Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 55-20 (2022) 451–456

Contact Tracing for Disease Containment: a Network-Based Analysis ⋆

Felix Gigler *,*** **Christoph Urach** ** **Martin Bicher** *,**

* *Institute of Information Systems Engineering, TU Wien, Favoritenstraße 11, 1050 Vienna, Austria (e-mail: felix.n.gigler@gmail.com)*

** *dwh GmbH, Neustiftgasse 57-59, 1070 Vienna, Austria (e-mail: christoph.urach@dwh.at)*

*** *AIT Austrian Institute of Technology GmbH, Giefinggasse 4, 1210 Vienna, Austria (e-mail: felix.gigler@ait.ac.at)*

Abstract: Since the outbreak of the COVID-19 pandemic in spring 2020, the concept of test, trace, and isolate (TTI) was used as a non-pharmaceutical intervention against further spreading of the disease. Hereby, recent contact partners of newly confirmed SARS-CoV-2 infected persons were identified and isolated along with the originally detected case to avoid potential secondary infections. While the policy is, given the compliance of the traced persons, generally deemed efficient, not much is known about network-specific impact factors.

In this work, we aim to evaluate the effectiveness of the TTI strategy when used (1) for diseases with different infectiousness levels and (2) on different contact networks. For the prior, we vary the infection probability per contact, for the latter, we analyse different clustering coefficients. Our goal is to test the validity of two hypotheses: First, we expect the policy to be more efficient if the infectiousness of the disease is small, since the time delay for isolating persons is crucial. Second, due to the implications of the friendship paradox, we expect the policy to be more effective if the clustering coefficient of the underlying contact network is high.

We make use of an agent-based network model consisting of three intertwined model parts: an epidemiological SEIR model, a quarantine model and a contact-tracing model. To test the hypotheses, the disease parameters and the clustering coefficient of the underlying contact network are varied.

The simulation results show that, indeed, tracing seems to have a slightly larger containment impact for networks with higher clustering, in particular for fast-spreading diseases. Yet, the effects are small compared to the impact of the infectiousness of the disease. Therefore, we find a significant decrease of the policy effectiveness the higher the transmission probability. The latter implies that the containment impact of tracing and isolating contacts becomes more efficient, if supported by additional measures that limit the infection probability or if applied in periods with low negative seasonality effects.

Keywords: COVID-19, SARS-CoV-2, epidemiology, contact-tracing, agent-based modelling, network modelling, discrete-event simulation, clustering

1. INTRODUCTION

Test, trace and isolate (TTI) is a well established countermeasure to reduce the chances of secondary infections and, as a result, the overall number of transmissions in an epidemic situation. Due to the spread of SARS-CoV-2 in 2020 and 2021 many countries in the world have established rigorous tracing programs to reduce the reproduction rate of the disease (84 countries by Sep 1st 2021 Ritchie et al. (2020)).

TTI involves four steps of which the first two are usually regarded as test and isolate (TI) strategy: (1) a new infected case is detected, (2) the case is isolated, (3) contact partners are traced, and (4) contact partners are isolated as well. Although it is well established that neither TI nor TTI alone are sufficient to contain a fast spreading pandemic Contreras et al. (2021); He et al. (2021); Bicher et al. (2021), these measures yet provide a proven reduction of the effective reproduction number which does not rely on limiting everyone's occupational and social live. Therefore, TI and TTI can be an important contribution to a well working containment strategy.

The timeframe for detecting and putting individuals under quarantine is crucial for TI and TTI strategies (Grantz et al. (2021); Smith et al. (2021)). The faster newly infected individuals can be isolated, the more secondary infections can be prevented. Usually some delay is involved consisting of time to detect and isolate the index case, time to find the contact partners, and time to notify and isolate them. Vice versa, keeping the delay constant, the policies

⋆ The study was partially funded by the Austrian Research Promotion Agency (FFG) Covid19 Emergency Call

10.1016/j.ifacol.2022.09.136

can prevent fewer secondary infections if the infectiousness of the disease was higher, since the serial interval, the average time between primary and secondary infection, decreases. Following this idea we would expect that TI and TTI strategies become less effective the higher the infectiousness. This also implies that the policies have a greater impact if they are supported by other measures that limit the virus transmission or when infectivity is reduced due to external influences (e.g. during summer) which we refer to as low seasonality (Liu et al. (2021)).

Moreover, both secondary infections and contact tracing depend on a person's contact network. Consequently, the structure of the contact network of the underlying population has direct implications for the spread of the disease and the effectiveness of TTI. For the prior, the authors in Lee et al. (2019) have already shown that a higher clustering coefficient of the underlying contact network causes lower disease prevalence. Implications of the clustering coefficient for the effectiveness of TI and TTI have not yet been investigated.

The famous friendship paradox supports the hypothesis that a high clustering coefficient of the underlying contact network increases the effectiveness of contact tracing. The friendship paradox Feld (1991) is an observation discovered by Scott Feld in the early 1990s and entails that your friends, on average, have more friends than you which can directly be transferred to contact partners. This seemingly paradoxical observation can be proven via probability/graph theory and becomes stronger the higher the friendship network is clustered. The paradox becomes more intuitive by considering the fact that you are more likely to be a friend of a person with more friends than of a person with few friends. Clearly, this idea can be directly transferred to epidemiological applications. Famously, strategies for specific vaccination programs are known since (at least) the early 2000s Holme (2004); McGail et al. (2022). Also case surveillance strategies exist Amaku et al. (2015).

Yet, so far, the paradox has not yet been investigated for its impact on TTI. Since potentially infectious contact partners are more likely to have more contacts than the index case, TTI might benefit from the paradox. Isolating contact partners would prevent more secondary infections in a highly clustered contact network than it would in a homogeneous one.

The goal of this work is to test the two stated hypotheses about the impact of infectiousness and network structure using an agent-based network model. Varying the disease parameters and the clustering coefficient of the contact network, we measure the benefit of using a TI or a TTI strategy and compare the results with a simulated disease outbreak without containment strategies.

2. METHODS

2.1 Model Definition

We use an agent based model where agents are members of a randomly generated social network, only interacting with connected individuals. These agent interactions are described as discrete events and are simulated in a Monte Carlo setting. Three different model parts, which allow application of different strategies, are combined to model the epidemic:

(1) An epidemiological SEIR-model which describes the general infection dynamics
(2) A quarantine module where TI counter measures are incorporated into the infection dynamics
(3) A tracing module which additionally allows for tracking secondary infections.

The three settings we focus on are the disease spread when no measures are being taken, when TI counter measures are employed or when TTI is applied.

Agents The agents in this model are described as nodes on a graph $g = (V, E)$, where V is the set of nodes and E is the set of edges between nodes. Two agents v_1, v_2 connected by an edge $\{v_1, v_2\} \in E$ are called *neighbours* and correspond to two individuals in the model who know each other and may have potentially infectious contacts.

Each agent has the attributes *InfectionState* and *ContactHistory*. The attribute *InfectionState* tracks the current state the agent is in with regards to the pathogen (susceptible, exposed,...) whereas the *ContactHistory* is a list of contacted agents that can be used for tracing.

Initialisation The graph g is initialised as an instance of a $G(n, p)$ Erdős-Rényi random network, meaning a network with n nodes where each pair of nodes is connected with probability p Gilbert (1959). Such Erdős-Rényi random networks are unrealistic models for social networks; scale-free networks such as the graph model of Barabási and Albert (1999) would better fit real networks. However, these scale-free models, as implemented in the library NetworkX Hagberg et al. (2008), allowed too little control over the observed clustering, making them unfit for our purposes.

At initialisation, only random networks which are connected graphs are accepted (i.e. there is a path between any two nodes in the graph). Disconnected networks are rejected in order to get comparable infection dynamics between different networks.

As observed in Oliveira et al. (2018), there are two common definitions of the clustering coefficient which are often used interchangeably, even though it can be shown that their values differ in some graphs. We use the definition of the *local clustering coefficient* $C(g)$ of a graph g provided in Oliveira et al. (2018): $C(g)$ is the average, measured over all nodes v of g, of the fraction of pairs of neighbours of v which are connected (and thus form a triangle with v).

The expected clustering of a $G(n, p)$ random network is p. After sampling the network we use a heuristic to artificially alter the clustering coefficient: Let τ be the target clustering coefficient value. Select two different nodes $a, b \in V$ at random, and let $d(a) \geq d(b)$. If $C(g) < \tau$, remove a random edge from b and move it to a (while assuring that the number of edges in the graph remain constant and the graph stays connected). This will lead to higher clustering around a, which in the majority of cases has a bigger effect than the loss of triangles around b, thus increasing $C(g)$. Vice versa, in the case of $C(g) < \tau$,

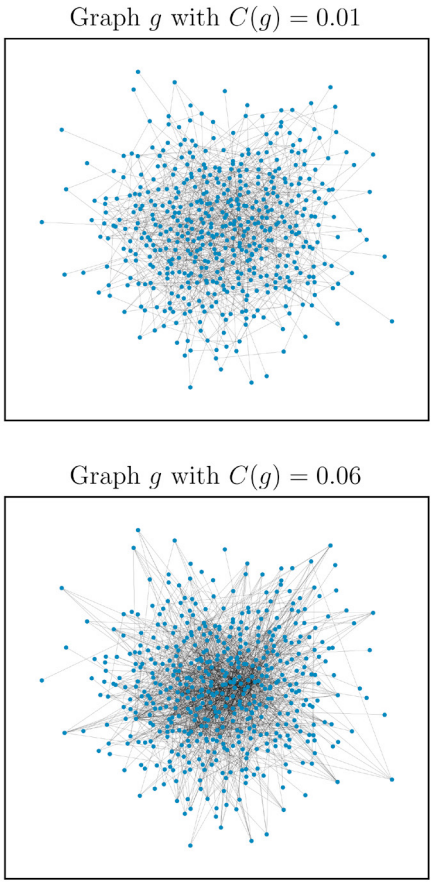

Fig. 1. Graph visualisation for different clustering coefficients.

moving an edge from a to b will decrease the clustering coefficient. The heuristic terminates once the clustering coefficient is sufficiently close to the target value. An example is visualised in Figure 1, where the same network is treated with the heuristic to create different amounts of clustering. It is noticeable that the second network has a higher degree of clustering, as the number of complete triangles is higher and the nodes appear to be much less homogeneously connected. The number of edges is the same in both networks.

All agents in the network initially have *InfectionState* ← *susceptible* (meaning they can contract the disease) and get an empty *ContactHistory*.

Disease Spread Module For interactions between agents, *discrete event simulation* (DES) is used. Temporal aspects of the infection dynamics are modelled as exponentially distributed random variables $t \sim Exp(\mu)$. The following event types and their event routines describe the base disease spread model:

(1) *infection*: The incubation period begins and the agent's *InfectionState* becomes *exposed*. An *infectious* event is scheduled after $t \sim Exp(t_i)$ days.
(2) *infectious*: The agent becomes infectious and two events are scheduled, a *contact* event after $t \sim Exp(t_c)$ days and a *recover* event after $t \sim Exp(t_r)$ days. *InfectionState* ← *infectious*.
(3) *contact*: All neighbouring agents are contacted. For each contact partner, a Bernoulli experiment with probability p_i decides if the contact leads to a new infection. If the Bernoulli experiment succeeds, a new *infection* event is scheduled for the contacted agent (at current time). In any case, a new *contact* event is scheduled after $t \sim Exp(t_c)$ days.
(4) *recover*: The agent recovers from the illness and *InfectionState* ← *recovered*. All scheduled contact events for this agent are cancelled.

A specific event is then identified with a tuple of the scheduled event time t, the event type and the unique ID of the acting agent. Scheduled events are saved in an event list and processed in order of their event time t. For efficiency, we use a *priority queue* as event list.

At simulation start, three agents are initially infected by scheduling corresponding infection events.

Quarantine Module In order to model TI counter measures, additional event types are necessary. We define additional event types:

(5) *quarantine*: The agent is isolated, triggering the assignment *InfectionState* ← *isolated*. An *end-quarantine* event is scheduled after a fixed time t_q. Until this event resolves, edges connecting the agent to others are deactivated. The *quarantine* event is additionally scheduled as part of the infection event routine described above, $t \sim Exp(t_d)$ days after infection with a certain probability p_d (meaning not all cases are detected).
(6) *end-quarantine*: The isolation is discontinued and edges are re-enabled.

These additional events, together with the base infection events described above, allow modelling the pandemic with TI counter measures.

Tracing Module Combining the events above with a new event routine *tracing* lets us simulate TTI infection dynamics:

(7) *tracing*: At time of the detection of an infection, this event is scheduled additionally. At completion it schedules quarantine events for each agent $v_i \in$ *ContactHistory* after $t_i \sim Exp(t_t)$ days. Each of these events get scheduled only with a certain probability p_t due to the difficulty of effective tracing.

Implementation The model is simulated in Python. The reproducible source code for all experiments conducted as part of this work can be found on GitHub [1].

2.2 Parameter Values and Experiment Design

The focus of this work lies on a qualitative analysis of the network's clustering on the effectiveness of TI and TTI counter measures. It is noted that the parameter values for this model are neither calibrated nor parameterized through systematic literature research, but are instead fixed to reasonably realistic values manually. The results should therefore be seen as purely qualitative examples

[1] https://github.com/figlerg/NetTraceSim

Table 1. Parameter values for base model

Parameter	Value	Unit	Description
n	500	#	number of agents
t_i	2	days	avg. incubation time
t_c	1	days	avg. time between contacts
t_r	10	days	avg. recovery time
t_d	6	days	avg. detection time
t_q	14	days	quarantine time
t_t	1	days	avg. tracing time
p_i	0.2	prob.	prob. of infection after contact
p_d	0.5	prob.	prob. of detection
p_t	0.75	prob.	prob. of successful tracing
p	0.01	prob.	edge prob. in random graph

in order to test the hypothesis in general networks. The parameter values can be seen in Table 1.

The experimental results are products of Monte Carlo simulations, where the resulting time series of multiple simulation runs are used to get an average system behaviour. Since the generation of random networks is slow in the library that is used, networks are only re-sampled every five iterations of the Monte Carlo simulation.

Variables of Interest We define two different variables of interest of the severity of an epidemic curve in order to evaluate the effectiveness of different strategies: The peak point prevalence of the pandemic α, defined as maximum fraction of infected individuals at any given time, as well as the overall period prevalence β, the fraction of individuals that were infected over the full course of the pandemic. When evaluating any specific counter measure, we use the ratio of the prevalence without counter measures to those with counter measures enabled. This gives the effectiveness in terms of an improvement factor when compared to the base simulations, where lower numbers mean a higher improvement.

Since α and β originate from the quotient of two stochastic simulation results, we need to emphasise that simply dividing the estimated peak and prevalence from the sample means of the Monte Carlo runs would result in a biased estimator. To avoid this bias and allow for quantitative uncertainty estimation, we defined the following process:

We calculate prevalence and peak of every experiment and calculate sample mean $\overline{\mu}$ and standard deviation $\overline{\sigma}$ over all Monte Carlo runs. Since the (point-wise) sample mean of the Monte Carlo simulation with N independent runs is approximately Gaussian distributed with mean $\overline{\mu}$ and standard deviation $\overline{\sigma}/\sqrt{N}$ (Central Limit Theorem), the quantities α and β can be estimated as the quotient of two normally distributed random variables. The corresponding ratio distribution is estimated numerically, leading to mean value and confidence intervals for the two variables of interest.

3. RESULTS

The effect of the clustering variation on the modelled epidemic curves (mean values of the Monte Carlo simulation) for $p_i = 0.2$ can be seen in Figure 2. Hereby, the baseline strategy without policies is shown in the uppermost subplot and the results for TI and TTI are displayed in the two subplots below. The curves show that the disease generally spreads more slowly on a highly clustered network than on

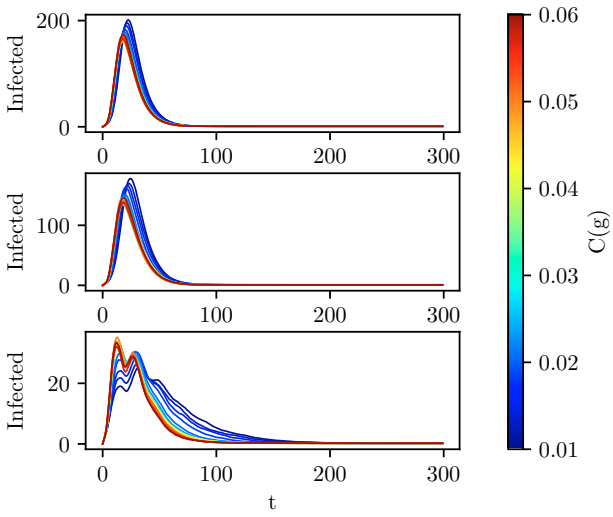

Fig. 2. Epidemic curves for $p_i = 0.2$ and varied clustering coefficients. The uppermost image displays the results without policies, below the results of the TI and TTI policy are shown.

a lower clustered one, which is a well known phenomenon. Moreover, the variables of interest are evaluated by determining peak and prevalence of the displayed curves and by dividing the ones from the results with the investigated policy by the corresponding ones from the baseline results. These are displayed in Figures 3 and 4.

4. DISCUSSION

The results displayed in Figure 3 show a clear trend with respect to the decrease of the policy impact when increasing the infectiousness. This observation is almost independent of the clustering structure of the network and fully confirms the first hypothesis. For the impact of the clustering coefficient displayed in Figure 4, the results are less conclusive. The expected trend towards a more effective policy for increasing clustering coefficient can only be seen for the TTI strategy but not for the TI strategy, which supports the validity of the hypothesis. Yet, for the TTI strategy only the results with higher infectiousness ($p_i \geq 0.2$) display the expected behaviour.

Comparing the results for different p_i in Figure 3, we find that the TTI policy can reduce the overall prevalence from about 75% with $p_i = 0.6$ to about 10% with $p_i = 0.05$, mostly independent of the network structure. The corresponding reductions on the peak are very similar. Following this trend, halving the infection probability would cause the policy to reduce the prevalence or peak height by an additional 20 to 30%. This observation has wide-ranged consequences on disease containment in reality. The results would imply that in European countries TTI could work significantly better during the summer months. In Engelbrecht and Scholes (2021) authors estimate that during the summer months the infectivity of the virus is up to $25 - 40\%$ smaller than in winter, which would make TTI about $10 - 20\%$ more effective. These ideas are also applicable to any other measure that reduces the infection probability, for example increasing hygiene and wearing face masks.

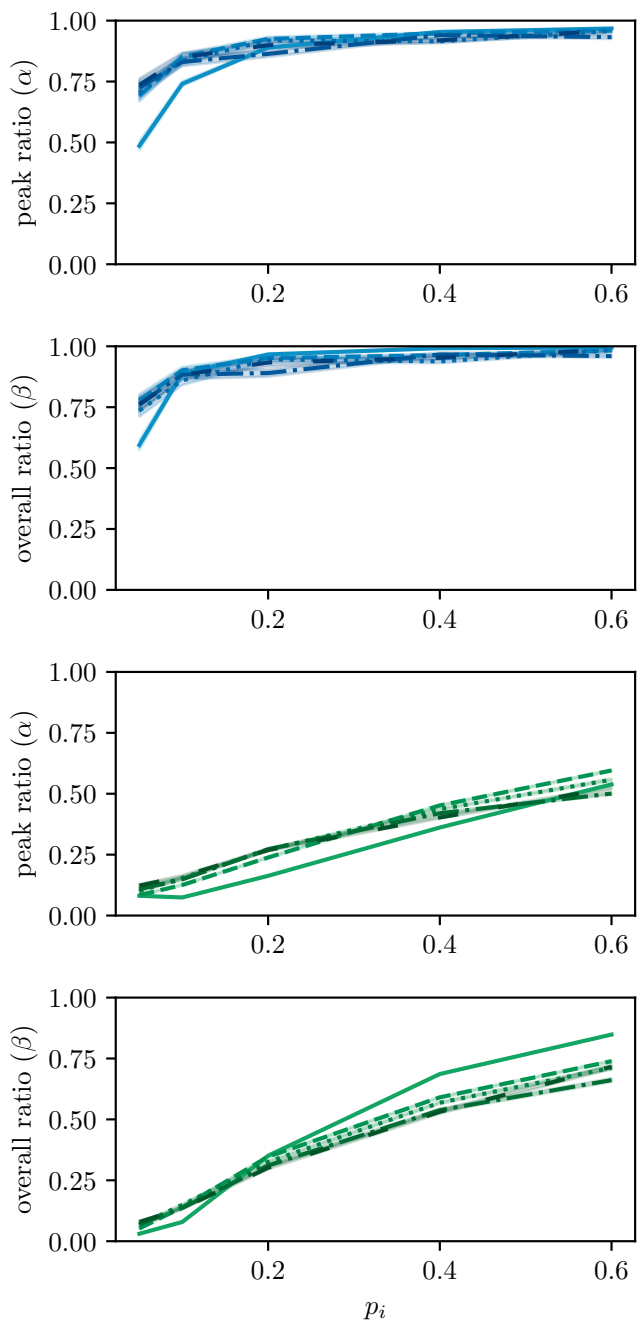

Fig. 3. Variables of interest for the TI and TTI strategy. We varied p_i between 0.05 and 0.6 and the clustering coefficient of the network between 0.01 (low clustered) and 0.06 (highly clustered). The shaded areas display a 90% confidence interval of the quantities with respect to the stochasticity of the simulation.

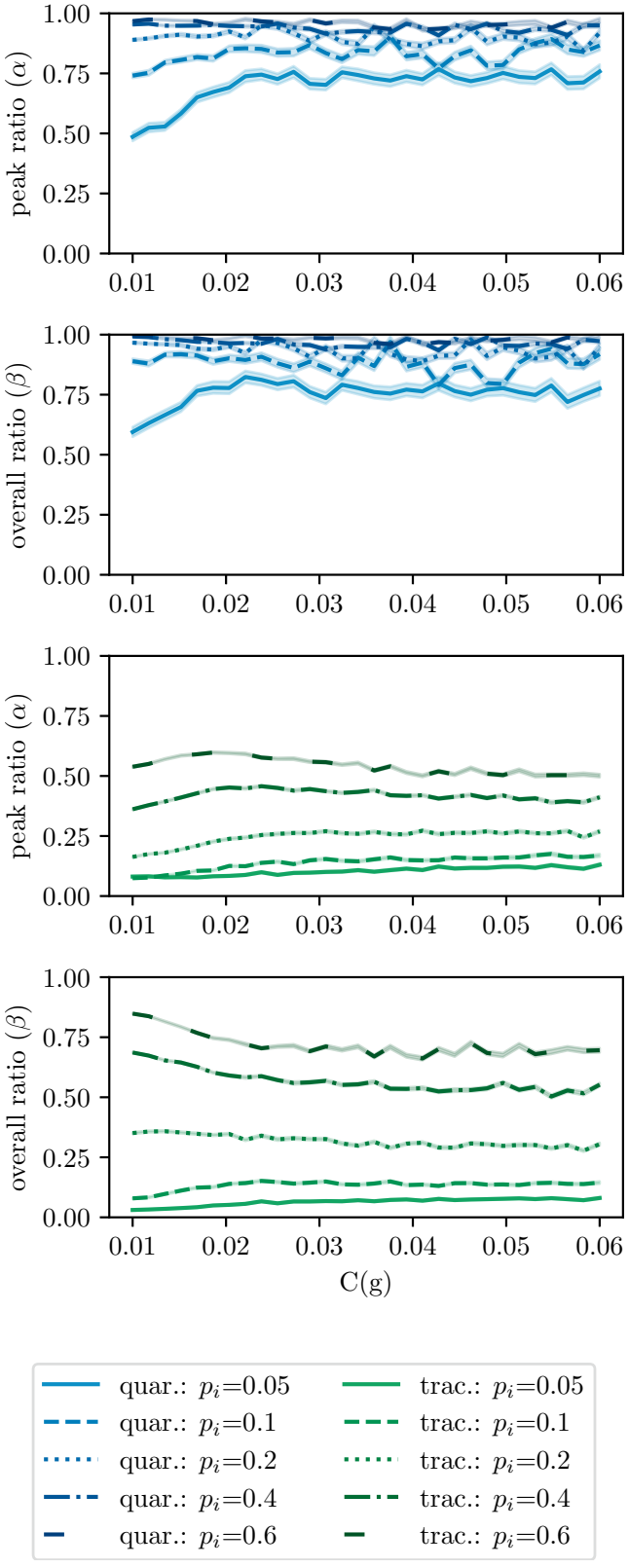

Fig. 4. Variables of interest for the TI and TTI strategy. We varied p_i between 0.05 and 0.6 and the clustering coefficient of the network between 0.01 (low clustered) and 0.06 (highly clustered). The shaded areas display a 90% confidence interval of the quantities with respect to the stochasticity of the simulation.

Considering the impact of the clustering coefficient, the results for the TTI strategy with $p_i > 0.1$ show a trend towards being more effective for larger clustering. The results with $p_i \leq 0.1$ surprisingly don't show this trend which might be due to the fact that the disease is already very close to extinction in these scenarios. Nevertheless, even for those results where the trend is visible, the slope of the curve is comparably small, in particular compared to the plots which vary the infection probability. Thus, the friendship paradox applies in these cases, but its impact on the effectiveness of the policy is very weak.

Interestingly, also the results for the TI strategy do not seem entirely independent of the clustering coefficient. The results show a decrease of the policy effectiveness when comparing a very loosely clustered network to a network with high clustering. The authors did not find a fully convincing argument to explain this behaviour, but it might originate from the nonlinear nature of the outcome variables for very high peaks, since the effect is particularly significant for low clustering and high infectiousness. Corresponding results for the TTI strategy show a similar but less distinct behaviour, which supports this idea.

The results of the study also imply that both TI and TTI are important measures to contain the epidemic and are capable of reducing prevalence of an epidemic outbreak. The latter, namely the variables of interest and the setup of the simulation experiment, also pose for the biggest limitation of this modelling and simulation study. Considering a synthetic contact network with a full epidemic outbreak of the disease is not a realistic situation since additional policies would be applied to prevent a full outbreak. Moreover, the baseline scenario is not a valid reference, since "no-policy" is not a valid alternative in the real system. In order to solve this problem, a similar strategy as shown in Bicher et al. (2021) could be applied instead. In this work, the effectiveness of contact tracing is quantified based on how many additional contact reduction policies need to be active at the same time to reach $R_{eff} = 1$. This strategy, yet, exceeded the scope of the present study. Considering the different setup, model complexity and defined outcome variables, the results of the two studies are not comparable. Generally, comparison with measured SARS-CoV-2 case data is difficult, since the metrics of the real contact network are not understood. With estimates for the clustering coefficient of real networks, a comparison e.g. between TTI effectiveness in different countries would be possible.

In this study we developed an agent-based simulation model for the purpose of comparing the impact of TI and TTI for different parameters. Hereby, we varied the infectiousness of the disease and the clustering coefficient of the underlying network. We conclude that both TI and TTI tend to become more effective, the less infectious the disease. The impact of clustering is less significant and only applies if infectiousness is also comparably high.

REFERENCES

Amaku, M., de Hildebrand Grisi-Filho, J.H., Negreiros, R.L., Dias, R.A., Ferreira, F., Neto, J.S.F., Cipullo, R.I., Marques, F.S., and Ossada, R. (2015). Infectious disease surveillance in animal movement networks: An approach based on the friendship paradox. *Preventive Veterinary Medicine*, 121(3), 306–313. doi: https://doi.org/10.1016/j.prevetmed.2015.08.002.

Barabási, A.L. and Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439), 509–512.

Bicher, M., Rippinger, C., Urach, C., Brunmeir, D., Siebert, U., and Popper, N. (2021). Evaluation of contact-tracing policies against the spread of sars-cov-2 in austria: An agent-based simulation. *Medical Decision Making*, 0272989X211013306.

Contreras, S., Dehning, J., Loidolt, M., Zierenberg, J., Spitzner, F.P., Urrea-Quintero, J.H., Mohr, S.B., Wilczek, M., Wibral, M., and Priesemann, V. (2021). The challenges of containing sars-cov-2 via test-trace-and-isolate. *Nature communications*, 12(1), 1–13.

Engelbrecht, F.A. and Scholes, R.J. (2021). Test for covid-19 seasonality and the risk of second waves. *One Health*, 12, 100202.

Feld, S.L. (1991). Why your friends have more friends than you do. *American journal of sociology*, 96(6), 1464–1477.

Gilbert, E.N. (1959). Random graphs. *The Annals of Mathematical Statistics*, 30(4), 1141–1144.

Grantz, K.H., Lee, E.C., D'Agostino McGowan, L., Lee, K.H., Metcalf, C.J.E., Gurley, E.S., and Lessler, J. (2021). Maximizing and evaluating the impact of test-trace-isolate programs: A modeling study. *PLoS medicine*, 18(4), e1003585.

Hagberg, A.A., Schult, D.A., and Swart, P.J. (2008). Exploring network structure, dynamics, and function using networkx. In G. Varoquaux, T. Vaught, and J. Millman (eds.), *Proceedings of the 7th Python in Science Conference*, 11 – 15. Pasadena, CA USA.

He, B., Zaidi, S., Elesedy, B., Hutchinson, M., Paleyes, A., Harling, G., Johnson, A.M., Whye Teh, Y., and group, R.S.D. (2021). Effectiveness and resource requirements of test, trace and isolate strategies for covid in the uk. *Royal Society open science*, 8(3), 201491.

Holme, P. (2004). Efficient local strategies for vaccination and network attack. *EPL (Europhysics Letters)*, 68(6), 908.

Lee, H.W., Malik, N., Shi, F., and Mucha, P.J. (2019). Social clustering in epidemic spread on coevolving networks. *Physical Review E*, 99(6), 062301. doi: 10.1103/PhysRevE.99.062301.

Liu, X., Huang, J., Li, C., Zhao, Y., Wang, D., Huang, Z., and Yang, K. (2021). The role of seasonality in the spread of covid-19 pandemic. *Environmental research*, 195, 110874.

McGail, A.M., Feld, S.L., and Schneider, J.A. (2022). You are only as safe as your riskiest contact: Effective covid-19 vaccine distribution using local network information. *Preventive medicine reports*, 27, 101787.

Oliveira, R.I., Ribeiro, R., and Sanchis, R. (2018). Disparity of clustering coefficients in the holme–kim network model. *Advances in Applied Probability*, 50(3), 918–943.

Ritchie, H., Mathieu, E., Rodes-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., and Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*. Https://ourworldindata.org/coronavirus.

Smith, L.E., Potts, H.W., Amlôt, R., Fear, N.T., Michie, S., and Rubin, G.J. (2021). Adherence to the test, trace, and isolate system in the uk: results from 37 nationally representative surveys. *bmj*, 372.