Size Effect in Cellular Automata Based Disease Spreading Model

Julianna Orzechowska, Dawid Fordon, and Tomasz M. Gwizdałła(✉)

Faculty of Physics and Applied Informatics, University of Łódź, Pomorska 149/153, 90-236 Łódź, Poland
tomgwizd@uni.lodz.pl

Abstract. In our paper we use the, recently proposed, model for simulating the process of disease spreading in the environment defined by the Cellular Automaton. The main effort goes to the analysis of the influence of cell size on the epidemic curves and other characteristics related to the studied process. We take into account some real data concerning the occupation in the city of Łódź, which has about 700000 inhabitants. The results show that by marshaling the parameters of simulation we can obtain explicitly different results. This comment applies to a lot of features like: the shape of epidemic curve, the total number of diseased or the amount of ill in particular areas/cells.

1 Introduction

The problem of modeling of the spreading of different illnesses in the populations is an interdisciplinary issue studied for many years and interesting due to the possibility of comparison of experimental and simulation results (see e.g. [1]). Basically, the majority of the proposed approaches base on the so-called **SEIR** model [2]. Its crucial part is the set of four first order differential equations describing the "velocity" of change of number of individuals in different groups, regarding to the state of illness. Some interesting versions of the use of **SEIR** model can be found in [3] where the seasonal increase of disease strength is studied in the frame of nonlinear transmission rate, [4] where the Principal Component Analysis is applied. We can mention also one of the review papers [5].

The Cellular Automata technique is also used to study the process of disease spreading. The main effort goes usually to model the transmission of disease through the boundaries of cells. The different approaches have been proposed by e.g.: Hoya White [6] where the slightly simplified **SIR** model with additional vaccination is studied on the system of equinumerous cells, Pfeifer et al. [7] who proposed a framework for study the different scenarios for Tyrol (Austria) conditions or quite new paper by Sharma et al. [8] where special attention is paid to the incubation process.

G. Mauri et al. (Eds.): ACRI 2018, LNCS 11115, pp. 146–153, 2018.
https://doi.org/10.1007/978-3-319-99813-8_13

2 Model

The model we use is based indeed on the one presented and accurately described in [9], therefore we do not pay special attention to present a lot of its details. We rather concentrate on the presentation of some basic ideas and formulas and later on weaknesses which can strongly influence the result.

The crucial observation following basic models devoted to study epidemics processes is the division of whole population into four groups defined as:

- Susceptible (**S**) - all people who can contract the disease
- Exposed (**E**) - people in the phase of incubation of illness
- Infectious (**I**) - individuals who are capable to transmit the disease
- Recovered (**R**) - individuals recovered who are permanently immune

In the model studied in the paper we do not use differential equations which correspond to some totalistic view on the problem but we adapt the Cellular Automata related approach [9]. As usually, we have to define the tuple containing the states, topology, neighborhood and the rule (transition function).

The set of states is certainly given by by 4-tuple $\{S, E, I, R\}$. In the original paper the topology is the result of the division of some geographical area into rectangular grid. Originally, as the considered one it was the territory of Poland divided into 36 rows and 36 columns. By combining these two factors we can obtain the mapping of number of individuals onto particular cells denoted by $\{S_{ij}, E_{ij}, I_{ij}, R_{ij}\}$, where $(i, j) \in C$ and C is the cellular space. The possibility of spreading of disease not only inside the cells but also into another ones is provided by the assumption about the possible transfers of individuals between different cells. In the paper we follow the original assumption that the transfer can take place only into cells in the Moore's neighborhood.

The transition function is based on some additional assumptions. When considering the particular disease we have to know the specific numbers defining the period when individual is in the exposed and infectious state. Following the original paper we denote then as a and b respectively and assume $a = 2$ and $b = 4$ [10] what corresponds to some statistical results of Infectious Period Distribution. The crucial problem is here the passage between states **S** and **E**, it means simply the chance to become ill. The probability of this process is given by the Eq. 1.

$$p_{ij}^t = \begin{cases} 0, & q_{ij}^t < 0 \\ 1, & q_{ij}^t > 1 \\ q_{ij}^t, & \text{elsewhere} \end{cases} \tag{1}$$

where the q_{ij}^t number is defined by:

$$q_{ij}^t = rnd\Big(1 - exp\Big(-\beta \frac{\sum_{k=1}^{b} l_{ij|k}^t + \sum_{(x,y)\in C} \sum_{k=1}^{b} I_{ij|k}^t (I_{x,y\to i,j|k}^t - I_{i,j\to x,y|k}^t)}{N_{ij}^t + \sum_{(x,y)\in C} (N_{x,y\to i,j}^t - N_{i,j\to x,y}^t)}\Big), c_v\Big) \tag{2}$$

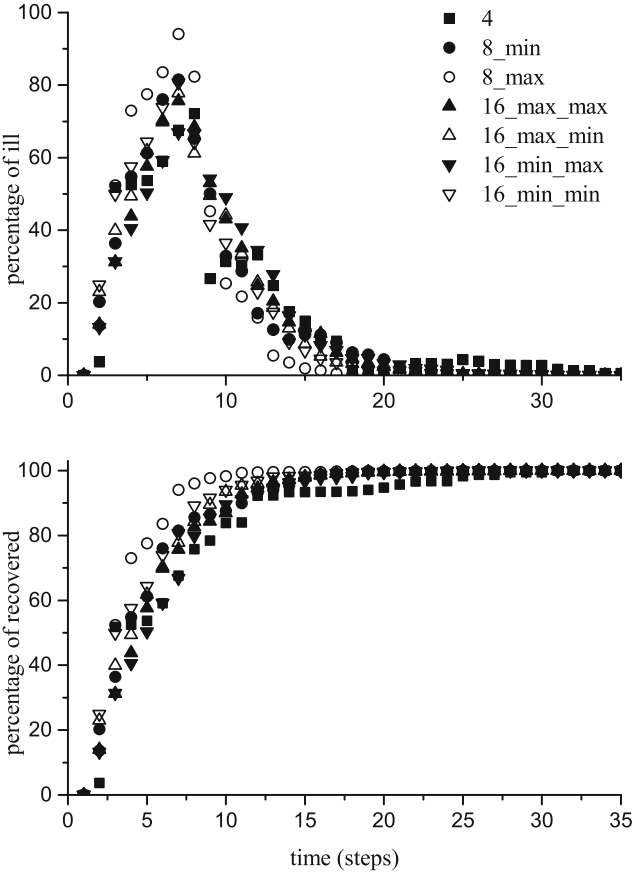

Fig. 1. Epidemic curves for parameters used in Holko [9] paper ($\beta = 0.6$, c_v = 0.5)

In the above equation *rnd* means sampling from the Gaussian distribution, β is called contact rate, pairs (i, j) and (x, y) describe he coordinates of cells on the rectangular grid, the vertical line in the subscript referring to the index of summation corresponds to the index of day, among the b in the infectious state, the individual stays in this state and t is the time index.

The model presented in [9] has several interesting features which need some discussion. We have e.g. to mention the large size of single cells (with edge of about 15 km) what leads to the fact that there are large differences between the occupation of particular cells. Also the division ratio of the studied area can be the property under consideration.

We select the smaller area, particularly the urban area of Łódź which shape is close to the square one. We are then able to easily find the amount of individuals in every cell. In order to do this we use the number of voters in the general election announced by the city council. Since this number is announced for irreg-

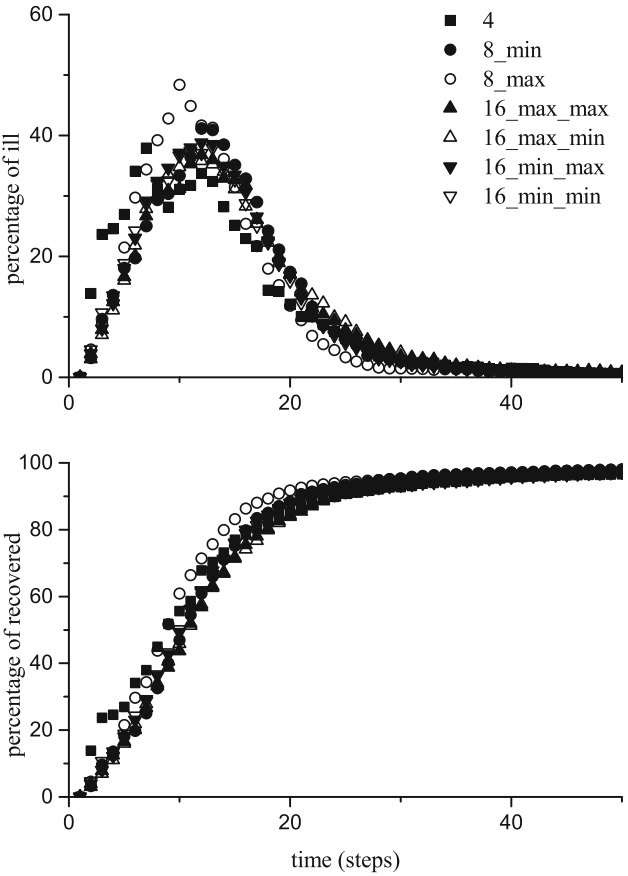

Fig. 2. Epidemic curves for lower c_v (c_v = 0.01)

ular parts of city called subdivision, by assuming that a distribution of people in every subdivision is uniform, we add to the given cell such a number of people which follows the surface percentage contribution of this cell in subdivision. Since the voters numbers is determined only for adults, the resulting numbers are then normalized in order to have in total 700 000 individuals. By using such an approximation we obtain for example from 1839 up to 106000 individuals in every cell of 4×4 grid.

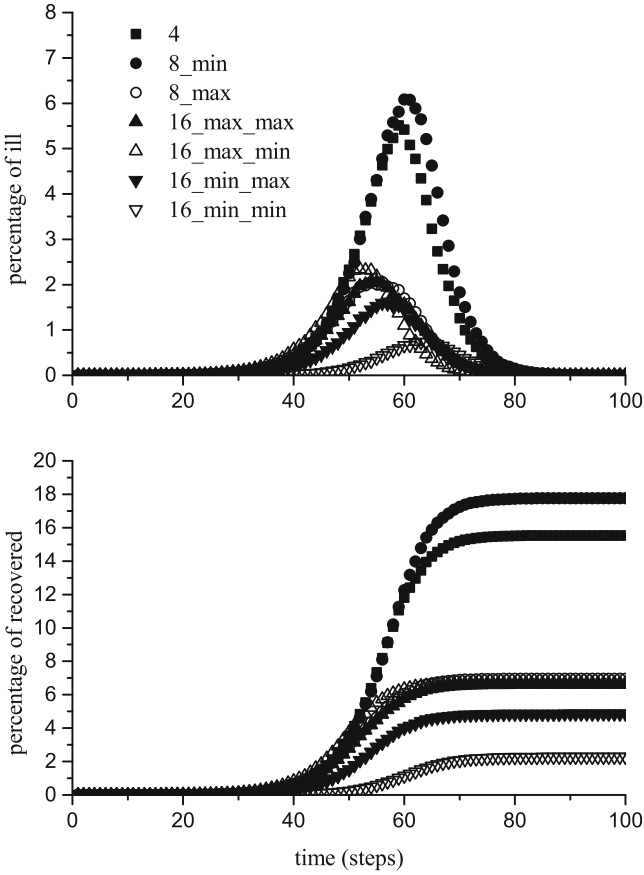

Fig. 3. Epidemic curves for variant c_v

3 Results and Conclusions

We decided to initialize the simulation with just one individual in the exposed state. The additional question we can try to answer is whether the location of "patient zero" influences the process of disease spread. These analysis we perform for more dense cells division. We decided to allocate "patient zero" in the cell $(1, 1)$ in the 4×4 grid, where the cell $(0, 0)$ is the westernmost and northernmost cell of grid. It means that our initial cell belongs to the group of cells with the relatively middle occupation however adjoining with the cells with low occupation as well as with the cell with the highest occupation in a system. When further dividing the cells into four smaller ones we consider the difference initializing the disease once in the cell with highest occupation among the subcells of cell $(1, 1)$ then in the cell with the lowest occupation among them. In the pictures they are denoted as "8_max" and "8_min" respectively. The same procedure is then adapted for the divided cell in 8×8 grid, so "16_max_max"

cell is the one with the highest occupation among four subcells of "8_max" while "16_max_min" is the one with lowest number of individuals among them. The same standard of notation is used for description of subcells of "8_min" cell.

Our first attempts are made for the parameters proposed in the original paper. Since the calculations were made for the set of values $\beta \in \{0.2, 0.4, 0.6\}$ we decided to use one of these pretested values and set $\beta = 0.6$. The selection of c_v value is also the effect of earlier suggestion and we choose the middle of the values used in the original paper and set $c_v = 0.5$. The results are presented in Fig. 1. It can be observed that there is no visible differences in the results of simulations performed for different sizes of grids. However, as opposed to the authors of original paper we present the epidemic curves on the percentage scale and not on the absolute one. This allow us to emphasize the fact that by using the model with given parameters we obtain the number of ill individuals encompassing whole population. It is well known that the typical local epidemy causes the illness of about 2–4% of population. Also the data concerning the most famous pandemy of 20th century [11–13] are not unambiguous. They show that the rate of ill was in the interval 10–30% and the mortality among the ill individuals was on the level of 10–30%. This leads to the conclusion that the numbers produced by the model are too high to describe the real case.

The first step we make is to decrease the c_v parameter. In the second attempt we use $c_v = 0.01$ which is close to the proposed $c_v = 0$ but introduces some dispersion of values. The Fig. 2 shows that this change does not lead to any substantial change. Certainly, the maximum of curves is shifted from about seventh to about tenth-twelfth day and the number of ill in the maximum declines from the totally unreal number of almost 100% to about 40–50% but the plot still do not correspond to the mentioned above features.

Finally we introduce the fundamental change into the model. We make c_v dependent on the average calculated as the first parameter of formula 2. The deviation is determined by simply multiplying the average by 0.1. Our idea is that the distribution can be ever wider when the number of ill is higher. We call ths case as "variant c_v" and present in Fig. 3. It can be observed that this change influences the results very strongly. We obtain the differences in the shape of epidemic curves as well the variable total number of ill. The crucial observation is that the decrease of cell size seems to decrease the epidemic curve. The points for 4×4 and 8×8 grids use to lay higher than for 16×16. The very interesting effect is the one that the height of epidemic curve peak is higher when starting the epidemy from the cell with lower occupation. This effect has to be explained by presenting the detailed curves for particular cells.

Some view on the run of simulation in particular groups is shown in Fig. 4. In the Figure, the number of people in particular phase of disease as well as the percentage in particular cell is presented for every cell and for the 7th day of simulation. The organization of rows and columns corresponds strictly to the organization of simulation and the cells number is 4×4. So, the upper row describes the northernmost part of the city and so on.

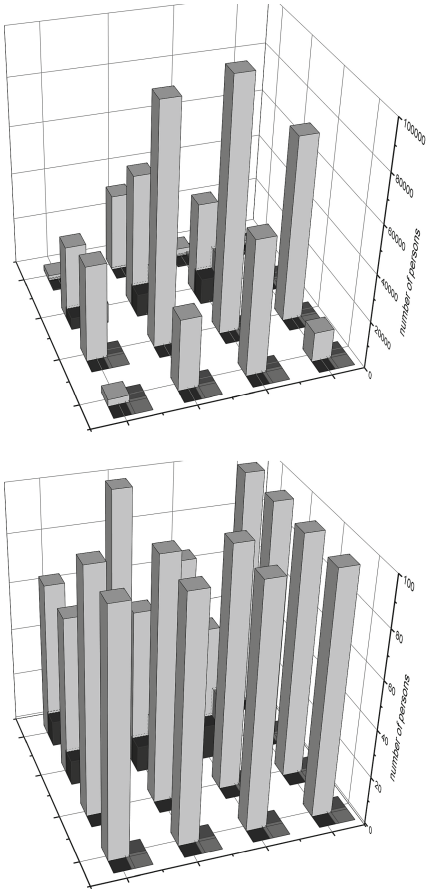

Fig. 4. The number of individuals in particular groups and cells in the same time - the 7th day. From up to down $c_v = 0.5$, $c_v = 0.01$, variant c_v From the lightest shade of grey: (S, R, E, I).

Only the results for variant c_v are presented. The results for absolute numbers shows clearly the high contribution of people in the **S** state for every cell in the simulation system. Especially when we compare it with the expected high number of individuals in the **E** and **I** states following the Figs. 1 and 2 for constant and larger c_v. Some more interesting information can be observed in the lower plot of Fig. 4 which shows the results for percentages. There are the lowest bars corresponding to the susceptible state in the northern part of the city, so the majority of ill concentrates in this region. Keeping in mind that a disease starts from the cell $(1, 1)$ so the cell second from up and second from left we can say that starting the illness in the sparsely populated part of the city we can rather easily limit it to rather confined area.

The presented results confirm that when using the Cellular Automata based model of disease spreading, even with the relatively simple mechanism of totalistic rules, we can generate different results corresponding to real process of the studied phenomena. In our opinion the next steps should be directed into several points, like: the further densifying of CA grid, the more realistic rules of disease transfer or the individualization of contacts between people what brings us closer to the mixed CA and agent oriented approach.

References

1. Guo, D., Li, K.C., Peters, T.R., Snively, B.M., Poehling, K.A., Zhou, X.: Multiscale modeling for the transmission of influenza and the evaluation of interventions toward it. Sci. Rep. **5**, 8980 (2015)
2. Aron, J.L., Schwartz, I.B.: Seasonality and period-doubling bifurcations in an epidemic model. J. Theor. Biol. **110**, 665–679 (1984)
3. Yi, N., Zhang, Q., Mao, K., Yang, D., Li, Q.: Analysis and control of an seir epidemic system with nonlinear transmission rate. Math. Comput. Model. **50**, 1498–1513 (2009)
4. Schimit, P., Pereira, F.: Disease spreading in complex networks: a numerical study with principal component analysis. Expert Syst. Appl. **97**, 41–50 (2018)
5. Keeling, M., Eames, K.T.D.: Networks and epidemic models. J. Roy. Soc. Interface **2**, 295–307 (2005)
6. Hoya White, S., Martin del Rey, A., Rodriguez Sanchez, G.: Modeling epidemics using cellular automata. Appl. Math. Comput. **186**, 193–202 (2007)
7. Pfeifer, B., et al.: A cellular automaton framework for infectious disease spread simulation. Open Med. Inform. J. **2**, 70–81 (2008)
8. Sharma, N., Gupta, A.K.: Impact of time delay on the dynamics of SEIR epidemic model using cellular automata. Phys. A: Stat. Mech. Appl. **471**, 114–125 (2017)
9. Holko, A., Medrek, M., Pastuszak, Z., Phusavat, K.: Epidemiological modeling with a population density map-based cellular automata simulation system. Expert Syst. Appl. **48**, 1–8 (2016)
10. Lloyd, A.L.: Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. Theor. Popul. Biol. **60**, 59–71 (2001)
11. Cartwright, F.F., Biddiss, M.D.: Disease and History, 2nd edn. Sutton Publishing, Stroud (2000)
12. Johnson, N., Mueller, J.: Updating the accounts: global mortality of the 1918–1920 spanish influenza pandemic. Bull. Hist. Med. **76**, 105–115 (2002)
13. Knobler, S., Mack, A., Mahmoud, A. (eds.): The Threat of Pandemic Influenza: Are We Ready? Workshop Summary. Institute of Medicine (US) Forum on Microbial Threats, National Academies Press (US) (2005)