УДК 004.94

***В.А. Минаев*[1], *Е.В. Вайц*[1], *А.В. Корячко*[2], *А.Э. Киракосян*[1]**
([1]МГТУ им. Н.Э. Баумана, [2]Рязанский государственный радиотехнический университет; e-mail: m1va@yandex.ru)

СИСТЕМНО-ДИНАМИЧЕСКОЕ МОДЕЛИРОВАНИЕ РАСПРОСТРАНЕНИЯ КОМПЬЮТЕРНЫХ ВИРУСОВ

Предлагается метод моделирования распространения компьютерных вирусов в сетях на основе SEIR- и PSIDR-моделей. Проведены имитационные эксперименты, позволяющие исследовать динамику числа уязвимых, инфицированных, латентных и "вылеченных" хостов сети.

Ключевые слова: *информационная безопасность, имитационное моделирование, распространение вирусов.*

V.A. Minaev, E.V. Vaitc, A.V. Koraychko, A.E. Kirakosayn

SYSTEM-DYNAMIC MODELING OF COMPUTER VIRUSES SPREADING PROCESSES

The method for simulation of computer virus spreading processes in networks based on SEIR- and PSIDR-models is proposed. Simulations have been carried out, allowing studying the number's dynamics of the vulnerable, infected, latent and "cured" network hosts.

Key words: *system-dynamic model, information security, simulation modeling, virus spreading, simulation experiment.*

Статья поступила в редакцию Интернет-журнала 11 апреля 2017 г.

Введение

Вирусные атаки представляют серьёзную угрозу для всех пользователей компьютерных сетей, включая защищённые сети МЧС России [1]. Своевременное обнаружение вредоносных программ и оперативное устранение последствий их воздействия играет большую роль в обеспечении информационной безопасности любой организационной структуры, поэтому исследование динамики распространения компьютерных вирусов по сети является актуальной задачей и для МЧС России [1-3].

В существующих научных работах по данной проблематике дано описание целого ряда математических моделей распространения компьютерных вирусов по сети [2-5]. В то же время пока недостаточное внимание уделяется имитационному моделированию, дающему исследователям широкий спектр возможностей для решения задач прогнозирования процессов заражения сетей вирусами [6].

Больше всего для имитационного моделирования указанных процессов подходят системно-динамические модели.

Системно-динамическое моделирование – направление в изучении сложных систем, исследующее их состояние во времени в зависимости от структуры элементов системы и взаимодействия между ними. Метод предложен Дж. Форрестером в 1950 годах. Моделируемые процессы отображаются в виде некоторой структуры, состоящей из накопителей – уровней, соединённых взаимосвязанными потоками, которые "перетекая", изменяют значение уровней [7].

Созданные до сегодняшнего дня модели динамики распространения компьютерных вирусов по сети, как правило, основываются на моделях эпидемических процессов [8, 9]. Самыми простыми моделями этого типа, являются SI-модель ("Susceptible – Infected model") и SIR-модель ("Susceptible – Infected – Removed model"). В данной статье рассмотрены их усложнённые модификации: SEIR-модель ("Susceptible – Exposed – Infected – Removed model") и PSIDR-модель ("Progressive Susceptible – Infected – Detected – Removed model").

Описание распространения компьютерных вирусов по сети на основе SEIR -модели

В SEIR-модели учитывается возможность того, что вирус может иметь некий "латентный период", во время которого он не наносит какого-либо вреда инфицированному узлу. Обычно вирус заражает уязвимый узел (*S*) до входа в свою латентную стадию. В течение латентного периода (*Ex*, *Exposed*) узел считается заражённым, но не распространяет вирус. Через некоторое время он становиться способным к заражению других хостов (*I*) и далее становиться "вылеченным" (*R*).

Построенная SEIR-модель (рис. 1) описывается следующей системой уравнений:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = OS(t) - SEx(t); \\ \dfrac{\mathrm{d}Ex}{\mathrm{d}t} = SEx(t) - ExI(t); \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = ExI(t) - IR(t); \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = IR(t); \\ SEx(t) = \dfrac{b \cdot S(t) \cdot I(t)}{n}; \\ ExI(t) = \dfrac{Ex(t)}{f}; \\ IR(t) = c \cdot I(t). \end{cases}$$

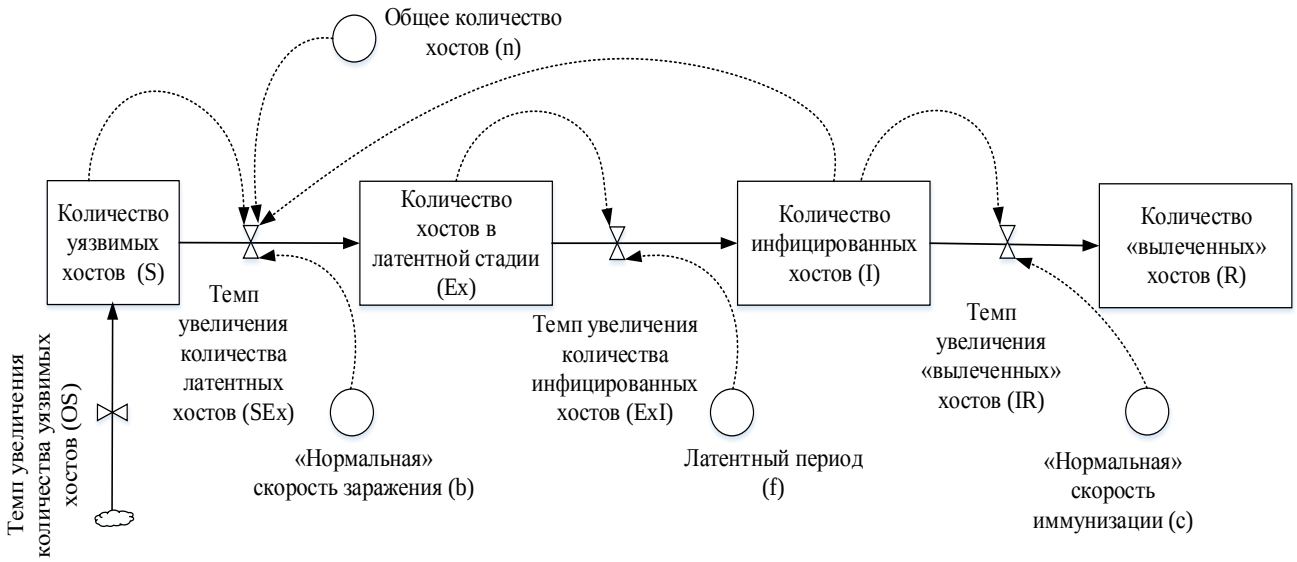


Рис. 1. Системно-динамическая модель распространения компьютерных вирусов по сети на основе SIER-модели

Расшифровка обозначений, используемых в SEIR- модели, приведена в табл. 1.

Таблица 1

Условные обозначения, используемые в модели

Условное обозначение элемента	Название элемента
S	Количество уязвимых хостов
Ex	Количество инфицированных узлов, находящихся в латентной стадии
I	Количество инфицированных хостов
R	Количество "вылеченных" хостов
n	Общее количество хостов в сети
OS	Темп увеличения новых уязвимых хостов (1/*час*)
Sex	Темп увеличения латентных хостов (1/*час*)
ExI	Темп увеличения инфицированных хостов (1/*час*)
IR	Темп увеличения "вылеченных" хостов (1/*час*)
b	"Нормальная" скорость заражения (*часть/час*)
f	Латентный период (*час*)
c	"Нормальная" скорость иммунизации (*часть/час*)

Понятие "нормальной" скорости введено Дж. Форрестером [7] – отношение числа заражённых или вылеченных хостов в день к общему числу уязвимых хостов.

Реализуем построенную модель распространения компьютерных вирусов по сети на основе SEIR-модели в программной среде Anylogic. Общий вид интерфейса модели показан на рис. 2.

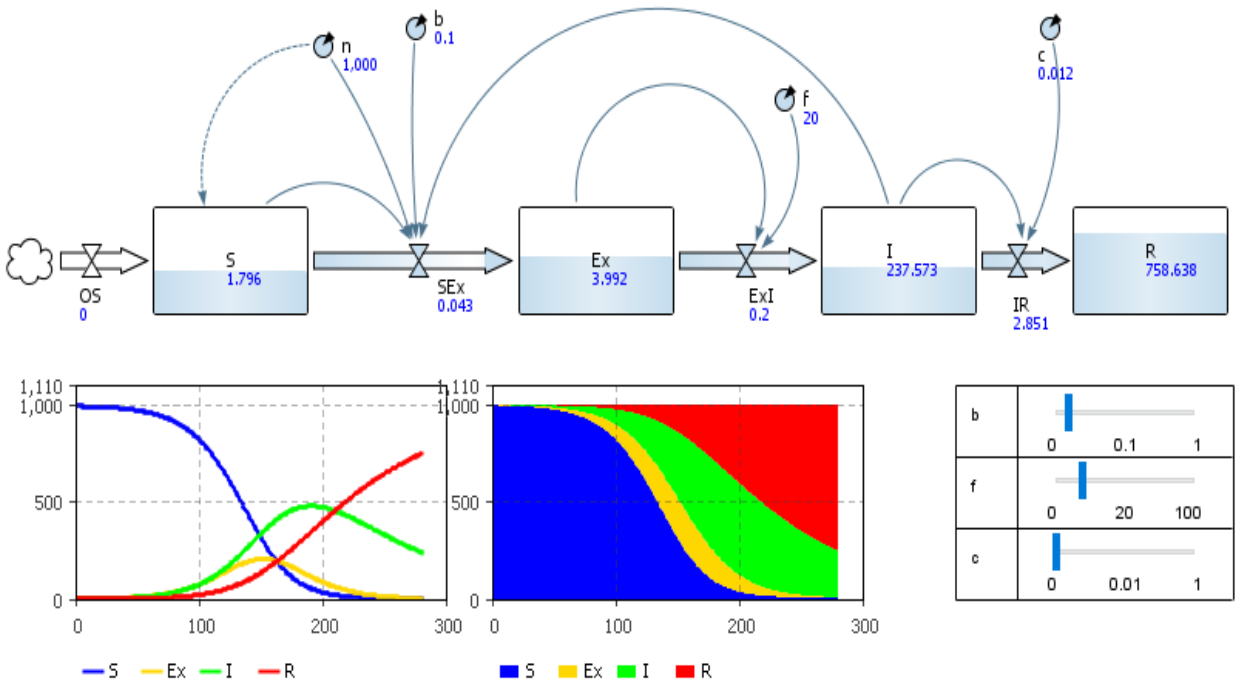


Рис. 2. Общий вид интерфейса системно-динамической модели распространения компьютерных вирусов по сети на основе SEIR-модели

Проведём имитационный эксперимент, в котором определим оптимальное значение "нормальной" скорости иммунизации при заданном ограничении на максимальное количество инфицированных хостов сети.

Математическая постановка такой задачи выглядит следующим образом:

$$\begin{cases} c \to \min; \\ I \le I_{\max}. \end{cases}$$

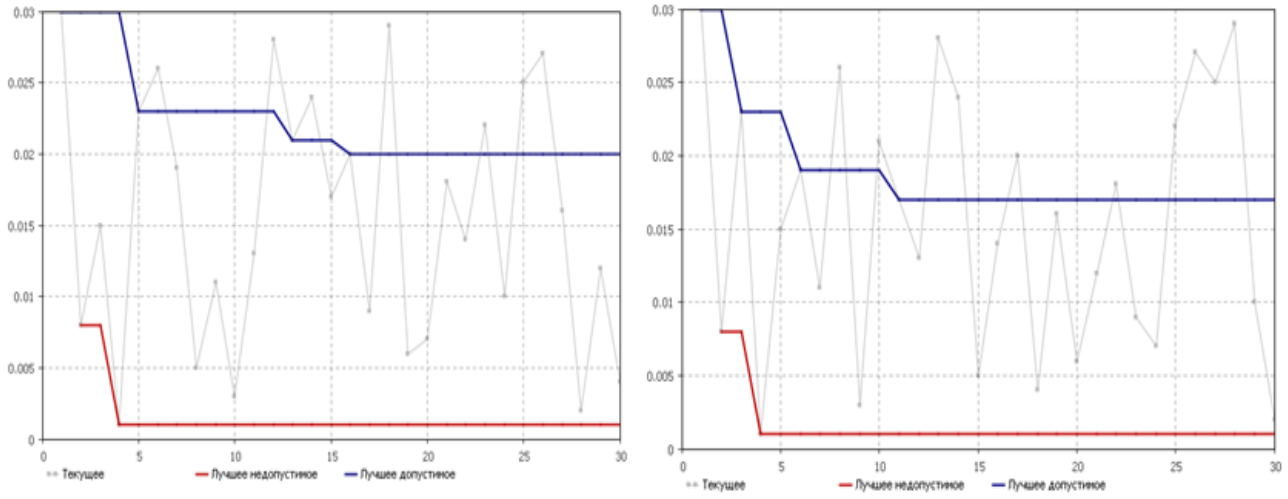
Диапазон возможных значений "нормальной" скорости иммунизации при проведении эксперимента: $c \in \{0{,}001; 0{,}03\}$, шаг – 0,001.

Начальные значения других параметров модели: $S(0) = n = 1000$; $I(0) = 2$; $R(0) = Ex(0) = 0$; $b = 0{,}1$; $f = 20$.

Оптимизационный эксперимент проведён с использованием встроенного в программу Anylogic специального алгоритма OptQuest с точными методами математической оптимизации, нейроннымие сетями и эвристическими подходами к поиску решений.

Приведём зависимости значений оптимизируемого параметра (c) от номера итерации при $I_{\max} = 47$ и $I_{\max} = 57$ (рис. 3).

На графиках красным цветом отображено лучшее недопустимое значение (полученное без учёта ограничений, наложенных на оптимизируемую модель), а синим – лучшее допустимое значение. Можно легко заметить, что с увеличением количества итераций значение оптимизируемого параметра стремится к наилучшему значению целевой функции.



а) *б*)

Рис .3. Зависимости значений оптимизируемого параметра от номера итерации при I_{max} = 47 (*а*) и при I_{max} = 57 (*б*)

По результатам эксперимента можно сделать вывод, что при ограничении количества инфицированных хостов (I_{max} = 47), минимальная "нормальная" скорость иммунизации должна составлять 0,02, а при ограничении количества инфицированных хостов (I_{max} = 57) минимальная "нормальная" скорость иммунизации должна составлять 0,017.

Описание модели распространения компьютерных вирусов по сети на основе PSIDR-модели

В PSIDR-модели предполагается, что эпидемические события разделены на два периода [1]:

• Предварительный период. Изначально вирус инфицирует один хост в сети. В течение определённого количества часов вирус распространяется по сети не замеченным большинством пользователей. В терминах PSIDR-модели эта фаза характеризуется "нормальной" скоростью заражения *b* без попыток излечения.

• Период отклика. Через некоторый период времени вирус обнаруживается на некоторых хостах. Осуществляется выделение его сигнатур и внесение их в базы антивирусного программного обеспечения. Неинфицированные узлы становятся невосприимчивыми к данному вирусу, а инфицированные хосты "вылечиваются" с некоторой частотой, зависящей от частоты обновления антивирусных баз. В рассматриваемой модели интервал времени, разделяющий эти периоды, обозначается τ.

Таким образом, PSIDR-модель предполагает, что течение эпидемии можно разбить на два периода: вначале система может находиться в двух состояниях $S \to R$, а по истечении времени $t = \tau$ система переходит в состояния $S \to I \to D \to R$ с возможностью прямого перехода между состояниями $S \to R$.

Построим системно-динамическую модель распространения компьютерного вируса по сети на основе PSIDR – модели (рис. 4).

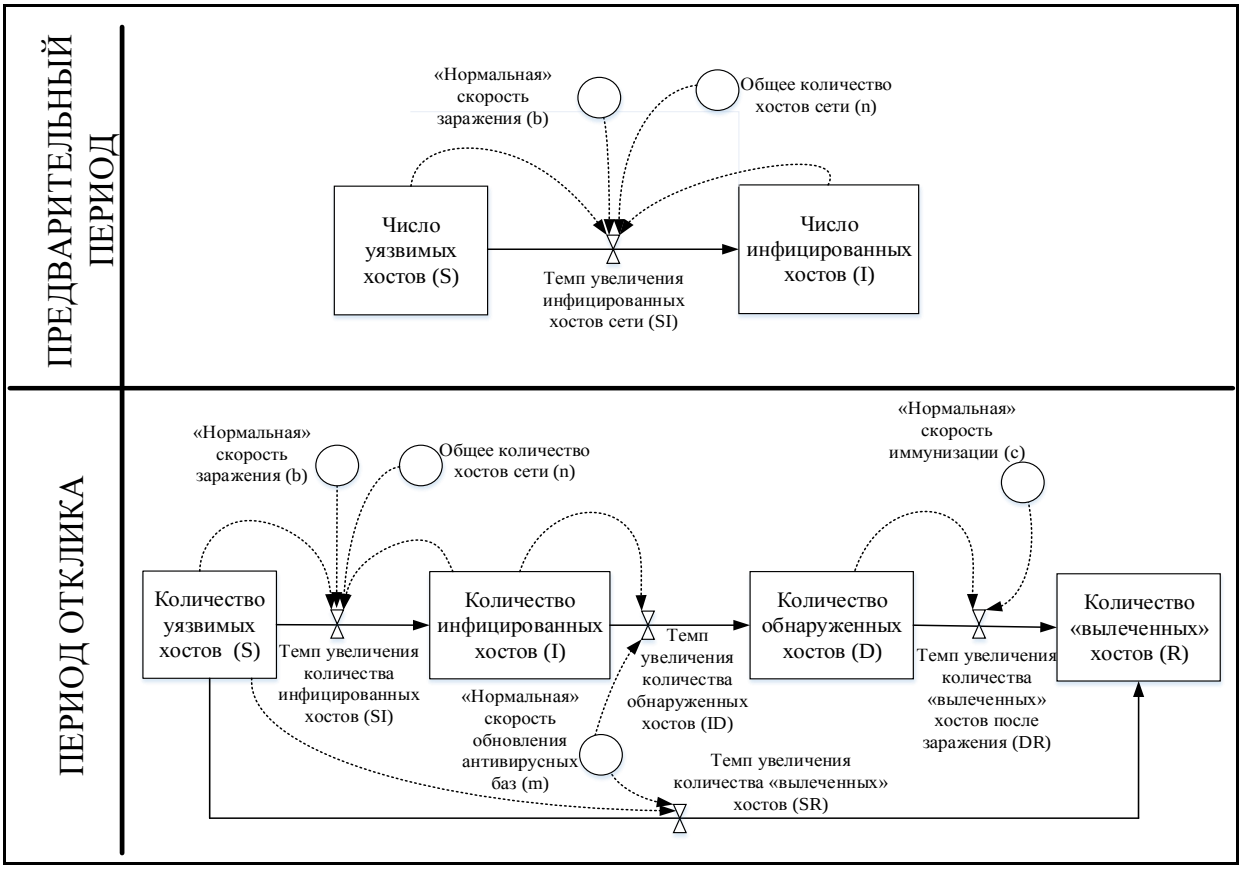


Рис. 4. Системно-динамическая модель распространения вируса по сети на основе PSIDR-модели

Приведём расшифровку обозначений, используемых в данной модели (табл. 2).

Таблица 2

Условные обозначения, используемые в модели

Условное обозначение элемента	Название элемента
n	Общее количество хостов сети
S	Количество уязвимых хостов
I	Количество инфицированных хостов
D	Количество обнаруженных хостов
R	Количество "вылеченных" хостов
SI	Темп увеличения количества инфицированных хостов (*час*)
ID	Темп увеличения количества обнаруженных хостов (1/*час*)
DR	Темп увеличения количества "вылеченных" хостов после заражения (1/*час*)
SR	Темп увеличения количества "вылеченных" хостов (1/*час*)
b	"Нормальная" скорость заражения (*часть/час*)
c	"Нормальная" скорость иммунизации (*часть/час*)
m	"Нормальная" скорость обновления антивирусных баз (*часть/час*)

В предварительный период система описывается следующими уравнениями:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = -SI(t); \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = SI(t); \\ SI(t) = \dfrac{b \cdot S(t) \cdot I(t)}{n}. \end{cases}$$

В период отклика система начинает описываться следующими уравнениями:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = -SI(t) - SR(t); \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = SI(t) - ID(t); \\ \dfrac{\mathrm{d}D}{\mathrm{d}t} = ID(t) - DR(t); \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = DR(t) + SR(t); \\ SI(t) = \dfrac{b \cdot S(t) \cdot I(t)}{n}; \\ ID(t) = m \cdot I(t); \\ DR(t) = c \cdot D(t); \\ SR(t) = m \cdot S(t). \end{cases}$$

Реализуем распространение вируса по сети на основе PSIDR-модели в программе Anylogic. Приведём общий вид интерфейса модели в предварительный период (рис. 5) и в период отклика (рис. 6).

Проведём два имитационных эксперимента, в которых исследуется влияние "нормальной" скорости заражения, "нормальной" скорости обновления антивирусных баз и "нормальной" скорости иммунизации сети на количество уязвимых, инфицированных, обнаруженных и "вылеченных" хостов сети.

Задачей данного эксперимента является исследование модели во время двух возможных периодов: предварительный период и период отклика.

Построим матрицу планирования экспериментов (табл. 3).

Таблица 3

Матрица планирования экспериментов

№ эксперимента	Значение фактора		Значение отклика			
	m	***c***	***S***	***I***	***D***	***R***
1	0	0	$S_1(t)$	$I_1(t)$	$D_1(t)$	$R_1(t)$
2	0,04	0,03	$S_2(t)$	$I_2(t)$	$D_2(t)$	$R_2(t)$

Начальные значения других параметров модели определим следующим образом: $S(0) = n = 1000$; $I(0) = 2$; $R(0) = 0$; $D(0) = 0$; $b = 0{,}15$.

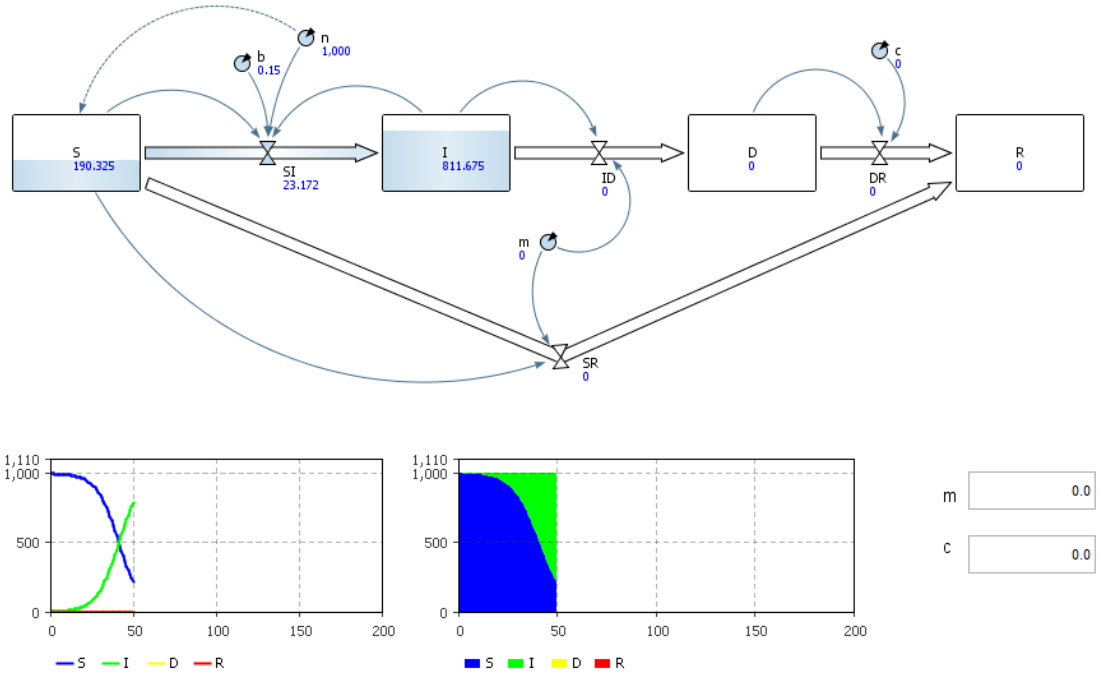


Рис. 5. Общий вид интерфейса системно-динамической модели распространения вируса по сети на основе PSIDR-модели в предварительный период

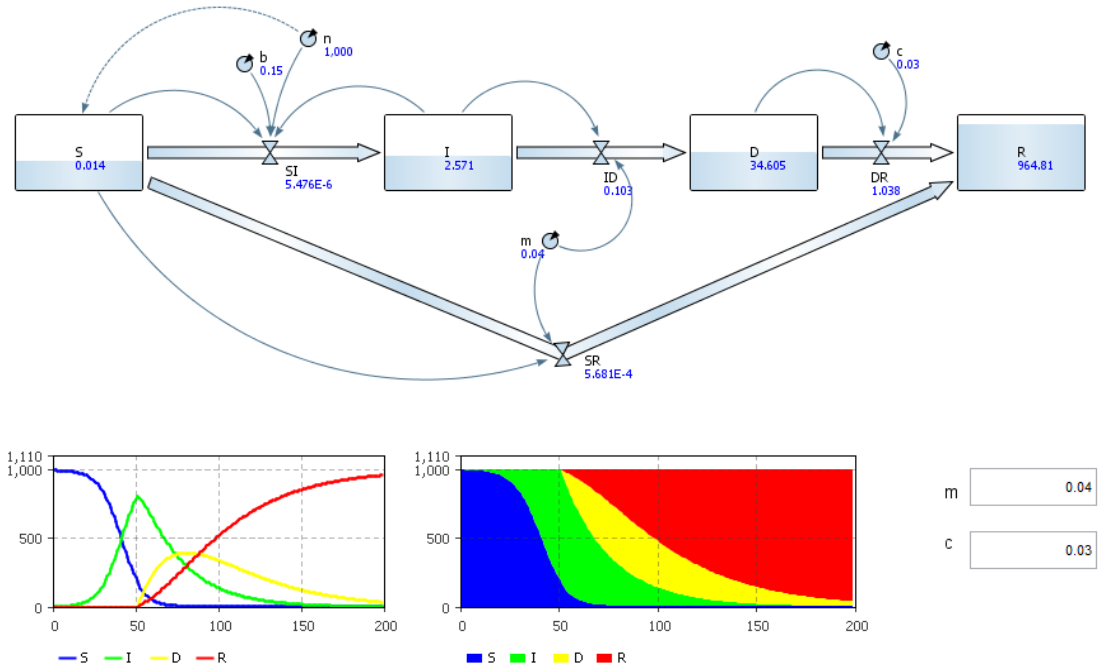
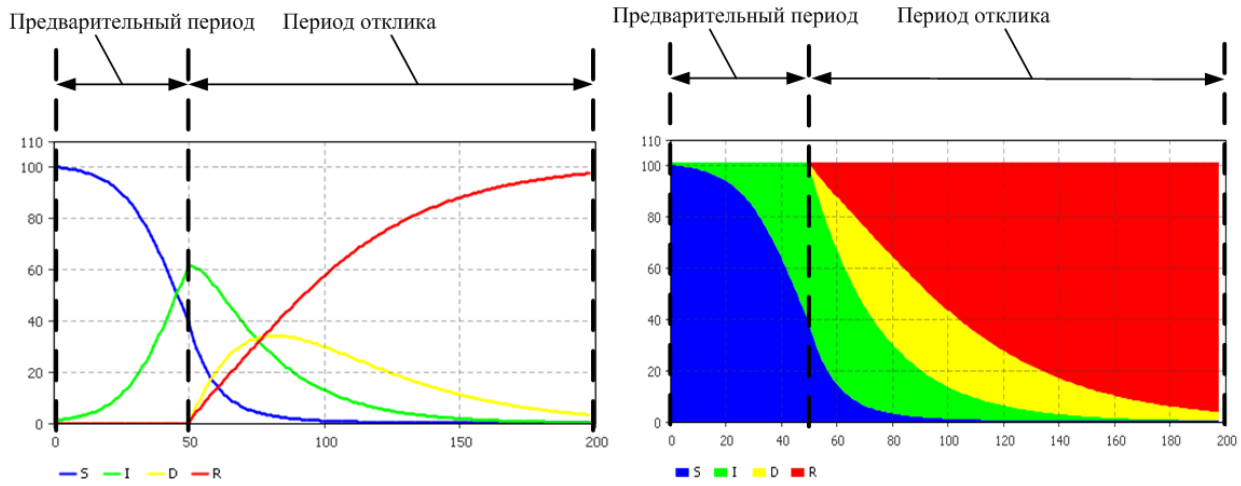


Рис. 6. Общий вид интерфейса системно-динамической модели распространения вируса по сети на основе PSIDR-модели в период отклика

Приведённые временные графики и временная диаграмма с накоплением (рис. 7) отражают динамику уязвимых, инфицированных, обнаруженных и "вылеченных" хостов сети во время двух периодов работы модели. Первый – предварительный период (при $t < \tau$) отражает только динамику уязвимых и инфицированных хостов сети, так как параметры m – "нормальная" скорость обновления антивирусных баз и c – "нормальная" скорость иммунизации в этот период равны нулю. Второй – период отклика (при $t \geq \tau$) отражает уже динамику всех состояний системы. В данном эксперименте длительность предварительного периода составила $\tau = 50$ *ч*.



а) *б*)

Рис. 7. Временные графики (*а*) и временная диаграмма с накоплением (*б*)

Выводы

1. Применение методов системно-динамического моделирования, предоставляет широкий спектр возможностей для исследования вирусных эпидемий в компьютерных сетях: решать задачи управления эпидемиями, прогнозирования их течения, определения оптимальных параметров противодействия и др.

2. Системно-динамические модели распространения компьютерных вирусов по сети, основывающиеся на эпидемических SEIR и PSIDR моделях, позволяют исследовать динамику "заражения" сети и выявлять степень влияния наиболее критичных факторов. Реализация построенных моделей в программной среде Anylogic позволяет наглядно отображать эпидемии компьютерных вирусов при различных значениях параметров модели.

3. Проведённые имитационные эксперименты позволяют прогнозировать динамику числа уязвимых, инфицированных, латентных, обнаруженных и "вылеченных" хостов сети в зависимости от различных значений параметров моделей, а также определять оптимальные значения этих параметров при заданных ограничениях характеристик распространения вирусов.

Литература

1. Концепция развития Интернет-ресурсов МЧС России до 2018 года. http://volga.mchs.ru/upload/site8/document_file/sAsGujDpTE.pdf.

2. Котенко И.В., Воронцов В.В. Аналитические модели распространения сетевых червей // Труды СПИИРАН. СПб.: Наука. 2007. С. 208-224.

3. Захарченко А. Черводинамика: причины и следствия // Защита информации. Конфидент. 2004. № 2. С. 50-55.

4. Семыкина Н.А., Шавыкина И.В. Математическая модель защиты компьютерной сети от вирусов // Программные продукты и системы / Software&Systems. 2016. Т. 29. № 4. С. 125-128.

5. Семенов С.Г., Давыдов В.В. Математическая модель распространения компьютерных вирусов в гетерогенных компьютерных сетях автоматизированных систем управления технологическим процессом // НТУ "ХПИ". Серия: Информатика и моделирование. Харьков. 2012. № 38. С. 163-171.

6. Аверенков В.И., Федоров В.П., Хейфец М.Л. Основы математического моделирования технических систем. М., 2011. 271 с.

7. Форрестер Д. Основы кибернетики предприятия (индустриальная динамика). М.: Прогресс, 1971. 340 с.

8. Тесалова О.Т., Минаев В.А., Кононенко В.И., Новиков Н.Ф. Моделирование динамики заболеваемости сифилисом с учётом влияния на параметры деятельности системы здравоохранения. 1983. № 2. С. 39-44.

9. Kephart J.O., White S.R. Directed-graph epidemiological models of computer viruses // Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy. Oakland, California. 1991. Pp. 343-359.

10. Минаев В.А., Сычев М.П., Вайц Е.В., Грачёва Ю.В. Математическая модель "хищник – жертва" в системе информационной безопасности // Информация и безопасность. 2016. Т. 19. № 3 (4). С. 397-400.

References

1. Kontseptsiia razvitiia Internet-resursov MChS Rossii do 2018 goda (The concept of development of Internet resources of EMERCOM of Russia until 2018). http://volga.mchs.ru/upload/site8/document_file/sAsGujDpTE.pdf.

2. Kotenko I.V., Vorontsov V.V. Analiticheskie modeli rasprostraneniia setevykh chervei (Analytical model of propagation of network worms) // Trudy SPIIRAN. SPb.: Nauka. 2007. Pp. 208-224.

3. Zakharchenko A. Chervodinamika: prichiny i sledstviia (Computer worms dynamics: causes and consequences) // Zashchita informatsii. Konfident. 2004. No 2. P. 50-55.

4. Semykina N.A., Shavykina I.V. Matematicheskaia model zashchity kompiuternoi seti ot virusov (Mathematical model of protecting computer networks from viruses) // Programmnye produkty i sistemy / Software&Systems. 2016. T. 29. No 4. Pp. 125-128.

5. Semenov S.G., Davydov V.V. Matematicheskaia model rasprostraneniia kompiuternykh virusov v geterogennykh kompiuternykh setiakh avtomatizirovannykh sistem upravleniia tekhnologicheskim protsessom (The mathematical model of the spread of computer viruses in heterogeneous computer networks of automated control systems of technological process) // NTU "KhPI". Seriia: Informatika i modelirovanie. Kharkov. 2012. No 38. Pp. 163-171.

6. Averenkov V.I., Fedorov V.P., Kheifets M.L. Osnovy matematicheskogo modelirovaniia tekhnicheskikh sistem (Foundations of mathematical modeling of technical systems). M., 2011. 271 p.

7. Forrester D. Osnovy kibernetiki predpriiatiia (industrialnaia dinamika) (Fundamentals of Cybernetics of the enterprise (industrial dynamics)). M.: Progress, 1971. 340 p.

8. Tesalova O.T., Minaev V.A., Kononenko V.I., Novikov N.F. Modelirovanie dinamiki zabolevaemosti sifilisom s uchetom vliianiia na parametry deiatelnosti sistemy zdravookhraneniia (Modeling the dynamics of syphilis morbidity taking into account the influence on the parameters of the health system). 1983. No 2. Pp. 39-44.

9. Kephart J.O., White S.R. Directed-graph epidemiological models of computer viruses (Directed-graph epidemiological models of computer viruses) // Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy. Oakland, California. 1991. Pp. 343-359.

10. Minaev V.A., Sychev M.P., Vaits E.V., Gracheva Iu.V. Matematicheskaia model "khishchnik – zhertva" v sisteme informatsionnoi bezopasnosti (Mathematical model of the "predator – prey" system of information security) // Informatsiia i bezopasnost. 2016. T. 19. No 3 (4). Pp. 397-400.