

ПРИБЛИЖЁННАЯ ОЦЕНКА СРЕДНЕГО ЧИСЛА ЗАРАЖЁННЫХ УЗЛОВ В МАРКОВСКОЙ МОДЕЛИ РАСПРОСТРАНЕНИЯ КОМПЬЮТЕРНЫХ ВИРУСОВ

А.О. Бельченко¹

магистрант, e-mail: artebel@mail.ru

А.А. Магазев²

д.ф.-м.н., профессор, e-mail: magazev@omgtu.ru

А.Ю. Никифорова²

аспирант, e-mail: skt-omgtu@mail.ru

¹Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

²Омский государственный технический университет, Омск, Россия

Аннотация. В работе рассматривается модель распространения компьютерных вирусов, являющаяся модификацией известной модели Рида–Фроста. Основным результатом статьи является приближённая рекуррентная формула для определения среднего числа заражённых узлов в произвольный момент времени. С помощью этой формулы описывается метод вычисления предельного значения среднего числа заражённых узлов, а также предлагается приближённый критерий для существования так называемого режима вымирания.

Ключевые слова: компьютерный вирус, марковская цепь, вирусная эпидемия, режим вымирания.

Введение

В настоящее время для борьбы с компьютерными вирусами имеется достаточно большое количество разнообразных средств и механизмов. Их эффективность, однако, во многом обусловлена тем, что они имеют дело с уже известными вредоносными программными средствами или их несложными модификациями. Принципиально же новые виды компьютерных вирусов могут стать серьёзной угрозой для пользователей вычислительных сетей. В связи с этим разработка и исследование математических моделей распространения «абстрактных» компьютерных вирусов является важным теоретическим инструментом, позволяющим прогнозировать или даже предотвращать возникновение масштабных вирусных эпидемий.

Одной из первых статей, посвящённых математическому моделированию эпидемий компьютерных вирусов с вычислительных сетей, является концептуальная работа Кефарта и Уайта [1]. В этой статье (по-видимому, впервые) был описан принцип построения соответствующих математических моделей, основанный на синтезе идей теории графов и математической теории распростра-

нения биологических эпидемий в человеческой популяции. Собственно после появления этой статьи и начала своё развитие математическая теория моделирования эпидемий компьютерных вирусов в сложных сетях. Среди наиболее свежих обзорных работ по данному направлению исследований мы отметим здесь статьи [2, 3].

Следует подчеркнуть, что подавляющее большинство математических моделей распространения эпидемий компьютерных вирусов основано на приближении, в рамках которого число заражённых узлов сети является непрерывной функцией времени (к этому типу относятся, например, различные варианты SIS и SIR моделей). Это позволяет привлечь для моделирования хорошо развитый аппарат дифференциальных уравнений, с помощью которого удаётся получить ряд довольно важных результатов о долгосрочных эпидемиологических перспективах. Типичная постановка задачи в таких моделях подразумевает нахождение точек равновесия соответствующей системы дифференциальных уравнений, асимптотическое поведение её решений и исследование различных режимов протекания эпидемии [4, 5]. Системы уравнений, возникающие в подобных моделях, как правило являются нелинейными обыкновенными дифференциальными уравнениями, содержащими большое число внешних параметров.

С другой стороны, авторами работы [6] было продемонстрировано, что подобные «непрерывные» модели не всегда успешно описывают динамику вирусных эпидемий на коротких временных промежутках. Для более детального описания процесса распространения вируса в сложных вычислительных сетях они предложили «дискретный» вариант модели, в которой эволюция эпидемии описывается с помощью некоторой марковской цепи с дискретным временем. Отметим, что подобный подход не является оригинальным: в соответствующей литературе периодически появляются статьи различных авторов с теми или иными вариантами представления эпидемиологической динамики в рамках теории марковских процессов (см., например, [7–9]).

Основная цель настоящей работы — дальнейшее исследование дискретной марковской модели, предложенной в [6]. Преимуществами этой модели является более тонкая оценка числа заражённых узлов в сети, а также возможность дальнейшей модификации за счёт увеличения числа возможных состояний узлов (в оригинальной версии модели их два). Указанная модель, однако, имеет серьёзный недостаток — быстрый рост вычислительной сложности при увеличении размеров сети. Так, например, для сети, содержащей n узлов, приходится оперировать (умножать и возводить в целую степень) матрицами, содержащими n^2 элементов, что при числе узлов $n \sim 10^6$ уже проблематично из-за потери точности. В этой связи мы выводим рекуррентную приближённую формулу для вычисления среднего числа заражённых узлов сети, использование которой не сопряжено с большими вычислительными трудностями и возможно даже при очень больших значениях n . Мы обсуждаем вопрос о применимости данной приближённой оценки и применяем её для вычисления асимптотической оценки числа заражённых узлов при $t \rightarrow \infty$.

1. Марковская модель распространения компьютерных вирусов в связных сетях

Опишем основные положения марковской модели распространения компьютерных вирусов в связных сетях, предложенной в работе [6].

Рассмотрим компьютерную сеть, состоящую из n узлов. Сеть мы будем представлять в виде связного ориентированного графа, рёбра которого ассоциированы с возможными каналами передачи вируса между отдельными узлами. Каждый узел может находиться в двух состояниях: *восприимчивый к заражению* (**S**) и *заражённый* (**I**). В рамках рассматриваемой модели время предполагается дискретным, т. е. любые изменения в сети происходят в заданные моменты времени, которые мы будем полагать целыми числами: $t = 0, 1, 2, \dots$

Обозначим через $i(t)$ и $s(t)$ число узлов, находящихся в момент t в заражённом и восприимчивом состояниях соответственно. Ясно, что $i(t) + s(t) = n$. Предполагается, что заражённые узлы могут оставаться таковыми или становиться восприимчивыми к заражению (излечиваться). Точно так же, восприимчивые узлы могут оставаться восприимчивыми или становиться заражёнными (инфицироваться). Для количественного описания данных переходов введём следующие параметры:

- β — вероятность передачи вируса от заражённого узла к восприимчивому (за один такт времени);
- δ — вероятность заражённому узлу стать восприимчивым (вылечиться) за один такт времени;
- c — *связность* сети, т. е. среднее относительное число соседей у узла. Отметим, что данный параметр связан со средней степенью $\langle k \rangle$ вершины в графе формулой $c = \langle k \rangle / n$. Если $c = 0$, сеть состоит только из изолированных узлов, если же $c = 1$, сеть представляет собой полный ориентированный граф.

Кроме того, мы введём ещё одну важную характеристику: вероятность μ восприимчивому узлу стать заражённым. Как мы увидим далее, эта величина не является постоянной; она зависит от β , c и числа заражённых узлов i .

На рис. 1 иллюстрируются возможные переходы между состояниями отдельного узла с указанием соответствующих вероятностей.

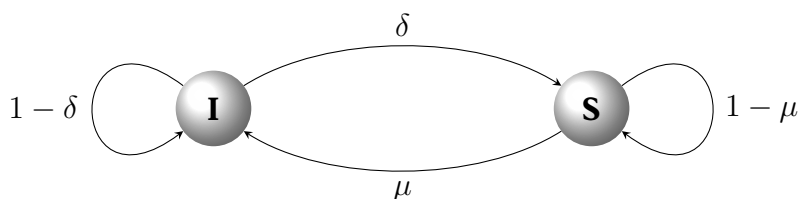


Рис. 1. Граф переходов между состояниями отдельного узла сети

Ассоциируем с нашей компьютерной сетью марковскую цепь, *состояния* которой определяются числом i заражённых узлов; всего имеется $n+1$ возможных состояний: $i = 0, 1, \dots, n$. Динамика данной цепи задаётся переходными вероятностями $\pi_{i,i'}$, т. е. вероятностями перейти сети из состояния i в состояние i' . В работе [6] показано, что эти вероятности вычисляются в соответствии с формулой

$$\pi_{i,i'} = \sum_{k=\max\{0, i-i'\}}^{\min\{i, n-i'\}} C_i^k \delta^k (1-\delta)^{i-k} C_{n-i}^{i'-i+k} \mu^{i'-i+k} (1-\mu)^{n-k-i'}, \quad (1)$$

где вероятность μ восприимчивому узлу стать заражённым в состоянии i даётся равенством

$$\mu = \sum_{k=0}^i C_i^k [1 - (1-\beta)^k] c^k (1-c)^{i-k}. \quad (2)$$

Здесь $C_i^k \equiv \frac{i!}{k!(i-k)!}$ — биномиальный коэффициент, равный числу сочетаний из i по k .

Если нам известна переходная матрица $\Pi = (\pi_{i,i'})$, составленная из всех возможных вероятностей (1), с её помощью мы можем вычислять различные интересующие нас величины. Например, мы можем вычислить вероятность $p_i(t)$ состояния i в момент времени t согласно формуле

$$p_i(t) = \sum_{i'=0}^n p_{i'}(0) \cdot \pi_{i',i}^{(t)},$$

где $\pi_{i,i'}^{(t)}$ — элементы матрицы Π^t , являющейся t -й степенью матрицы Π . Здесь $p_i(0)$ представляют собой вероятности состояний цепи в начальный момент времени $t = 0$. Обычно предполагают заданным *начальное число заражённых узлов* $i_0 = i(0)$ так, что

$$p_i(0) = \begin{cases} 1, & i = i_0, \\ 0, & i \neq i_0. \end{cases}$$

Зная вероятности $p_i(t)$, мы также можем оценить *среднее число заражённых узлов* в момент t — величину, которая будет играть основную роль в настоящем исследовании:

$$I(t) = \sum_{i=0}^n i \cdot p_i(t). \quad (3)$$

Перед тем как перейти к дальнейшему анализу этой модели, сделаем несколько замечаний.

Рассматриваемая нами модель распространения вирусов представляет собой модификацию хорошо известной модели Рида–Фроста [10], предложенную Л. Ридом и В. Фростом в 1920-х гг. для описания биологических эпидемий. Авторы этой модификации приняли некоторые дополнительные допущения, связанные с особенностями эпидемий компьютерных вирусов [6]; многие из

этих допущений заимствованы из концептуальной работы Кефарта и Уайта [1]. В то же время, в отличие от Кефарта и Уайта, в статье [6] была предложена более тонкая оценка вероятности μ заражения восприимчивого узла; в [1] эта вероятность линейно зависела от числа заражённых узлов: $\mu = \beta ci$. Как нетрудно видеть, данная оценка получается из формулы (2) её разложением в ряд Тейлора и отбрасыванием всех членов выше первого порядка (см. ниже). Таким образом, согласно модели, предложенной в [6], вероятность заражения с увеличением числа заражённых узлов i растёт медленнее, чем это предполагалось Кефартом и Уайтом.

Авторами работы [6] также было отмечено, что рассматриваемая марковская модель распространения компьютерных вирусов может быть сравнительно легко обобщена на случай m различных состояний узлов при предположении, что переходы между этими состояниями происходят по кольцевому графу [11]. Ими также была высказана возможность дальнейших модификаций, в которых параметры модели являются функциями состояний марковской цепи (так же, как это имеет место для вероятности заражения μ восприимчивого узла). Подобный уровень обобщения может приводить к более сложной нелинейной динамике распространения вирусов.

2. Приближённая оценка среднего числа заражённых узлов

При исследовании особенностей распространения эпидемий компьютерных вирусов основной интерес представляет динамика изменения числа заражённых узлов «в среднем». Иными словами, возникает необходимость изучения среднего числа заражённых узлов I как функции времени t . Однако в случае больших сетей, т. е. при $n \gtrsim 10^5$, прямой способ вычисления $I(t)$ в соответствии с формулой (3) затруднён, так как приходится возводить в целую степень матрицу переходных вероятностей $\Pi = (\pi_{i,i'})$ размера $n \times n$, что довольно затратно. Кроме того, при нахождении элементов $\pi_{i,i'}$ этой матрицы в соответствии с равенством (1) требуется вычислять биномиальные коэффициенты C_i^k для больших значений i и k , что также неудобно с вычислительной точки зрения. В настоящем разделе мы опишем альтернативный способ, с помощью которого можно получить *приближённую* оценку для среднего числа заражённых узлов, минуя сложные и затратные вычисления.

Для начала упростим формулу (2) для вероятности μ восприимчивому узлу стать заражённым. Для этого мы раскроем квадратные скобки в этой формуле:

$$\mu = \sum_{k=0}^i C_i^k c^k (1-c)^{i-k} - \sum_{k=0}^i C_i^k (1-\beta)^k c^k (1-c)^{i-k}. \quad (4)$$

Первое слагаемое — это биномиальное разложение единицы, так как $1 = (c + 1 - c)^i$. Второе слагаемое перепишем в соответствии с формулой для бинома

Ньютона в виде:

$$\sum_{k=0}^i C_i^k [(1-\beta)c]^k (1-c)^{i-k} = [(1-\beta)c + 1-c]^i = (1-\beta c)^i.$$

Подставляя последнее выражение в формулу (4), получаем

$$\mu = 1 - (1 - \beta c)^i. \quad (5)$$

Полученное упрощённое выражение для μ позволяет сразу же сделать ряд полезных замечаний, касающихся этой величины. Во-первых, из равенства (5) следует, что величина μ зависит от β и c не по отдельности, а в комбинации βc . Это наблюдение означает, что модель обладает симметрией относительно преобразований вида $\beta \rightarrow \lambda\beta$, $c \rightarrow c/\lambda$, $\lambda > 0$ (отметим, что β и c входят в формулу (1) только через μ). Это позволяет сузить число входных параметров модели с четырёх (β , c , δ , i_0) до трёх ($\alpha = \beta c$, δ , i_0). Во-вторых, необходимо подчеркнуть факт *монотонного возрастания* μ с увеличением числа заражённых узлов i , причём в пределе $i \rightarrow \infty$ мы получаем $\mu \rightarrow 1$. Для конечных сетей, однако, i не может превышать n , так что для сети с n узлами максимальное значение вероятности μ будет равно $\mu_{\max} = 1 - (1 - \beta c)^n$. В-третьих, при $0 < \beta c < 1$ мы имеем разложение в ряд Тейлора

$$\mu = \beta c i - \frac{i(i-1)}{2!} \beta^2 c^2 + \frac{i(i-1)(i-2)}{3!} \beta^3 c^3 + O(\beta^4 c^4),$$

из которого, в частности, следует, что в линейном приближении $\mu \approx \beta c i$. Именно такую оценку для μ и предлагали Кефарт и Уайт в своей работе [1].

Получим теперь удобное рекуррентное равенство для приближённой оценки $I^*(t)$ среднего числа заражённых узлов в момент времени t .

Допустим, что в некоторый момент времени t среднее число заражённых узлов в сети известно и равно $I^*(t)$. Тогда вероятность восприимчивому узлу стать заражённым в следующий момент времени $t+1$ будет равна

$$\mu = 1 - (1 - \beta c)^{I^*(t)}.$$

Отсюда следует, что в момент $t+1$ «в среднем» будет дополнительно заражено $\mu [n - I^*(t)] = [1 - (1 - \beta c)^{I^*(t)}] [n - I^*(t)]$ узлов (из тех, которые были восприимчивыми в момент t). С другой стороны, часть узлов, инфицированных в момент t , вылечатся, т. е. станут восприимчивыми в момент $t+1$; среднее число таких узлов будет равно $\delta I^*(t)$. Таким образом, среднее число заражённых узлов в момент $t+1$ ожидается равным

$$I^*(t+1) = (1 - \delta) I^*(t) + [n - I^*(t)] [1 - (1 - \beta c)^{I^*(t)}]. \quad (6)$$

Рекуррентное соотношение (6) естественным образом дополняется начальным условием

$$I^*(0) = i_0, \quad (7)$$

означающим, что в начальный момент времени $t = 0$ в сети имеется ровно i_0 заражённых узлов.

Насколько хорошо построенная нами рекуррентная последовательность $I^*(t)$ аппроксимирует истинную функцию $I(t)$, получаемую в соответствии с формулой (3)? Строгий анализ этой проблемы является довольно сложной задачей, поэтому в настоящей работе мы ограничимся лишь несколькими качественными замечаниями, подтверждаемые численными экспериментами.

1. Во-первых, отметим довольно очевидный факт: с ростом размера сети n разница между $I(t)$ и $I^*(t)$ уменьшается, т. е. для больших сетей $I^*(t)$ весьма неплохо аппроксимирует среднее число заражённых узлов.
2. Второе наблюдение заключается в том, что при фиксированных n и i_0 разница $\Delta I(t) = |I(t) - I^*(t)|$ слабо, но всё же зависит от параметров $\alpha = \beta c$ и δ . Эта разница особенно существенна вблизи тех значений α и δ , при которых наблюдается так называемый *режим вымирания* — режим, в котором $I_\infty = \lim_{t \rightarrow \infty} I(t) = 0$. Полученное нами приближение $I^*(t)$, напротив, в этом режиме демонстрирует ненулевое предельное значение I_∞^* , причём расхождение с I_∞ может достигать относительной величины порядка 20 %.
3. Наконец, наше третье замечание касается начального условия (7): расхождение между $I(t)$ и $I^*(t)$ довольно чувствительно по отношению к начальному значению i_0 и с его ростом быстро уменьшается. По-видимому, это связано с тем фактом, что для рассматриваемой нами марковской модели предельное значение I_∞ поначалу довольно сильно зависит от i_0 , но с ростом этой величины эта зависимость исчезает (обычно это происходит уже при $i_0 > 0.1n$) [6]. Вместе с тем оценка $I^*(t)$ демонстрирует ещё большую нечувствительность к i_0 . Иллюстрация этого замечания приведена на рис. 2, из которого видно, что имеется существенное расхождение между $I(t)$ и $I^*(t)$ при малых i_0 , которое, однако, быстро уменьшается с ростом начального числа заражённых узлов.

В заключение этого раздела отметим, что, несмотря на сделанные выше замечания, приближение $I^*(t)$ довольно хорошо «работает» на достаточно широком диапазоне входных значений модели и может быть использовано для её дальнейшего изучения. В частности, в следующем разделе мы проиллюстрируем применение рекуррентной формулы (6) для оценки предельного значения I_∞ — характеристики, представляющей большой интерес с точки зрения долгосрочных прогнозов динамики вирусной эпидемии.

3. Предельное значение среднего числа заражённых узлов и режим вымирания

В предположении, что существует предел рекуррентной последовательности (6), перейдём в этом равенстве к пределу при $t \rightarrow \infty$. Используя обозначе-

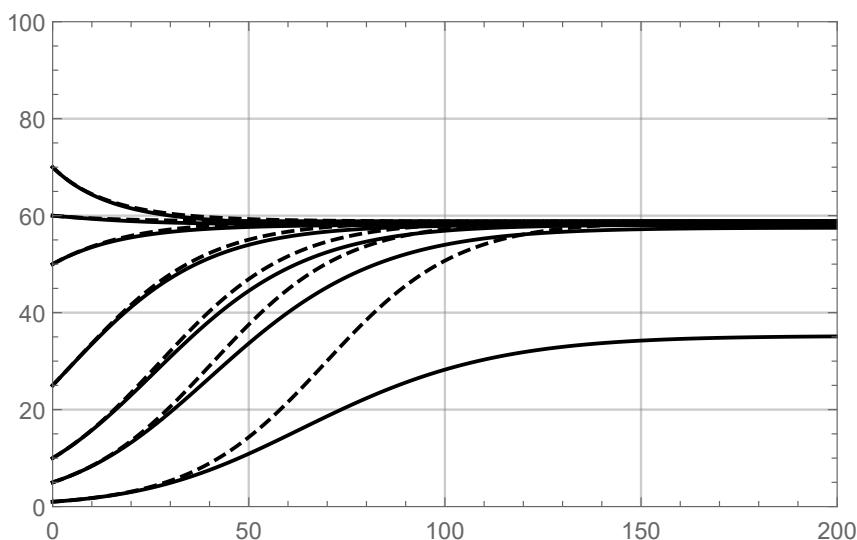


Рис. 2. Графики функций $I(t)$ (сплошные линии) и $I^*(t)$ (прерывистые линии) для различных начальных значений i_0 ($= 1, 5, 10, 25, 50, 60, 70$); $n = 100$, $\beta c = 0,001$, $\delta = 0,04$

ние $I_\infty^* = \lim_{t \rightarrow \infty} I^*(t)$, мы получаем

$$I_\infty^* = (1 - \delta) I_\infty^* + [n - I_\infty^*] [1 - (1 - \beta c)^{I_\infty^*}]. \quad (8)$$

Сразу же обратим внимание на тот факт, что решения этого уравнения не зависят от начального значения i_0 , хотя исходная рекуррентная последовательность (6) эту зависимость (неявно) предполагала. В предыдущем разделе мы уже указывали на эту особенность, также иллюстрируемую численными экспериментами.

Для дальнейшего анализа уравнения (8) введём новый параметр $\gamma = (1 - \beta c)^n$, а также вспомогательную переменную $x = I_\infty^*/n$, представляющую собой относительное число заражённых узлов при $t \rightarrow \infty$. После этого уравнение (8) принимает вид

$$\delta x = (1 - x)(1 - \gamma^x). \quad (9)$$

Отметим, что данное уравнение рассматривается относительно неизвестного x , область значений которого представляет собой единичный отрезок $[0, 1]$. Кроме того, будем полагать, что $\delta > 0$.

Уравнение (9) является трансцендентным уравнением и, по-видимому, в общем случае не решается в аналитическом виде. С другой стороны, его можно легко решить численными методами, например методом бисекций или методом Ньютона. Некоторую информацию о характере корней этого уравнения, однако, можно получить и без явного его решения.

Очевидно, что число $x = 0$ всегда является корнем уравнения (9), однако в большей степени нас будут интересовать его ненулевые корни из отрезка $[0, 1]$, так как таким корням отвечают ненулевые значения I_∞^* . В то же время отсутствие таких корней означает наличие *режима вымирания*, в котором $I_\infty^* = 0$ для любых значений начального числа заражённых узлов i_0 .

Обозначим $f(x) = \delta^{-1}(1-x)(1-\gamma^x)$. Отметим, что функция $f(x)$ бесконечно дифференцируема всюду на числовой прямой. В частности, её вторая производная имеет вид

$$f''(x) = \frac{\gamma^x}{\delta} \ln \gamma (2 - (1-x) \ln \gamma)$$

и является отрицательной на единичном отрезке. Это значит, что функция $f'(x)$ монотонно убывает на отрезке $[0, 1]$ и, так как $f'(0) = -\ln(\gamma)/\delta > 0$ и $f'(1) = (\gamma - 1)/\delta < 0$, имеется *единственная* точка $\xi \in (0, 1)$, такая, что $f'(\xi) = 0$. Таким образом, функция $f(x)$ монотонно возрастает на полуинтервале $[0, \xi)$ и монотонно убывает на полуинтервале $(\xi, 1]$.

Учитывая указанное поведение функции $f(x)$, мы можем теперь дать ответ на вопрос, когда уравнение (9) не имеет на единичном отрезке корней, кроме нулевого. Так как (9) эквивалентно уравнению $x = f(x)$, ненулевых корней не будет, если и только если касательная к функции $f(x)$ в точке $x = 0$ лежит *под* прямой $y = x$, так только в такой ситуации графики функций $y = f(x)$ и $y = x$ не пересекаются на полуотрезке $(0, 1]$. Уравнение касательной к функции $f(x)$ в точке $x = 0$ имеет вид

$$y = -\frac{\ln \gamma}{\delta} x,$$

откуда критерий отсутствия ненулевых корней уравнения (9) на отрезке $[0, 1]$ принимает вид $-\ln \gamma / \delta \leq 1$ или в исходных обозначениях:

$$(1 - \beta c)^n \geq e^{-\delta}. \quad (10)$$

Данное неравенство представляет собой условие, при котором должен иметь место режим вымирания — режим, в котором эпидемия всегда будет затухать на бесконечно больших временах. В частности, пороговое значение $\delta_{\text{кр}} = -n \ln(1 - \beta c)$ отделяет режим вымирания от режима, в котором эпидемия никогда не затухнет.

Так как $I^*(t)$ представляет собой лишь приближённую оценку к истинному значению $I(t)$ среднего числа заражённых узлов, неравенство (10) также является всего лишь приближением к настоящему условию возникновения режима вымирания. В действительности оно является весьма грубым, так как в реальности режим вымирания начинает наблюдаться при значениях δ , меньших, чем предсказываемое пороговое значение $\delta_{\text{кр}}$. С другой стороны, если условие (10) выполнено, то режим вымирания всегда имеет место; это также подтверждается наблюдениями.

Для иллюстрации сформулированных замечаний приведём результаты одной серии численных экспериментов, осуществлённых нами с применением системы символьных вычислений Wolfram Mathematica. Мы взяли исходную марковскую модель с числом узлов $n = 100$ и параметром $\alpha = \beta c = 0,00606$. Отметим, что пороговое значение $\delta_{\text{кр}}$, предсказываемое нашим приближением, в этом случае равно $\delta_{\text{кр}} = 0,60784$. Далее для двенадцати различных значений параметра δ мы вычислили значение I_∞^* с помощью решения уравнения (9),

а затем для каждого из этих значений δ и пяти различных значений i_0 начального числа заражённых узлов мы вычислили соответствующие истинные предельные значения I_∞ стандартными алгоритмами теории марковских цепей. Полученные результаты приведены в таблице 1.

Таблица 1. Сравнение значений I_∞ и I_∞^* для $n = 100$ и $\beta c = 0,00606$

δ	I_∞					I_∞^*
	$i_0 = 1$	$i_0 = 5$	$i_0 = 10$	$i_0 = 25$	$i_0 = 50$	
0,001	99,6580	99,7790	99,7790	99,7790	99,7790	99,7806
0,010	96,6112	97,8012	97,8012	97,8012	97,8012	97,8176
0,050	83,7142	89,2378	89,2379	89,2379	89,2379	89,3396
0,100	69,0677	79,0778	79,0808	79,0808	79,0808	79,2680
0,200	44,2002	60,1127	60,2082	60,2084	60,2084	60,6590
0,300	0,00000	0,00000	00,00000	0,00000	0,00000	43,7863
0,400	0,00000	0,00000	0,00000	0,00000	0,00000	28,3589
0,500	0,00000	0,00000	0,00000	0,00000	0,00000	14,1530
0,600	0,00000	0,00000	0,00000	0,00000	0,00000	0,99239
0,700	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,800	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,900	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000

Анализ таблицы 1 подтверждает высказанные нами выше утверждения. При значениях δ , много бóльших, чем $\delta_{кр}$, значение I_∞ довольно хорошо аппроксимирует истинное значение I_∞ , особенно с ростом i_0 . При приближении δ к $\delta_{кр}$ наблюдается существенное расхождение I_∞^* и I_∞ : режим вымирания начинает наблюдаться при бóльшем значении δ , чем это предсказывается пороговым значением $\delta_{кр} = 0,60784$ (мы наблюдаем его уже при $\delta = 0,3$). Наконец, при значениях $\delta \geq \delta_{кр}$ оба подхода предсказывают наличие режима вымирания и дают одни и те же результаты.

Заключение

В настоящей статье мы предложили приближённую оценку среднего числа заражённых узлов сети с помощью рекуррентного соотношения (6). По сравнению с прямым подходом, известным из теории марковских цепей, наш метод является менее трудоёмким с вычислительной точки зрения, хотя цена подобного выигрыша состоит в потере точности, особенно вблизи тех значений параметров модели, для которых наблюдается переход от режима вымирания к режиму постоянного наличия эпидемии.

ЛИТЕРАТУРА

1. Kephart J.O., White S.R. Directed-Graph Epidemiological Models of Computer Viruses // Proceedings of the IEEE Symposium on Security and Privacy. Oakland, California, 1991. P. 343–359.
2. Pare P.E., Beck C.L., Basar T. Modeling, estimation, and analysis of epidemics over networks: An overview // Annual Reviews in Control. 2020. V. 50. P. 345–360. DOI: <https://doi.org/10.1016/j.arcontrol.2020.09.003>.
3. Zino L., Cao M. Analysis, prediction, and control of epidemics: A survey from scalar to dynamic network models // IEEE Circuits and Systems Magazine. 2021. V. 21, N. 4. P. 4–23. DOI: 10.1109/MCAS.2021.3118100.
4. Zhu Q., Yang X., Ren J. Modeling and analysis of the spread of computer virus // Communications in Nonlinear Science and Numerical Simulation. 2012. V. 17, N. 12. P. 5117–5124. DOI: <https://doi.org/10.1016/j.cnsns.2012.05.030>.
5. Lu-Xing Y., Yang X., Liu J. [et al.] Epidemics of computer viruses: a complex-network approach // Applied Mathematics and Computation. 2013. V. 219, N. 16. P. 8705–8717. DOI: <https://doi.org/10.1016/j.amc.2013.02.031>.
6. Billings L., Spears M.W., Schwartz I.B. A unified prediction of computer virus spread in connected networks // Physics Letters A. 2002. V. 297, N. 3–4. P. 261–266. DOI: [https://doi.org/10.1016/S0375-9601\(02\)00152-4](https://doi.org/10.1016/S0375-9601(02)00152-4).
7. Van Mieghem P., Omic J., Kooij R. Virus spread in networks // IEEE / ACM Transactions On Networking. 2009. V. 17, N. 1. P. 1–14. DOI: 10.1109/TNET.2008.925623.
8. Далингер Я.М., Бабанин Д.В., Бурков С.М. Математические модели распространения вирусов в компьютерных сетях различной структуры // Информатика и системы управления. 2011. № 4(30). С. 3–11.
9. Van de Bovenkamp R., van Mieghem P. Survival time of the susceptibleinfected-susceptible infection process on a graph // Physical Review E. 2015. V. 92, N. 3. P. 032806. DOI: <https://doi.org/10.1103/PhysRevE.92.032806>.
10. Hoppensteadt F.C. Mathematical Methods of Population Biology. Cambridge : Cambridge University Press, 1982. 160 p.
11. Spears W., Billings L., Schwartz I. Modeling viral epidemiology // Naval Research Laboratory. 2001.

AN APPROXIMATE EVALUATION OF THE INFECTED NODES NUMBER FOR A MARKOV MODEL OF VIRUSES SPREADING

A.O. Belchenko¹

Master's Degree Student, e-mail: artbel@mail.ru

A.A. Magazev²

Dr.Sc. (Phys.-Math.), Professor, e-mail: magazev@omgtu.ru

A.Yu. Nikiforova²

Ph.D. Student, e-mail: skt-omgtu@mail.ru

¹Dostoevsky Omsk State University, Omsk, Russia

²Omsk State Technical University, Omsk, Russia

Abstract. In this work, we consider a model of computer virus spreading, which is a modification of the well-known Reed–Frost model. The main result of the paper is an approximate recurrence formula for determining the mean number of infected nodes at any moment of time. Using this formula, we describe a method for calculating the limit value of the mean number of infected nodes, and then we derive an approximate criterion for the existence of the so-called extinction regime.

Keywords: computer virus, Markov chain, virus epidemic, extinction regime.

REFERENCES

1. Kephart J.O. and White S.R. Directed-Graph Epidemiological Models of Computer Viruses. Proceedings of the IEEE Symposium on Security and Privacy, Oakland, California, 1991, pp. 343–359.
2. Pare P.E., Beck C.L., and Basar T. Modeling, estimation, and analysis of epidemics over networks: An overview. Annual Reviews in Control, 2020, vol. 50, pp. 345–360. DOI: <https://doi.org/10.1016/j.arcontrol.2020.09.003>.
3. Zino L. and Cao M. Analysis, prediction, and control of epidemics: A survey from scalar to dynamic network models. IEEE Circuits and Systems Magazine, 2021, vol. 21, no. 4, pp. 4–23. DOI: [10.1109/MCAS.2021.3118100](https://doi.org/10.1109/MCAS.2021.3118100).
4. Zhu Q., Yang X., and Ren J. Modeling and analysis of the spread of computer virus. Communications in Nonlinear Science and Numerical Simulation, 2012, vol. 17, no. 12, pp. 5117–5124. DOI: <https://doi.org/10.1016/j.cnsns.2012.05.030>.
5. Lu-Xing Y., Yang X., Liu J. [et al.] Epidemics of computer viruses: a complex-network approach. Applied Mathematics and Computation, 2013, vol. 219, no. 16, pp. 8705–8717. DOI: <https://doi.org/10.1016/j.amc.2013.02.031>.
6. Billings L., Spears M.W., and Schwartz I.B. A unified prediction of computer virus spread in connected networks. Physics Letters A, 2002, vol. 297, no. 3–4, pp. 261–266. DOI: [https://doi.org/10.1016/S0375-9601\(02\)00152-4](https://doi.org/10.1016/S0375-9601(02)00152-4).
7. Van Mieghem P., Omic J., and Kooij R. Virus spread in networks. IEEE, ACM Transactions On Networking, 2009, vol. 17, no. 1, pp. 1–14. DOI: [10.1109/TNET.2008.925623](https://doi.org/10.1109/TNET.2008.925623).

8. Dalinger Ya.M., Babanin D.V., and Burkov S.M. Matematicheskie modeli rasprostraneniya virusov v komp'yuternykh setyakh razlichnoi struktury. Informatika i sistemy upravleniya, 2011, no. 4(30), pp. 3–11. (in Russian)
9. Van de Bovenkamp R. and van Mieghem P. Survival time of the susceptibleinfected-susceptible infection process on a graph. Physical Review E, 2015, vol. 92, no. 3, pp. 032806. DOI: <https://doi.org/10.1103/PhysRevE.92.032806>.
10. Hoppensteadt F.C. Mathematical Methods of Population Biology. Cambridge, Cambridge University Press, 1982, 160 p.
11. Spears W., Billings L. and Schwartz I. Modeling viral epidemiology. Naval Research Laboratory, 2001.

Дата поступления в редакцию: 17.03.2022