Application of Data-Driven Measures for Impeding COVID-19 Spread at an Academic Institution

Minhao Zhang[a], Geng Zhang[a], Shengxi Zhang[a], Zian Liu[a], Vitaly Ford[a] and Victoria Turygina[b]

[a] *Arcadia University, 450 S Easton Rd, Glenside, PA 19038, U.S.A.*

[b] *Ural Federal University, Prospekt Lenina, 51, Yekaterinburg, Sverdlovskaya oblast', Russia, 620075*

Abstract

Many U.S. universities are embracing the hybrid teaching modality thanks to the start of the COVID-19 vaccinations and availability of online teaching tools. This work presents a continuation of our previous research, in which we analyzed and developed a methodology to inhibit COVID-19 spread on a university campus. We simulate the virus spread on campus, comparing *SIR* and *SEIR* models, and examine how different course policies can affect the number of infected students. We demonstrate that we can achieve a safer environment on campus by moving a certain number of courses with the highest centrality values. Additionally, we analyze how the student flow rate can help reduce the R_0 value representing the metric of how many other people an infected individual could infect. This work also presents the simulation analysis of the opened public places on campus and the application of the sensitivity analysis to develop the most efficient approach determining the exact courses that need to be moved online. We conclude with the recommendations and analysis results.

Keywords

COVID-19, simulation, university campus, centrality parameter, hybrid teaching

1. Introduction

Many U.S. universities are embracing the hybrid teaching modality such as HyFlex [1] thanks to the start of the COVID-19 vaccinations and availability of online teaching tools. Previously, in the paper titled "Analysis and Methodology of Inhibiting COVID-19 Spread on a University Campus" [2], we analyzed and developed a methodology to address the issue of the COVID-19 spread on campus. In this paper, we first compare the *SIR* [3] and *SEIR* [4] models and then simulate the virus spread using the *SEIR* model. One of our major goals is to identify specific course policies that would create a safer environment and facilitate decision-making processes at the leadership level of the Arcadia [5] and other universities.

In [2], we developed the methodology based on the Degree Centrality, Closeness Centrality, and Betweenness Centrality parameters that were aggregated in the rectified centrality value representing the number of connections, distance, and connectivity strength in the social network (courses-students) - all in one. In this work, we continue building upon that knowledge

IVUS2021: Information Society and University Studies 2021, April 23, 2021, Kaunas, Lithuania

mzhang_01@arcadia.edu (M. Zhang); gzhang_02@arcadia.edu (G. Zhang); szhang_01@arcadia.edu (S. Zhang); zliu@arcadia.edu (Z. Liu); fordv@arcadia.edu (V. Ford); v.f.volodina@urfu.ru (V. Turygina)

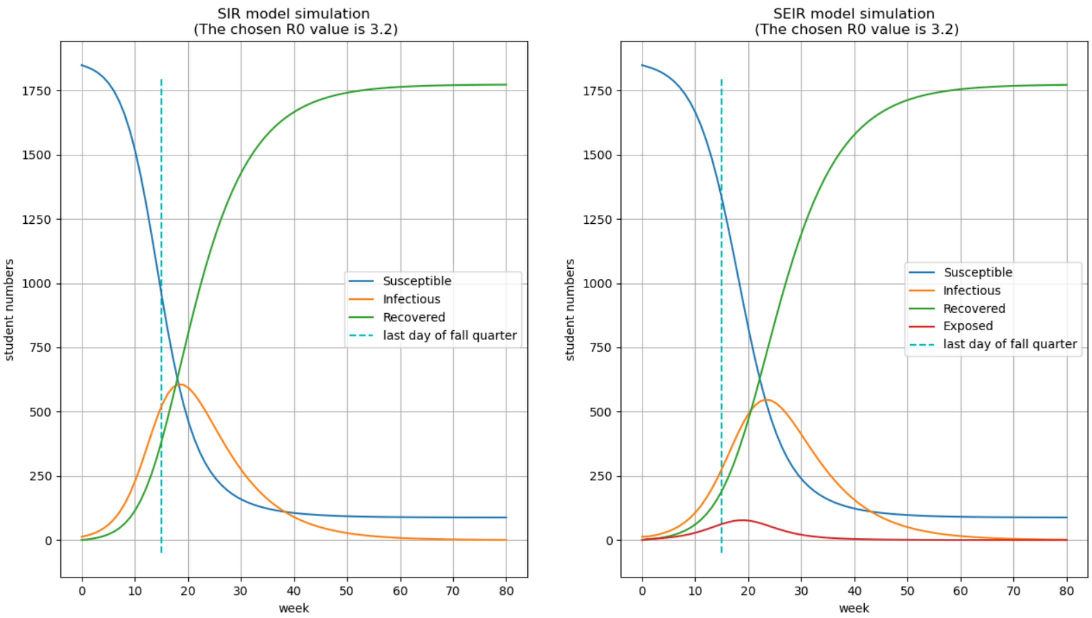

Figure 1: *SIR* (left) and *SEIR* (right) model comparison.

and develop the procedures that help us make policies that provide a valuable input on which courses should be moved online, what to do with the public places like dining halls, and what the most important parameters are (through the sensitivity analysis in section 3) that affect the virus spread on campus.

2. Simulation

2.1. Model definition

We compared two different simulation methods and selected the *SEIR* model [4] according to the characteristics of COVID-19. The advantage of the *SEIR* model is that the incubation factor is taken into account in the simulation process. In figure 1, we present the comparison of the simulation results of *SEIR* and *SIR* [3] models with R_0 [6] set to 3.2. By using *SEIR*, we can obtain more useful information by the end of the simulated semester (specifically, the number of students exposed) compared to the *SIR* model.

2.2. Control group simulation

Our control group represents the number of students being infected, assuming that the semester is fully online. The data we use come from the CDC [7], and we apply the local region's rate of infection in the simulation of the control group (figure 2).

The control group simulation results demonstrate that after 15 weeks in the fully online fall semester, the total number of infected students could be 374. In the following section, we are

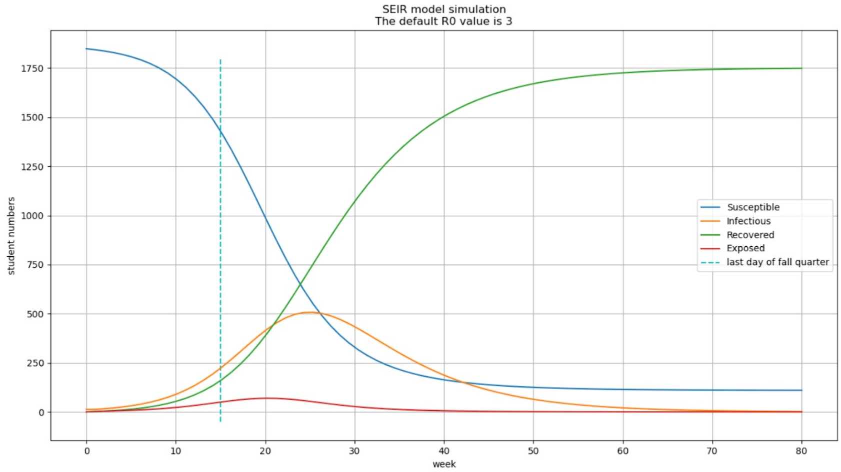

Figure 2: The control group simulation.

Table 1
The number of infected people based on different policies, under four scenarios.

<table>
<tr><th></th><th>Optimistic Scenario</th><th>Normal Scenario</th><th>Pessimistic Scenario</th><th>Online Semester</th></tr>
<tr><td>Without removing any courses</td><td>120</td><td>217</td><td>456</td><td rowspan="3">374</td></tr>
<tr><td>Removing five courses with the highest rectified centrality</td><td>101</td><td>176</td><td>364</td></tr>
<tr><td>Removing ten courses with the highest rectified centrality</td><td>91</td><td>156</td><td>316</td></tr>
</table>

going to compare the control group's results with the face-to-face semester under our proposed measures to move only certain courses online.

2.3. Simulation scenarios and outcomes

In this section, we simulate the virus spread in the pessimistic, normal, and optimistic scenarios (public places like the dining hall are assumed to be closed). To match our simulation as close as possible with the complexities of the real situation on campus, we focus on the pessimistic configuration as it describes the worst-case scenario. The following diagrams are built based on the pessimistic scenario with the initial R_0 set to 3.2. The results of the normal and optimistic scenarios are also shown in table 1.

We performed several simulation experiments and would like to share the results under 3 different policies presented in figures 3, 4, and 5.

1. Policy #1: none of the courses are moved online and all students meet face-to-face.
2. Policy #2: five courses with the highest centrality are moved online.
3. Policy #3: ten courses with the highest centrality are moved online.

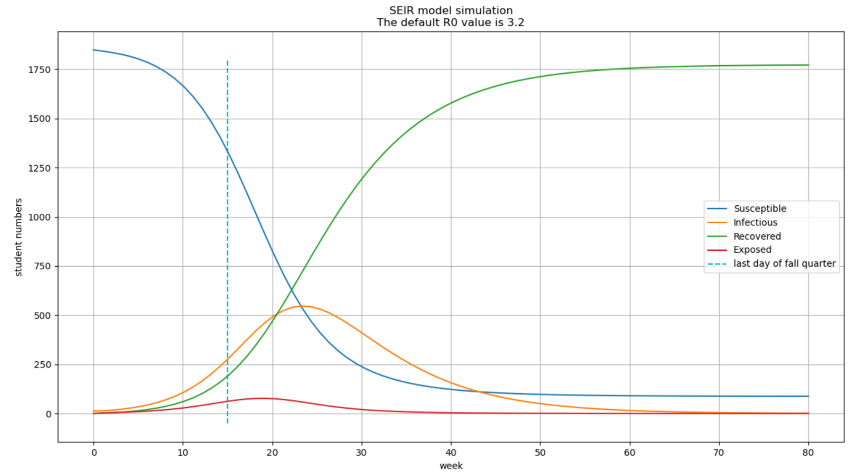

Figure 3: Policy #1: none of the courses are moved online and all students meet face-to-face.

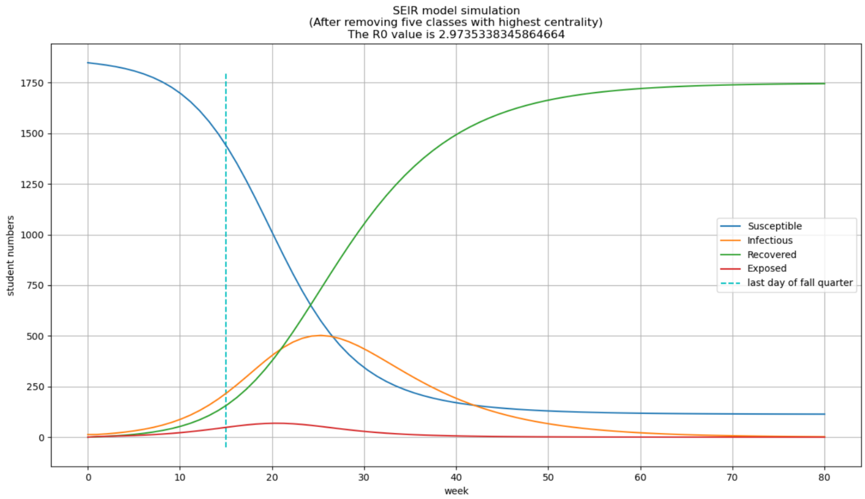

Figure 4: Policy #2: five courses with the highest centrality are moved online.

It can be seen that even under the most pessimistic circumstances, the number of infected people can be reduced by moving five courses with the highest sensitivity online compared with the control group. After moving 10 high sensitivity courses online, the number of infected students decreased by 15.5%, from 374 to 316.

Figure 1 demonstrates the different outcomes of moving a certain number of courses online in a pessimistic scenario relative to the control group (fully online semester).

We also simulated a special situation by closing certain classrooms every day with the highest student flow rate (figure 6). The expected R_0 value decreased by 38.63%, from 3.2 (default) to 1.96, and the total number of infected students reduced to 115. We propose to move such classes to other available classrooms, resulting in the lowest infection rate with the number of

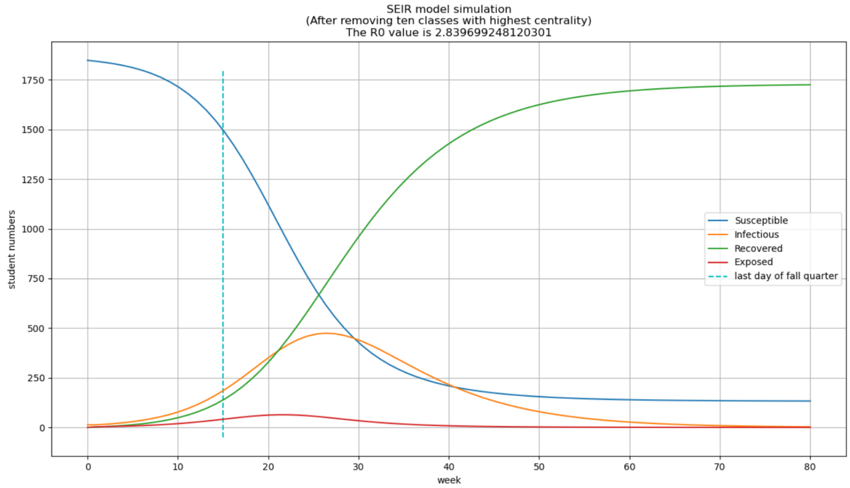

Figure 5: Policy #3: ten courses with the highest centrality are moved online.

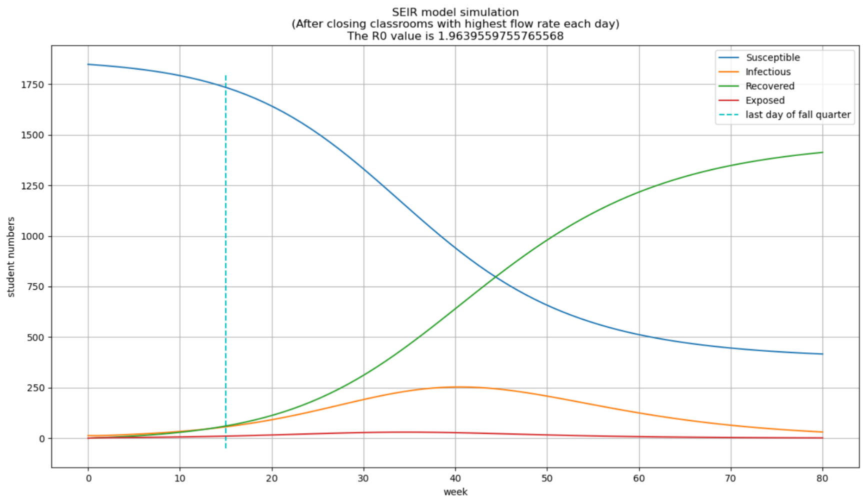

Figure 6: Closing classrooms with the highest flow rate.

infected students to be in the range from 115 to 374, depending on how many of those classes can be relocated.

2.4. Simulation analysis of opened public places

One of the issues that are not addressed in the social network graph is the issue of students' dining options. Considering that there are only a few places at Arcadia that could provide food for students, their closure could become an inconvenience for everyone. However, opening those buildings may bring students together again and speed up the spread of the virus, which does not conform to the notion of reducing the number of face-to-face courses to protect the

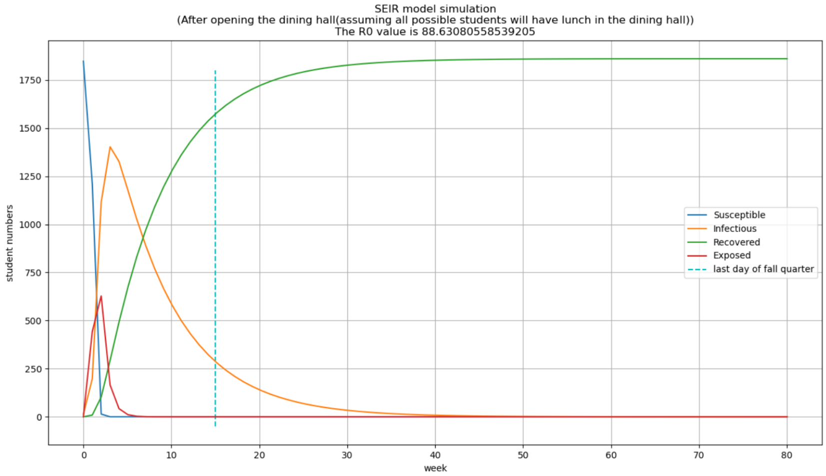

Figure 7: Simulation of opening the dining halls.

health of students. Consequently, we also simulated the spread of the virus in the school if we did not close public places like the dining hall (assuming all possible students have lunch). For example, we studied each student's schedule to determine if they could have lunch in the canteen at noon (there are around 650 such students), and then we added it as a "lunch course" in the social network to simulate the place where students may eat together. The expected value of R_0 increased by 2,769% (from 3.2 (default) to 88.63), meaning that almost all the students would be infected. Figure 7 presents the simulation of opening dinning halls without moving any courses online.

To solve the dining problem, meals can be delivered to each student's room or takeout meals without contact could be arranged. Also, while processing the data, we were able to determine the number of students that would possibly dine on campus at a given time because, for instance, they had a break between the morning and afternoon classes. Based on this information, Arcadia could estimate the outdoor seating needs at a given time and provide a sufficient number of tables and chairs on campus, taking into consideration that not all students dine in public places on campus.

Additionally, other public places like gyms can set up isolation zones for each part of the equipment and limit the number of people using the gym at a time.

3. Sensitivity Analysis

The purpose of this project is to reduce the number of infected students in a face-to-face semester as much as possible and make sure that the risk of being infected on campus is lower than the one in the local region, making it a safer educational environment. At the same time, we hope to change the curriculum (shift a certain number of courses online) as little as possible. Therefore, we focus on developing the most efficient approach to determine the courses that need to be moved online. To better understand our results from section 2, we performed a

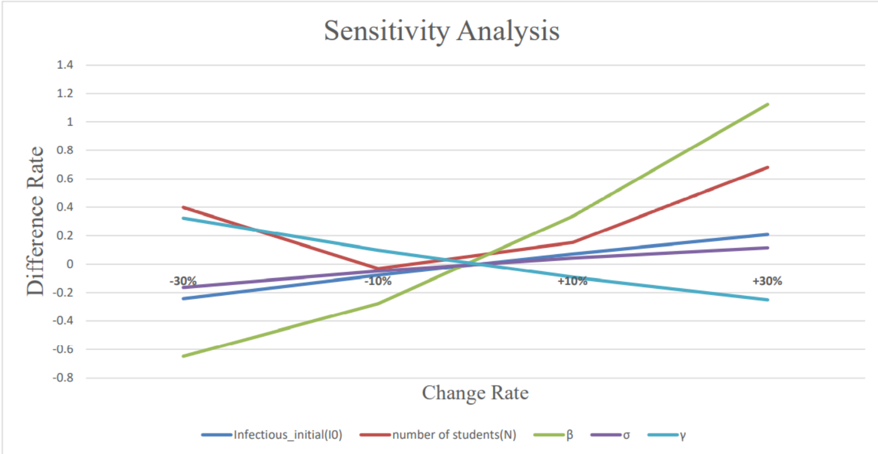

Figure 8: Sensitivity analysis of simulation parameters.

sensitivity analysis to determine the changes and correlations in our model depending on the input variables simulating the virus transmission.

3.1. Sensitivity results for parameters in R_0 simulation

We compared the effects of each parameter on the results with parameter changes of -30%, -10%, 10%, and 30%, and presented them on figure 8, in which:

- β is the speed rate at which students change from being "susceptible" to being "infectious" (i.e., the infection rate). Its value can be roughly calculated by the product of R_0 and the recovery rate.
- σ is the speed rate at which students change from being "exposed" to being "infectious" (i.e., the incubation rate). Its function is defined as $1/latent\ period$.
- γ is the speed rate at which students change from being "infectious" to being "recovered" (i.e., the recovery rate). Its function is defined as $1/recovery\ period$.

Table 2 shows the exact changes of each parameter and the respective impact on the simulation outcomes.

Based on the figure 8 and table 2, the β infection rate has the greatest impact on our virus transmission model, and it confirms the fact that it is crucial to reduce the infection rate. The β infection rate can only be reduced by wearing masks and following the safety guidelines. In our model, β is already at a very low level, so it is difficult to decrease it even more. But its sensitivity is still of great significance to our model because it explains the importance of controlling the infection rate.

Table 2 also depicts a situation when the β parameter increases only by 30% but the number of infected students grows by 119%, meaning that their relationship is not linear. Thus, it is even more important to keep the β parameter as low as possible.

Table 2
Sensitivity analysis results.

	Parameters	Initial value	Change rate	Value after the change	Number of infected students after the change	Number of infected students before the change	Difference value	Difference rate
0	Infectious initial (IO)	12.80	-30%	8.96	346	456	-110	-0.24
1	Infectious initial (IO)	12.80	-10%	11.52	421	456	-35	-0.08
2	Infectious initial (IO)	12.80	+10%	14.08	489	456	33	0.07
3	Infectious initial (IO)	12.80	+30%	16.64	551	456	95	0.21
4	Number of students (N)	1861	-30%	1302	638	456	182	0.40
5	Number of students (N)	1861	-10%	1674	442	456	-14	-0.03
6	Number of students (N)	1861	+10%	2047	526	456	70	0.15
7	Number of students (N)	1861	+30%	2419	766	456	310	0.68
8	$beta$ (β)	0.46	-30%	0.32	160	456	-296	-0.65
9	$beta$ (β)	0.46	-10%	0.41	330	456	-126	-0.28
10	$beta$ (β)	0.46	+10%	0.50	610	456	154	0.34
11	$beta$ (β)	0.46	+30%	0.60	968	456	512	1.12
12	$sigma$ (σ)	1.35	-30%	0.94	382	456	-74	-0.16
13	$sigma$ (σ)	1.35	-10%	1.21	434	456	-22	-0.05
14	$sigma$ (σ)	1.35	+10%	1.48	476	456	20	0.04
15	$sigma$ (σ)	1.35	+30%	1.75	509	456	53	0.12
16	$gamma$ (γ)	0.14	-30%	0.1	604	456	148	0.32
17	$gamma$ (γ)	0.14	-10%	0.13	501	456	45	0.10
18	$gamma$ (γ)	0.14	+10%	0.16	414	456	-42	-0.09
19	$gamma$ (γ)	0.14	+30%	0.19	342	456	-114	-0.25

The σ parameter can be simply understood as the number of days that a student can continuously infect others before being tested (and later quarantined) or beginning to show symptoms. Therefore, it is particularly important to detect asymptomatic or latent patients as early as possible, directly reducing the average number of infections per patient. Therefore, it is necessary for the schools to conduct regular tests when possible. With enough testing equipment, smaller campuses like Arcadia could efficiently conduct virus tests, reflecting an advantage of school opening because students can be constantly organized and tested. Regular testing is comparatively more difficult to implement in the wider community, often leading to family and community clustering infections.

4. Discussion and Conclusion

Based on the simulation results, sensitivity analysis, and rectified centrality formula, we can answer the following questions that we defined in the previous work [2].

Question #1: "would schools be able to stay open during COVID-19 by moving high sensitivity classes and students online?" The simulation results demonstrate that it is possible, as long as the school follows the policies we proposed and all assumptions we made are true. Even under a pessimistic R_0 set to 3.2, the number of infected students at Arcadia could be 316 after moving 10 classes with the highest centrality online, which is still lower than 374 (the predicted value of infected students made if the semester is fully online). But this result is based on the assumption that students would follow safety practices, reduce other forms of contact as much as possible in addition to basic academic activities, and all students, faculty, and staff could be tested before returning to and while being on campus. We found that even if all students take classes online, they will still be exposed to an environment where R_0 equals 3 and 374 students could get infected at the end of the semester. Therefore, returning to school based on the proposed policies can achieve a lower infection rate and provide better protection to students compared with the fully online semester. However, it should be noted that the school's dining places and gyms should be closed, otherwise almost all students would be infected.

Question #2: "How many students would be infected compared to the model without converting classes to an online format?" If we do not convert any courses to the online format, the number of infected students in the optimistic scenario would be 120, in the normal scenario it would be 217, and in the worst-case pessimistic scenario it would be 456.

4.1. Supplemental statistical data evaluation

Some of the notable characteristics of the data that we worked with are presented below.

- The average and maximum numbers of students that one student may be exposed to while taking classes on campus throughout a week are: max = 175; $mean$ = 66.
- Courses with the respective values of different degrees of centrality were shown in a table in [2]. We used these data to decide which courses to move online, reducing the connections among all students based on the values of rectified centrality in section 2.
- The average number of students per class is 12 which is lower than usual.
- Based on the flow rate analysis, we identified several high flow rate classrooms for every day of the week.
- There are a total of 61,186 pairs of students who are directly connected.

4.2. Recommendations

Our final recommendations based on the simulation results and analysis are presented below.

- It is possible to reduce the R_0 value by moving high sensitivity courses online.

- The dining places should be closed. We proposed alternative options in section 2.4.
- Temperature-measuring devices [8] that can remotely determine body temperature are needed in every room. As schools plan to reopen during the pandemic, non-contact temperature assessment devices will play an important role in every classroom. As part of the first check to identify people who have an elevated body temperature, these devices can help students, faculty, and staff conduct preliminary self-examinations every day, ensuring that patients with symptoms can be isolated in time.
- Instructors should avoid running multiple classes one after another in the same classroom to prevent a potential virus accumulation over time. In section 2.3, the data showed that R_0 could decrease from 3.2 to 1.96 by determining the classrooms with a high flow rate and moving some of those courses into the classrooms with a low flow rate.
- Student gatherings (clubs, events) should only be allowed in masks outside of buildings.
- All faculty and staff meetings should be moved online.
- Weekly virus testing should be conducted. If it is not feasible to conduct it for everyone, a random sampling of students from different departments could be a viable option.

Acknowledgements

We would like to thank the Computer Science and Math Department, Arcadia University, for providing support and help during this research.

References

[1] B. Beatty, Designing the HyFlex World-Hybrid, Flexible Classes for all Students, in: Association for Educational Communication and Technology International Conference, Dallas, TX, 2006.

[2] M. Zhang, G. Zhang, S. Zhang, Z. Liu, V. Ford, V. Turygina, Analysis and Methodology of Inhibiting COVID-19 Spread on a University Campus, in: accepted, 26th International Conference Information Society and University Studies - IVUS, 2021.

[3] D. Smith, L. Moore, et al., The SIR model for spread of disease: the differential equation model, Convergence (2004).

[4] A. Grant, Dynamics of COVID-19 epidemics: SEIR models underestimate peak infection rates and overestimate epidemic duration, MedRxiv (2020).

[5] Arcadia University, Arcadia university, 2021. URL: https://arcadia.edu.

[6] B. Ridenhour, J. M. Kowalik, D. K. Shay, Unraveling R_0: Considerations for public health applications, American journal of public health 108 (2018) S445–S454.

[7] L. Matrajt, T. Leung, Evaluating the effectiveness of social distancing interventions against COVID-19, MedRxiv (2020).

[8] U.S. FDA, Non-contact Temperature Assessment Devices, 2021. URL: https://www.fda.gov/medical-devices/.