

УДК 51-7:614.4

Имитационная SI-модель распространения эпидемии в замкнутой популяции

Сыромясов А. О., Мальков В. Е.

Национальный исследовательский
Мордовский государственный университет им. Н.П. Огарёва

Аннотация: В статье изучается простейшая компартментальная SI-модель эпидемии, в которой популяция делится всего на две группы: восприимчивые и зараженные индивиды. Построен клеточный автомат, позволяющий реализовать дискретный вариант указанной модели. Описаны численные эксперименты, в которых с помощью этого автомата имитируется распространение коронавируса на массовых мероприятиях и в населенных пунктах. Показано, что расчеты согласно дискретному и непрерывному вариантам SI-модели приводят к практически одинаковым результатам, а предложенная методика проведения эксперимента позволяет достаточно точно определить параметры непрерывной модели. Кроме того, установлено, что эти параметры существенно зависят от стратегии перемещения индивидов в замкнутой области обитания популяции.

Ключевые слова: SI-модель, математическая модель эпидемии, клеточный автомат, имитационное моделирование.

1. Введение

Наиболее распространенным методом, применяемым при моделировании эпидемий, служит использование компартментальных моделей, в которых вся популяция разбивается на непересекающиеся подмножества в зависимости от эпидемиологического статуса входящих в них индивидов. Простейшая такая модель носит название SI, поскольку любой индивид в ней может иметь одно из двух состояний: восприимчивый к болезни (S – *susceptible*) или уже зараженный (I – *infected*). Далее эта идея развивается в моделях SIR, SEIR и т.д.

Динамика численности каждой из групп, выделенных в какой-либо модели, как правило, задается дифференциальными уравнениями; пионерские работы в этой области были проведены У. Кермаком и А. Маккендриком [1]. Так, модель SI может быть описана системой

$$\frac{dI}{dt} = \beta SI, \quad \frac{dS}{dt} = -\beta SI, \quad S + I = N,$$

где I и S – численность зараженных (больных) и здоровых индивидов, соответственно, t – время. Постоянная величина N есть общая численность популяции, коэффициент β также постоянен, а начальная численность зараженных $I(0) = I_0$ известна. Смысл приведенных уравнений заключается в том, что скорость распространения инфекции пропорциональна количеству встреч восприимчивых индивидов с зараженными, а значит, произведению величин S, I . Переходя к единственному уравнению относительно $I(t)$, можно записать:

$$\frac{dI}{dt} = \beta I(N - I), \quad I(0) = I_0. \tag{1}$$

На модель (1) накладываются следующие очевидные ограничения:

1. Численность популяции не меняется вследствие рождения или смерти индивидов (например, по причине того, что популяция изучается в течение короткого времени).

2. Ареал, в который помещена популяция, замкнут. Иными словами, изменение численности N в связи с миграцией также исключено.

3. Численность популяции достаточно велика. Поскольку численности групп заболевших и здоровых индивидов суть дискретные величины, описывать их динамику дифференциальными уравнениями не вполне корректно. Чтобы допускаемая при этом ошибка была небольшой, необходимо, чтобы приращения численности групп, обозначенные дифференциалами dI и dS , были малы по сравнению с самими I и S .

Такого рода модели просты для понимания и исследования. Однако обыкновенные дифференциальные уравнения не учитывают локальных особенностей передвижения индивидов внутри ареала, а заранее определить входящие в них коэффициенты (такие, как β , а в более сложных моделях – и другие) может быть затруднительно.

Альтернативным способом моделирования эпидемии является использование клеточных автоматов (КА). При таком подходе все занятое автоматом поле делится на отдельные клетки с конечным числом состояний, причем состояния эти меняются с течением времени по определенным правилам [2]. Именно этот подход развивается в настоящей работе.

Конкретным применением построенного клеточного автомата служит модель распространения COVID-19 в населенном пункте или на массовом мероприятии, поэтому далее вместо словосочетания «поле, занятое КА» используется короткий термин «город».

Конечной целью работы автомата служит определение коэффициента β в уравнении (1).

2. Дискретная SI-модель эпидемии и ее реализация клеточным автоматом

Итак, «город» представляет собой квадрат $K \times K$ клеток, в который помещены N индивидов (здоровых и уже зараженных). На каждом шаге работы КА индивиды случайным образом перемещаются в соседние клетки по горизонтали, вертикали или диагонали; они могут также остаться на месте. Иными словами, перемещение каждого члена популяции происходит в рамках так называемой окрестности Мура порядка 1 (рис. 1).

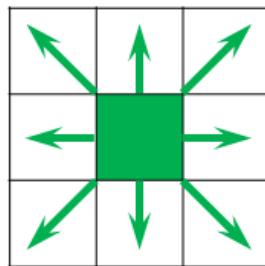


Рис. 1. Возможные перемещения индивида при работе КА.

Если в одной клетке оказались здоровый и зараженный индивиды, разыгрывается случайная величина, так что с заданной вероятностью p здоровый индивид заболевает.

Указанный автомат был реализован с помощью фреймворка Java FX. Это позволило на персональном компьютере изучать динамику эпидемии в «городе» с площадью в несколько сотен тысяч клеток и «населением» в несколько десятков тысяч индивидов.

Отдельно остановимся на применимости описанного КА для исследования модели (1).

Очевидно, в этом автомате выполнены первые два требования, наложенные выше на SI-модель: постоянство численности популяции и замкнутость ареала. Последнее требование (большая численность популяции) должно отслеживаться в рамках каждого конкретного численного эксперимента.

Пусть теперь I_n – численность заболевших на n -м шаге работы автомата. Повторяя рассуждения о том, что приращение количества инфицированных пропорционально числу встреч больных индивидов со здоровыми, получим дискретный аналог (1):

$$I_{n+1} - I_n = \gamma I_n (N - I_n), \quad (2)$$

причем величина I_0 имеет тот же смысл, что и выше, и вновь известна заранее.

Если предположить, что значение функции $I(t)$ в (1) изменяется в последовательные моменты времени t_n , разделенные достаточно малым интервалом Δt , то указанное соотношение может быть переписано в виде

$$I(t_n + \Delta t) - I(t_n) \approx \beta I(t_n) (N - I(t_n)) \Delta t.$$

Последнее равенство соответствует разностному уравнению (2), если принять, что шаги работы автомата совершаются в моменты t_n , $I_n = I(t_n)$, и положить

$$\gamma = \beta \Delta t. \quad (3)$$

Таким образом, динамика численности инфицированных индивидов в непрерывной SI-модели описывается с помощью КА, если интервал Δt между шагами работы КА мал.

3. Описание численных экспериментов

При проведении численных экспериментов предполагалось, что сторона одной клетки равна 1.5 м (социальная дистанция при COVID-19), а интервал между шагами работы КА составляет $\Delta t = 1$ с. Тогда скорость движения представителей популяции по клеткам оказывалась равной 1.5 м/с (при перемещении по горизонтали или вертикали, см. рис. 1) или 2.1 м/с (при движении по диагонали), т.е. лишь немногим более скорости пешехода.

Общая численность «городского населения» была принята равной $N = 1000$, а начальное количество инфицированных I_0 считалось равным 10, 20 и 30. Тем самым, было выполнено последнее из приведенных ранее требований к непрерывной SI-модели: численность популяции оказалась достаточно велика.

В качестве K рассматривались два значения: 338 и 84. Первая величина при заданном N обеспечивает «плотность населения», примерно равную плотности населения г. Саранска (около 3900 чел. на кв. км). Действительно, реальная площадь

такого «города» составляет $(1.5 \cdot 338)^2 = 257049 \text{ м}^2$, что при населении в 1000 человек и дает искомую плотность. Второй вариант K позволяет смоделировать массовое мероприятие (концерт, митинг) на Советской площади того же города (около 16000 м^2).

Предполагалось, что $p = 1.5\%$ или $p = 70\%$; это установленные Роспотребнадзором вероятности передачи вируса COVID-19 от носителя к незараженному человеку при условии, что они оба находятся в масках или без них соответственно.

Наконец, автомат проделывал 600 шагов, т. е. моделировались 10 минут жизни популяции. Исходными данными для определения параметра β служили I_0 , а также полученные в ходе работы КА величины I_{120} , I_{240} , I_{360} , I_{480} и I_{600} . Поскольку 120-секундные интервалы, через которые запоминались значения I , многократно больше выбранного Δt , то уравнение (1) с хорошей точностью приближается соотношением (2), а в силу (3) $\beta = \gamma$.

Поскольку на каждом шаге работы КА направление движения любого индивида выбирается произвольно, то и получаемые I_n являются случайными величинами. Чтобы упорядочить этот процесс, при каждом наборе параметров (K, p, I_0) описанная 600-шаговая имитация производилась несколько раз (до 80), а затем результаты на каждом контрольном шаге усреднялись. Таким образом получались величины $\langle I_{120} \rangle$, $\langle I_{240} \rangle$, ..., $\langle I_{600} \rangle$, равные средней численности зараженных через 120, 240, ..., 600 шагов после начала эксперимента.

В альтернативной серии экспериментов параметры K , p , I_0 оставались прежними, но режим движения по «городу» был иным: вместо перемещения на одну из клеток окрестности Мура положение каждого индивида на следующем шаге менялось абсолютно случайным образом, так что КА «засеивался» здоровыми и зараженными клетками. Остальные этапы эксперимента не менялись. Если в одной клетке оказывались восприимчивый и инфицированный индивиды, то с вероятностью p здоровый индивид заражался. По итогам нескольких «пробегов» автомата (продолжительностью 600 шагов каждый) усреднялись величины I_{120} , I_{240} , ..., I_{600} .

Далее два описанных варианта работы КА будут называться режимами «перемещения» и «засеивания», соответственно.

4. Анализ результатов эксперимента

Аналитическое решение (1) известно:

$$I(t) = \frac{NI_0 e^{\beta Nt}}{N - I_0 + I_0 e^{\beta Nt}}. \quad (4)$$

Зная N , I_0 , а также принимая в качестве значений $I(t)$ в моменты $t = 120, \dots, 600$ с усредненные величины $\langle I_{120} \rangle, \dots, \langle I_{600} \rangle$, можно из (4) приближенно определить β . Для этого может быть использован метод наименьших квадратов (так, в пакете Wolfram Mathematica он реализован стандартной функцией FindFit).

Поскольку перечисленные $\langle I_n \rangle$ по-прежнему случайны, предварительно необходимо:

1. Убедиться в их устойчивости относительно изменения количества «пробегов» автомата T (*trials*), по которым ведется усреднение.

2. Оценить степень отклонения I_n от их средних значений.

В качестве иллюстрации в табл. 1 приведена зависимость $\langle I_n \rangle$ от T , полученная при $K = 338$, $p = 70\%$, $I_0 = 20$ для режима «перемещения» (при других наборах начальных данных, а также для режима «засеивания» ситуация сходна).

Таблица 1. Зависимость $\langle I_n \rangle$ от T

T	$\langle I_{120} \rangle$	$\langle I_{240} \rangle$	$\langle I_{360} \rangle$	$\langle I_{480} \rangle$	$\langle I_{600} \rangle$
10	32.3000	50.3000	73.9000	104.4000	143.7000
20	32.7000	49.3500	72.7500	103.0000	140.7000
40	34.0500	52.9000	78.2000	109.1750	149.6000
80	33.7125	53.1375	78.0750	109.6875	150.6000

Как видно, при увеличении T средние значения I_n стремятся к некоторому единому значению. Далее в качестве $\langle I_n \rangle$ рассматриваются величины, полученные при $T = 80$.

Попытка построить доверительные интервалы для $\langle I_n \rangle$ оказалась неудачной: очевидная гипотеза о нормальном распределении I_n , проверенная с помощью χ^2 -критерия, была отвергнута уже на уровне значимости 0.9. Поэтому в качестве меры разброса числа зараженных была принята стандартная ошибка среднего SE (*standard error*), равная $SE = s/\sqrt{T}$. Здесь s – выборочное среднее квадратическое отклонение исследуемой величины.

Вычисленные при обработке результатов эксперимента значения SE составляют единицы процентов по сравнению как с самими $\langle I_n \rangle$, так и с их «чисто случайными» компонентами $\langle I_n \rangle - I_0$. Так, для приведенного выше набора начальных параметров и режима «перемещения» величина SE для $\langle I_{600} \rangle$ равна 2.59, в то время как $\langle I_{600} \rangle - I_0 = 130.60$. Это позволяет считать колеблемость $\langle I_n \rangle$ при $T = 80$ достаточно малой.

Проведенная подготовительная работа позволяет определить β при любом наборе исходных параметров и для обоих режимов движения. Результаты приведены в табл. 2; прочерк означает, что соответствующий эксперимент не проводился.

Полученные значения β обладают следующими отличительными особенностями:

1. Практически полная независимость от I_0 . Это закономерно, ибо в (1) коэффициент β дифференциального уравнения не связан с начальными данными задачи Коши.

2. Существенная зависимость от способа движения. Коэффициенты, полученные для режима «перемещения», могут в разы отличаться от величин для режима «засеивания».

Кроме того, очевидно, что $\beta(p, K)$ возрастает при увеличении p и убывает с ростом K ; последний факт объясняется тем, что с увеличением K уменьшается плотность населения, а значит, и интенсивность встреч зараженных индивидов со здоровыми.

На рис. 2 показано соответствие экспериментальных данных (красные точки) и логистической функции (4) с подобранным β при $K = 338$, $p = 70\%$, $I_0 = 20$ (сплошная линия). Как видно, это соответствие является вполне удовлетворительным.

Отметим, что при $K = 84$ логистическая кривая быстро выходит на насыщение при обоих режимах движения даже если $p = 1.5\%$. Иными словами, если среди 1000 участников массового мероприятия на городской площади будет хотя бы 10 зараженных, то через 10 минут практически все остальные также будут заражены (в условиях, когда им позволено свободно перемещаться). Разумеется, в реальности передача

Таблица 2. Зависимость β от параметров КА

K	p	I_0	βN («перемещение»)	βN («засевание»)
338	70%	10	0.00358	0.00860
338	70%	20	0.00367	0.00867
338	70%	30	0.00359	0.00873
338	1.5%	10	0.00014	–
338	1.5%	20	0.00014	–
338	1.5%	30	0.00013	–
84	70%	10	0.04507	–
84	70%	20	0.04886	–
84	70%	30	0.05106	–
84	1.5%	10	0.01626	0.02063
84	1.5%	20	0.01707	0.02088
84	1.5%	30	0.01724	0.02066

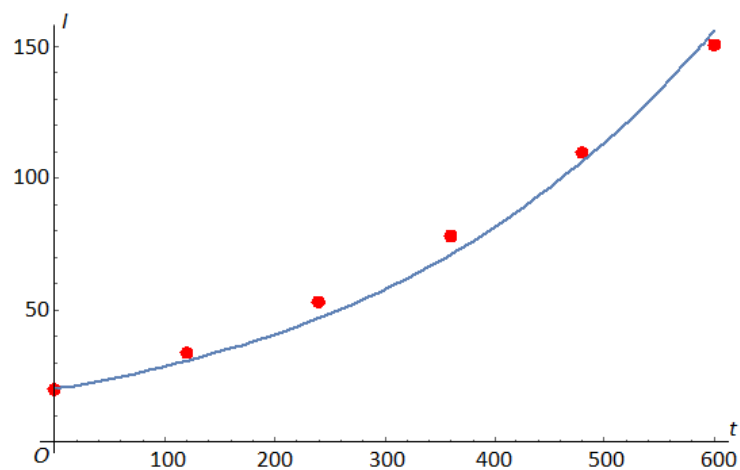


Рис. 2. Соответствие экспериментальных данных аналитическому решению.

инфекции не происходит мгновенно, должен пройти некий инкубационный период и т. д. Тем не менее, этот предельный случай позволяет выработать рекомендации по проведению массовых мероприятий во время эпидемии: помимо соблюдения масочного режима и экспресс-тестирования на заболевание при входе на мероприятие желательно ограничить подвижность участников.

5. Заключение

В настоящей работе показано, что клеточный автомат может быть эффективным инструментом моделирования эпидемии, причем для эмуляции небольшого населенного пункта (несколько тысяч жителей) может быть достаточно использовать персональный компьютер. В частности, использование КА позволяет воспроизвести поведение компартментальных моделей, заданных обыкновенными дифференциальными уравнениями.

Конкретным примером данного подхода в статье служит моделирование распространения COVID-19; исходные данные в численном эксперименте приближены к реальности. Получено, что скорость распространения заболевания зависит не только от вероятности передачи инфекции при непосредственном контакте и от плотности расселения индивидов в ареале, но и от характера их передвижения.

Указанный метод в дальнейшем может совершенствоваться. Во-первых, число состояний индивида в модели может быть увеличено, а заболеванию (после контакта здорового индивида с зараженным) может предшествовать инкубационный период. Во-вторых, соседние клетки могут быть разделены непроницаемыми границами, воспроизводящими городскую топографию (стены домов, заборы). В-третьих, если моделируемый промежуток времени велик, то может быть учтена различная активность индивидов в зависимости от времени суток. В-четвертых, скорость индивида может зависеть от его местоположения: так, если занятая им клетка символизирует участок автодороги, то его скорость вдоль дороги растет.

Литература

1. Kermack W. O., McKendrick A. G. A Contribution to the Mathematical Theory of Epidemics // Proc. Roy. Soc. Ser. A. 1927. V. 115, N 772. P. 700–721.
2. Von Neumann J. Theory of Self-Reproducing Automata. University of Illinois Press, 1966.

MSC 92D30, 68Q80, 34A34

Imitational SI-model of epidemic spread in closed population

A. O. Syromyasov, V. E. Malkov

National Research Ogarev Mordovia State University

Abstract: The article explores simplest compartmental SI-model of epidemic, where all the individuals in a population are divided into two groups: susceptible (S) and infected (I) ones. Cellular automaton is built that allows to emulate discrete version of this model. Numerical experiments imitating disease propagation using the automaton are described in the paper. It is shown that the calculations corresponding to discrete and continuous versions of SI-model lead to almost the same results and that supposed experimental procedure allows to find parameters of continuous model quite accurately. Moreover, it is established that these parameters essentially depend on strategy of individuals movement in a closed area of population inhabitation.

Keywords: SI-model, mathematical modelling of epidemic, cellular automata, imitational modelling.

References

1. W.O. Kermack, A.G. McKendrick, A Contribution to the Mathematical Theory of Epidemics, *Proc. Roy. Soc. Ser. A*, 1927, V. 115, N 772, P. 700–721.
2. J. Von Neumann, Theory of Self-Reproducing Automata, University of Illinois Press, 1966.