

ANALYSIS OF SIR EPIDEMIC MODELS

MUHAMMAD HANIF DURAD¹, MUHAMMAD NAVEED AKHTAR²

¹ Department of Computer and Information Science (DCIS),
Pakistan Institute of Engineering and Applied Sciences (PIEAS),
P.O. Nilore, Islamabad, Pakistan

hanif@pieas.edu.pk

² Department of Computer and Information Science (DCIS),
Pakistan Institute of Engineering and Applied Sciences (PIEAS),
P.O. Nilore, Islamabad, Pakistan

naveed@pieas.edu.pk

Received October 2013

ABSTRACT. *Modeling epidemics has played a vital role in predicting the impact of a disease in a population and the ability to overcome it. The objective of this paper is to introduce the concept of agent based modeling classification for SIR epidemic model, suggest a new discrete-time agent based SIR model and formally document Gillespie's Algorithm based SIR epidemic model. Furthermore a computational analysis of chemical master equation and Gillespie's algorithm based SIR models has been performed showing that the use of latter results in considerable reduction in execution time.*

Keywords: SIR epidemic model, agent based SIR model, Gillespie's Algorithm based SIR model, Chemical Master Equation.

1. Introduction. Modeling the spread of diseases is the area of biology termed as mathematical epidemiology. The real life diseases may be described by using differential equations as their first formulation. Epidemic models are based on dividing the infected population into a few sections, each containing individuals that are identical in terms of their status [1]. SIR model divides the community in three groups:

- Susceptible: the individual who has no immunity may become infected
- Infected: the individuals who are infected and can transmit the infection to susceptible persons
- Recovered: the individuals who are immune and don't affect the transmission dynamics

It is conventional to denote the number of individuals in each of these sections as S, I and R, respectively. The whole population is $N = S + I + R$ remains constant. The SIR model depends on two constant parameters, the infection rate α and the recovery rate β , it customary to express it as.

$$S' = -\alpha SI; \tag{1}$$

$$I' = \alpha SI - \beta I; \tag{2}$$

$$R' = \beta I; \tag{3}$$

The equations above are coupled nonlinear differential equations representing the SIR model and can be solved by numerical finite difference methods. However the SIR model has been implemented by two more stochastic techniques in this paper. The rest of this paper is structured as follows: In section 2 related work is discussed, in section 3 a new SIR Model Classification is proposed, in section 4 the implemented SIR models are discussed. Finally in section 5, the paper is concluded.

2. Related Work. A number of variants of SIR models can be found in literature [1]. These include SIS (disease with no immunity); SEIR and SEIS (with an exposed 'E' period between being infected and

becoming infective). Quarantine class (Q) could destabilize the disease endemic equilibrium in SIS and SIR models, introducing SIQS and SIQR models respectively. The dynamics of deterministic and stochastic epidemic models are analyzed and compared in [2]. In fact a lot of work has been done on SIR epidemic models; however the following limitations were found.

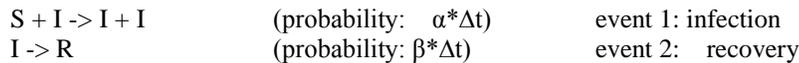
1. The concept of agent based modeling hasn't been introduced formally for epidemic models.
2. Few models such as agent based [3] and Gillespie model has been suggested in literature but have not been implemented properly for SIR model.
3. Inadequate computational analysis various SIR implementations has been performed.

In the nutshell, it is believed that this paper will overcome few of above limitations and broadens the picture of various SIR models to quantitatively capture dynamics of disease spread.

3. Suggested SIR Model Classification. SIR models can be classified as stochastic and non-stochastic. The basic finite difference model is an example of non-stochastic implementation. Stochastic models have been classified as discrete-time Markov chain (DTMC), Continuous-time Markov chain (CTMC) solving chemical master equation (CME), and stochastic differential equation (SDE) model assuming epidemic as Wiener process. In our opinion this classification in [2] doesn't capture the notion of agent-based modeling, specifically for agent based [3] and Gillespie algorithm SIR Models. The latter two models in our opinion are equally stochastic but also agent-based models. While implementing these two models uniform random number generator acts as an agent which decides the next step in our simulation. Hence a new classification scheme can be suggested for SIR models. The former algorithms namely finite difference, DTMC, CTMC and SDE can be termed as **None-Agent based models (NABM)** and the later agent based [3] and Gillespie Algorithm models can be called as **Agent-based models (ABM)**.

4. Discussed SIR Models. Basic reproduction number $R_0 = \alpha N / \beta > 1$ is the essential requirement for the spread of diseases. It is assumed $N = 100$, $I = 1$, $\alpha = 0.01$ and $\beta = 0.1$ for each set of 1000 simulations in next sections. Here, three SIR models namely Discrete Agent Based Model (DABM), CTMC model and Gillespie's Algorithm based SIR model have been discussed.

4.1. Discrete Agent Based Model (DABM). The idea has been given the name as agent based model by Shlomo Ta'asan [3]. Godfrey Beddard [4] calls it discrete Monte Carlo simulation. We term it as discrete agent based model (DABM) to differentiate it with simple DTMC model. The SIR dynamics is expressed in reaction forms as:

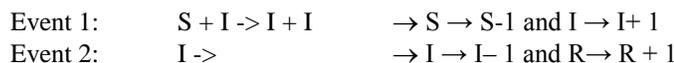


We pick Δt small enough such that only one event will happen so that the actions are time homogenous; the probabilities of events are as:

$$P(S+I \rightarrow I + I) = \alpha * S * I * \Delta t \quad \rightarrow \text{event 1 happens just once during } \Delta t$$

$$P(I \rightarrow R) = \beta * \Delta t * I \quad \rightarrow \text{event 2 happens just once during } \Delta t.$$

Since probabilities are < 1 , Δt should be small enough. More precisely, we want the sum of all probabilities to be less than 1. Each event changes the variables of the model S, I, R.



Here

$$P(S \rightarrow S - 1 \ \& \ I \rightarrow I + 1) = \alpha * S * I * \Delta t \text{ and } P(I \rightarrow I - 1 \ \& \ R \rightarrow R + 1) = \beta * \Delta t * I.$$

A natural question is how to proceed for such two types of events? The answer is to use a uniform random number generator as an agent to fix on the next step in simulation. That is why Godfrey Beddard [4] calls it

discrete Monte Carlo simulation. Using the parameters $N = 100$, $\alpha = 0.01$ and $\beta = 0.1$ and # of simulations = 1000, the results are shown as Fig. 1.

The solid lines in Figure 1 represent the mean values of S, I and R whereas the dotted line above and below each solid line show the upper and lower limits of the variables for each 1000 simulation. The mean values S, I and R are compared to ode45 in the following Figure 2.

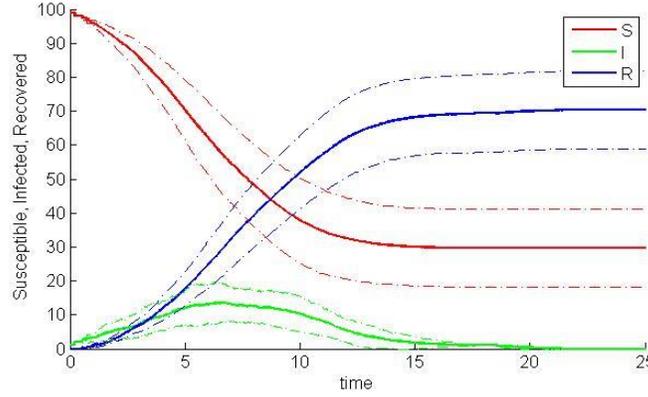


Figure 1. Values of S, I and R using for DABM Model

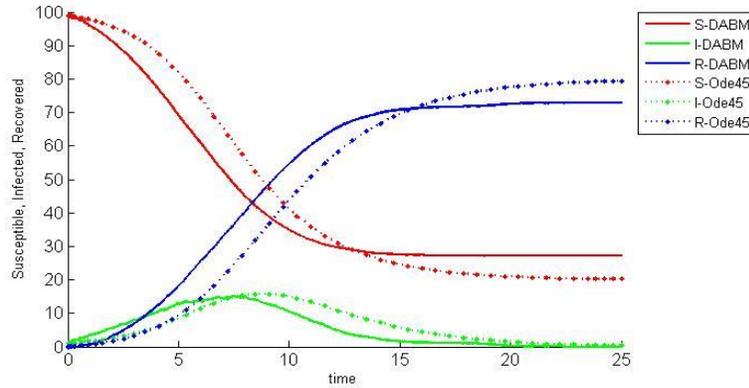


Figure 2. Comparison of the values of S, I and R using ode45 and DABM Model

It can be seen from Figure 2 the results of DABM model show similar tendency with respect to ode45. It must be noted DABM model is stochastic in nature but has similar trend as exact ode45 curves.

4.2. Chemical Master Equation. In this model, epidemic progression is defined on a continuous time scale, $t \in [0, \infty)$, but the states $S(t)$, $I(t)$, and $R(t)$ are discrete random variables, i.e. $S(t), I(t), R(t) \in \{0, 1, 2, \dots, N\}$. Also $S(t) + I(t) + R(t) = N$. The model is also known as Continuous-time Markov chain (CTMC) in literature [2]. There is no latent period so that infected individuals are also infectious. Only two of the random variables are independent. Assume the transition probabilities are:

$$\text{Prob}\{\Delta S(t) = i, \Delta I(t) = j | (S(t), I(t))\} = \begin{cases} \alpha S(t) I(t) \Delta t + o(\Delta t) & (i, j) = (1, -1) \\ \beta I(t) \Delta t + o(\Delta t) & (i, j) = (0, -1) \\ 1 - [\alpha S(t) I(t) + \beta I(t)] \Delta t + o(\Delta t) & (i, j) = (0, 0) \\ o(\Delta t) & \text{otherwise} \end{cases}$$

when $I(t) = -1$, then $R(t) = 1$.

Assume that the initial population distribution is $(S(0), I(0)) = (s_0, i_0)$, where $s_0 + i_0 = N$, $s_0 \geq 0$, $i_0 > 0$. Let $p_{(i,j)}(t) = \text{Prob}\{S(t)=i, I(t)=j\}$; then the state probabilities are solutions of the forward Kolmogorov equations:

$$\begin{aligned} \frac{dp_{(i,j)}(t)}{dt} = & \alpha(i+1)(j-1)p_{(i+1,j-1)}(t) + \beta(j+1)p_{(i,j+1)}(t) \\ & - [\alpha ij + \beta(j+1)]p_{(i,j)}(t); \end{aligned} \quad (4)$$

where $i=0,1,2,\dots,N$, $j=0,1,2,\dots,N-i$, $i+j \leq N$. In the biology literature, this system of equations is termed the Chemical Master Equation (CME).

If (i,j) lies outside this range, the probabilities are assumed to be zero. For example, for $j=0$,

$$\frac{dp_{(i,0)}(t)}{dt} = p_{(i,1)}(t), \quad \text{and} \quad \frac{dp_{(N,0)}(t)}{dt} = 0 \quad (5)$$

for $i=0,1,2,\dots,N-1$. The $N+1$ states $(i,0)$ are closed, $i=0,1,2,\dots,N$. No transition occurs out of one of these states.

The forward Kolmogorov equations can be expressed in vector form as:

$$P'(t) = P(t)Q \quad (6)$$

where $P(t) = p_{(i,j)}(t); i, j \in S$ where S is the state space of the model consisting of all the possible states the model may adopt, and q_{ij} is the transition rate from state i to state j . The solution (when S is bounded; hence trivially satisfied if finite) is:

$$p(t) = p(0) e^{(Qt)} = p(0) \sum_{n=0}^{\infty} \frac{Q^n t^n}{n!} \quad (7)$$

There exists a variety of methods for solving these equations, for example [5]. We will consider three approaches to solving this equation for CME models of infectious disease.

The solution to Eq. 7 involves matrix exponential. First approach is to evaluate the simple matrix exponentiation (using `expm` in MATLAB), the second method is taking exponential is via Krylov subspace methods, as for example implemented in the algorithm of EXPOKIT. The third most recent method worth mentioning is implicit Euler (IE) method which exploits the structure of the master equation [5]. All these techniques have been implemented by using a 2.67 GHz Intel Core™ i5 processor, 4 GB RAM, running MATLAB 2012B for various population sizes and the results are shown in Fig. 3 along with Gillespie's Algorithm based SIR model to be discussed in next section.

4.3. Gillespie's Algorithm based SIR model. It was Gillespie who justified the use of CTMC model to solve chemical master equation [4]. Assuming S, I, R as chemical entities, we can simulate the SIR model. System state is represented by individual count in each category and is treated as a discrete random variable. Two discrete events are required to be considered for simple SIR Model as:

1. The infection rate is $\alpha * S * I$, and it result in $S \rightarrow S-1$ and $I \rightarrow I+1$.
2. The recovery rate is $\beta * I$, and it result in $I \rightarrow I-1$ and $R \rightarrow R+1$.

Different schemes can be used to implement this model; however this paper uses Gillespie Monte Carlo method (Algorithm 1) suggested in [4]. It can be proved that though the stochastic processes considered here follow exponential distribution but ratios which could be simulated by using uniform random number generator [4].

Algorithm 1 Gillespie's direct method

```
1: Initialize  $S, I, R$  and set  $t \leftarrow 0, t_f \leftarrow$  target time.
2:  $a_1 \leftarrow \alpha SI$ 
3:  $a_2 \leftarrow \beta I$ 
4:  $a_0 \leftarrow a_1 + a_2$ 
5: while  $t \leq t_f$  and  $(S > 0$  or  $I > 0)$  do
6:   Generate two uniform random numbers  $U_1$  and  $U_2$ .
7:    $\Delta t \leftarrow -\ln(\frac{U_1}{a_0})$ 
8:    $t \leftarrow t + \Delta t$ 
9:   if  $U_2 < \frac{a_1}{a_0}$  then
10:     $S \leftarrow S - 1$ 
11:     $I \leftarrow I + 1$ 
12:   else
13:     $I \leftarrow I - 1$ 
14:     $R \leftarrow R + 1$ 
15:   end if
16: end while
```

The Gillespie's Algorithm based SIR model along with other exact methods for CME, have been implemented for various population sizes, the results as shown in Figure. 3.

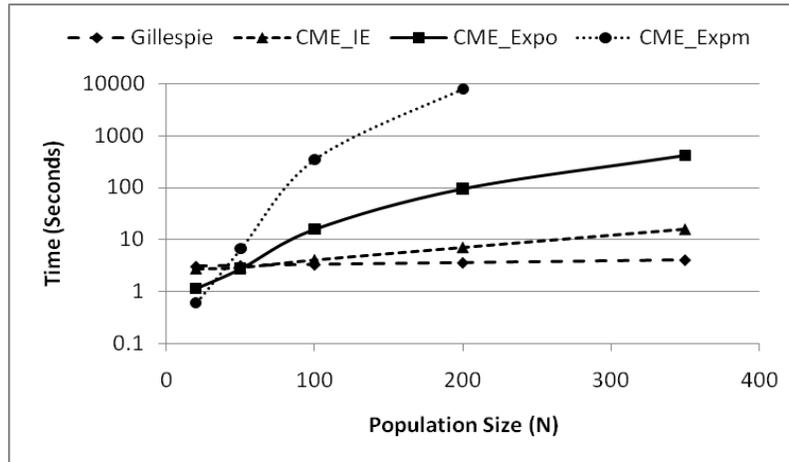


Figure 3. Comparison Timing analysis of the CME based and Gillespie's Algorithm based SIR model

It can be seen from Figure. 3 the execution time for the CME based exact methods is much larger than Gillespie's Algorithm based SIR models for population sizes of about 100 and above. However the implicit Euler (IE) shows a good promise as compared to other CME exact methods. Figure. 3 implies that CME based exact methods will be very difficult to simulate for large population of billions.

5. Discussion and Future Work. Modeling of the dynamics of a disease even for the simplest epidemics is computationally daunting task. It was observed that even with recent faster CPUs; solving CME based CTMC model for large population sizes is intractable problem. No doubt the rescue is Gillespie's Algorithm based SIR model which can solve the problem very efficiently.

In future, we are interested to simulate the other SIR variants such as Simple SIS, SIS-vital, SIR, SIR-vital, SIRS and SIRS-vital for Discrete Agent Based Model (DABM) and Gillespie's Algorithm based SIR models.

REFERENCES

- [1] Earn, D. J. (2008). A light introduction to modelling recurrent epidemics. In *Mathematical epidemiology* (pp. 3-17). Springer Berlin Heidelberg.
- [2] Allen, L. J., & Burgin, A. M. (2000). Comparison of deterministic and stochastic SIS and SIR models in discrete time. *Mathematical biosciences*, 163(1), 1-33.
- [3] Information on <http://www.math.cmu.edu/math/faculty/taasan.html>
- [4] Beddard, G. S. (2009). *Applying Maths in the Chemical and Biomolecular Sciences: An Example-based Approach*. Oxford University Press.
- [5] Jenkinson, G., & Goutsias, J. (2012). Numerical integration of the master equation in some models of stochastic epidemiology. *PLoS one*, 7(5), e36160.