Agent-based modeling of disease spreading in a scenario of bio-terrorism

Ella Guiladi, Kevin Rylander, Emma Svensson and Haodong Zheng
Chalmers University of Technology, Gothenburg, Sweden

January 12, 2020

Abstract

The natural and intentional spread of infectious diseases through a population is a process of great importance to the modern, overpopulated and globalized society. Traditional epidemiological models do not take the individual information into account, thus have restricted ability to apprehend the intrinsic complexity of the disease spreading process. In order to overcome this limitation, agent-based models have been suggested as a successful approach to model complex disease spreading. In this paper an agent-based model for infectious disease is proposed for a greater understanding of disease spreading through a network of human contacts. By developing a base model, i.e an epidemiological model, into a more complex agent-based model, that model can be extended in the form of additional features to the environment as well as added behavioral changes to and differences between agents. It was shown that adding travellers to the simulations and lowering the threshold for when people exhibit a panic response, increased the likelihood of a population wide-spread disease. In conclusion, the natural response to panic and flee an area with potential danger, in this case has the undesired effect of helping a disease spread faster. The extensions of the model were made with the aim to make the model more realistic and to adjust it more to scenarios of intentional disease spreading, i.e bio-terrorism.

Introduction

As the global population is constantly increasing, leading to an overpopulated world, the threat of natural virus spreading and bio-terrorism becomes increasingly imminent. In order to predict such behaviour computational modelling has come to play an important part in mapping disease spreading with the aim to restrict the outbreak as much as possible [1].

Bio-terrorism is the use of microorganisms or toxins in production of weapons causing death and disease among humans, animals and plants, which has been used since the 14th century [2]. As with natural epidemics, the biological attacks develop within complex social systems. Since it is not always clear how to detect or respond to a disease outbreak, whether it was deliberate or not, it is of great interest to develop tools that detect the disease spread. By simulating the disease spreading in socially connected groups, it is possible to investigate the different detection and response actions [3].

Epidemic modelling in form of differential equations were developed in the 1920s. These models assume that the population is perfectly mixed, with people moving from being susceptible, to being infected to lastly recover or die [1]. In these different health states, there exist no diversity among the population as well as no diversity in their behaviour. Even though these are strongly idealizing assumptions, this approach results in some important discoveries such as "herd immunity" [1]. However, these models are not suited to represent complex social networks and interactions between the different individuals, who adapt their behaviours based on the magnitude of the disease [4].

Agent-based models, ABMs, have a different approach, which is able to embrace this complexity [5]. ABMs are practically artificial social societies where every single individual is presented as an agent in the form of a distinct software object who can for example travel and sleep. The model is applied in a way that makes it possible to track the agents interaction and movement in the virtual space, thereby enabling comparison with real epidemic data after several runs. The agents can interact directly with one another and can in that sense be made to behave similar to real people having distinct behaviours [6].

In order to simulate real-world disease spreading an agent-based model was used in this report, since it describes complex societies well. For the purpose of evaluating effects of different features a basic SIR model was compared to a more complex SEIR model with additional features and agent behaviours [7], [8]. The paper shows that factors like the amount of people moving around and exhibiting realistic human responses to outbreaks, leads to a faster disease spreading and makes it easier for the disease to become population wide-spread.

Simulation

A disease spreading model was developed from a baseline model, through extensions in the form of additional features to the environment and added behavioral differences between agents. This was done for two reasons, to make the model more complex and realistic and also to adjust the model more to a scenario of bio-terrorism. Both the base model and all aspects of the extended model are discussed in detail in the following sections.

Base model

As a baseline, an agent-based version of the SIR model has been simulated, where agents have independent health statuses and move randomly in a grid environment with fixed boundaries. According to the SIR scheme, see Figure 1, susceptible agents can become infected through contact with other infected agents, based on proximity and an infection rate, β. After this they recover and become immune to the disease after some time step, based on a recovery rate, γ. Here a grid environment was used where infected individuals infect every susceptible individual on the same site based on the infection rate. The simulation was initialized such that all individuals in the most dense site are infected and run until there were no infected individuals left.
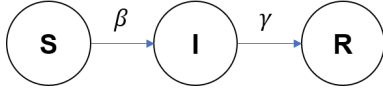

Figure 1: Scheme of the SIR model.

Extension of the base model

Since the base model is very basic and idealized the following extensions were made in order to make it more realistic. First, a time-delay was added to the progression of the disease, i.e. an incubation time. This corresponds to the SEIR model where individuals first become exposed without showing symptoms and later, after some incubation time starts showing symptoms. This incubation time is set accordingly every time an agent becomes exposed

$$t_{\mathrm{i}} = \Theta(\tilde{t}_{\mathrm{i}} + r_{\mathrm{normal}}) + 1, \qquad (1)$$

where $\Theta(x)$ is the Heaviside step function, $\tilde{t}_{\mathrm{i}}$ is the desired mean incubation time and $r_{\mathrm{normal}} \sim N(0, 10)$. The scheme of the SEIR model as defined in this section can be found in Figure 2.

For the purpose of simulating a bio-terrorism scenario a virus for which exposed as well as symptomatic individuals are contagious was chosen. In addition, the virus was made deadly by a death rate, δ, adding a fifth health state, deceased. This also, more closely resemble how a virus used in a bio-terrorism attack might operate in reality. Agents now go from susceptible to exposed to in-

fected to recovered or deceased. Behavioural differences between agents in different health states were also introduced. Such that symptomatic individuals move slower compared to others, simulating their sickness.
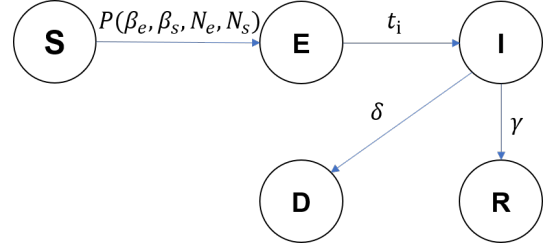

Figure 2: Scheme of the SEIR model with death rate.

In the extended model each time step can be considered an hour. To this, a day and night cycle was added where day time is 16 hours long and night time is 8 hours long. During day time agents move according to previously stated rules but during night time all movement stops. However, the disease development still progresses throughout the night, adding an extra layer of complexity combined with the time-delay of the disease.

As another feature in the extended model, the grid environment was divided into four subareas through inner boundaries, each with periodic boundary conditions that confine agents to stay inside the given area. This feature comes with airports, one in each subarea, through which the agents can travel between the different areas. The implementation was made so that every individual that walks onto the position of an airport has a given probability to stay in that location or travel to one of the other airports.

In contrast to the base model the extended model is meant to imitate a more dense environment, inspired by the overpopulation of the world. Thus, more agents are used together with a smaller dimension of the environment. This way every grid-area should be thought of as its own community with inner dynamics. Here, rather than infect every susceptible in a site, every individual is infected based on the number of exposed and symptomatic individuals at the site according to the following expression,

$$\mathrm{P(infection)} = 1 - (1 - \beta_{\mathrm{e}})^{N_{\mathrm{e}}}(1 - \beta_{\mathrm{s}})^{N_{\mathrm{s}}}, \qquad (2)$$

where β_{e} and β_{s} are the infection rates for exposed and symptomatic individuals and N_{e} and N_{s} are the number of respective individuals. This equation is used as an approximation of the inner dynamics within one site. It implies that every individual is in contact with all others in the same site, within one time step. The inner dynamics could be further developed in the future.

With regards to these changes, a second behavioural feature was also added to the agents. In a relatively short time frame, such as a few weeks, most people tend to stay in the community where they live and work. Only

a few travel outside their community. Therefore the population of agents was divided between "travellers", which move around according to the previously described rules, and "inhabitants", which stay in the community of origin. This division is done based on a traveller ratio, τ, of the desired amount of travellers.

Lastly, the following behavioral trait was added to the model. If an outbreak occurs in a site, every non-symptomatic inhabitant panics and flees the site. During a panic the individuals run away until they find a new location, free from an outbreak, where they settle down again. The panic is based only on the proportion of symptomatic individuals at the given site, and how this proportion relates to a set panic threshold. A population with a high panic threshold will not panic and the opposite for a population with low panic threshold.
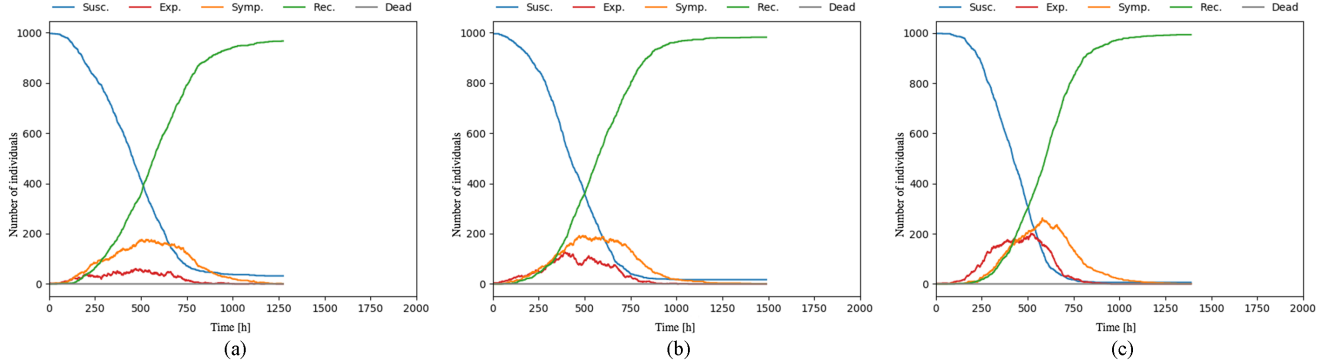

(a)   (b)   (c)

Figure 3: Simulations on baseline model but with added time-delay until showing symptoms, in the form of an incubation time drawn according to Equation 1 with $\tilde{t}_i$ = (a) 24 h, (b) 48 h and (c) 72 h. Both exposed and symptomatic individuals are contagious.

Results

The results of this paper are divided between a parameter search and a deeper exploration of the effects of, and relation between, traveller ratio τ and panic threshold. In the parameter search each new feature has been added one by one to the baseline model, for the purpose of checking the functionality and getting an intuition for the parameter settings. More feature have then been combined in order to find the best parameters settings for the final analysis.

Parameter search

Throughout the tests done on the baseline model, moving probability $d = 0.8$, infection rate $\beta = 0.6$ and recovery rate $\gamma = 0.01$ was used. The early tests showed that in a sparse population even a small death rate doesn't allow for any spreading of the disease. With a dense population, the disease becomes wide-spread much faster compared to a sparse population. The baseline model with the day and night cycle feature produces less wide-spread result compared to the original baseline model. This is expected since the infected have more time to recover and less time to spread the disease to new individuals during night time. However, the night time could also increase the chance of infected agent infecting susceptible agents in the same site, since they stay together for a longer time.

Figure 3 presents the proportions of individuals in each health state throughout simulations with mean incubation times of 24, 48 and 72 hours added to the baseline model. The proportion of susceptible individuals decreases and stagnates at some point close to zero for all incubation times. However, with longer incubation time the disease reaches more of the population. For the longest incubation time, 72 hours, the disease becomes population wide-spread, meaning the number of susceptible agents reaches 0 at the end of the simulation. It can also be seen that the number of exposed and symptomatic individuals increase with increasing incubation time.

Results from when inner boundaries and airport have been added to the model can be seen in Figure 4. Here, the feature of subareas and airports have been added to the baseline model on top of the other features; day and night cycle, incubation time of 48 hours and death rate 0.01. The simulations were run on a more dense population consisting of 4000 agents in a 50 by 50 grid, evenly divided into 4 quarters, each with its own periodic boundaries. In the figure three different runs are presented, with different probabilities of travelling given that the airport is entered, the top one shows a scenario with a 5% chance of travelling, the middle one shows a 1% chance and the bottom one shows a 0.1% chance.

It can be seen that with only a very small chance of travelling the disease doesn't spread to any areas other than the one where the outbreak started. This is clear
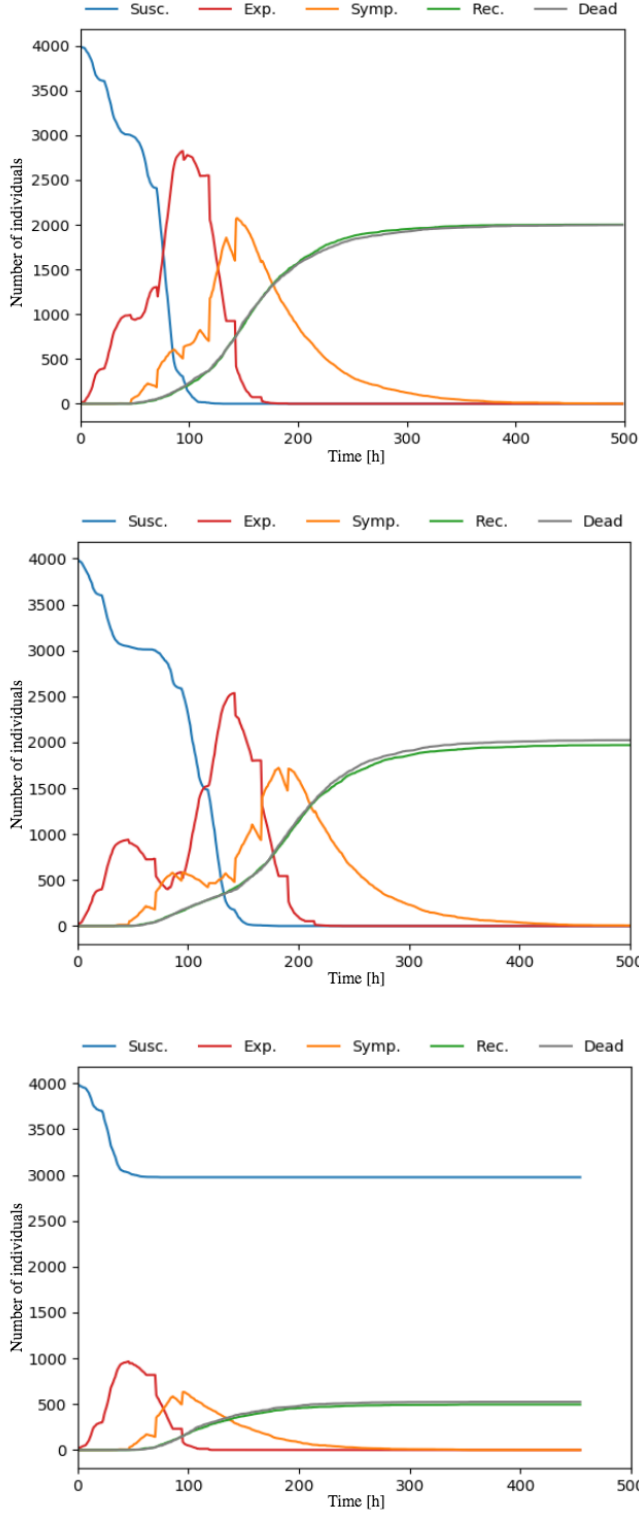
Figure 4: Simulations on the baseline model but with inner boundaries and airports, 4000 agents in a 50x50 grid, day and night cycle, incubation time of 48 hours and death rate 0.01. Top panel shows simulation with 5% chance to travel to one of the other airports given that an agent enters the airport, middle panel shows 1% chance to travel and bottom panel shows 0.1% chance to travel.

since the total number of infected, i.e. sum of recovered and dead, reaches a quarter of the total population by the end of the simulation. With a higher chance of travelling through the airports the disease always reached the entire population eventually. However, in the middle panel of Figure 4, the progression of the disease can be seen to run its course first through the subarea where the outbreak occurred, during the first 100 hours, and later spread to the remaining areas and run its course there. With the highest chance to travel tested, the disease spread fairly quickly to all the other areas.

Deeper analysis of the effects of traveller ratio and panic threshold

From the parameter search above the following parameters were chosen for the final tests regarding the ratio of travellers and the panic threshold. The incubation time was set to 48 hours. A moving probability of 0.2 for symptomatic individuals and 0.9 for others were chosen. Recovery rate $\gamma = 0.01$, infection rate for exposed $\beta_{e} = 0.15$ and infection rate for symptomatic $\beta_{s} = 0.3$ were used. And the change to travel to another airport given that an agent entered an airport was set to 50%.

These settings were used with an even higher population density consisting of 10000 agent in a 50x50 grid. Note the slightly higher moving probability for non-symptomatic individuals, which is due to the fact that now only "travellers" and panicked "inhabitants" are moving. Recall that β_{e} and β_{s} are used to determine the probability of infecting a susceptible individual at a site according to Equation 2. The probability to travel through an airport was set higher than what was tested in the results shown in Figure 4. The reason for this is that here most of the ones moving are travellers which should be able to go through the airports. A time limit for the simulations was set to 720 hours, i.e. 30 days.

Firstly, simulations were done on two scenarios, with and without travellers. In addition to the parameter settings mentioned above, a death rate of 0.01 was used. In Figure 5, two screen shots from the simulations for the two scenarios are shown. Without travellers, the disease can be seen to have spread in a manner similar to how a forest fire spreads, because an infected spot can only infect its adjacent sites through the panic response. Thus, the effect of the panic is that exposed agents spread out to other sites, thereby unintentionally helping the disease to spread. However, with travellers in the model, infected spots can instead be seen to be relatively far away from each other, and as a result the disease spreads faster and further away. Note especially the differences in time steps for the two screen shots shown in Figure 5. With travellers the disease had spread to most of the population in 170 hours whilst without travellers it took 700 hours for it to spread remotely as far.
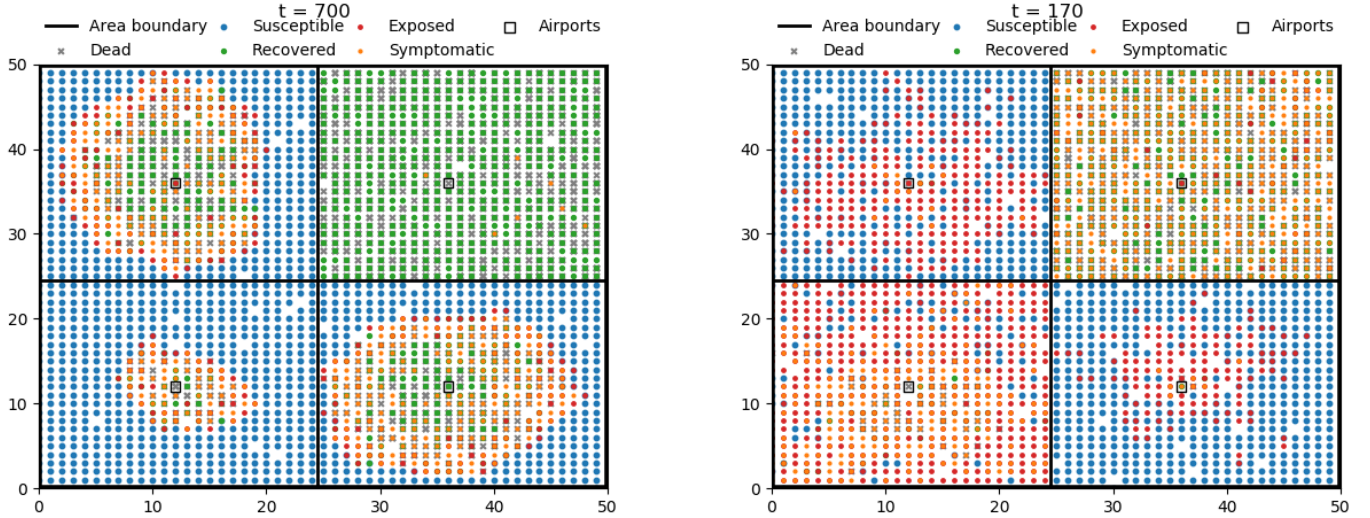
Figure 5: Visualizations from simulations of the model with all features and behaviours added. In both simulations the panic threshold was set to 20%. The left panel shows a case where there are no travellers, only inhabitants. Here, the panic is the only thing spreading the virus which then spreads more like a forest fire. In the right panel, on the other hand, travellers are introduced with a ratio of 10%.

As a measure of how fast spreading and wide reaching the disease is depending on the ratio of travellers and the panic threshold two things were explored. For these remaining tests it should be noted that the death rate was set to zero. This is due to the focus on the spreading of the disease, not the aftermath of it and thus deceased individuals have no purpose. First, the evolution of the proportion of susceptibles throughout simulations for different ratios of travellers, τ, was explored. This was done for multiple panic thresholds and the result for panic threshold 0.4 can be seen in Figure 6.
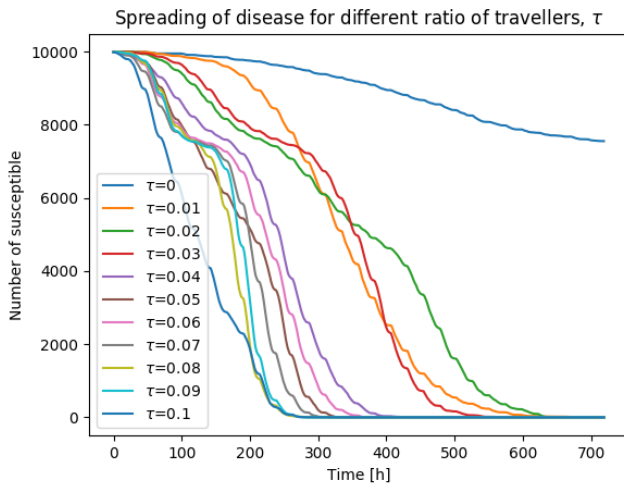

Figure 6: Number of susceptible individuals throughout simulations done with different traveller ratios, τ, for a panic threshold of 40%.

In general a higher ratio of travellers help spread the disease faster, which is clear from the figure since the slopes become increasingly steep as τ increases. For the extreme case of no travellers, $\tau = 0$, the disease never spreads outside the subarea of origin, but for the remaining cases the disease reaches a population wide-spread within the span of the simulation, 720 hours. Note that in many of the cases where the disease reaches the entire population the spread through one or more of the subareas is still visible as bumps in the slope. As an example the slope for $\tau = 0.02$ spreads in the subarea of origin approximately during time 0-200 hours, then spreads to another subarea up to time approximately 400 hours. After this point the remaining subareas get affected simultaneously. Note that the day and night cycles, of 24 hours, are also visible in the figure as the the small bumps that occur throughout most of the slopes.

To dig deeper into the effect of travellers and panic, simulations were performed on more cases with different traveller ratios and panic thresholds. The result is shown on Figure 7, each point on the 3D surface were obtained by averaging over 20 runs. From the figure, it's clear that, for a fixed panic threshold, as the traveller ratio increases, the number of susceptible agents left at terminal time tends to decrease, which means the disease is more wide-spread with more travellers. On the other hand, if the traveller ratio remained unchanged, increasing the panic threshold would result in more susceptible agents at termination time. This implies if the agents were less likely to panic towards epidemic outbreaks, the disease would be less wide-spread. In other words, more travellers and easier-panic population would lead to wider
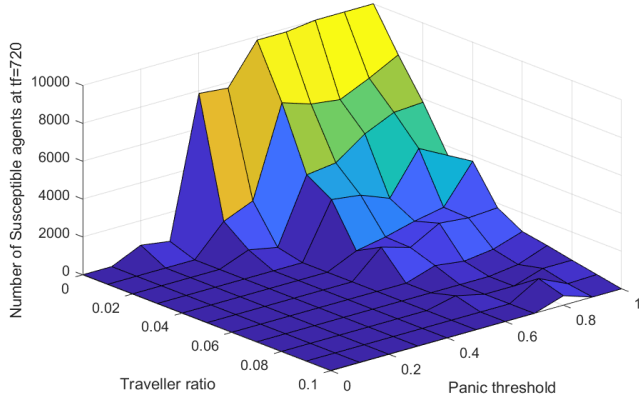
disease spreading.

Figure 7: Number of susceptible individuals at tf = 720 (hours), averaged over 20 simulations, for traveller ratio $\tau = 0, 0.01, 0.02, ..., 0.08, 0.09, 0.1$, $panic threshold = 0, 0.1, 0.2, ..., 0.8, 0.9, 1$.

As the second measure of the speed and reach of the disease, the time until the number of susceptibles reached zeros and the number of susceptibles at the end of the simulations were calculated. Recall that no individuals can go from any health state back to susceptible. Therefore, these two measures complement each other. Whilst one of the two measures take values in between its limits the other takes values close to its limit. For example if the number of susceptible never reached 0 within the time span of the simulation, or in other words the time until wide-spread disease is 720 hours, the final number of susceptible takes some value in between 0 and 10000. These measures were found for different ratios of travellers and different panic thresholds, and the results for panic threshold 0.4 and 0.7 are shown in Figure 8. Here, the results are averaged over 20 simulation.

Figure 8 shows that a higher panic threshold, meaning less panic, allows for a higher ratio of travellers before the disease reaches the entire population within the time span. It can also be seen that larger traveller ratios corresponds to the disease becoming population wide-spread faster. Another interesting point to note about the results regarding the traveller ratio, is that the effect that increasing the traveller ratio has on the speed is less and less prominent with increasing traveller ratio. This implies that travellers are very effective in helping the disease spread. Since both an increased ratio of travellers and a scenario with more panic, lower panic threshold, implies that more people are moving around, a conclusion can be drawn that more movement helps spread the disease faster to larger areas.

Discussion

From Figure 5 it is shown that when there are no travellers at all, the disease spreading is similar to the spreading of a forest fire, since the disease spread is solely based on the effect of panic in that case. The outcome of implementing a panic parameter shows that by increasing the panic in a population, the individuals in the population moves away from the area of infection, regardless of their health state. As expected this results in a wider spreading of the disease. However, this is an interesting point, since in the case of a bio-terrorist attack panic might be unavoidable, at least if the population is aware of the attack/danger. Similarly, the results from Figures 6 and 8 presents that when increasing the traveller ratio to about 10% of the population, the disease becomes population wide-spread. It is also shown that the effect of an increasing traveller ratio becomes less prominent as the traveller ratio is increased. This implies that more movement corresponds to faster and more wide reach of a disease, i.e the travellers are effective in spreading of the disease.

In the proposed model, death rate denotes the probability of a symptomatic agent turning its health state to a deceased state at each time step, which was implemented in Figures 4 and 5 for all agents invariant over time. In reality, this might not be the case since people are not likely to die at the beginning of the symptomatic phase. Since a disease can develop in the human body, one can assume that the longer the symptoms lasts, the more likely the patient would be to die of the disease. Thereby, another more realistic way to model the death rate, would be to implement the death of individuals as an increasing function of the time that they have been infected.

Moreover, it is clear from Figure 3 that with longer incubation times, i.e. infected individuals staying exposed and not showing symptoms for a longer period of time, the disease also becomes more wide spread. This is fairly obvious but this has an even larger effect when considering the added panic response. In this case longer incubation time also mean that susceptible individuals will stay in the vicinity of exposed individuals for longer. Thus, when eventually enough individuals become symptomatic, more people will have become unknowingly exposed spreading the disease even more during the inevitable panic.

This aspect makes the threat of bio-terrorism even more dangerous as viruses used in these kinds of attacks are likely to be synthetically manufactured to have longer incubation times.
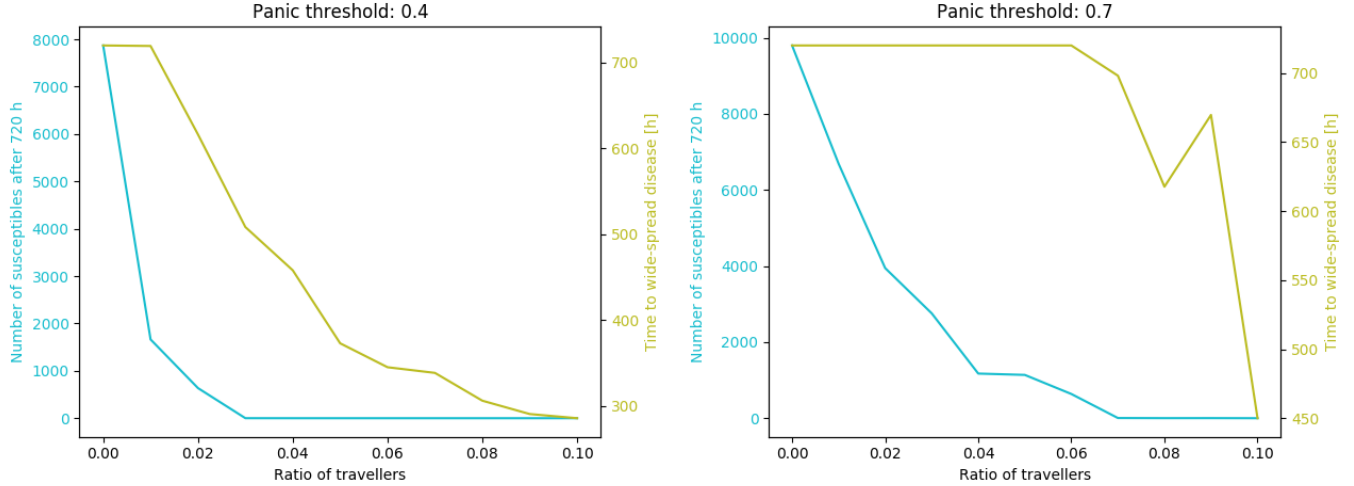
Figure 8: Measures of the speed and reach of the disease, averaged over 20 simulations, for travel ratio $\tau = 0, 0.01, 0.02, ..., 0.08, 0.09, 0.1$. Done for panic threshold 0.4, shown in the left panel, and 0.7, in the right panel. The right axis in each figure refers to the time until the simulated population had zero susceptible individuals left, i.e. the time until the disease had reach the entire population. The left axis refers to the number of susceptibles left at the end of each simulation, i.e. after 720 hours.

Drawbacks in the model

In this project the inner dynamics within one site was approximated by a simple formula which assumed that every agent was in contact with all other agents in the same site within 1 hour. This assumption is rather unrealistic and might lead to a simplified result. The population is also initialized uniformly and randomly over a 50 by 50 grid map for every experiment. If one wants to relate the model to the real world then concepts of landscape and cities would have had to be introduced. Other aspects should also taken into account, such as the fact that most people cluster around the urban regions.

Furthermore, the movement of travellers are approximated entirely by random walks. This might not be a fully realistic approximation since the movement of individuals are more directed, as it is driven by different factor such as personalities and occupations. Under the scenario of a bio-terrorism attack, one thought is that some travellers might actively move around a lot for the purpose of spreading the disease as much as possible. This behaviour could have been modelled different to ordinary travellers, for example by having them walk in such a way as to cover as much distance as possible instead.

In the model, the only form of transportation that is considered is by instantaneous travel through the 4 implemented airports. Transportation in real life is more complicated and diverse, since it is generally more common to use other form of transportation on a daily basis. To investigate the effect of transportation in the spreading of the disease, modifications should be made to the present model. Moreover, the inhabitants should have some additional movement introduced, i.e the chance of going out of their community and then come back after some time.

Future work

The following are some ideas for how this project could be expanded and interesting features that could be investigated as well as some additional purposes that the model could serve given certain extensions.

Additional added features to the developed model An interesting feature that could be added would be to make a distinction between environments such as residential areas, work and study areas, and entertainment areas with the population distributed according to reality [9]. These environments would contain different population densities, leading to the spreading of the disease having different width depending on the location of the origin of the disease.

It is also interesting to assign different attributes and behaviors to agents, i.e the general process that connects the information that agents can pick up from its environment and interactions to its decisions and actions. In order to reflect a real society, one can divide the population into different socioeconomic groups, including people with different professions as well as students and unemployed citizens [10]. In that case, the movement of an agent as a procedure (daily routine) rather than a random process could be modeled. Agents in different groups would have their corresponding routines. For example, students have to go to school on weekdays. Based

on this, the disease spreading could be simulated in a community with different infrastructures and functional areas.

Prediction model by Recurrent Neural Network
An idea for expanding the project is to train a recurrent neural network with long short-term memory units on the simulations in order to predict the disease spreading in each time step [11]. This prediction would be solely based on the location and the symptomatic status of the individuals. The purpose with this kind of training would be to be able to use the result on real-world scenarios by evaluating the possibilities of how wide spread a disease is and thereby enabling action to contain the spread.

Preventative strategies from Reinforcement learning agent OpenAI has successfully managed to perform simulations where multiple agents learn strategies in the game of hide and seek through reinforcement learning [12]. In their simulations both the hiders and the seekers managed to find advanced strategies and even abused glitches in the environment.

Artificial intelligence is known to have the capability of finding new strategies to given problems. This could be applied in the simulation in this project by introducing a few "intelligent agents" that are able to perform some form of response actions. For instance, these actions could be that the agents have limited amount of vaccines that they could inject others with, confinement tactics and the possibility to move other agents around. Furthermore, the problem of containing a disease might also give new insights into how to deal with bio-terrorism.

Reinforcement learning has a wide and similar use in different complex systems, such as fire predictions [13]. Since reinforcement learning makes decisions as the simulation is running, prior knowledge of the simulation as a base for the decisions isn't necessary. This makes it possible to react immediately to changes in a complex environment that are impossible to predict, such as spreading of a disease [14].

Conclusion

To conclude, the results of this paper can be directly tied into a scenario of bio-terrorism. During such an attack it is a natural response for people to panic and try to flee the area. It is also realistic that some proportion of a population will always travel outside their community even during an active threat of an attack or outbreak. It has shown in this paper that both these factors contribute to, and worsen the effects, of an outbreak. Thus it is crucial for societies to have policies set in place, to restrict people in the case of a crisis. However, even with for example travel restrictions in place it is often difficult to control a population. According to our result it is also easy to imagine why governments might to try to hide

information about these kinds of events from the public. All for the purpose of controlling the panic, and in doing so containing the outbreak at the expense of a few but for the good of many.

References

[1] J. M. Epstein, "Modelling to contain pandemics," *Nature*, vol. 460, no. 7256, p. 687, 2009.

[2] M. Pal, M. Tsegaye, F. Girzaw, H. Bedada, V. Godishala, and V. Kandi, "An overview on biological weapons and bioterrorism," *Am. J. Biomed. Res*, vol. 5, pp. 24–34, 2017.

[3] K. M. Carley, D. B. Fridsma, E. Casman, A. Yahja, N. Altman, L.-C. Chen, B. Kaminsky, and D. Nave, "Biowar: scalable agent-based model of bioattacks," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 36, no. 2, pp. 252–265, 2006.

[4] J. Parker and J. M. Epstein, "A distributed platform for global-scale agent-based models of disease transmission," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 22, no. 1, p. 2, 2011.

[5] E. Frias-Martinez, G. Williamson, and V. Frias-Martinez, "An agent-based model of epidemic spread using human mobility and social network information," in *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing.* IEEE, 2011, pp. 57–64.

[6] J. M. Epstein and R. Axtell, *Growing artificial societies: social science from the bottom up.* Brookings Institution Press, 1996.

[7] Y. Moreno and A. Vazquez, "Disease spreading in structured scale-free networks," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 31, no. 2, pp. 265–271, 2003.

[8] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, V. Colizza, L. Isella, C. Régis, J.-F. Pinton, N. Khanafer, W. Van den Broeck *et al.*, "Simulation of an seir infectious disease model on the dynamic contact network of conference attendees," *BMC medicine*, vol. 9, no. 1, p. 87, 2011.

[9] L. Perez and S. Dragicevic, "An agent-based approach for modeling dynamics of contagious disease spread," *International journal of health geographics*, vol. 8, no. 1, p. 50, 2009.

[10] C. Macal and M. North, "Introductory tutorial: Agent-based modeling and simulation," in *Proceedings of the Winter Simulation Conference 2014*. IEEE, 2014, pp. 6–20.

[11] C. Sathler and J. Luciano, "Predictive modeling of dengue fever epidemics: A neural network approach," 2017.

[12] B. Baker, I. Kanitscheider, T. Markov, Y. Wu, G. Powell, B. McGrew, and I. Mordatch, "Emergent tool use from multi-agent autocurricula," 2019.

[13] M. A. Wiering and M. Dorigo, "Learning to control forest fires," 1998.

[14] W. J. Probert, S. Lakkur, C. J. Fonnesbeck, K. Shea, M. C. Runge, M. J. Tildesley, and M. J. Ferrari, "Context matters: using reinforcement learning to develop human-readable, state-dependent outbreak response policies," *Philosophical Transactions of the Royal Society B*, vol. 374, no. 1776, p. 20180277, 2019.