G_p| - G_r) / C_r & \text{else if } |p_i - G_p| > G_r \\ |p_i - G_p| / G_r & \text{otherwise} \end{cases} \quad (2)$$

where  $G_p$  is the center position of the nearest gathering,  $G_r$  is its radius,  $C_r$  is a constant distance from the circumference of affected area.

## 2.3 Pathogenesis

Fig. 2 illustrates the individual pathogenesis that includes changes of severeness and infectiousness with events such as infection, showing symptom, turning to recovery and death. The health condition of each individual takes one of the five states; 1. susceptible, 2. asymptomatic, 3. symptomatic, 4. recovered, and 5. died. A susceptible individual may be infected in a probability when he/she approaches to an infected individual within a fixed distance. The symptomatic severeness progresses in a speed depending on the patient's characteristic, it reaches the onset of sickness, and then passes away when the severeness goes beyond each of the threshold. When the period of the turning point allocated to each individual has passed before his/her death, the severeness reduces in a constant speed toward the full recovery. A recovered individual is assumed to obtain immunity against the infection, but the effect disappears to turn the health state back to be susceptible after a fixed period.

The contagious strength of infected individual is zero until a fixed length of delay passes from the infection, and then it increases up to the peak level before it reaches the level of onset.

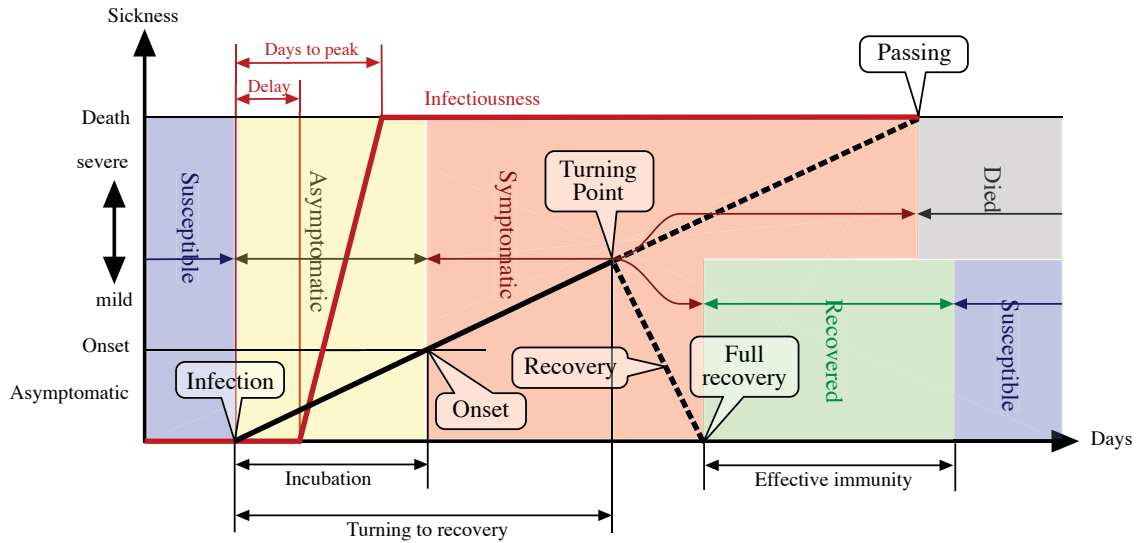
## 2.4 Countermeasures

The models of social distancing, mobility restriction, gathering restriction, quarantine, and contact tracing are implemented as follows, together with tests described in the next section.

*Social distancing* is adjusted using a repulsion force to avoid a mutual collision among individuals as described in the section of Individual behavior, with two parameters, magnification bias of the force and rate of obedience.

*Mobility restriction* is adjustable using a distribution of the distances for long distance movement and the frequency as described in section 2.2.1.

*Gathering restriction* is adjustable using a distribution of the size, strength and frequency as described in section 2.2.2.



**Fig. 2:** Pathogenesis. The patient’s sickness follows the black thick line from the left bottom. It turns to go lower at the turning point when the patient starts recovery. Otherwise he/she is going to die toward the upper boundary. The red thick line shows the path of infectiousness that indicates how strongly the patient infects the virus to the others.

The individuals whose test results are positive move to the hospital for *quarantine*. They return back to the original position in the field when they get fully recovered.

*Contact tracing* is implemented by managing a list of contacts for each individual in the field. When it encounters an individual in the list again, the time stamp associated with the individual identifier is revised. The items in the list older than 14 days are removed. The contact is not always recorded but done with a pre-defined probability.

## 2.5 Tests

Tests are important to identify who are infected for both ways to go to and return from the hospital or other places for separation and care. The tests are conducted for the subjects selected from three groups with a probability assigned for each group; symptomatic, susceptible, and contacted individuals. A subject of susceptible individual can be assumed that he/she shows similar symptom as the target disease but caused by something else. The result is determined with sensitivity, the ratio of correct detection among infected subjects, and specificity, the ratio of negative results among susceptible and recovered subjects, of the test. The other characteristics, such as delay of the conduct and delay of the result, are also adjustable. It is not necessary to specify the probability for contacts other than the probability to record them into the lists of contacts, because the resulted probability is the multiplication of these two values and it is meaningless to separate them for this simulation.

## 3 CONTROL AND MONITOR

### 3.1 Software architectures

The softwares we developed include three different scales; Web application, macOS application, and Web server.

#### 3.1.1 Web application

The first prototype was developed in March, 2020 using p5.js [3], a javascript library for creative coding. It has basic functionality of graphics and parameter settings available on popular web browsers running on the personal computer and smartphone. Because of the restriction of computing performance and storage, it manages only four or five thousands individuals as the population. The scale is not enough but it will be useful to provide an opportunity for many people to have the usage experience. It is accessible from the authors’ web site [4].

#### 3.1.2 Mac application

The version of macOS application software has almost full set of functionality except the batch job management. The user is allowed to save the parameter and scenario settings as a document in the local or cloud storage to open it later. By employing a multi-thread parallelism, it manages 40,000 individuals on an older machine such as MacBook Pro 2013, and 100,000 individuals in high-performance machine, such as Mac Pro and Mac mini 2020, for smooth animation in realtime. The binary code is downloadable from the authors’ web site [4]. Fig. 3 shows an example of screen shot that runs on a population of 10,000 individuals.

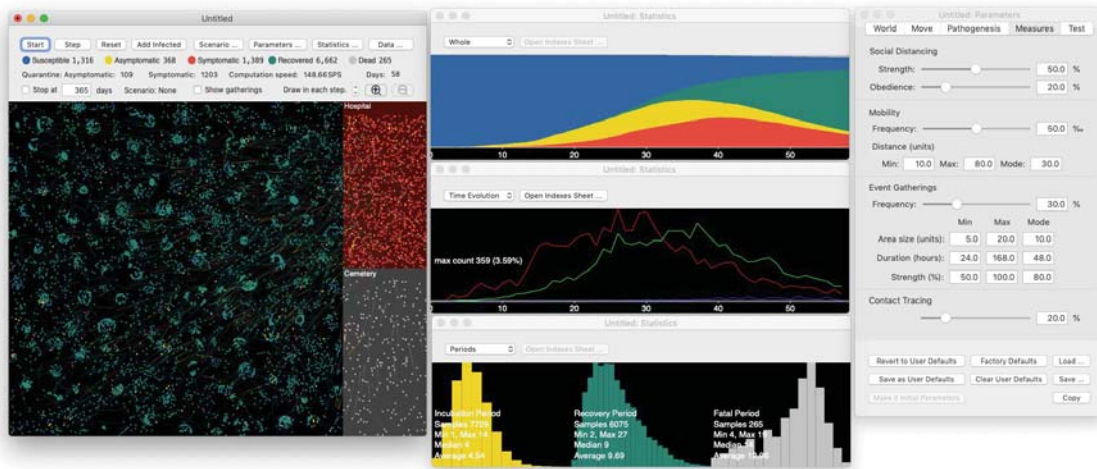


Fig. 3: A sample screen shot of macOS application version.

### 3.1.3 Web server

The web server version runs as a background job that accepts requests from HTTP client software, typically a web browser. It also provides a set of codes written in HTML and javascript, so that the user can access the simulator through graphical user interfaces on a web browser just as a web service.

For the purpose of statistic analysis, it also has a functionality to manage multiple batch job with different parameter settings and different random number sequences. We need to wait for hours but it can complete a simulation process of 365 days with a population of hundreds thousands people using a machine with many-cores CPU under a settings of 16 simulation steps per day. It is possible for the user to download the numerical data in CSV format including several indexes after the jobs are completed.

It is useful to prepare a command script that controls a set of batch jobs for a variation of settings without a web browser. The author conducted several times of simulation by this type of method using a shell script including a command, `curl`, to request starting a batch job and collecting the results. In such cases, the parameters and the scenario are coded in JSON format, that the user can organize by the version either of Mac application or web service.

### 3.2 Statistical indexes

It is necessary to record a number of indexes in order to investigate what has happened in the simulation to compare with the events in the real world. The indexes include the number of susceptible, asymptomatic, symptomatic, recovered, and died individuals; the number of hospitalization and returns; the number of tests; the daily number of test positive, test-positive ratio; and so on.

### 3.3 Scenario

To suppress a tragic epidemic of deadly disease, it is usually important for people to cooperate in the countermeasure organized by the national and local governments, such as lockdown. Simulation of epidemic dynamics combined with a scenario, a schedule of countermeasures as the response against the aspects of spread, will be helpful for both the measure planner and the citizens. It is also useful to revise the model in order to reduce the gap with the reality by comparing the simulation results and the real events that have already passed.

A scenario is represented as a form of array of two types of elements, *condition* and *action*. The elements are principally interpreted from head to tail in the order of arrangement in the array.

The condition itself is also constructed by *conditional expression* and *control*. The conditional expression is constructed by predicates, comparison between the value of index and constant; and logical operation, conjunction, disjunction, and negation. The indexes include not only those described in 3.2 but also the number of elapsed days from the starting point of the simulation process. When the system interprets a condition element, it waits until the result of conditional expression becomes true, then it interprets the control. The control may be a demand to jump to another element indicated by the index number attached, or just to move to the next element.

The action may be a demand to add a number of infects, or a demand to assign a new value to a parameter. In the former case, the system randomly selects the indicated number of susceptible individuals to turn their health states to be asymptomatic. In the latter case, the system changes the parameter value to the value given as the argument.

### 3.4 Graphical animation

The simulator provides graphical animations of both people's movements and the time series of indexes.

The world in which the individuals are living is a two-dimensional space. The whole population is observable when the size is smaller than 40,000, but it is difficult for larger population due not only to the tiny size of individuals but also to the computational cost for rendering many individuals. The Mac application version has a convenient functionality that allows the user to magnify the world view to observe a part of area and to explore in the space by drag operation of the pointing device. Such observation is helpful to monitor what types of events are happening in the simulation. It is also provide a useful visualization for real people to help their understanding on the dynamics of epidemic.

The user is allowed to open a number of windows each of which graphically shows the indexes by the user's choice. For the time evolution of indexes, the scale of horizontal axis representing the time is automatically adjusted as the simulation process is going on. The histograms showing distributions of incubation period, recovery period, fatal period, and the number of infects by a single infected individual are also included in the choice.

### 3.5 Batch jobs

In a type of simulation including a stochastic process, it is important to examine a number of different processes for the same parameter settings in order to induce a statistically meaningful consequence. In such a case, a phenomenon happening in a single process might be an occasional case.

To satisfy this requirement, the web server version accepts a request to conduct a number of processes using a POST method of HTTP, by giving information as the *payload* including the number of processes, the longest duration in days, initial parameter set, scenario, and a set of index names to be collected after the completion. The information is represented in JSON format, so that the user can import it from Mac application version and it is easy for the user to edit it. The server responds a unique alpha-numerical character string as the identifier for submitted job, so that the user use this identifier for queries on the job. The submitted jobs are enqueued as they are processed in a manner of first come first served. The number of simulation processes executable in parallel is specified as a command option given when the server process is launched.

After the submission of batch jobs, the server responds to a query asking the processing status including information on how many simulation processes are completed, on going, and waiting. After the job is completed, the server becomes ready for a query to download the results in a form of compressed archive containing a number of text files of CSV format.

## 4 RESULT EXAMPLE

Targeting COVID-19, we set the parameters according to the literatures such as [5, 6, 7] for pathogenesis and tests. It is difficult to correctly identify the real values for formalizing the mathematical model, because our observation is not perfect. For example, the length of incubation period is one of the important feature, but it is possible to gather the data only for the patients who were detected as positive. We might miss many other cases on who did not report as positive because of false negative on test and recovery before showing severe symptom.

Fig. 4 shows an example of the results related to the effects of long-distance travel on the epidemic dynamics. We conducted ten times of simulation processes of 200 days for each combination of the population size, 10,000, 40,000 and 90,000 individuals; and travel frequency, 0 to 10% stepping by 1%. The number of processes is  $3 \times 11 \times 10 = 330$  in total. The number of simulation steps per day is 16, that means each process is  $16 \times 200 = 3200$  steps. The top figure *a* illustrates the average time sequences of the percentage of infected in the population of 40,000 individuals. It shows a similar figure as SIR model when the travel frequency is higher than 3%, that is, more individuals are infected and the spread ends earlier when people are traveling more frequently. However, as the middle figure *b* shows, the spread also ends earlier when the frequency is lower than 2%. It is possible to assume that this phenomenon is caused by the probability of remaining infected individual when the percentage is low enough. As shown in the bottom figure *c*, the early end of spread happens in lower frequency when the population size is large. It suggests that it is relatively difficult to stop the spread in the large city comparing to the small town.

## 5 CONCLUDING REMARKS

A number of researchers have been challenging to build their own multi-agent-based simulators from different interests and perspective, such as [8, 9, 10]. The authors has developed an individual-based simulator employing a technique to simulate a collective behavior of multi-agent roaming in a two dimensional continuous space. It is a unique feature of this system that the agents move smoothly, that might not be essential from a view point of experts in epidemiology. However, the authors think it is useful to help our intuitive understanding on the dynamics of epidemic, not only for the tactics planner but also ordinary citizens.

An example of the simulation results is described in the above section 4 that revealed a difference between the epidemic dynamics for small and large population. The other results are available from the project Web site [4], such as a one year simulation on interaction between spreads and countermeasures.

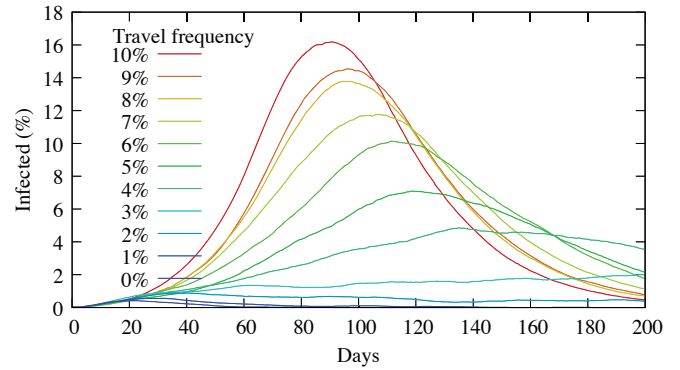
The authors are hoping that this project will contribute to reduce any tragic situation caused by COVID-19 and menace by unknown infectious diseases in future.

## ACKNOWLEDGMENT

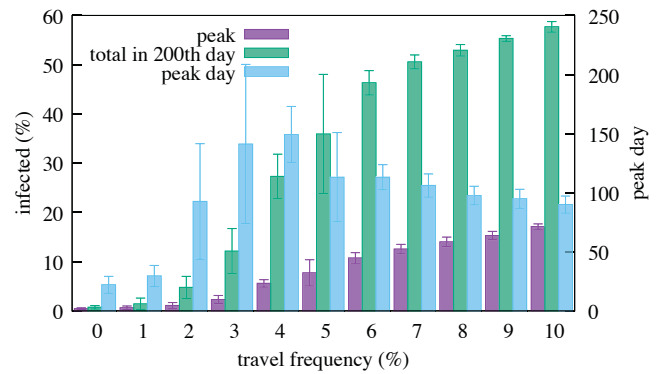
The authors thank professors of Soka University, Shoko Nishihara, Sayaka Takase, and Isamu Okada; and the committee members of governmental project for Development of Simulation using AI and related technologies, Hiroaki Kitano and Atsushi Shinjo; who gave us valuable suggestions through on-line discussions. This research was supported by the Cabinet Secretariat of Japan Government.

## REFERENCES

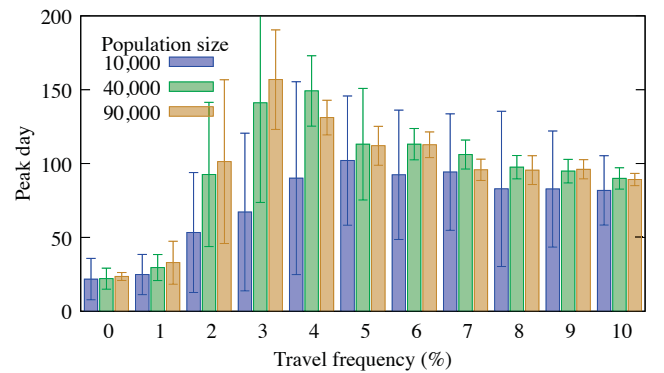
- [1] K. W. Ogilvy and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society A*, vol. 115, no. 772, pp. 700–721, 1927.
- [2] T. Unemi and H. Matsumoto, "Loversflow v2 : an individual-based evo-eco simulator on sexual dimorphism – a challenge toward evolutionary aesthetics," in *Proceedings of the 23rd International Symposium on Artificial Life and Robotics*, Beppu, Japan, January 2018, pp. 295–300.
- [3] Processing Foundation. p5.js. <https://p5js.org>
- [4] T. Unemi. Simepidemic1: An individual-based epidemic simulator. <http://www.intlab.soka.ac.jp/~unemi/SimEpidemic1/info/>
- [5] N. Sethuraman, S. S. Jeremiah, and A. Ryo, "Interpreting diagnostic tests for SARS-CoV-2," *JAMA*, vol. 323, no. 22, pp. 2249–2251, 12/14/2020 2020. <https://doi.org/10.1001/jama.2020.8259>
- [6] X. He *et al.*, "Temporal dynamics in viral shedding and transmissibility of COVID-19," *Nature Medicine*, vol. 26, no. 5, pp. 672–675, 2020. <https://doi.org/10.1038/s41591-020-0869-5>
- [7] R. Wölfel *et al.*, "Virological assessment of hospitalized patients with COVID-2019," *Nature*, vol. 581, no. 7809, pp. 465–469, 2020. <https://doi.org/10.1038/s41586-020-2196-x>
- [8] S. Kurahashi, "Estimating effectiveness of preventing measures for 2019 novel coronavirus diseases (COVID-19)," *Transactions of the Japanese Society for Artificial Intelligence*, vol. 35, no. 3, pp. D–K28 1–8, 2020.
- [9] N. Hoertel *et al.*, "A stochastic agent-based model of the SARS-CoV-2 epidemic in france," *Nature Medicine*, vol. 26, no. 9, pp. 1417–1421, 2020. <https://doi.org/10.1038/s41591-020-1001-6>
- [10] B. Wilder *et al.*, "Modeling between-population variation in covid-19 dynamics in hubei, lombardy, and new york city," *Proceedings of the National Academy of Sciences*, vol. 117, no. 41, pp. 25 904–25 910, 2020. <https://www.pnas.org/content/117/41/25904>



a. Average percentage of infected in the population of 40,000 individuals.



b. Average percentage of infected at the peak day and cumulative number at 200th day, and peak day.



c. Comparison of the peak day among different sizes of population. and peak day.

**Fig. 4:** Simulation results on the effect of long-distance travel.