An Agent-Based Model in Activity-Driven Network of COVID-19 Epidemic using Mobility and Infection Data in Tokyo 2020

Suggesting Practical Vaccine Strategies and Vaccine Passport

Kazumoto Takayanagi Graduate School of Science and Technology University of Tsukuba Tokyo, Japan e-mail: k.takayanagi1882@gmail.com

Abstract—In situations where the pandemic of Coronavirus Disease 2019 (COVID-19) has been destroying the daily lives of global human community, a model that reliably predicts the spread of infection within society would be extremely helpful for a variety of purposes. This paper presents an agent-based model over temporal networks that are fitted to real mobility data reported in Tokyo. The parameters of the model are inferred via approximate Bayesian computation to ensure that the model represents well the observed infection data. Through the simulations using this model, we demonstrate a comparison of the effectiveness of different vaccination strategies.

Keywords-agent-based modeling; activity-driven network; approximate Bayesian computation; vaccination strategies; vaccine passport.

I. INTRODUCTION

Since early in 2020, the pandemic of COVID-19 has been devastating both the daily lives and the economic activities in human communities globally. For over a year, many governmental authorities have taken strong measures such as lockdowns of cities, to restrain the spread of infection of Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2) within society. While these measures are effective in mitigating the epidemic of infection, their adverse impacts toward social and economic activities are significant. It is difficult to make political decisions that optimally balance reducing the number of infected people and maintaining social activity among individuals. Although vaccines can be a 'game-changer' in these circumstances, their supply is relatively abundant only in developed countries and has been extremely insufficient in developing nations until the middle of 2021. A social simulation model that may help predict the spread of virus infection under certain conditions would be remarkably helpful in this global context.

Agent-Based Modeling (ABM) is a method which is excellent at incorporating the heterogeneity of characteristics within agents as well as interactions and feedbacks among agents [1]. Because a well-designed ABM can serve as a tool to represent nonlinear dynamics emerging from the bottom of social system, it allows us to elucidate the mechanisms underlying complex macro phenomenon observed in society without reducing it to individual micro features. The Setsuya Kurahashi Graduate School of Science and Technology University of Tsukuba Tokyo, Japan e-mail: kurahashi.setsuya.gf@u.tsukuba.ac.jp

spreading of infectious disease within society is one of the phenomena that have received a great deal of attention from researchers devoted to this modeling [2]. Many studies have already been performed on the COVID-19 pandemic using this method for a variety of purposes [3-10]. We have developed an agent-based model validated with observed infection data to help predict the courses of the COVID-19 epidemic under various conditions.

Temporal network is a modeling framework in which we explicitly consider the times when edges are active between nodes [11]. Because specifying who is connected to whom at each time is a significant determinant of how epidemic process behave within a society, temporal networks have frequently been utilized in models that aim at representing in detail the spreading process of infectious disease [12]. The present paper focuses on a particular kind of temporal network that is called activity-driven network where each node is characterized by an activity rate a [13][14]. In activity-driven networks, this activity rate a encodes the probability of the node to generate links with other nodes at each timestep. While the behavioral restrictions have been repeatedly imposed on individuals in the pandemic situation, this framework of activity-driven network allows us to represent the successive changes of people's real activities in the model. To the best of our knowledge, we are the first to have validated the modeling of activity-driven networks by fitting agents' activities to actual community mobility data provided by Google [15] in order to examine the epidemic of COVID-19 over temporal networks.

Since the ABM is an attempt to model complex phenomena with a high degree of freedom, it is crucial to validate the modeling using observed data so that the modeling is considered as reliable. Parameter inference that fits the parameters of a model to actual data allows us to exploit ABMs for trustworthy analysis. However, ABMs do not have explicit forms for their likelihood functions due to the intrinsic complexity of ABM [16]. *Approximate Bayesian Computation* (ABC) is a flexible Likelihood-Free Inference (LFI) method for posterior and is one of the widely accepted approaches to infer the parameters of ABM [17][18]. The parameter inference in ABC is to infer the values of parameters that yield a simulator's output that agrees with observed data. The ABC algorithm for parameter inference of ABM is typically as follows [16][19],

- 1. Sample parameter θ_i from prior $\pi(\theta)$.
- 2. Simulate $f(\theta_i)$ by running simulator f with θ_i .
- 3. Reject θ_i based on the metrics of comparison between $f(\theta_i)$ and the observed data *X*.
- 4. Repeat 1-3 until a sufficiently large number of samples are obtained.

We infer the parameters of our ABM by ABC using the observed infection data so that the model is sufficiently validated.

The contributions of this paper are threefold. Firstly, we have developed an ABM over activity-driven networks where agents' activities are fitted to actual mobility data, which enables us to investigate the effects of governmental restrictions imposed on people's behavior. We also incorporate the heterogeneity of immune response to viruses among agents, a key factor in the virus spreading process. Secondly, we have inferred the posterior distributions of the model parameters via approximate Bayesian computation using real infection data. Thus, our model with estimated parameter values represents the observed data accurately and can serve as a reliable tool to predict the spread of infection under various conditions. Finally, we have demonstrated several promising vaccination strategies through the 'wouldbe' simulations carried out under virtual conditions, which shows a comparison of the effectiveness between those strategies. In addition, we investigated the possibility of 'vaccine passport' scenario via simulations that focus on the association between social activity and vaccination as well as the effects of vaccination on the epidemic.

The remainder of this paper is constructed as follows. Section II describes related work. In Section III, we describe our agent-based model in detail. In Section IV, the results of simulations by our model are demonstrated using various graphs and some implications are derived from the results. In Section V, several vaccine strategies are presented and the comparison of effectiveness between those strategies is provided through simulations. Finally, in Section VI, conclusion of this paper and future work are described.

II. RELATED WORK

Significant efforts have already been devoted to modeling the spreading process of COVID-19 to forecast the epidemic of the viruses within community as well as to suggest effective measures to prevent the pandemic. Hoertel et al. [3] proposed an ABM of the epidemic of COVID-19 in France that is well fitted to observed data in order to predict the potential impact of certain post-lockdown measures on the spreading of the epidemic. Rossetti et al. [4] presented an agent-based framework that organizes the population in five statuses and incorporates activity-driven networks to simulate the effects of public interventions on the unfolding of epidemic. Silva et al. [5] developed an ABM of the COVID-19 epidemic to assess the social and economic effects of certain scenarios with several social-distancing interventions. Aleta et al. [6] built a detailed agent-based model of SARS-CoV-2 transmission in the Boston metropolitan area using mobility data to examine the impact of testing, contact tracing, and household quarantine on the epidemic. Nishi et al. [7] used agent-based simulations of a network-based infection model to investigate network intervention strategies for mitigating the spread of infection while maintaining economic activities. Kano et al. [8] proposed an agent-based model of COVID-19 that accounts for economic activities and examined the tradeoff between health and economic damage through the simulations. Li et al. [9] used large-scale agent-based simulations to study the effectiveness of a nationwide vaccine campaign under different conditions related to vaccine efficacies and other factors. Moghadas et al. [10] developed an agent-based model of COVID-19 transmission to compare the impact of two different vaccination strategies, i.e., to vaccinate more individuals with the first dose and delay the second dose, or to continue with the recommended two-dose series.

To the best of our knowledge, few prior studies have provided an agent-based model integrated with activity-driven networks fitted to both mobility and infection data, which we propose in this paper.

III. MODEL

We generate an artificial society composed of N agents which represent individuals socially associated with others in temporal networks, which we describe in detail below.

A. Agent-based SEIR model

While the epidemic of virus infection in this society spreads over time, each agent belongs to one of four statuses, i.e., Susceptible, Exposed, Infectious, or Recovered (SEIR) at each timestep t = 1, ..., T.

A susceptible agent linked in a network with an infectious agent at timestep t gets infected according to the probability P that is obtained by multiplying the transmission rate Tr and the susceptibility s of the susceptible agent. P is calculated as,

$$P = Tr \cdot s. \tag{1}$$

An infected agent does not transmit the viruses to others during the incubation period and remains exposed. By the probability α , an exposed agent becomes an infectious one who may infect others with the viruses. An infectious agent recovers by the probability β and acquires continuous immunity to COVID-19. A recovered agent no longer transmits the viruses to others.

Since the strength of the immune response to viruses varies from person to person [20], our model presumes the heterogeneity of susceptibility to viral infection among agents. A susceptibility value s is assigned to each agent according to the Gaussian distribution, which is supposed plausible in this case [20].

B. Activity-driven networks fitted to real data

Activity-driven networks in which agents interact with others are created according to the algorithm as follows [14].

1. At each time step t, the network starts with N disconnected agents.

- 2. Each agent *i* becomes active with probability a_i and an active agent creates *m* links with *m* other randomly selected agents.
- 3. At the next time step t + 1, all the links in the network are deleted and the process returns to 1.

As observations in different real-world networks suggest [21], the activity value a is heterogeneous within agents and approximately follows a power law distribution described as

$$P(a) = Ba^{-\gamma}, \tag{2}$$

where *B* is a constant and γ denotes a scaling exponent.

In the present model, the values of activity *a* of all agents are fitted to real mobility data so that activity values change over time in accordance with the change of mobility observed in real data. The data report the *mobility* which indicates how visitors to or time spent in categorized places change within a specific region of a country. Categorized places include 'retail and recreation', 'grocery and pharmacy', 'parks', 'transit station', 'workplaces', and 'residential'. The reported mobility is shown as a positive or negative percentage since the mobility for the report date is compared with the baseline which indicates a normal value for that day of the week [15].

We have selected three categorized places that are likely to be affected by the 'state of emergency declaration', i.e., 'retail and recreation', 'transit station', and 'workplaces', because we focus on the period when the declaration was enacted. Figure 1 indicates the successive change of mobility values observed in the three selected places in Tokyo, Japan from February 15th until June 30th (137 days).

Using the mobility values reported from these three places, activity $a_{i,t}$ for agent *i* at timestep *t* is calculated as,

$$a_{i,t} = a_i \cdot M_t \tag{3}$$

where M_t denotes the average of the mobility values reported from the three categorized places at timestep *t*.

C. Settings and conditions of simulation

The number of agents N in our model is 100,000, which represents approximately one-hundredth of Tokyo's population. One timestep in the model corresponds to one day, so the simulation progresses for 137 timesteps, which represent the period from February 15th to June 30th. Figure 2



Figure 1. Change of activity in three categorized places from Feb 15^{th} to Jun 30^{th} in Tokyo.



Figure 2. Course of number of daily positive cases from Feb 15^{th} to Jun 30^{th} in Tokyo.

shows the course of the daily number of positive cases with SARS-CoV-2 during the same period in Tokyo [25]. Before running simulations, the values of both activity *a* and susceptibility *s* are allocated to all agents based on the distributions described before in this section. The values of the parameters to be estimated from the simulation results, i.e., Tr, α , and β are initially given by sampling from the prior distributions. Table I summarizes the hyperparameters of distributions from which the parameter values are sampled.

A sequence of 137 edge-lists, each of which represents the links between agents in the activity-driven network at each timestep, is generated in advance following the algorithm presented in Section III *B*. The simulation starts with one infectious agent at the first timestep.

We have encoded all the models using Python Numpy, Scikit-learn, and other toolkits.

IV. RESULTS

Simulations were carried out 3,500 times in parallel on 30 CPUs yielding a total of 105,000 results over approximately 60 hours. A series of numbers showing the course of the number of newly infected agents at each timestep was obtained as the result of each simulation running with a particular set of parameter values.

We scored the result of each simulation based on the Mean Squared Error (MSE) calculated using the observed data in Tokyo described in Figure 2. With respect to the observed data, we assumed that only 10% of the actual number of daily infected individuals was reported officially, i.e., the reporting rate is 10% [22]. By using the sets of parameter values taken from the top 1,000 simulations scored by MSE (acceptance rate = 1%), the posterior distributions of parameters were inferred. Figure 3 displays the inferred distributions for parameters Tr, α , and β .

As to β , which is the recovery rate, the inference does not appear to have converged sufficiently since the distribution seems uniform. However, the mean value of distribution (i.e.,

TABLE I. HYPERPARAMETERS OF DISTRIBUTIONS

Parameter	Distribution	Hyperparameter
Activity a	$Ba^{-\gamma}$	$B = 1, \gamma = 2.2$
Susceptibility s	$N(\mu, \sigma^2)$	$\mu=0.5, \sigma=0.2$
Transmission rate Tr	U(<i>a</i> , <i>b</i>)	a = 0.5, b = 1.0
Probability α	U(<i>a</i> , <i>b</i>)	a = 0.1, b = 0.4
Probability β	U(<i>a</i> , <i>b</i>)	a = 0.1, b = 0.4



0.25) coincides adequately with the length of infectious period that empirical research suggests [23]. The median time of infectious period suggested in research is 4.1 days, so it is the inverse of the mean value of distribution estimated for recovery rate. With respect to α which is the probability of becoming infectious, the distribution appears to gradually converge around the value 0.27, although the shape of the distribution is not evident enough. The mean of the distribution could have converged to approximately this value, had we carried out far more simulations. The inverse of the value 0.27 coincides with the median time of incubation period, i.e., 4-5 days, as suggested by empirical research [24]. As for Tr, i.e., transmission rate, the distribution seems to converge moderately between 0.9 and 1.0, while its shape is somewhat ambiguous. If the mean of the prior for Tr had been larger and many more simulations had been run, the shape of distribution could have been more obvious.

To validate our model, we compared the simulation results with real data, in terms of the 7-day average number and the cumulative number of newly infected agents, by calculating Pearson correlation coefficient. As for the 7-day average, the top 10 simulation results scored by MSE reached approximately 0.90 (p<1e-40) in Pearson correlation coefficient, as indicated in Figure 4. For the cumulative number, the top 10 results evaluated by MSE reached as high as 0.993 (p<1e-100) in Pearson correlation coefficient. Therefore, we appear to have sufficiently validated our model through parameter inference that fits the model to the observed data, and we have succeeded in identifying the corresponding parameter values. The values identified for parameters *Tr*, *a*, and *β* are 0.99, 0.26, and 0.31, respectively.

With these parameter values identified to represent the actual data, we performed additional simulations under imaginary conditions. The simulations suggest the results that could have been realized when people's activity had not been restricted. Without restriction on people's activity, the number of daily infected cases would have reached over 150,000 and

more than half of the people could have been infected, leading to *herd immunity*. However, on the other hand, as Figure 5 shows, the cumulative number of social links in the networks fitted to real mobility data is 33% less than that in virtually generated activity-driven networks in which the activities of agents do not decline. The decrease of links in the networks can be considered as the economic and social cost to restrain the spread of infection with COVID-19.

Although it will be extremely difficult to evaluate the effectiveness as well as legitimacy for measures like 'state of emergency declaration', predictions using a well-validated simulation model may help do that more rationally.

V. VACCINE STRATEGIES

While vaccines can arguably be a 'game-changer' in the pandemic, their supply is still inadequate globally. Examination of the effectiveness of various vaccine strategies using a model well-fitted to the observed data will be practically needed in this context.

A. Comparison between various strategies

Utilizing virtual activity-driven networks in which agents' activities are not restricted, we demonstrate a comparison of the effectiveness of three vaccination strategies through simulations of our model with parameter values identified in Section IV.

We examine the following three vaccine strategies:

- 1) Random: vaccinating randomly chosen agents at each timestep.
- 2) Priority for highly susceptible people: preferentially vaccinating the most susceptible 30% of agents, e.g., elderly, then randomly.
- Priority for agents relatively active in networks: preferentially vaccinating the most active 30% of agents, then randomly.



Figure 4. Top 10 results (blue, dashed) scored by MSE, observed data (red), daily(left), 7-day average(center), cumulative(right).



In addition, we run simulations under the following different conditions:

- a) Vaccine availability: 1,000 shots (1% of the people) or 500 shots (0.5% of the people) at each timestep.
- b) Vaccine efficacy: reducing the viral susceptibility of the agents by 80% or 30% (e.g., against variants).

The availability of vaccines and their efficacy may affect the evaluation of the performance of a vaccine strategy, so simulations are performed with these different conditions in mind. We run simulations over activity-driven networks with no governmental restrictions on the agents' activities. The vaccination starts from the 40th timestep in these simulations.

Figure 6 displays the courses of the number of newly infected agents when each of three strategies is performed under four different conditions. It is observed that preferential vaccination for agents who are more active in the networks is the most effective strategy under all conditions. It is also apparent that the difference in effectiveness between the three strategies is eminent when vaccines are abundant and effective. As the speed of vaccination slows down, or the effects of vaccines against viruses decline, the three strategies differ less evidently in their performance.

Figure 7 shows the cumulative number of infected agents among the highly susceptible group (e.g., elderly) when each of the three strategies is undertaken under four different conditions. Though it may be counterintuitive, the number of infected agents among the highly susceptible ones is lower when active agents are prioritized than when highly susceptible agents are vaccinated first. Taking these observations together, it can be implicated that preferentially vaccinating for more active individuals in the network is a promising strategy, though it may be practically difficult to carry out.

B. Vaccine passport

Once people are vaccinated, they will probably be enthusiastic to resume their activities they have been forced to abandon during the pandemic. 'Vaccine passport', which allows vaccinated people to restart their business as well as leisure activities is practically needed. However, can vaccine passport really achieve both restraint of virus spreading and stimulation of social activity? It will make sense to give a clue to this question by running simulations with a model that represents well the observed data.

Therefore, we compared a vaccine passport scenario with other scenarios with respect to the mitigation of infection as well as the decrease of social links. We examined the following three scenarios.

- 1) Vaccine passport: agents are allowed to increase their activity by 20% after being vaccinated.
- 2) Sustained restriction on activity: agents are forced to reduce their activities even after vaccination.
- No restriction on activity: no restriction on agents' activities is executed.

Simulations of our model are carried out over activitydriven networks where the activities of all agents are initially reduced by 5%. The vaccination starts from the 40th timestep.

In terms of the three scenarios, Figure 8 shows the courses of cumulative number of infected agents among the highly susceptible group (left) and the comparison of total number of social links generated in the networks (right).

When compared with no restriction scenario, vaccine passport scenario reduces the number of infected agents by 73% while it only loses social links by 4.8%. Sustained restriction scenario reduces the spread of infection more efficiently while decreasing social links more significantly.



Based on these observations, it may be suggested that vaccine passport definitely helps the activity of society resume while mitigating the spread of infection considerably. Thus, provided that the vaccine's efficacy is sufficient as expected among individuals and health care system is unlikely to be overloaded, vaccine passport could be a promising means to improve the situation of COVID-19.

VI. CONCLUSION AND FUTURE WORK

We have developed an agent-based model of COVID-19 infection with activity-driven networks that are fitted to actual mobility data. We inferred the parameters of our model via approximate Bayesian computation with 105,000 results of simulations. Through additional simulations under certain conditions, we also examined the effectiveness of several vaccination strategies and suggested promising one.

In future work, we will attempt to infer posterior of parameters more sufficiently by using deep learning [19] as summary statistics in approximate Bayesian computation.

ACKNOWLEDGMENT

We appreciate Google for providing useful data.

REFERENCES

- U. Wilensky and W. Rand, "An introduction to agent-based modeling: Modeling natural, social, and engineered complex systems with Netlogo," MIT Press ISBN978-0-262-73189-8 2015.
- [2] H. Arduin, M. Domenech de Cellès, D. Guillemot, L. Watier, and L. Opatowski, "An agent-based model simulation of influenza interactions at the host level: insight into the influenza-related burden of pneumococcal infections," BMC Infectious Diseases 17:382 2017.
- [3] N. Hoertel et al., "A stochastic agent-based model of the SARS-CoV-2 epidemic in France," nature medicine letters vol.26 pp. 1417-1421 Sep. 2020.
- [4] G. Rossetti, L. Milli, S. Citraro, and V. Morini, "UTLDR: an agent-based framework for modeling infectious diseases and public interventions," working paper arXiv:2011.05606v1 [cs.SI] 11 Nov. 2020.
- [5] 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 agentbased model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions," Chaos, Solitons and Fractals 139 110088 2020.
- [6] A. Aleta et al., "Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19," nature human behavior vol 4 pp. 964-971 Sep. 2020.
- [7] A. Nishi et al., "Network interventions for managing the COVID-19 pandemic and sustaining economy," PNAS vol.117 no.48 pp. 30285-30294 Dec. 1 2020.
- [8] T. Kano, K. Yasui, T. Mikami, M. Asally, and A. Ishiguro, "An agent-based model of the interrelation between the COVID-19 outbreak and economic activities," Proc.R.Soc.A 477:20200604.
- [9] J. Li and P. Giabbanelli, "Returning to a Normal Life via COVID-19 Vaccines in the United States: A Large-scale Agent-Based Simulation Study," JMIR Med Inform vol.9 iss.4 e27419 pp. 1-19 2021.



Figure 8. Cumulative number of infected agents among highly susceptible ones(left), cumulative number of links in the networks(right).

- [10] S. M. Moghadas et al., "Evaluation of COVID-19 vaccination strategies with a delayed second dose," PLOS Biology 19(4):e3001211.
- [11] P. Holme and J. Saramäki, "Temporal networks," Phys. Rep. vol.519 iss.3 pp. 97-125 Oct. 2012.
- [12] N. Masuda and P. Holme, "Introduction to temporal network epidemiology," in N. Masuda and P. Holme(editors), "Temporal Network Epidemiology," Springer ISBN 978-981-10-5286-6 2017.
- [13] M. Karsai, N. Perra, and A. Vespignani, "Time varying networks and the weakness of strong ties," Sci.Rep.4:4001 2014 doi: 10.1038/srep04001.
- [14] M. Karsai and N. Perra, "Control strategies of contagion processes in time-varying networks," in N. Masuda and P. Holme(editors), "Temporal Network Epidemiology," Springer ISBN 978-981-10-5286-6 2017.
- [15] Google COVID-19 Community Mobility Reports https://www.google.com/covid19/mobility/ [retrieved: June 2021].
- [16] E. v. d. Vaart, M. A. Beaumont, A. S. A. Johnston, and R. M. Sibly, "Calibration and evaluation of individual-based models using Approximate Bayesian Computation," Ecological Modelling 312 pp. 182-190 2015.
- [17] M. A. Beaumont, "Approximate Bayesian Computation," Ann.Rev.Stat. Appl 2019.6 pp. 2.1-2.25.
- [18] B. Lambert, A. L. MacLean, A. G. Fletcher, A. N. Combes, M. H. Little, and H. M. Byrne, "Bayesian inference of agentbased models: a tool for studying kidney branching morphogenesis," J.Math.Biol. 76 pp. 1673-1697 2018.
- [19] M. Kesson, P. Singh, F. Wrede, and A. Hellander, "Convolutional neural networks as summary statistics for approximate Bayesian computation," a preprint arXiv:2001.11760v4 [stat.ML] 14 Sep 2020.
- [20] J. L. Schultze and A. C. Achenbrenner, "COVID-19 and the human innate immune system," Cell 184, April 1, 2021.
- [21] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, "Activity driven modeling of time varying networks," Sci.Rep.2:469 2012 doi:10.1038/srep00469.
- [22] G. Kobayashi, S. Sugasawa, H. Tamae, and T. Ozu, "Predicting intervention effect for COVID-19 in Japan: state space modeling approach," BioScience Trends Advance Publication doi: 10.5582/bst.2020.03133.
- [23] H. Y. Cheng, S. W. Jian, D. P. Liu, T. C. Ng, W. T. Huang, and H. H. Lin, "Contact Tracing Assessment of COVID-19 Transmission Dynamics in Taiwan and Risk at Different Exposure Periods Before and After Symptom Onset," JAMA Intern Med. 2020:180(9) pp. 1156-1163.
- [24] Centers for Disease Control and Prevention (CDC) in the United States, "Clinical Presentation: incubation period," https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinicalguidance-management-patients.html [retrieved: June 2021].
- [25] COVID-19 Information Website Tokyo Metroplitan Gov. https://stopcovid19.metro.tokyo.lg.jp/en/ [retrieved: June 2021].