

# A SIS Epidemiological Model Based on Cellular Automata on Graphs

M.J. Fresnadillo<sup>1</sup>, E. García<sup>1</sup>, J.E. García<sup>1</sup>, A. Martín<sup>2</sup>, and G. Rodríguez<sup>3</sup>

<sup>1</sup> Dpto. Medicina Preventiva, Salud Pública y Microbiología Médica  
Facultad de Medicina, Universidad de Salamanca  
C/Alfonso X "El Sabio" sn, 37007-Salamanca, Spain

{jofrema, engarsan, joegas}@usal.es

<sup>2</sup> Department of Applied Mathematics  
E.P.S. de Ávila, Universidad de Salamanca  
C/Hornos Caleros 50, 05003-Ávila, Spain  
delrey@usal.es

<sup>3</sup> Department of Applied Mathematics  
E.P.S. de Zamora, Universidad de Salamanca  
Avda. Cardenal Cisneros 34, 49022-Zamora, Spain  
gerardo@usal.es

**Abstract.** The main goal of this work is to introduce a new SIS epidemic model based on a particular type of finite state machines called cellular automata on graphs. The state of each cell stands for the fraction of the susceptible and infected individuals of the cell at a particular time step and the evolution of these classes is given in terms of a local transition function.

**Keywords:** Epidemiological model, cellular automata, SIS model.

## 1 Introduction

The public health issues have a lot of importance in our society, particularly viral spread through populated areas. Epidemics refer to a disease that spreads rapidly and extensively by infection and affecting many individuals in an area at the same time. In this way, the most recent worrying epidemic was the Severe Acute Respiratory Syndrome (SARS) outbreak in Asia. Infectious disease accounts for 29 of 96 major causes of human morbidity and mortality listed by the World Health Organization and the World Bank, and 25% of global deaths (over 14 million deaths annually). Consequently, since the publication of the first modern mathematical epidemic models in the first years of XX century (see [6,9]), several mathematical models to study the dynamics of epidemics have been appeared in the literature.

Traditionally, mathematical models are based on differential equations. Nevertheless, this approach has some drawbacks since they do not take into account spatial factors such as population density, they neglect the local character of the spreading process, they do not include variable susceptibility of individuals, etc.

As a consequence, this can lead to very unrealistic results, such as, for example, endemic patterns relying on very small densities of individuals, which are called “atto-foxes” or “nano-hawks” (see [8]). Other mathematical models are based on a particular type of discrete dynamical systems called cellular automata (see, for example [2,5,7,10,11]). These simple models of computation eliminate the last mentioned shortcomings, and are specially suitable for computer simulations. Roughly speaking, cellular automata (CA for short) are a special type of finite state machines capable to simulate physical, biological or environmental complex phenomena. Consequently, several models based on such mathematical objects have been proposed to simulate growth processes, reaction-diffusion systems, self-reproduction models, epidemic models, forest fire spreading, image processing algorithms, cryptographic protocols, etc. (see, for example, [12,13]). Specifically, a two-dimensional CA is formed by a two-dimensional array of identical objects called cells which can be disposed in a rectangular, triangular or an hexagonal lattice (called cellular space). These cells are endowed with a state that changes in discrete steps of time according to a specific rule. As the CA evolves, the updated function (whose variables are the states of the neighbor cells) determines how local interactions can influence the global behaviour of the system.

Usually, mathematical models to study epidemic spreading are divided into three types: SIS models, SIR models and SEIR models, depending on the classes in which the population can be classified. The model introduced in this paper deals with SIS epidemic diseases (for example the group of those responsible for the common cold), that is, the population is divided into susceptible individuals ( $S$ ) and infected individuals ( $I$ ). The susceptible individuals are those capable to contracting the disease whereas the infected individuals are those capable of spreading the disease. For a SIS model, infected individuals return to the susceptible class on recovery because the disease confers no immunity against reinfection. Moreover, some assumptions will be common to all models: (1) The disease is transmitted by contact between an infected individual and a susceptible individual; (2) There is no latent period for the disease, hence the disease is transmitted instantaneously upon contact; (3) All susceptible individuals are equally susceptible and all infected individuals are equally infectious; (4) The population under consideration is fixed in size. This means that no births or migration occurs, and no deaths are taken into account.

The main goal of this work is to introduce a new SIS model to simulate the spread of a general epidemic based on cellular automata on graph. Specifically, in the proposed model, the state of each cell stands for the fraction of the susceptible and infected individuals of the cell at a particular time step. The local transition function is a function involving the states of the neighbor cells and other parameters such as the virulence of the epidemic, the rate of recovered infected individuals, etc. Moreover, as is mentioned above, the standard paradigm for cellular automata states that the topology of the cellular space is given in terms of a regular rectangular or hexagonal lattices. Nevertheless, in this paper we will consider a more efficient topology to model an epidemic disease, which is given by an undirected graph where its nodes stand for the cells of the cellular automata.

There are several CA-based algorithms to simulate a SIS epidemic model (see, for example [1,3,4]). The standard paradigm of these models states that each cell stands for an only one individual. Unfortunately, there are few models considering more than one individual in each cell (see for example [5]). We think that this new paradigm is more accurate than the other one in order to obtain more realistic simulations. The main advantage of the model presented in this paper over the model introduced in [5] is the use of graph tology and a more realistic transition function involving new parameters as the portion of susceptible individuals that moves from one cell to another one.

The rest of the paper is organized as follows: In Section 2 the basic theory about cellular automata on graphs is provided; the proposed model is introduced in Section 3; the analysis of the model is shown in Section 4; and, finally, the conclusions and the future work are presented in Section 5.

## 2 Basic Theory of Cellular Automata on Graphs

A graph  $G$  is a pair  $(V, E)$  where  $V = \{v_1, v_2, \dots, v_n\}$  is a ordered non-empty finite set of elements called nodes (or vertices), and  $E$  is a finite family of pairs of elements of  $V$  called edges. Two nodes of the graph,  $v_i, v_j \in V$ , are said to be adjacent (or neighbors) if there exists an edge in  $E$  of the form  $(v_i, v_j)$ . We consider undirected graphs, that is,  $(v_i, v_j) = (v_j, v_i) \in E$ . A graph  $G$  is called simple if there is not two edges of  $G$  with the same ends and no loops exist, *i.e.* edges whose start and end is located at the same node.

The neighborhood of a node  $v \in V$ ,  $N_v$ , is the set of all nodes of  $G$  which are adjacent to  $v$ , that is:  $N_v = \{u \in V \text{ such that } (v, u) \in E\}$ . The degree of a node  $v$ ,  $d_v$ , is the number of its neighbors.

A cellular automaton on an undirected graph  $G = (V, E)$  is a 4-uple  $\mathcal{A} = (V, S, N, f)$  where: The set  $V$  defines the cellular space of the CA such that each node stands for a cell the cellular automaton.  $S$  is the finite set of states that can be assumed by the nodes at each step of time. The state of the node  $v$  at time step  $t$  is denoted by  $s_v^t \in S$ . These states change accordingly to the local transition function  $f$ .  $N$  is the neighborhood function which assigns to each node its neighborhood, that is:

$$N: V \rightarrow 2^V, v_i \mapsto N(v_i) = N_{v_i} = \{v_{i_1}, v_{i_2}, \dots, v_{i_{d_v}}\}$$

Note that the neighborhoods of the nodes are, in general, different from others. The local transition function  $f$  calculates the state of every node at a particular time step  $t + 1$  from the states of the its neighbors at the previous time step  $t$ , that is,  $s_v^{t+1} = f\left(s_{v_{i_1}}^t, s_{v_{i_2}}^t, \dots, s_{v_{i_{d_v}}}^t\right) \in S$ , where  $N_v = \{v_{i_1}, v_{i_2}, \dots, v_{i_{d_v}}\}$ .

## 3 The SIS Mathematical Model

In the mathematical epidemiological model introduced in this paper, the population is divided into two classes: those who are susceptible to the disease and

those who are infected to the disease. Moreover, the population is located at city centres which stand for the nodes of a graph  $G$ . If there is some type of transport connection (by car, train, airplane, etc.) between two of these cities, the associated nodes are connected by an edge. The following assumptions are also made:

1. The population of each node remains constant over time, that is, no births or deaths are taking into account (it is a SIS model without vital dynamics). Moreover, the population distribution is inhomogeneous: Let  $P_u$  be the number of individuals of the node  $u \in V$ , and set  $P = \max \{P_u, u \in V\}$ .
2. The transmission of the disease (that is, the passing of the disease from an infected individual to a susceptible individual) is through direct physical contact: touching an infected person, including sexual contact.
3. The population (susceptible and infected people) are able to move from its node to another one and return to the origin node at every step of time.

Since the model introduced in this work is a SIS model, then the state of the node  $u \in V$  at time step  $t$  is:  $s_u^t = (S_u^t, I_u^t) \in Q \times Q = S$ , where  $S_u^t \in [0, 1]$  stands for the fraction of susceptible individuals of the node  $u$  at time  $t$ , and  $I_u^t \in [0, 1]$  stands for the fraction of infected individuals of the node  $u$  at time  $t$ . Consequently, the transition function of the CA is as follows:

$$\begin{aligned}
 s_u^t &= f \left( s_{v_1}^{t-1}, \dots, s_{v_{d_u}}^{t-1} \right) = (S_u^t, I_u^t) \\
 &= \left( (d \circ f_S) \left( s_{v_1}^{t-1}, \dots, s_{v_{\gamma_u}}^{t-1} \right), (d \circ f_I) \left( s_{v_1}^{t-1}, \dots, s_{v_{\gamma_u}}^{t-1} \right) \right)
 \end{aligned}$$

where  $d$  is a suitable discretization function.

The ground where the epidemic is spreading is modeled as a weighted graph where each node stands for a city or a town, and the arc between two nodes represents the connection between the corresponding cities. In this sense, the connection factor between the nodes  $u$  and  $v$  is the weight associated to the arc  $(u, v) \in E$  and it is denoted by  $w_{uv}$ . It depends on the transportation capacity of the public and non-public transport. Consequently

$$w_{uv} = \frac{h_{uv}}{\max \{h_{xy}, \forall x, y \in V\}} \in [0, 1],$$

where  $h_{uv}$  is the total amount of population which move from  $u$  to  $v$  during a time step.

The evolution of the number of infected individuals of the node  $u \in V$  is as follows: The infected individuals of  $u$  at time step  $t$  is given by the sum of:

1. The infected individuals at the previous time step which have not been recovered from the disease.
2. The susceptible individuals which have been infected during the time step. In this case we have to take into account the recovery rate  $r \in [0, 1]$ . These new sick individuals of  $u$  can be infected both by the infected individuals of  $u$  or

by the infected individuals of the neighbor nodes of  $u$  which have moved to  $u$  during the time step. In the first case, only the rate of transmission,  $p \in [0, 1]$ , is involved, whereas in the second case we have to consider the connection factors between the nodes, and the population and movement factor of each node. Moreover we also consider the susceptible individuals of  $u$  moved to a neighbor node during the step of time and infected in this neighbor node by its corresponding infected individuals; in this case  $\eta_u \in [0, 1]$  yields the portion of moved susceptible individuals from  $u$  to its neighbor nodes.

Then, the mean-field equation for infected individuals is the following:

$$\begin{aligned}
 f_I \left( s_{v_1}^{t-1}, \dots, s_{v_{\gamma_u}}^{t-1} \right) &= (1 - r) I_u^{t-1} + p(1 - \eta_u) S_u^{t-1} I_u^{t-1} \\
 &+ p(1 - \eta_u) S_u^{t-1} \sum_{v \in N_u} \frac{P_v}{P} w_{vu} I_v^{t-1} \\
 &+ p\eta_u S_u^{t-1} \sum_{v \in N_u} (1 - w_{vu}) I_v^{t-1}. \tag{1}
 \end{aligned}$$

On the other hand, the susceptible individuals of each node is given by the difference of the susceptible individuals of the node at the previous time step and the susceptible individuals which have been infected as is mentioned above. Note that, as a simple calculus shows:  $I_u^t + S_u^t = I_u^{t-1} + S_u^{t-1} = P_u$ , and consequently equation the equation which rules the evolution of the susceptible is as follows:

$$f_S \left( s_{v_1}^{t-1}, \dots, s_{v_{\gamma_u}}^{t-1} \right) = 1 - I_u^t.$$

Moreover, since  $f_S \left( s_{v_1}^{t-1}, \dots, s_{v_{\gamma_u}}^{t-1} \right) \in [0, 1]$  and  $f_I \left( s_{v_1}^{t-1}, \dots, s_{v_{\gamma_u}}^{t-1} \right) \in [0, 1]$ , then a discretization function  $d: [0, 1] \rightarrow Q$  must be used in order to get a finite state set. In our case, the discretization function used is the following:

$$d: [0, 1] \longrightarrow Q, x \mapsto d(x) = \frac{[100 \cdot x]}{100}$$

where  $[m]$  stands for the nearest integer to  $m$ . As a consequence,  $Q = \{0, 0.01, \dots, 1\}$ . Then, the system of equations governing the evolution of the two classes of population is:

$$\begin{cases}
 I_u^t = d \left( (1 - r) I_u^{t-1} + p(1 - \eta_u) S_u^{t-1} I_u^{t-1} \right. \\
 \left. + p(1 - \eta_u) S_u^{t-1} \sum_{v \in N_u} \frac{P_v}{P} w_{vu} I_v^{t-1} + p\eta_u S_u^{t-1} \sum_{v \in N_u} (1 - w_{vu}) I_v^{t-1} \right) \\
 S_u^t = d(1 - I_u^t)
 \end{cases}$$

### 4 Analysis of the Model

One of the most important question in a mathematical epidemiological model is the study of the possibility of the eradication of disease. In relation with every mathematical epidemiological model, it is very important to determine under

what circumstances the epidemic occurs. Taking into account the intrinsic characteristics of our model, we will demand two conditions: (1) The epidemic disease must spread among the nodes of the graph; and (2) The infected population grows.

The initial conditions in the study are the following: At time step  $t = 0$ , we will consider only one node, for example  $u \in V$ , with infected individuals:

$$I_u^0 > 0, S_u^0 = 1 - I_u^0, I_v^0 = 0, S_v^0 = 1 \quad \forall v \in N_u.$$

#### 4.1 Spreading from an Initial Infected Cell to Another

First of all we will show the necessary condition for epidemic spreading from the node  $u$  to its neighbor  $v \in N_u$ , at the next step of time  $t = 1$ . Thus, it occurs if  $I_v^1 \geq q$  where  $q = \min \{x \in Q - \{0\}\}$ . As the unique node with infected population at time  $t = 0$  is  $u$ , then taking into account (1), it yields:

$$I_v^1 = p(1 - \eta_v) \frac{P_u}{P} w_{uv} I_u^0 + p\eta_v (1 - w_{uv}) I_u^0 \geq q.$$

As a consequence:

$$I_u^0 \geq \frac{q}{p(1 - \eta_v) \frac{P_u}{P} w_{uv} + p\eta_v (1 - w_{uv})}.$$

This equation must hold for every neighbor nodes of  $u$ , then the following result holds:

**Theorem.** *The epidemic disease spreads from node  $u$  to its neighbor nodes if the following condition holds:*

$$I_u^0 \geq \frac{q}{p(1 - \eta) \frac{P_u}{P} w + p\eta(1 - w)},$$

where:  $\eta = \max \{\eta_v, v \in N_u\}$ ,  $w = \max \{w_{uv}, v \in N_u\}$ .

#### 4.2 Growth of Infected Population

Now we will study what conditions that must be held to get a growth of the infected population in a node  $u$ . We have to distinguish two cases: (1) There not exist infected individuals from neighbor nodes to  $u$ ; (2) There exist such infected individuals.

1. In the first case it is  $I_u^{t+1} > I_u^t$ , that is:

$$(1 - r) I_u^t + p(1 - \eta_u) S_u^t I_u^t > I_u^t.$$

As a consequence the growth occurs if:

$$S_u^t > \frac{r}{p(1 - \eta_u)}.$$

2. In the second case, the inequality  $I_u^{t+1} > I_u^t$  gives:

$$I_u^t < (1 - r) I_u^t + p(1 - \eta_u) S_u^t I_u^t + p(1 - \eta_u) S_u^t \sum_{v \in N_u} \frac{P_v}{P} w_{vu} I_v^t + p\eta_u S_u^t \sum_{v \in N_u} (1 - w_{vu}) I_v^t$$

which occurs if:

$$S_u^t > \frac{r}{p(1 - \eta_u) + p(1 - \eta_u) \sum_{v \in N_u} \frac{P_v}{P} w_{vu} \frac{I_v^t}{I_u^t} + p\eta_u \sum_{v \in N_u} (1 - w_{vu}) \frac{I_v^t}{I_u^t}}$$

### 4.3 Computer Simulations

In this example, for the sake of simplicity we will suppose that the epidemic is spreading over  $n = 10$  cities,  $v_1, \dots, v_{10}$ , forming a complete graph  $K_{10}$ . In this example, we will consider the following initial configuration:

$$S_{u_1}^0 = 0.9, I_{u_1}^{0.1} = 0.1, S_{u_i}^0 = 1, I_{u_i}^0 = 0, 2 \leq i \leq n.$$

That is, there is only one node at time  $t = 0$  with infected population. Moreover, the parameters used are  $p = 0.25, r = 0.8, \eta_{u_i} = 0.2, 1 \leq i \leq 6$ . Moreover, let us suppose that the population of each node is the same:  $P_{u_i} = 100$  with  $1 \leq i \leq 6$ , and also the transport capacity between two nodes is the same:  $w_{u_i u_j} = 1$  for  $1 \leq i, j \leq 6$ . Note that this example deals with an homogeneous-symmetric case.

In Figure 1 the evolution of the total number of infected and susceptible individuals is shown. If we set  $p = 0.15$  instead of  $p = 0.25$ , the number of infected and susceptible individuals also remains constant with time, but in this case the number of susceptible is greater than the number of infected individuals.

## 5 Conclusions and Future Work

In this work a new mathematical model to simulate the spreading of an epidemic is introduced. It is based on the use of cellular automata on graphs endowed with

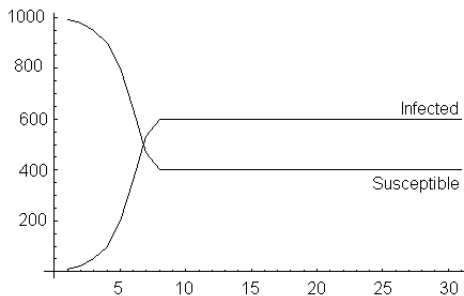


Fig. 1. Evolution of the total number of infected and susceptible individuals

a suitable local transition function. The state of each cell is considered to be the portion of its population which is infected at each time step. The analysis of the model proposed in this paper seems to be in agreement with the results obtained for other mathematical models not based on discrete event systems, such as ODEs or PDEs. Future work aimed at designing a more complete CA-based epidemic model involving additional effects such as the population movement, virus mutation, etc. Furthermore, it is also interesting to consider non-constant connections factors and also the effect of migration between the cells must be considered.

**Acknowledgments.** This work has been partially supported by Consejería de Sanidad, Junta de Castilla y León (Spain).

## References

1. Ahmed, E., Elgazzar, A.S.: On some applications of cellular automata. *Physica A* 296, 529–538 (2001)
2. Beauchemin, C., Samuel, J., Tuszynski, J.: A simple cellular automaton model for influenza A viral infections. *J. Theor. Biol.* 232, 223–234 (2005)
3. Boccaro, N., Cheong, K.: Critical behaviour of a probabilistic automata network SIS model for the spread of an infectious disease in a population of moving individuals. *J. Phys. A-Math. Gen.* 26, 3707–3717 (1993)
4. Fuentes, M.A., Kuperman, M.N.: Cellular automata and epidemiological models with spatial dependence. *Physica A* 267, 471–486 (1999)
5. White, S.H., del Rey, A.M., Sánchez, G.R.: A model based on cellular automata to simulate epidemic diseases. In: El Yacoubi, S., Chopard, B., Bandini, S. (eds.) ACRI 2006. LNCS, vol. 4173, pp. 304–310. Springer, Heidelberg (2006)
6. Kermack, W.O., McKendrick, A.G.: Contributions to the mathematical theory of epidemics, part I. *Proc. Roy. Soc. Edin. A* 115, 700–721 (1927)
7. Martins, M.L., et al.: A cellular automata model for citrus variegated chlorosis. *Physica A* 295, 42–48 (2001)
8. Molisson, D.: The dependence of epidemic and population velocities on basic parameters. *Math. Biosci.* 107, 255–287 (1991)
9. Ross, R.: *The prevention of malaria*, 2nd edn. Murray, London (1911)
10. Satsuma, J., et al.: Extending the SIR epidemic model. *Physica A* 336, 369–375 (2004)
11. Sirakoulis, G.C., Karafyllidis, I., Thanailakis, A.: A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. *Ecol. Model.* 133, 209–223 (2000)
12. Toffoli, T., Margolus, N.: *Cellular Automata Machines: A New Environment for Modeling*. MIT Press, Cambridge (1987)
13. Wolfram, W.: *A New Kind of Science*. Wolfram Media Inc. (2002)