A Comparative Analysis of CA Model and ODE Model for SIR-type Epidemics

A MATLAB – based Solution to ARGESIM Bench-mark C17 'Spatial Dynamics of SIR-Type Epidemic'

Anita Gerstenmayer[1*], Florian Miksch[2]

[1]inst. of Analysis and Scientific Computing, Vienna University of Technology,
Wiedner Haupstrasse 8-10, 1040 Vienna, Austria; *Anita.Gerstenmayer@tuwien.ac.at
[2]dwh Simulation Services, Neustiftgasse 57-59, 1070 Vienna, Austria

Simulation Notes Europe SNE 23(2), 2013, 107 - 110
DOI: 10.11128/sne.23.bn17.10193
Received: February 20, 2013; Revised: May 25, 2013;
Accepted: June 30, 2013;

Abstract. This paper presents an extended and modified approach and solution to ARGESIM benchmark C17 'Spatial Dynamics of SIR-type Epidemic, which is comparing modelling and simulation approach with ODEs (McKendrick's SIR model) and with cellular automata (LCGA - Lattice Gas Cellular Automaton). Model implementations (ODE and CA) are directly programmed in MATLAB, using the MATLAB ODE solvers, and programming the CA update by means of vector and matrix manipulations. The contribution analyses and documents the differences between ODE solutions and aggregated CA solutions, mainly investigating dependencies on initial values. And on the other hand , the contributions presents 'similarities' between ODE solutions and CA solutions by balancing spatial inhomogenities in the CA dynamics using equally distributed populations.

1 Modeling

1.1 ODE model

The differential equations model is based on Kermack and McKendrick's SIR model and consists of three equations for the numbers of susceptible (S), infected (I) and recovered (R) individuals as functions of the time t:

$$S'(t) = -\gamma S(t)I(t)$$
$$I'(t) = \gamma S(t)I(t) - \delta R(t)$$
$$R'(t) = \delta R(t)$$

The infection rate γ can be expressed by $\gamma = \alpha \frac{c}{N-1}$ as stated in Benchmark C17. The recovery rate δ can simply be expressed by the provided parameter β. This model was simulated using Matlab's numerical ODE-solver `ode45`, which is based on an explicit Runge-Kutta (4,5) method.

1.2 LGCA model

The LGCA was implemented in Matlab using a three dimensional $n \times n \times 6$ array for the cellular automaton (n^2 cells with 6 sub cells each). Each array entry takes one of the four following values representing its state: 0 (empty), 1 (susceptible), 2 (infected) or 3 (recovered).

For the initialization the given initial numbers of individuals are distributed randomly across the LGCA. In each time step, the three phases of the update rules are carried out. For the infections, recoveries and collisions a loop among all cells is made. First, the number of infected individuals in the cells is calculated, as it is necessary for computing the infection probability. Then the recoveries and infections are executed using the probabilities α_c, for infection of a susceptible particle from an infected particle and β_c for recovery of an infected particle. Next it is checked whether the cell is in one of the possible configurations for collisions and the reflections of the particles take place.

The movement phase was implemented according to the FHP-I rules, particles on the boundary reenter the LGCA on the opposite side. For each time step, the number of individuals in the three states is calculated as a result. The model parameters α_c, β_c and n are identified with provided system parameters according to the specifications of Benchmark C17.

2 Differences in ODE and CA Solutions

The aim of this first investigation is to compare the simulations obtained by the two modeling approaches varying the parameters I_0 (and correspondingly S_0), α and β to find settings for similar and different results and analytical explanations for these findings. The other three parameters, the population N = 100 000 as well as the initially recovered persons $R_0 = 0$ and the number of contacts C, are kept constant.

The results for both models in the first parameter setting, in which the initial number of infected people is rather high and α and β quite low, are nearly indistinguishable, as can be seen in Figure 1.

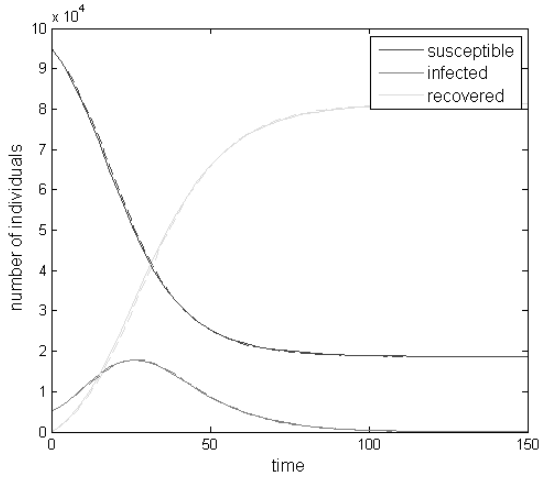

Figure 1: $S(t)$, $I(t)$ and $R(t)$ simulated with differential equations (solid line) and LGCA (dashed line) for $I_0 = 5000$, $\alpha = 0.1$ and $\beta = 0.1$.

In the second setting (Figure 2), where only I_0 is diminished by a factor 100, the outbreak of the epidemic is delayed by approximately 50 time steps in comparison to the first setting. Qualitatively the behavior of both models is still the same, but the LGCA epidemic occurs slightly later with a lower maximum of infected individuals.

The increase of α and β in the third setting (Figure 3) leads to a faster outbreak and shorter duration of the epidemic. There are considerable differences between the results for both models. The maximal value of I is about four times as high for the differential equations model than for the LGCA and the duration of the epidemic is about twice as long for the LGCA.

Also, with the LGCA model a lot more individuals stay susceptible for all time.

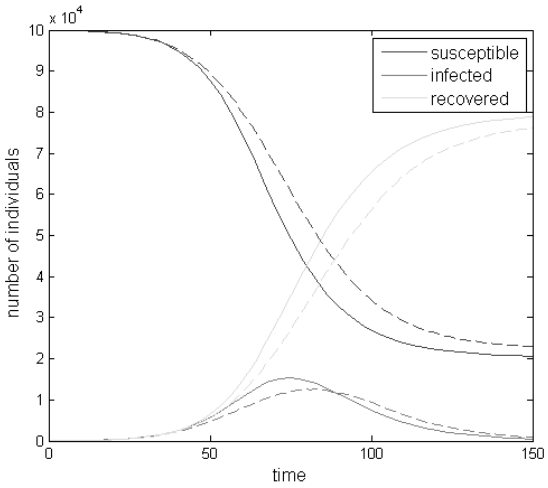

Figure 2: $S(t)$, $I(t)$ and $R(t)$ simulated with differential equations (solid line) and LGCA (dashed line) for $I_0 = 50$, $\alpha = 0.1$ and $\beta = 0.1$.

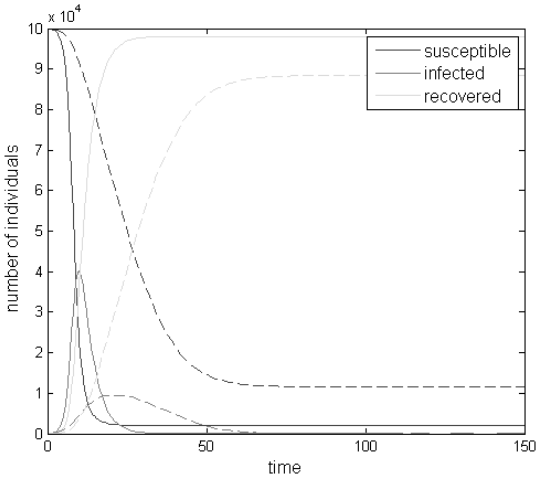

Figure 3: $S(t)$, $I(t)$ and $R(t)$ simulated with differential equations (solid line) and LGCA (dashed line) for $I_0 = 50$, $\alpha = 0.6$ and $\beta = 0.3$.

In the last setting (Figure 4) I_0 is set to the very low number of only five individuals. For the differential equations model this just results in a small time delay compared to the third simulation. The LGCA model on the other hand responds quite differently, the curve of infected individuals is so flat that an outbreak of the epidemic is hardly visible and R and S increase and decrease very slowly.

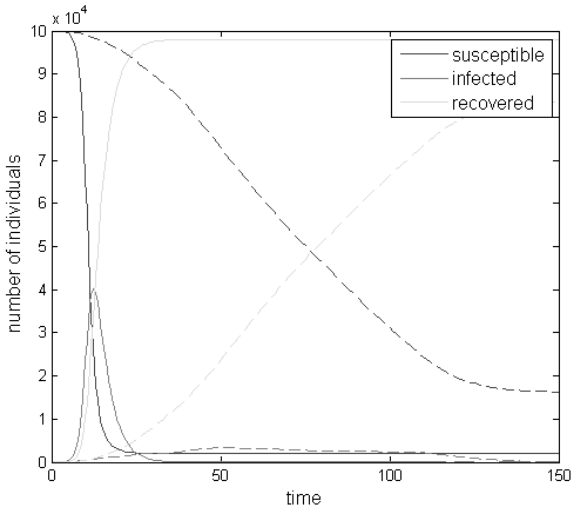

Figure 4: $S(t)$, $I(t)$ and $R(t)$ simulated with differential equations (solid line) and LGCA (dashed line) for $I_0 = 5$, $\alpha = 0.6$ and $\beta = 0.3$.

One conclusion that can be drawn from these results is that the LGCA reacts much more sensitively to changes of the initial values I_0 and S_0 than the continuous model. For the differential equations model the changed initial values merely seem to shift the outbreak of the epidemic to the left or right, the shape of the curves hardly changes. This does not hold for the LGCA. The smaller the initial number of infected individuals, the flatter the graph for I becomes due to the spatial distribution of the individuals. If just a small number of infected individuals is placed in the LGCA, the propagation of the disease takes very long because only the few individuals in the neighborhood of the infected ones can get infected, but as the speed of recovery is not affected, I is kept down and reaches a lower maximum.

So the results for high values of I_0 are similar in both models, but for low initial values of infections the LGCA epidemic curve is a lot flatter. This shows that in some cases the LGCA does not provide a sufficiently homogeneous mixing of infected and susceptible particles, and hence the simulation results differ from the ODE.

Also the infection parameter α can cause different behavior for the two models. For low values of α both models behave alike. This can be explained analytically by calculating the expected number of new infections per time unit in the LGCA.

The probability of an infection is given by $\sum_{i=0}^{5} q_i(1-(1-\alpha)^i)$, where q_i is the probability of i places in the cell being occupied by infected individuals (given by a hypergeometric distribution), and the probability of an infection of a susceptible individual in this cell is $(1-(1-\alpha)^i)$. For values of α close to zero it is reasonable to approximate this term with the Taylor expansion of first order. This yields

$$\sum_{i=0}^{5} q_i\,(1-(1-\alpha)^i) \approx \alpha \sum_{i=0}^{5} q_i i = \alpha I \frac{C}{N-1} \qquad (1)$$

Multiplied with the number of susceptible individuals we have that the expected number of infections equals $\alpha SI\frac{C}{N-1}$, which is consistent with the differential equations. For higher values of α the Taylor approximation is not sufficiently accurate, it overestimates the infection probability $(1-(1-\alpha)^i)$ and thus the actual number of infections per time step in the LGCA is much lower than for the simulation with differential equations.

The recovery parameter β has the same influence on both models, as the expected number of infected individuals in the LGCA who recover per time step equals βI which corresponds to the differential equations. Also the recoveries are not dependent in any way on the spatial distribution. Higher values of β lead to smaller epidemic outbreaks in both models.

3 Balancing Spatial Inhomogenities in CA Model

In this investigation a modified version of the LGCA model, in which all individuals are distributed randomly across the LGCA after each time step, was implemented and compared to the previous results.

For the implementation, the movement and collision parts of the LGCA code were simply replaced by a random permutation of all array entries. This was achieved by reshaping the three-dimensional array into a vector, using Matlab's random permutation function `randperm` to permute its entries, and then converting the vector back into an $n \times n \times 6$ array.

All three models were simulated with the parameter values $S_0 = 99\,900, I_0 = 100, R_0 = 0$ and $\beta = 0.5$.

This time α and C were varied, for the first simulation $\alpha = 0.075$ and $C = 4$ were used, for the second $\alpha = 0.3$ and $C = 1$. The results are shown in figure 5.

It can be observed that the differential equations model yields the exact same result for both simulations. This is due to the fact that only the product of C and α is relevant for this model in form of the infection rate

$$\gamma = \alpha \frac{C}{N-1}.$$

As this product equals 0.3 in both settings, there cannot be any difference in the simulation results of this deterministic model.

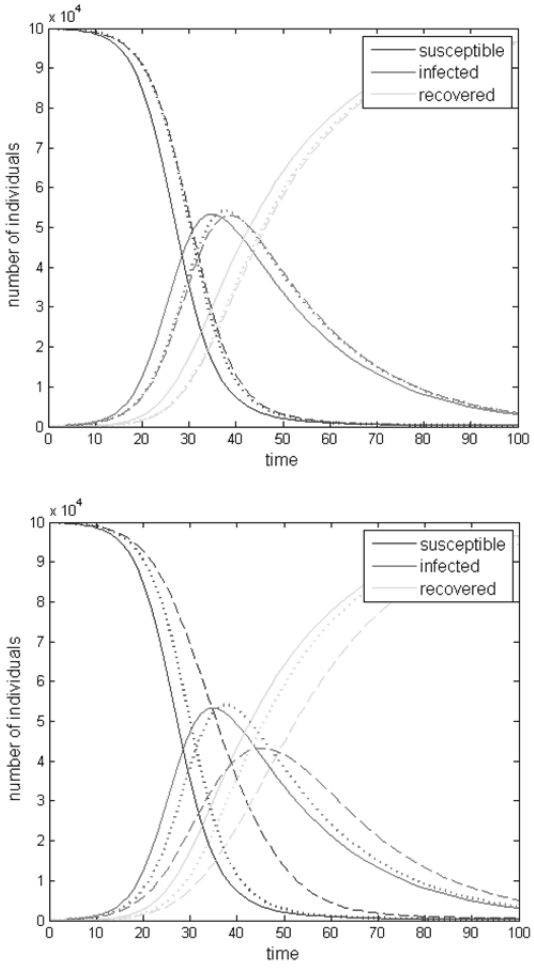

Figure 5: S(t), I(t) and R(t) simulated with differential equations (solid line), LGCA (dashed line) and LGCA with random distribution (dotted line) for C=4 and α=0.075 (upper panel) and for C=1 and α=0.3 (lower panel).

In the first setting, both LGCA versions lead to very similar results, and also match the differential equation simulation rather closely, with approximately the same maximal number of infected individuals. A small delay in the outbreak of the epidemic is probably caused by the relatively small number of initially infected individuals $I_0 = 100$ in a population of $N = 100\,000$.

The second simulation gives more interesting results. As already observed in the previous task, the higher value for α leads to a quite different behavior of the LGCA. Here the epidemic occurs later and with a lower maximum of infected individuals than obtained with the differential equations model. The results for the modified LGCA however have not changed with the parameter α. The random placement of all individuals after each time step riddens the model of the spacial inhomogeneities that would otherwise form and delay the outbreak of the epidemic.

4 Conclusions

It can be concluded from the previous tasks and simulations that for some situations, especially those with a low infection rate and a high number of infected individuals at time zero, both models behave very much alike. In some cases however, the locality of the LGCA leads to spatial groupings of the infected individuals and thus slows down the speed of the epidemic in comparison to the differential equations model. This occurs for high values of α and low values of I_0. By distributing all individuals randomly across the LGCA after each time step the influence of this locality can be prevented, but still a difference remains between the discrete calculations and the continuous model.

Model sources

ODE models are implemented with MATLAB ODE solvers, and the CA propagation is directly programmed in MATLAB using vector and matrix feature. All MATLAB m-files and a short file documentation can be downloaded (zip format) by EUROSIM societies' members from SNE website, or are availably from the author.

References

[1] Hötzendorfer H, Popper N, Breitenecker F. Temporal and Spatial Evolution of SIR-type Dynamics – ARGES-IM Comparison C17 – Definition. *SNE Simulation News Europe*. 2004; 14(2): 42-44.