

Network-Based Analysis of Early Pandemic Mitigation Strategies, Solutions, and Future Directions

Pegah Hozhabrierdi, Raymond Zhu*, Maduakolam Onyewu*, and
Sucheta Soundarajan

Syracuse University, Syracuse NY 13244, USA
{phozhabr,rzhu09,monyewu,susounda}@syr.edu

Abstract. Despite the large amount of literature on mitigation strategies for pandemic spread, in practice, we are still limited by naïve strategies, such as lockdowns, that are not effective in controlling the spread of the disease in long term. One major reason behind adopting basic strategies in real-world settings is that in the early stages of a pandemic, we lack knowledge of the behavior of a disease, and so cannot tailor a more sophisticated response. In this study, we design different mitigation strategies for early stages of a pandemic and perform a comprehensive analysis among them. We then propose a novel community-based isolation method and show its efficacy in reducing the speed of the spread by a large margin as compared to current methods. We also show that the test-trace-isolation strategy can outperform lockdown and random test-trace in reducing the economic impact and spread of the disease if combined with k -hop neighborhood ranking. The novelty of our work lies in using network structural properties (local and global) to design a strategy for the early stages of a pandemic. Our results encourage further investigation into community-based mitigation strategies and shed more light on the differences between current methods of choice in practical setting.

Keywords: Pandemic · Network-based Mitigation strategy · Economic Impact.

1 Introduction

In response to the COVID-19 pandemic, governments across the world have attempted a variety of strategies to mitigate the spread of disease, including lockdowns, contact tracing, and others. However, there has been little analysis on the relative merits of such strategies; and because these strategies have negative effects on the economy and the morale of the people, it is critically important to understand their efficacy. Although this work is inspired by the COVID-19 pandemic, our analysis can apply to any contagious disease. Because the response

* Equal contribution

to a pandemic depends on whether it is in the early stages (no vaccination available) or later stages (vaccination available), we focus on early pandemic mitigation strategies in correspondence with the COVID-19 situation.

In this work, we analyze variants of the pandemic mitigation strategies practiced in the real world- e.g. lockdown and test-trace-isolate- from a network perspective. Inspired by the new psychology findings on the correlation between community membership and pandemic response [1–3], we also offer a community-based mitigation strategy and demonstrate its efficacy in comparison. To the best of our knowledge, we are the first to offer a network-based comparison of practical mitigation strategies in the early stages of a pandemic. To evaluate each strategy, unlike the majority of related work in this field, we consider both the speed of spread and economic impact as cost factors. For example, a mitigation strategy such as a total lockdown might have the best performance in terms of controlling the spread of disease if prolonged for long enough time until the discovery of the vaccine. However, the devastating economic impacts of such a decision makes this strategy inefficient in the real world. On the other hand, the “Do Nothing” strategy, which relies on herd immunity (see Section 3), results in less economic impact (at least in the early stages), but does nothing about the spread of the disease. An ideal mitigation strategy should offer a trade-off between these two losses. We allocate a budget to each strategy to count for the economic impact and report both the spread and budget spent for each strategy simulation.

To ensure that our results are generalizable to other contagious diseases, we enforce only general assumptions about the nature of the disease and cost of battling the spread (see Section 2). We use the SIRD epidemic model for our simulations and only consider the budget spent on isolation strategies. We validate each strategy on a set of 10 real-world social networks (see Section 4.2). To have a close approximation of human-human contact behavior, these networks are chosen based on the method of data collection and the meaning of connections between two individuals. We also consider a set of online social networks that are frequently used in the disease spread literature [4–6] and, in some cases, have been shown to give a close-enough approximation of real-world social networks [5]. Our results show the superiority of the test-trace-isolation strategy if combined with k-hop neighborhood ranking (specifically for $k = 1$). We also confirm the theoretical results from psychology studies on the impact of community membership in reducing the spread of the disease and show the further direction in adopting such strategies.

2 Problem Statement

We model the population as a simple undirected graph $G(V, E)$, where individuals are represented by nodes (V) and connections between them (E) represent physical contact. We use undirected edges due to the nature of physical contact, for which a directed relationship does not bear any meaning. We also consider unweighted and un-attributed graphs, as attribute information is not easy to gather in a real-world setting and in the practical strategies discussed below. However,

our simulations can easily be extended to attributed or weighted graphs. For example, the strength of an edge can be considered as the frequency or length of the contact, where a higher value increases the probability of infection spread. In the following discussion, we discuss different models of disease spread and the reasons behind our choice of the SIRD model. We will also discuss our method of extending the model to count for the economic impact.

2.1 Viral Spread Modeling

Previous work on mathematical modeling of viral spread can be grouped into two categories of (1) general spread models and (2) virus-specific spread models. The former includes famous models such as SIR, SIRD, SIS, SIER, and SIRS [7–9]. The virus-specific models have been proposed for viruses observed in real world and consider the specific properties of a certain virus into the modeling of the spread [10–12]. The focus of our study is on the effectiveness of different mitigation strategies for an unknown pandemic scenario (i.e., a pandemic whose specific behavior and potential remedies are unknown). It is known that battling a new pandemic heavily relies on adopting a proper mitigation strategy in its early stages [13]. In these early stages, our knowledge of the nature of the virus is very limited, and the virus-specific strategies require prior knowledge gained from time-consuming clinical trials. Thus, the general models with little to no conditions on virus-specific behavior are more practically applicable in the early stages of a pandemic. For our study, we choose the SIRD model due to its minimal assumptions on the nature of a fatal spread, which is explained below.

2.2 SIRD Epidemic Model

Given a *closed* community in which the population is fixed (no birth, no migration, and no death due to causes irrelevant to the disease under study), the SIRD model assumes four possible states for each individual in the community at each timestamp: Susceptible (never contaminated by the virus), Infectious (contaminated and can spread the virus), Recovered (recovered from contamination and can no longer spread the virus), and Dead (due to infection). The possible transition between states and their respective probabilities are depicted in figure 1. The three parameters of this model are α , β , and γ that indicate infection, recovery, and mortality rate respectively.

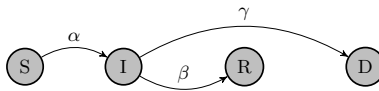


Fig. 1: SIRD state transitions. Parameters α , β , and γ indicate infection, recovery, and mortality rate respectively.

2.3 Budget Allocation

The exact modeling of a pandemic’s economic impact is a complicated problem and requires a comprehensive study on its own. [14–16] However, we still can introduce a simplified measure of cost for comparison between different mitigation strategies. As we try to minimize the number of isolated individuals while reducing the rate of spread, we have to compensate for the portion of the population that is under quarantine (either compulsorily or voluntarily) to make isolation practical and possible without threatening the well-being of families and individuals. We treat this compensation as a required budget for each isolation strategy. Ultimately, an ideal isolation strategy should use a small compensation budget while minimizing the peak number of the infectious population over time. A smaller amount of budget spent also indicates isolated individuals, implying less possibility of economic impact due to work-force perturbation.

3 Mitigation Strategies

Two of the most important problems in the early stages of a pandemic are (1) the capacity of healthcare centers and (2) economic consequences [14–17]. An optimal mitigation strategy seeks to reduce the occupancy of hospitals (lower the number of infected) while maintaining the productivity of the society to eliminate economic impacts. However, these two objectives often bear conflicting interests. So far, the strategies for lowering the number of infected individuals practiced in real-world setting have negatively affected the economic well-being of the society. A current example is the **lockdown strategy** adopted by many countries (such as the USA, Spain, and Italy) in 2020 to mitigate the COVID-19 spread¹. Interestingly, lockdown does not offer an optimal solution to either of the objectives above. First, lockdown leads to a *second wave* of spread and has to be implemented in several phases to be effective in lowering the burden on the healthcare system [18]. Second, it is shown (both in theory and practice) that lockdown strategy causes severe damages to the economy [17].

To trade-off between the need for isolation and economic prosperity, [17] suggests employing a **Test-Trace-Isolate strategy (TTI)**. This method puts the focus on the neighborhood of the individuals with positive test results (infected). According to [19], the countries who employed the TTI strategy against COVID-19 were able to combat the spread more successfully than those who followed *herd immunity*² or full containment (lockdown) strategy. This, however, mainly considers the medical benefits of the mitigation. The cost-effectiveness of TTI (economical aspect) heavily depends on its implementation [20]. For example, how do we choose whose neighborhood to trace? Is it the people who show symptoms or those who have tested positive? Furthermore, how many people in the candidate’s neighborhood should we isolate and how big should the size of this

¹ https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdowns

² Herd Immunity is an epidemiological concept and is defined as “the percentage of people with protective immunity needed in a population to stop the propagation of an infectious agent” [19]. This strategy, although seemingly giving an optimal solution to economic impact of the spread, results in a devastating death toll in the population.

neighborhood be? Tracing and isolating steps of TTI are costly and, if implemented in a naïve way, it can be less efficient than lockdown strategy. In this study, we examine three different strategies for TTI. These methods all use local neighborhood information. We consider random and centrality-based TTI with tracing radius up to k -hops away from the infected node (for $k \in \{1, 2\}$). The details of each method are discussed in Section 4. Note that due to small-world property of social networks, for k values higher than two, we capture almost all of whole network, which is counter-intuitive for TTI strategy.

In a pandemic, human behavior plays as important of a role as properties of the virus (if not a more important role) [1]. Such behavior is directly connected to psychological traits of individual’s personality [2]. An interesting relevant observation is that shared community membership increases the speed of the spread [3]. These psychological observations are rarely considered in the study of pandemic spread. We argue that by considering these findings, we can improve isolation (and possibly vaccination) strategies. As such, we propose a **Community-based Isolation strategy (CI)** and show its effectiveness in comparison to lockdown and TTI strategies. The results of CI are presented to show that using the community membership of individuals as an isolation strategy indeed reduces the speed of spread. At first glance, this method might not be as practical as lockdown or TTI, owing to the fact that community membership is only partially known and tracing the memberships can be even more costly than TTI (as shown in Section 4). However, the results of our experiments show that community-based isolation surpasses all other methods in reducing the spread of the disease without the disadvantage of a second wave. We argue that further approximation of community information and isolation of only bridge nodes (those who contribute the most to the spread from one community to the next) can improve the cost associated with CI strategy.

4 Experiments

4.1 Assumptions

As mentioned before, for lockdown and TTI, we do not enforce any disease-specific information on the model. Additionally, we do not assume the presence of network structural data that are hard or impossible to obtain in real-world setting. For example, we do not assume that we have global information for nodes or edges (e.g. shortest path-based centralities, diameter of the network, or spectral properties). We only have the information on the neighborhood of each individual (as obtained through individual surveys in real world). It is noteworthy that we consider the neighborhood information to be incomplete and any new information added to our data will improve the result of the simulations.

For CI, we assume the community membership of individuals is known. In the real-world, the communities can be considered at different levels; from a club membership level up to county and state levels. Considering that not all of our datasets have ground-truth communities, we obtain membership through Louvain partitioning of the graph that maximizes the modularity.

Data	Type	Edge Meaning	$ V $	$ E $	Avg. Deg.
Infectious (INF)	Human Interaction	Contact	410	2,765	13.49
Hyptertext2009 (HX9)	Human Interaction	Contact	113	2,196	38.87
Haggle (HAG)	Human Interaction	Contact	315	2,899	18.41
Adolescent Health (AH)	Human Social	Friendship	2,539	10,455	8.23
Residence Hall (RH)	Human Social	Friendship	217	1,839	16.95
Physicians (PHY)	Human Social	Trust	241	923	7.66
Jazz Musicians (JAZ)	Human Social	Collaboration	198	2,741	27.69
Pretty Good Privacy (PGP)	Online Contact	Interaction	10,680	24,316	4.55
Facebook NIPS (FBN)	Online Social	Friendship	2,888	2,981	2.06
Hamster Full (HAM)	Online Social	Friendship	2,426	16,630	13.71

Table 1: Contact datasets for spread simulation

4.2 Data

Recent studies show the importance of using real-world human-human interaction data to account for the influence of human behavior in the simulation of a spread [1]. We chose seven real-world datasets that have been collected based on physical human interaction/connections in real world. These datasets, although the best resource for real-world interactions, are generally small due to the cost of data collection. As such, many studies tend to use online social networks as approximate behavior of the users in physical world. [5] showed that online behavior approximation for some online networks such as Facebook is close to physical behavior. To both confirm their results for other online social networks and consider the result of our simulation on larger networks, we also consider an additional three larger datasets from online social networks. All of our 10 datasets are chosen based on the nature of their contact (edge meaning). These datasets and their general statistics are shown in Table 1. All datasets are publicly available in the Konect repository [21].

4.3 Strategies

In this section, we briefly go over the implementation of each mitigation strategy mentioned in Section 3. The hyperparameters are the same among all strategies (such as quarantine compensation, duration of quarantine, duration of disease, and parameters α , β , γ). In all simulations, we start with only one infectious node chosen at random. For each model, we repeat the simulation for 100 different starting node and report the average among the 100 trials. These hyperparameters are presented in table 2.

- **Do Nothing (DN)**: Although not exactly a mitigation strategy, DN can be used as the baseline to compare the performance of other methods against it. It is a simple SIRD model that reaches the peak of infection quickly and fades away quickly as well (due to herd immunity).
- **Lockdown**: The duration of the lockdown is fixed at 14 days, and it starts after the detection of the first infectious sample. We randomly choose 90% of the population for lockdown and compensate all of them according to *Daily*

Hyperparameter	Value
Sickness Duration	7 days
Mortality Rate	0.04
Immunity Time	15 days
Probability of Infection	0.2
Daily Reward (Compulsory Isolation)	\$100
Daily Reward (Volunteer Isolation)	\$50
Volunteer Isolation Probability	0.5

Table 2: Hyperparameters chosen for all mitigation strategies when applicable.

Reward (Compulsory Isolation) in Table 2. After the lockdown is lifted, the disease spreads according to the SIRD model and we expect the same peak as in DN but shifted over time.

– **TTI:**

- *K-hop Ranking for $k \in \{1, 2\}$:* Prior to the simulation, each node is ranked according to the size of its k -hop neighborhood. For example, if the rank of a node is 16, it means this node, if infectious, can potentially contaminate 16 other individuals. In the tracing stage of TTI, we choose to forcibly isolate neighborhood of an infected node who have a ranking above 90 percentile of all rankings in the graph. We also choose the neighborhood with ranking above 80 percentile (and less than 90 percentile) as candidates to voluntarily quarantine themselves. Both of these groups (forced quarantine and volunteer quarantine) are compensated but with different amounts (see *Daily Reward* for compulsory and volunteer isolation in Table 2). We assume the candidates for volunteer isolation accept the offer 50% of the times. Note that both of these thresholds can be chosen upon trial and error and does not require global information on the graph.
- *Random Ranking:* In the tracing stage of TTI, we randomly choose candidates from the k -hop neighborhood ($k \in \{1, 2\}$) of an infected node to isolate. The isolated nodes are compensated as in lockdown.

- **CI:** We obtain community memberships through Louvain partitioning [22]. At each timestamp, we isolate an entire community if the portion of infected nodes within the community is greater than a threshold, i.e. $\frac{I_c}{|V_c|} > T$ where $|V_c|$ is the population within community c and T is a hyperparameter. We report the results for $T \in \{0.1, 0.2, \dots, 0.9\}$. The isolated members are compensated as in lockdown.

4.4 Results

The results of the DN, Lockdown, and TTI strategies are shown in Figure 2. In this figure, the y axis depicts the peak of the infectious population normalized by the overall population. In all datasets, the best performance is achieved through the k -hop neighborhood strategy. In 7/10 datasets, the 2-hop strategy achieves a better performance but, in most cases, is very close to that of 1-hop. However, looking at the required budget in Figure 3, it is evident that the choice of 2-hop

neighborhood comes with a greater cost, especially for HX9 dataset. The special case with HX9 dataset shows the limitation of considering neighborhoods over 1-hop of the infectious nodes: if the average shortest paths in the network is too small, the 2-hop neighborhood captures the entirety of the network and, in practice, gives a less optimal result than lockdown or even DN strategy. Considering this trade-off between cost and peak, we can conclude that 1-hop TTI strategy is the best practical strategy among the rest in real-world scenarios.

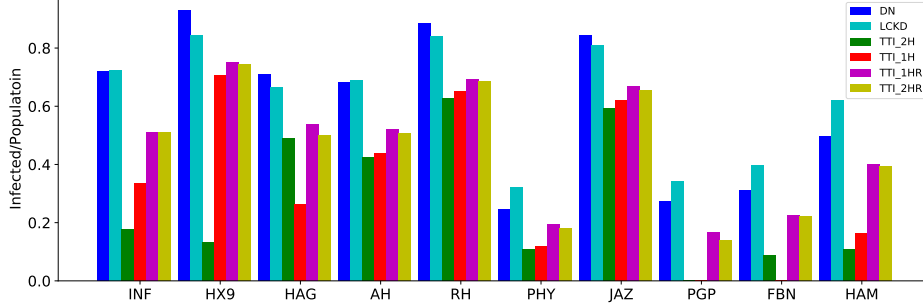


Fig. 2: Proportion of infected individuals over 100 trials of simulation for each mitigation strategy among our datasets. DN, LCKD, and TTI abbreviate *Do Nothing*, *Lockdown*, and *Test-Trace-Isolate* strategies. TTI suffixes: 1H and 2H represent k-hop ranking, 1HR and 2HR represent random ranking within k-hop neighborhood.

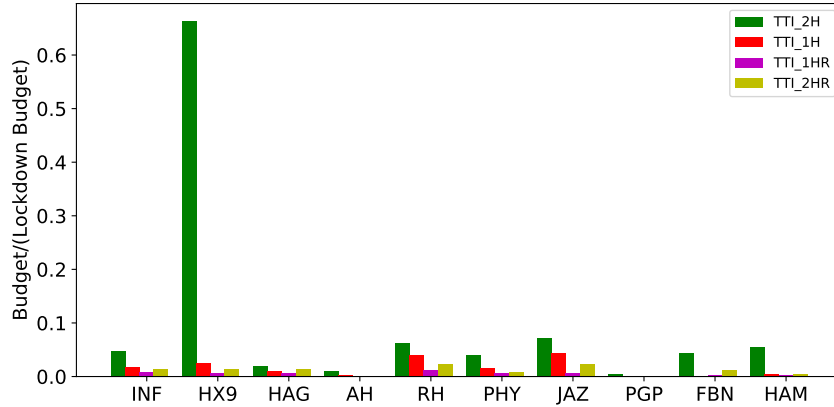


Fig. 3: The budget spent on isolation strategies. The budget is normalized by lockdown budget as the baseline.

Table 3 summarizes the results for CI-based isolation for different threshold values. CI, surprisingly, does a much better job at reducing the infectious peak

than any of the other methods (compare the proportion of infected in Table 3 with those in Figure 2). This confirms the suggestions from psychology literature that mitigation strategies based on community membership can result in a better control over the speed of the spread. However, the thresholding is very important for CI. As seen from the table, for higher thresholds, CI generally comes with a much higher budget than TTI, and unlike the previous methods, surpasses the lockdown budget in many instances. However, keeping the threshold below 0.4 offers a considerable reduction in speed with a reasonable budget. On the other hand, we argue that our CI strategy of choice is a naïve way of implementing a community-based method and it still performs better than other strategies we discussed in reducing the speed. The superiority of CI in reducing spread shows that designing an optimal community-based strategy for further alleviation of the economic impact is a promising research direction. Moreover, community information can be local and noisy (through individual self-reported or publicly known memberships such as geographical proximity in a region). Our effort is to encourage more research on community-based mitigation strategies rather than brute-force methods such as lockdown or naïve TTI. Although k-hop and community-based methods seem to require extra effort for tracing the impact, they are still practical in real-world. Our results show that with local approximation of network’s structure, we still can obtain solutions that reduce both the physical and economic impact of the pandemic in a global scale.

Data	0.1	0.3	0.5	0.7	0.9
Infectious (INF)	0.02 (0.28)	0.11 (0.74)	0.23 (0.91)	0.28 (1.23)	0.72 (1.63)
Hyptertext2009 (HX9)	0.19 (1.23)	0.45 (0.07)	0.69 (0.25)	0.72 (0.05)	0.88 (0.78)
Haggle (HAG)	0.06 (0.58)	0.21 (0.95)	0.31 (1.12)	0.35 (0.9)	0.63 (0.16)
Adolescent Health (AH)	0.03 (1.42)	0.17 (0.44)	0.21 (1.15)	0.39 (1.63)	0.39 (0.002)
Residence Hall (RH)	0.12 (0.31)	0.23 (0.42)	0.44 (0.15)	0.57 (0.27)	0.76 (0.69)
Physicians (PHY)	0.07 (0.15)	0.11 (8.00)	0.17 (12.16)	0.21 (15.11)	0.27 (0.02)
Jazz Musicians (JAZ)	0.19 (0.19)	0.26 (0.19)	0.46 (0.17)	0.54 (0.20)	0.74 (0.53)
Pretty Good Privacy (PGP)	0.01 (0.27)	0.06 (0.62)	0.09 (0.64)	0.09 (0.62)	0.04 (0.54)
Facebook NIPS (FBN)	0.02 (0.45)	0.03 (0.96)	0.11 (0.74)	0.35 (0.15)	0.35 (0.00)
Hamster Full (HAM)	0.1 (0.54)	0.19 (1.01)	0.21 (1.3)	0.41 (91.89)	0.6 (0.03)

Table 3: Community-based isolation strategy for different thresholds (only five values are shown due to space constraints). The number in parenthesis is the budget spent normalized by lockdown budget. The number outside is the peak of the number of infected individuals normalized by the network size. The bold values show thresholds for which the best trade-off between cost and peak of infection is achieved.

5 Conclusion

In this study, we focused on several mitigation strategies practice in real world and provided two strategies that can lead to better performance. We showed

(1) The addition of 1-hop neighborhood ranking to TTI can enhance its performance considerably both in cost and reducing the speed of spread, and (2) community membership directly influences the speed of spread with a reasonable cost, and is a promising direction for designing new mitigation strategies. Our results encourage further research on TTI-based and Community-based methods as powerful tools to substitute the current naïve solutions to the pandemic mitigation strategies in early stages.

References

1. D. Weston, K. Hauck, and R. Amlôt, "Infection prevention behaviour and infectious disease modelling: a review of the literature and recommendations for the future," *BMC public health*, vol. 18, no. 1, p. 336, 2018.
2. B. J. Cowling, D. M. Ng, D. K. Ip, Q. Liao, W. W. Lam, J. T. Wu, J. T. Lau, S. M. Griffiths, and R. Fielding, "Community psychological and behavioral responses through the first wave of the 2009 influenza a (h1n1) pandemic in hong kong," *The Journal of infectious diseases*, vol. 202, no. 6, pp. 867–876, 2010.
3. T. Cruwys, M. Stevens, and K. H. Greenaway, "A social identity perspective on covid-19: Health risk is affected by shared group membership," *British Journal of Social Psychology*, 2020.
4. J. Cannarella and J. A. Spechler, "Epidemiological modeling of online social network dynamics," *arXiv preprint arXiv:1401.4208*, 2014.
5. T. Kuchler, D. Russel, and J. Stroebel, "The geographic spread of covid-19 correlates with structure of social networks as measured by facebook," tech. rep., National Bureau of Economic Research, 2020.
6. S. Ma, L. Feng, and C.-H. Lai, "Mechanistic modelling of viral spreading on empirical social network and popularity prediction," *Scientific reports*, vol. 8, no. 1, pp. 1–10, 2018.
7. P. Holme, "Model versions and fast algorithms for network epidemiology," *arXiv preprint arXiv:1403.1011*, 2014.
8. T. Britton, "Stochastic epidemic models: a survey," *Mathematical biosciences*, vol. 225, no. 1, pp. 24–35, 2010.
9. C. Anastassopoulou, L. Russo, A. Tsakris, and C. Siettos, "Data-based analysis, modelling and forecasting of the covid-19 outbreak," *PloS one*, vol. 15, no. 3, p. e0230405, 2020.
10. E. Bonyah and K. O. Okosun, "Mathematical modeling of zika virus," *Asian Pacific Journal of Tropical Disease*, vol. 6, no. 9, pp. 673–679, 2016.
11. V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, and A. Vespignani, "Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions," *PLoS Med*, vol. 4, no. 1, p. e13, 2007.
12. B. Ivorra, M. R. Ferrández, M. Vela-Pérez, and A. Ramos, "Mathematical modeling of the spread of the coronavirus disease 2019 (covid-19) taking into account the undetected infections. the case of china," *Communications in nonlinear science and numerical simulation*, p. 105303, 2020.
13. Y. B. de Bruin, A.-S. Lequarre, J. McCourt, P. Clevestig, F. Pigazzani, M. Z. Jeddi, C. Colosio, and M. Goulart, "Initial impacts of global risk mitigation measures taken during the combatting of the covid-19 pandemic," *Safety Science*, p. 104773, 2020.
14. M. I. Meltzer, N. J. Cox, and K. Fukuda, "The economic impact of pandemic influenza in the united states: priorities for intervention," *Emerging infectious diseases*, vol. 5, no. 5, p. 659, 1999.
15. A. Atkeson, "What will be the economic impact of covid-19 in the us? rough estimates of disease scenarios," tech. rep., National Bureau of Economic Research, 2020.
16. B. N. Ashraf, "Economic impact of government interventions during the covid-19 pandemic: International evidence from financial markets," *Journal of Behavioral and Experimental Finance*, vol. 27, p. 100371, 2020.
17. G. Bonaccorsi, F. Pierri, M. Cinelli, A. Flori, A. Galeazzi, F. Porcelli, A. L. Schmidt, C. M. Valensise, A. Scala, W. Quattrocioni, et al., "Economic and social consequences of human mobility restrictions under covid-19," *Proceedings of the National Academy of Sciences*, vol. 117, no. 27, pp. 15530–15535, 2020.
18. C. N. Ngonghala, E. A. Iboi, and A. B. Gumel, "Could masks curtail the post-lockdown resurgence of covid-19 in the us?," *Mathematical biosciences*, vol. 329, p. 108452, 2020.
19. H. Brüßow, "Covid-19: Test, trace and isolate-new epidemiological data," *Environmental Microbiology*, 2020.
20. R. S. D. Initiative, "Test, trace, isolate." <http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm>. Accessed: 2020-08-21.
21. J. Kunegis, "Konect: the koblenz network collection," in *Proceedings of the 22nd International Conference on World Wide Web*, pp. 1343–1350, 2013.
22. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.