

Visualization of the COVID-19 Spread in Universities and the Strategies of Flattening the Curve by GNNs

Huaze XIE[†], Yuanyuan WANG^{††}, Da LI[†], and Kawai YUKIKO^{†,††}

[†] Division of Frontier Informatics, Kyoto Sangyo University
Kamigamo Motoyama, Kita Ward, Kyoto, 603-8555, Japan

^{††} Graduate School of Sciences and Technology for Innovation, Yamaguchi University
2 Chome-16-1 Tokiwadai, Ube, Yamaguchi 755-0097, Japan

^{†††} Division of Information Engineering, Osaka University
1-1 Yamadaoka, Suita, Osaka 565-0871, Japan

E-mail: [†]x.huaze2019@outlook.com, ^{††}y.wang@yamaguchi-u.ac.jp, ^{†††}{lida,kawai}@cc.kyoto-su.ac.jp

Abstract We consider the reopening of university after the Coronavirus disease (COVID-19) pandemic or influenza as complications were contained, and the spread of COVID-19 on campus. In this work, we simulate and visualize the spread of COVID-19 after the first infected person appears on the university campus. We propose the dynamic transmission process of potentially infected students from two paths: enclosed spaces (face-to-face lectures, restaurants, bus, and other public transportation) and student social activities (group discussion, outdoors face-to-face talking). In the first place, we train a set of COVID-19 transmission data through GNN (Graph Neural Network) to get the spreading efficiency and transmission model of the virus indoors and outdoors. Next, we define a graph structure that includes a spatial-time network to describe the spread of COVID-19 and apply it in data visualization. Furthermore, we discuss the spread curve of COVID-19 on campus based on spatial-time graph structure. Finally, we discuss the COVID-19 spread with three strategies and predict the transmission curves of these strategies by GNNs. Our results show that after the first COVID-19 patient appeared on the campus, the spread of the virus can be effectively controlled in a week later by separating the social activities of students on campus and the dense indoor groups.

Key words COVID-19 Spread Process, Spatio Temporal Data Mining, Epidemic Diffusion Modeling, Message Passing Neural Network Framework

1 Introduction

A few months after the outbreak of the coronavirus disease 2019 (COVID-19) in Wuhan, the epidemic has been brought under control in some countries, and restrictions on public places have been gradually lifted. Therefore, we try to put forward feasible suggestions for university restart (face to face teaching) from three aspects: classroom, social, and diet. A research (Per Block,..., etc) describe and evaluate the effectiveness of three distancing strategies designed to keep the curve flat and aid make the spread of the COVID-19 controllable [1]. And Block's social networks have a guiding role in our research for students to return to school. However, Block's three strategies are not applicable, only work in a virtual environment with 500 and 4000 cases, and there is no significant difference in the results. In this article, we fit the real COVID-19 data to a Poisson likelihood (ML) model to complete a priori graph structure data set [12], and

combine the seat distribution data of classrooms and school restaurants with the student's class schedule, import it into a graph convolutional neural network (GCNN) with susceptible infected recovered (SIR) process training model [16]. We discuss the efficiency of the COVID-19 transmission and fit the process time curve, the widely utilized method of physical isolation virus transmission is quantitatively analyzed and applied to the student spread.

The COVID-19 pandemic is caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). And SARS-CoV-2 is spread primarily via respiratory droplets during close face-to-face contact [9]. Transmission may also occur via aerosols (smaller droplets that remain suspended in air), but it is unclear if this is a significant source of infection in humans outside of a laboratory setting [15]. Research of the different contact methods of various groups found that the contact infection methods include [3]: (1) family: live together (18.07%), meal together

Table 1 The datasets of medical text description

	Features	Description	Granularity
8*Labels	Infected	The cumulative number of PCR tests positive person (including recovered and dead).	Regional, provincial
	Not Infected	The cumulative number of PCR tests negative person.	Regional, provincial
	Susceptible	Not infected persons, but have close contact with infected.	Regional, provincial
	Asymptomatic	Infected person without obvious symptoms.	Regional, provincial
	Treatment	The cumulative number of infected in hospital.	Regional, city
	Severe	The cumulative number of infected in ICU ward.	Regional, provincial
	Recovery	The cumulative number of recovered.	Regional, provincial
	Death	The cumulative number of dead.	Regional, city

(11.75%); (2) friends (including neighbors): outdoor conversation (20.00%), dinner/meeting/entertainment (12.50%), take the same transportation (4.55%); (3) general population: the same hall (1.94%), the same supermarket (0.56%). We apply these contact infection methods as a Markov process transmission hyperparameter to the SIR process [5] in GCNN to describe the virus transmission efficiency in the message passing networks (MPNN) framework [13].

In this article, we import the COVID-19 medical text set and regional monitoring data into the graph structure in GNN and modify it into the epidemic model SIR. We evaluate the proposed approach with a case study that analyses the COVID-19 outbreak in universities. Our training data is the Japan official COVID-19 data from February 8th to June 20th and compares the forecasts of the trained model with real data on the epidemic.

2 Method Background

In this section, we introduce some basic concepts on which our method relies. Including GCN and MPNN in GNN, we expect to complete the learning model through text data mining and Spatio-Temporal data analysis of COVID-19, and apply it to the classic epidemic spread model SIR to predict the early trends of COVID-19. Since our research is expected to be able to control the trend in the early stage of disease transmission, we have made improvements to the epidemic transmission model.

2.1 Graph Neural Network

The graph G in the graph neural network is composed of nodes and edges: $G = (V, E)$, and the goal of GNNs is to understand the representation $h_v \in R^m$ of each node v . The representation of each node is donated from the features of the node, the edge of connection nodes, the neighbor representation and the neighbor of the nodes:

$$h_v = f(x_v, x_{co[v]}, h_{ne[v]}, x_{ne[v]}) \quad (1)$$

Where f is a local transaction function, and this graph neural network based on the cyclic structure makes the entire graph converge and predicts through the propagation of

node information. However, this GNNs do not set independent parameters for the edges, which means the feature of the edges cannot be learned. The gated graph sequence neural networks (GGNNs) improved the problems from GNNs through a new status update function:

$$h_v^{t+1} = GRU \left(h_v^t, \sum_{u \in ne[v]} W_{edge} h_u^t \right) \quad (2)$$

And the evaluation of GGNNs shows that encoding the schema structure improves our parser accuracy from 33.8% to 39.4%, dramatically above the current state of the art, which is at 19.7%. In this article, we describe the characteristics of COVID-19 medical text through a graph structure, and we classify these according to the features and labels in Table 1.

2.2 Graph Convolution Neural Network

For GCNNs, the realization of convolution using graph Fourier transform like spectral convolution. It means the Laplacian operator in the spectral-domain is derived by using the Laplacian matrix of a graph. The formula of graph convolution is derived by analogy with convolution in Euclidean space in the spectral-domain. For a graph $G = (V, E, W)$, V is a finite set of $|V| = n$ nodes, E is a set of edges among nodes and $W \in R^{n \times n}$ is a weighted adjacency matrix representing the weights of edges. An input vector $x \in R^n$ is seen as a signal defined on G with x_i denotes the spectral information of node i .

For the spectral-domain convolution of an undirected graph, the relationship between the feature function f and the convolution kernel g can be expressed as follows:

$$(f * g) = F^{-1}[F[f] \odot F[g]] \quad (3)$$

$$(f *_G g) = U(U^T f \odot U^T g) = U(U^T g \odot U^T f) \quad (4)$$

And we can put $U^T g$ as an learnable kernel write as g_θ , the final convolution formula is:

$$o = (f *_G g)_\theta = U g_\theta U^T f \quad (5)$$

Suppose the hidden state of layer l is $h^l \in R^{N \times dl}$, the status

of the spectral-domain convolutional layer is updated to:

$$h_{:,j}^{l+1} = \sigma \left(U \sum_{i=1}^{d_l} \Theta_{i,j}^l U^T h_{:,i}^l \right) \quad (6)$$

2.3 Message Passing Neural Network

Message passing networks (MPNN) is a formal framework for spatial convolution [6]. It decomposes spatial convolution into two processes: message transmission and status update. And a study improved the efficiency of information sampling and message transmission with graph sample and aggregate (GraphSage) [7]. The GCNN layer can be summarized as:

$$\mathbf{x}_i^{(k)} = \sum_{j \in \mathcal{N}(i) \cup \{i\}} \frac{1}{\sqrt{\deg(i)} \cdot \sqrt{\deg(j)}} \cdot \left(\Theta \cdot \mathbf{x}_j^{(k-1)} \right) \quad (7)$$

where the neighboring \mathcal{N} node features are first transformed by a weight matrix Θ , normalized by their degree, and finally summed up. With $\mathbf{x}_i^{(k-1)} \in R^F$ denoting node features of node i in layer $(k-1)$ and the normalization coefficients are derived by the node degrees $\deg(i)$ for each node i which gets transformed to $1/(\sqrt{\deg(i)} \cdot \sqrt{\deg(j)})$ for each edge $(j, i) \in \mathcal{E}$. The edge convolutional layer can be summarized as:

$$\mathbf{x}_i^{(k)} = \text{ADD}_{j \in \mathcal{N}(i)} h_{\Theta} \left(\mathbf{x}_i^{(k-1)}, \mathbf{x}_j^{(k-1)} - \mathbf{x}_i^{(k-1)} \right) \quad (8)$$

where h_{Θ} denotes a random transmission obey the SIR process, and the aggregation method of GraphSage is "ADD".

2.4 Epidemic Diffusion Model

Through the GNN model learning the process of COVID-19 transmission, we can understand some general behaviors of this virus. However, since our main purpose is to develop a predictive model in order to understand the key factors that impact the COVID-19 transmission, I'll move on to one of the most famous epidemic diffusion models: SIR.

Infectious disease modeling aims to understand whether or how a virus can spread among a population. Compartmental models are the most used mathematical tool to accomplish the task, they divide the population into compartments and define how people may progress between them. In most cases, these models can be described through ordinary differential equations and rely on the estimation of various epidemiological parameters (e.g. the contact and recovery rates) representing the interaction between compartments.

For the general SIR model of the dynamic epidemic transmission process, transmission remains local when infection rates are below critical values, and this growth is similar to the growth of percolation clusters, which are governed by the "dynamical percolation" universality class (finished clusters are in the same class as static percolation, while growing clusters have additional dynamic exponents). In asynchronous models, the individuals are considered one at a time, as in kinetic Monte Carlo or as a "Stochastic Lattice Gas."

SIR is a simple model that considers a population that

belongs to one of the following states:

- (1) Susceptible (S): The individual hasn't contracted the disease, but she can be infected due to transmission from infected people;
- (2) Infected (I): This person has contracted the disease;
- (3) Recovered/Deceased (R): The disease may lead to one of two destinies: either the person survives, hence developing immunity to the disease or the person is deceased.

Susceptible people may become infectious depending on the contact rate β they have with other ill people, as well as infectious ones may heal (or die) after the removal time $T_r = 1/\gamma$. Since the goal of our method is to be able to suppress the transmission of COVID-19, our interest is focused on the short term, we will consider that people develop immunity (in the long term, immunity may be lost and the COVID-19 may come back within a certain seasonality like the common flu) and there is no transition from recovered to the remaining two states. With this, the differential equations that govern the system are:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \quad (9)$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I \quad (10)$$

$$\frac{dR}{dt} = \gamma I \quad (11)$$

All parameters are usually estimated considering the characteristic of the disease and optimized to fit at best the observed data. Forecast predictions are usually accomplished by making hypotheses of how external factors (e.g. governmental countermeasures, vaccine distribution) affect the above-mentioned parameters.

3 Method

In this section, we discuss the entire spread of COVID-19 and the flow of data in the model. Our goal is to describe a complete COVID-19 propagation process through the graph neural network and the message passing network framework in graph convolution. However, due to the limitation of data set and time, we only focus on the process of COVID-19 spreading in the early stage (from February to June). In the graph structure information of our GNN model, each node represents an individual, region, or city, and the edge information collects the spread of epidemics, and we believe that this spread depends on the movement of individuals. In addition, the dynamic spatio-temporal data contained in the graph structure allows us to visualize the entire propagation process.

3.1 Spatio-Temporal Data Mining

In this article, we describe two datasets: COVID-19 medical text set and regional monitoring data set. Both need to

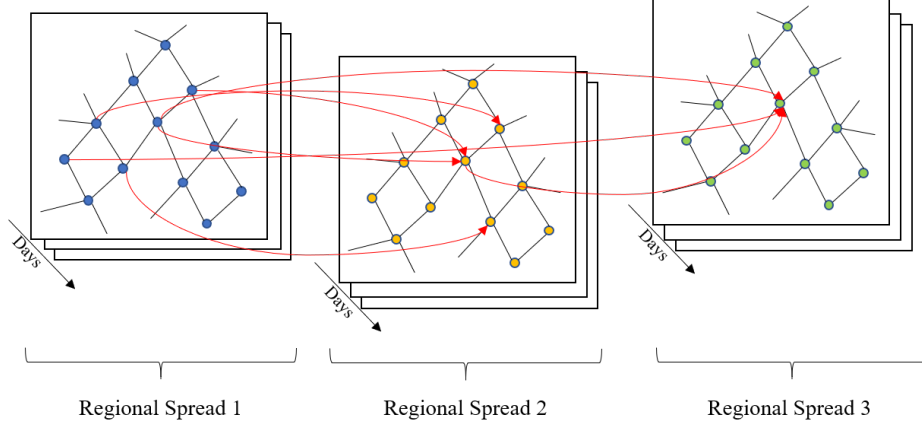


Figure 1 The Graph Structure with COVID-19 medical text set and regional monitoring data set nodes, and Spatio-Temporal edges

discuss time validity and spatial validity. Therefore, we propose a spatio-temporal analysis method based on the graph structure of GNN that analyzes COVID-19 related data, including:

Time validity: During our research on COVID-19-related data, we found that both medical text sets and monitoring data sets are time-sensitive, which is specifically reflected in the changes of labels and features. For instance, when the first infected sample appears in the graph structure data set composed of a certain group or region, because the SARS-CoV-2 virus has an incubation period, it is difficult to monitor and change the individual label (uninfected change to infected) at the initial stage. And the characteristic representation of the data in the graph structure has hysteresis. Therefore, it is necessary to add the average incubation period(5 days) of the virus in the monitoring process.

Spatial validity: In the process of spatial transmission of the data we obtain, we need to consider the relationship between nodes in the graph structure, and this relationship is expressed by the activities of individuals in the area. And this characteristic is often associated with time validity, specifically at the population level, the characteristics of the population in a geographical area are not easy to change, so we only need to observe the virus spread in a fixed area (county, city, province) process. However, due to the occurrence of regional-to-regional transmission (usually depending on the traffic between cities), we use some government departments' data on traffic flow to monitor the mobility between regions (depending on government policies and time Relevance) and describe the relevance between each graph structure.

3.2 Graph Structure

In the previous chapter, we introduced the graph neural network and graph convolution architecture. In this section, we first classify the COVID-19 medical text information ac-

cording to the labels and features given in Table 1. The features in constructs a set of graph structure data. The nodes summarize the main features in each medical text, and the side information describes the relationships that exist between the texts.

We introduce the representation of medical text data and regional monitoring data set in the spatio-temporal graph structure in Figure 1. And the definition of graph structures including:

The finite sequence of directed weighted graphs $G_i = (N_i, E_i, W_i)$, where $i \in R_T$, and $R_T = [1, 2, 3, 4, \dots, T]$. N_i donates the nodes from data sets and all the graphs in the sequence share the same nodes. E_i donates the spatio and temporal features in same regional spread structures with a time line and different regional spread structures with different time lines. W_i donates the weights of spatio and temporal graph sets.

We use $x_i^k, i \in R^l$ denotes the feature vector of our datasets where $n^k \in N$. $A_i, i \in R_T$ denotes the sequence of adjacency matrices, where A_i refers to the graph G_i . $X_i, i \in R_T$ is the sequence of the matrices of feature vectors, $X_i \in R^{l \times l}$ denotes the matrix of feature vectors at time i . And $Y_i, i \in R_T$ is the sequence of dependent variables' matrices we want to observe.

Our method means the feature vectors $x_i^k, i \in R^l$ consists of the features which characterize the regional spread graph at the time step $i \in R_T$. In the feature of graph structure with the region as the unit, the dynamic spatio-temporal attributes are the number of inhabitants and the density of population are considered to be static features since their changes during the observation period are imperceptible; On the other hand, examples of dynamic features are the mean radius of movements of individuals within the place and percentage of individuals who spend more than three hours outside their home, which vary according to the

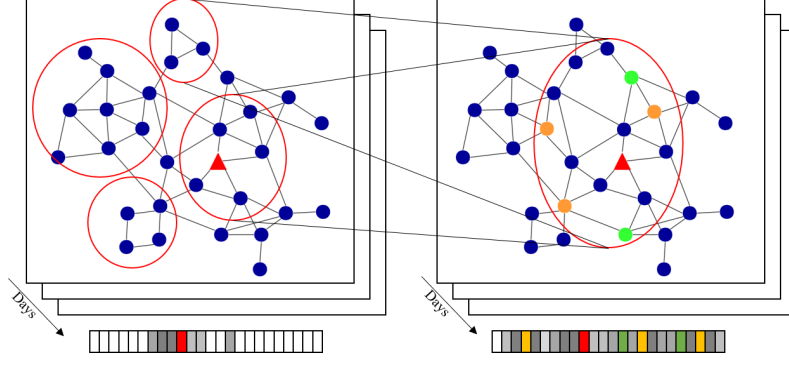


Figure 2 The Visualization of dynamic Spatio-Temporal regional spread processing

restrictive measures imposed by local authorities.

3.3 Training

We utilize the spectral graph convolution model to enhance the node feature state and connecting the output of each layer with the embedding learned from the time node feature, our prediction model can be expressed as [10]:

$$H_0 = \text{mlp}(x_t | x_{t-1} | \dots | x_{t-d}) \quad (12)$$

$$H_{l+1} = \sigma(\hat{A}H_l W_l) \parallel H_0 \quad (13)$$

$$P = \text{mlp}(H_s) \quad (14)$$

where H represents the hidden state at layer l , \hat{A} is the spectral normalized adjacency matrix, W is the learned weight matrix at layer l , \parallel is the concat operator, and σ is a non-linearity (in our case, a Relu). A visualization shown in Figure 2 discuss this process. The first embedding: H_0 is simply the output of a 'mlp'(Multi-layer Perceptron) over the node's temporal features x at time t reaching back d days, while the final prediction is the output of a 'mlp' over s spatial hops. We also discuss the situation where the same individual appear in different groups (red circle) in the entire region during training, and led to the accelerate spread of this region in Figure 2 with orange and green nodes.

3.4 SIR Model Modification

After the training model, we need to make the description process of the graph structure conform to the medical model of the epidemic. We define the graph sets $G_i, i \in R_T$ be a sequence of T graphs, and $A_i, i \in R_T$ and $X_i, i \in R_T$ the related sequences of adjacency matrices and feature matrices, respectively. $Y_i, i \in R_T$ be the sequence of label matrices where the i -th matrix refers to the day following that of the corresponding feature matrix, i.e. the contact rate between individuals at time t affects the future epidemic trends ($t_0 > t$). Then, the task we aim to solve consists in learning a predictive function f such that [8]:

$$f((G_i)_{i \in R_T}, (A_i)_{i \in R_T}, (X_i)_{i \in R_T}, (Y_{i-1})_{i \in R_T}) = y_i \quad (15)$$

Next, the spread trend we get can be applied to the SIR

model to describe the spread efficiency of COVID-19 on the timeline. Which includes the the number of Susceptible (S_t^r), Infected (I_t^r) and Recovered (R_t^r) individuals. Supposing that $R_0^r = 0$ and $S_0^r = P^r - I_0^r$, P is the number of inhabitants of graph spread region r and I is the initial number of infections (provided by the user), the evolution of the epidemic is caught by the following equations:

$$S_{t+1}^r = S_t^r - \frac{\beta_t^r S_t^r I_t^r}{P^r} \quad (16)$$

$$R_{t+1}^r = R_t^r + \gamma^r I_t^r \quad (17)$$

where β_t^r is the contact rate returned by the Contact Rate Estimator module, γ^r is the recovery rate of graph spread region r . After modifying the epidemic spread model SIR, we simulate student activities on the university campus and made an overall prediction. Since our research and training objectives are mainly focused on the early stages of the spread of the virus, we will describe the early stages of the COVID-19 spread on university campuses in the next section.

4 Case Prediction and Strategies

In this section, we introduce a case study of students' spread simulation in universities. We utilize the spatial graph convolution model with message passing network framework to simulate the COVID-19 transmission and fit the process time curve. We focus on the early stages of the COVID-19, and discuss some strategies for flattening the curve.

4.1 University Spread Simulation

The COVID-19 spread curve describes the spread over time after an infected person appears in a group. This spread curve is also used to find the first infected person in a group [13]. We utilize the COVID-19 spread curve to describe the infection process and propose the response plan for restarting the university campus. For COVID-19, the incubation period is 1 to 2 weeks [11]. And in this article, we import the median incubation period: 5.1 days to describe the spread processing of a group on a university campus [11].

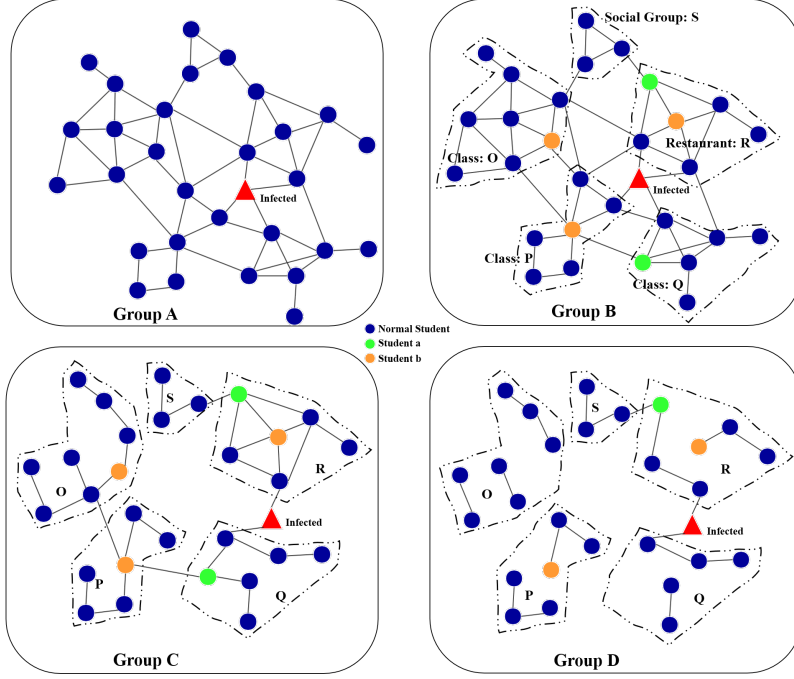


Figure 3 The COVID-19 Spread Processing with Different Responses on University Campus.

We discuss the SARS-CoV-2 spread processing from the first infected student to this student is detected to be infected.

Group A in Figure 3 is a normal spreading process. We use the infection curve A (shown in Figure 4) to describe the number of newly infected students in this group changes over time. From the infection curve A (shown in Figure 4), the first infected student appears in stage 3, the peak of the newly infected students reach in stage 7, and the newly infected student disappears in stage 12. In the spreading process A, we import the median incubation period: 5.1 days in Figure 4. After stage 7 (black dotted line), the infected students will be confirmed in Group A, but in this stage, 72.3% of students in Group A have been infected.

On university campuses, the COVID-19 spreading process is even worse. Group B of Figure 3, introduces the actual spread model on a university campus. Group B includes an infected student (red triangle), class (O, P, Q), restaurant group (R), and social activity group (S). And the blue nodes represent normal students, the green and orange nodes represent student "a" and "b", respectively.

We discuss the class spreading process with student seats. If the student's seat is not fixed, the spread situation will be much worse, and the neighbors of the infected student are random. In the free-spreading stage (from the first infected to the end of the incubation period shown in Figure 4), the speed of spread in class is much faster than students with fixed seats. We discuss this situation with an experiment in the next section.

Similarly, we discuss the same students to appear in different classes shown in Group B. Student "a" appears in Class Q and Restaurant R, and student "b" appears in Class O, P, and Restaurant R. In the spread curve of Group B (red curve), the first infected student appears in stage 3, the peak of the newly infected students reach in stage 6, and the newly infected student disappears in stage 12. And more than 83.3% of students in Group B are infected until the first infected is confirmed. Compared with Group A, the peak of the newly infected student in Group B comes earlier. It means the SARS-CoV-2 is more efficient in the free-spreading stage of Group B after the first infected student in this group is detected.

4.2 Flatten Curve

If we can flatten the SARS-CoV-2 spread curve with some strategies, we can control the transmission efficiency of the virus in the free-spreading stage. A work introduced the benefits and costs of flattening the curve for COVID-19 [14]. According to the Centers for Disease Control and Prevention (CDC) data in the United States, the prevalence of people aged 18-29 years is the highest (1231710 cases), but the fatality rate is much lower than that of people over 30 years old (792 cases, 0.5%). A work introduced the hospital critical care capacity limits and mortality from SARS-CoV2 cases in US counties by flattening the curve [2]. In our university transmission model, we fixed student seats in Group C and forced students to maintain social distance and wear masks. This series of physical protection measures are introduced

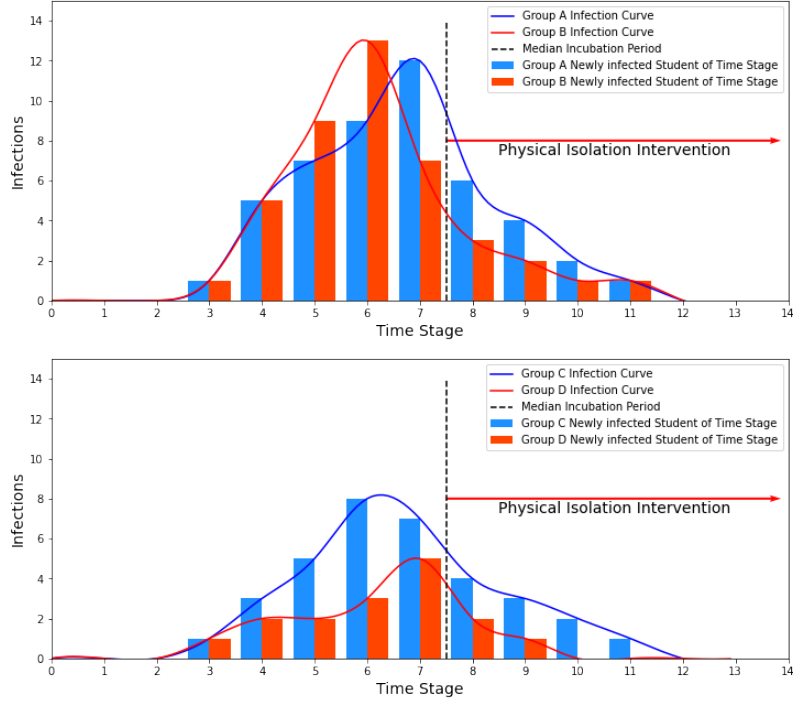


Figure 4 The COVID-19 Spread Curve for Group: A, B, C, D.

in a research with a non-medical environment [4]. The blue spreading curve shown in the bottom bar Figure 4 discuss the process of Group C shown in Figure 3 in our model of universities. Compare with group B, group C adopts physical protection measures, and the spread path of class is reduced. In Group C, the first infected student appears in stage 3, the peak of the newly infected students reach in stage 6, and the newly infected student disappears in stage 12. About 48.9% of students in Group C are infected until the first infected is confirmed.

In addition, we limit the number of students in the classroom and restaurant in Group D, the first infected student appears in stage 3, the peak of the newly infected students reach in stage 7, and the newly infected student disappears in stage 10. Compare with Group C, the peak of the newly infected students in Group D is postponed, and the spread of COVID-19 is shortened. About 28.9% of students in Group D are infected until the first infected is confirmed.

In our proposed COVID-19 transmission model on a university campus, fixing student seats, limiting the number of indoor students, adopting strict social distance and physical protection measures can reduce the SARS-CoV-2 transmission efficiency in the free-spreading stage.

5 Results

The simulation training results shown in Figure 5, including Open Group (blue curve) and Closed Group (red curve) represent the change in the number of infected students throughout the university over time. In Open Group,

the first infected student appears in stage 3, the peak of the newly infected students reach in stage 13, and the newly infected student disappears in stage 20. More than 8.67% of students in Open Group are infected until the first infected is confirmed. And without the intervention of physical isolation measures during the free-spreading stage, 214 students would be infected eventually. In Closed Group, the first infected student appears in stage 3, the peak of the newly infected students reach in stage 11, and the newly infected student disappears in stage 17. About 4.67% of students in Closed Group are infected until the first infected is confirmed. And without the intervention of physical isolation measures during the free-spreading stage of Closed Group, 95 students would be infected eventually. Compare with Open Group and Closed Group, since the number of students in the experiment has reached 600, COVID-19 cannot spread on a large scale under the premise of implementing the first strategy. From the prediction results, the Closed Group is more resistant to COVID-19 and has a smaller spreading range. Before the physical isolation measures are implemented, the number of infected students in Closed Group less than in Open Group.

6 Conclusion

In this research, we analyzed the medical text and regional monitoring data set of COVID-19 through graph neural network and spatio-temporal dynamic data. We revised the obtained prediction model to the classic epidemic spread model SIR and applied it to simulate the overall distribution predic-

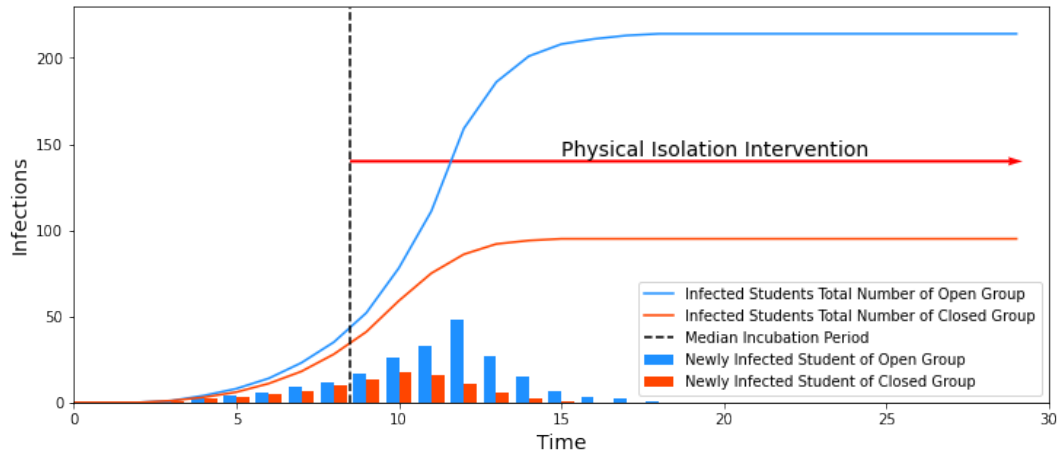


Figure 5 The Results of a Comparative Experiment Between Closed Group and Open Group.

tion of infected persons on university campuses. In addition, we also proposed a strategy of fixing the seats of students in the classroom to slow down the speed of transmission and flatten curve in the group region. The simulation analysis results show that our social distancing strategies can reduce the risk of COVID-19 transmission after school reopen. In the future work, we will utilize the data of a certain city to complete the prior data model of the location, and promote the SIR message-passing network with the local students' social networks for GCNN training.

Acknowledgment

This work was partially supported by JSPS KAKENHI Grant Numbers JP17K12686, 19H04118, and JP20H00584.

References

- [1] Block, P., Hoffman, M., Raabe, I.J., Dowd, J.B., Rahal, C., Kashyap, R., Mills, M.C.: Social network-based distancing strategies to flatten the covid-19 curve in a post-lockdown world. *Nature Human Behaviour* pp. 1–9 (2020)
- [2] Branas, C.C., Rundle, A., Pei, S., Yang, W., Carr, B.G., Sims, S., Zebrowski, A., Doorley, R., Schluger, N., Quinn, J.W., et al.: Flattening the curve before it flattens us: hospital critical care capacity limits and mortality from novel coronavirus (sars-cov2) cases in us counties. *medRxiv* (2020)
- [3] Chen, Y., Wang, A., Yi, B., Ding, K., Wang, H., Wang, J., Shi, H., Wang, S., Xu, G.: Epidemiological characteristics of infection in covid-19 close contacts in ningbo city. *Zhonghua liu Xing Bing xue za zhi= Zhonghua Liuxingbingxue Zazhi* **41**(5), 667–671 (2020)
- [4] Chu, D.K., Akl, E.A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H.J., El-harakeh, A., Bognanni, A., Lotfi, T., Loeb, M., et al.: Physical distancing, face masks, and eye protection to prevent person-to-person transmission of sars-cov-2 and covid-19: a systematic review and meta-analysis. *The Lancet* (2020)
- [5] Cooper, I., Mondal, A., Antonopoulos, C.G.: A sir model assumption for the spread of covid-19 in different communities. *Chaos, Solitons & Fractals* **139**, 110057 (2020)
- [6] Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., Dahl, G.E.: Neural message passing for quantum chemistry. *arXiv preprint arXiv:1704.01212* (2017)
- [7] Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. In: *Advances in neural information processing systems*. pp. 1024–1034 (2017)
- [8] Kapoor, A., Ben, X., Liu, L., Perozzi, B., Barnes, M., Blais, M., O'Banion, S.: Examining covid-19 forecasting using spatio-temporal graph neural networks. *arXiv preprint arXiv:2007.03113* (2020)
- [9] Klompas, M., Baker, M.A., Rhee, C.: Airborne transmission of sars-cov-2: theoretical considerations and available evidence. *Jama* (2020)
- [10] La Gatta, V., Moscato, V., Postiglione, M., Sperli, G.: An epidemiological neural network exploiting dynamic graph structured data applied to the covid-19 outbreak. *IEEE Transactions on Big Data* (2020)
- [11] Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S., Reich, N.G., Lessler, J.: The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine* **172**(9), 577–582 (2020)
- [12] Liu, Y., Gayle, A.A., Wilder-Smith, A., Rocklöv, J.: The reproductive number of covid-19 is higher compared to sars coronavirus. *Journal of travel medicine* (2020)
- [13] Shah, C., Dehmamy, N., Perra, N., Chinazzi, M., Barabási, A.L., Vespignani, A., Yu, R.: Finding patient zero: Learning contagion source with graph neural networks. *arXiv preprint arXiv:2006.11913* (2020)
- [14] Thunstrom, L., Newbold, S., Finnoff, D., Ashworth, M., Shogren, J.F.: The benefits and costs of flattening the curve for covid-19. Available at SSRN 3561934 (2020)
- [15] Wiersinga, W.J., Rhodes, A., Cheng, A.C., Peacock, S.J., Prescott, H.C.: Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (covid-19): a review. *Jama* **324**(8), 782–793 (2020)
- [16] You, C., Deng, Y., Hu, W., Sun, J., Lin, Q., Zhou, F., Pang, C.H., Zhang, Y., Chen, Z., Zhou, X.H.: Estimation of the time-varying reproduction number of covid-19 outbreak in china. *International Journal of Hygiene and Environmental Health* p. 113555 (2020)