УДК 004.49.5

В.В. Давыдов

Национальный технический университет «ХПИ», Харьков

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ РАСПРОСТРАНЕНИЯ КОМПЬЮТЕРНЫХ ВИРУСОВ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМ ПРОЦЕССОМ

Проведено сравнительное исследование математических моделей распространения компьютерных вирусов, выявлены их основные особенности, достоинства и недостатки. Построены сравнительные графики зависимости количества зараженных узлов от времени функционирования компьютерной системы при распространении эпидемии. Сделан вывод о целесообразности использования модели PSIDR как основу новых разработок и усовершенствования модели распространения компьютерных вирусов в автоматизированных системах управления технологическими процессами с учетом связности и топологии сети.

Ключевые слова: *автоматизированная система управления технологическим процессом, компьютерные вирусы, модель PSIDR.*

Введение

Постановка задачи. Угроза заражения компьютерных систем злоумышленным программным обеспечением (ПО) (например, компьютерными вирусами) представляет ощутимую угрозу не только персональным ЭВМ, но и компьютерным системам, используемым в различных областях жизнедеятельности. Особую опасность злоумышленное ПО представляет процессу функционирования компьютерных систем критического применения (систем атомных станций, управления транспортом, управления безпилотными летательными аппаратами и др.).

Существующие средства защиты автоматизированных систем управления технологическими процессами (АСУТП) не всегда оперативно справляются с эпидемиями компьютерных вирусов. Поэтому разработка и внедрение новых средств обнаружения и защиты систем управления на производстве, способных предотвратить или сдержать эпидемию на ранних стадиях является актуальной научной задачей.

Анализ литературы [1, 2, 4] показал, что в настоящее время существует множество подходов математического моделирования компьютерных систем, основой которых являются теории связи, массового обслуживания, нейронных сетей, нечеткой логики и др.

Однако, среди всего многообразия математических моделей для описания компьютерных вирусов чаще всего используют биологические подходы моделирования [3].

В данной статье предлагается провести анализ и сравнительное исследование существующих биологических моделей распространения компьютерных вирусов.

Основная часть

В настоящее время известно несколько разновидностей математических моделей распространения компьютерных вирусов, разработанных на основе биологических подходов, отличаются между собой областью ограничения и условиями применения в реальных технических системах. Среди них можно выделить следующие модели:

SI (Suspected-Infected),

SIR (Suspected-Infected-Recovered),

SEIQR (Suspected-Exposed- Infected-Quarantined-Recovered),

PSIDR (Progressive Suspected-Infected-Detected-Recovered).

Проведенные исследования показали, что существующие модели распространения компьютерных вирусов на данный момент имеют ряд недостатков:

1) не учитывают связность компьютерной сети;

2) не учитывают временные задержки как внутри каждой компьютерной локальной сети, так и на «мостах».

1. Модель SI. Модель SI [1, 2] характеризуется наличием двух типов объектов управления: зараженные (I) и не зараженные (S).

Характерной особенностью SI модели является пренебрежение антивирусным ПО, что приводит к необратимости эпидемического процесса в компьютерных системах.

Обобщенная структура компьютерной системы на основе модели SI может быть представлена с помощью выражения:

$$N = S(t) + I(t), \qquad (1)$$

где S(t) – количество уязвимых объектов; I(t) – количество зараженных объектов

Динамическое изменение характеристик данной модели описывается с помощью системы:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\dfrac{\beta I(t)}{N} S(t); \\ \dfrac{dI(t)}{dt} = \dfrac{\beta I(t)}{N} S(t); \\ \dfrac{dI(t)}{dt} + \dfrac{dS(t)}{dt} = 0; \end{cases} \quad (2)$$

$$\beta = V_s \cdot \frac{N}{N_{ip}}, \quad (3)$$

где β – частота заражения; N_{ip} –размер общего адресного пространства (для сетей IPv4 $N_{ip} = 2^{32}$); V_s – средняя скорость сканирования вирусом сети; N – количество заражаемых компьютеров

Кривые зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SI представлены на рис. 1.

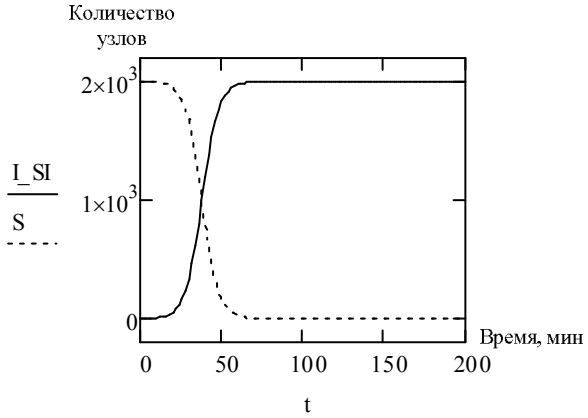

Рис. 1. График зависимости изменения количества зараженных и не зараженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SI, при коэффициенте заражения $\beta = 0.2$

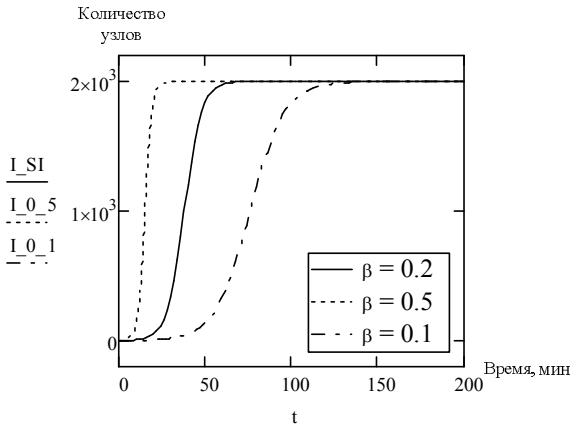

Рис. 2. Графики кривых зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SI при варьировании коэффициента заражения β

В связи с отсутствием антивирусного ПО, как фактора, влияющего на процесс распространения компьютерных вирусов, эпидемия не может угаснуть.

Кривые зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SI при разных коэффициентах заражения β представлены на рис. 2.

Анализ графиков на рис. 2 показал, что коэффициент заражения β прямо-пропорционально влияет на скорость распространения вируса.

Проведенные исследования модели SI показали, что введение первоначальных ограничений не позволяет адекватно описать процесс распространения в реальных компьютерных системах с гибридной топологией.

Кроме того, пренебрежение наличием антивирусного ПО ограничивает применение математической модели SI стадией заражения.

2. Модель SIR. Модель SIR [1, 2] характеризуется наличием трех типов объектов управления: зараженные (I), не зараженные (S), вылеченные объекты, обладающие иммунитетом (R).

Обобщенная структура компьютерной системы на основе модели SIR может быть представлена с помощью выражения:

$$N = S(t) + I(t) + R(t), \quad (4)$$

где S(t) – количество уязвимых объектов; I(t) – количество зараженных объектов; R(t) – количество вылеченных объектов, обладающие иммунитетом.

Динамическое изменение характеристик данной модели описывается с помощью системы:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\dfrac{\beta I(t)}{N} S(t); \\ \dfrac{dI(t)}{dt} = \dfrac{\beta I(t)}{N} S(t) - \delta I(t); \\ \dfrac{dR(t)}{dt} = \delta I(t); \\ \dfrac{dS(t)}{dt} + \dfrac{dI(t)}{dt} + \dfrac{dR(t)}{dt} = 0, \end{cases} \quad (5)$$

где β – частота заражения; δ – частота лечения, «скорость иммунизации»

Данная модель подразумевает что эпидемия возможна лишь при $\beta > \delta$.

Кривые зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SIR представлены на рис. 3.

Кривые зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели *SIR* при разных коэффициентах заражения β и лечения δ представлены на рис. 4.

Для анализа графиков рис. 4, введем коэффициент γ:

$$\gamma = \frac{\beta}{\delta}. \qquad (6)$$

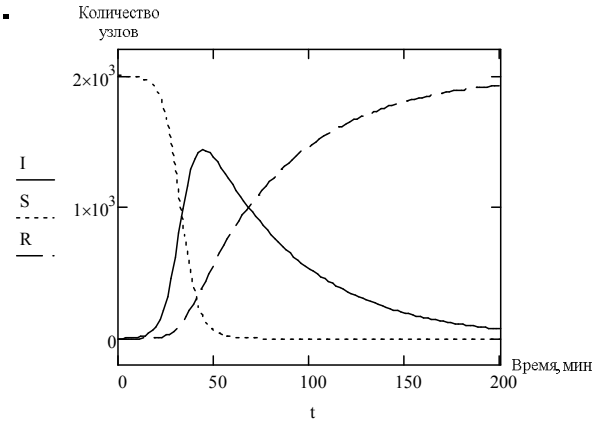

Рис. 3. Графики зависимости изменения количества зараженных, незараженных и вылеченных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SIR

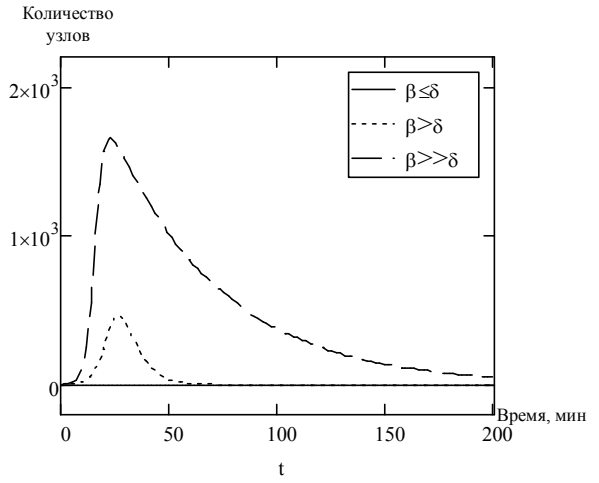

Рис. 4. Графики зависимостей изменения количества зараженных узлов от времени в условиях распространения эпидемии по модели SIR при варьировании коэффициента заражения и лечения

Анализ зависимости изменения количества заражения узлов от времени функционирования компьютерной системы показал:

1) при $\gamma \le 1$ эпидемии не происходит;

2) увеличение коэффициента γ в 10 раз приводит к увеличению количества зараженных объектов в 5 раз, а длительность эпидемии увеличивается примерно в 2 раза.

Проведенные исследования модели SIR показали, что введение дополнительного типа объекта управления и учет возможного фактора лечения, позволило повысить точность конечного результата в условиях наличия обновляемого антивирусного ПО.

Однако проблемы, связанные с отсутствием учета топологических особенностей компьютерных сетей в данной модели не устранены. Кроме того, в реальных условиях для лечения компьютерных систем существует необходимость идентификации и локализации злоумышленного ПО. Данная процедура требует определенных (от доли секунды до десятка часов) временных затрат. Данный фактор в модели SIR не учитывается, что также снижает область применения данной модели.

3. SEIQR модель. Модель SEIQR [4] модель является развитием SIR модели. Она характеризуется наличием 5 типов объектов управления:

зараженные (I),

не зараженные (S),

вылеченные объекты, обладающие иммунитетом (R),

зараженные, но не распространяющие инфекцию (E),

объекты, находящиеся в карантине (Q).

Во время инкубационного периода злоумышленное ПО не наносит вреда инфицированному узлу (вирус находится в латентном состоянии и не способен заражать другие объекты).

Обобщенная структура компьютерной системы на основе модели SEIQR может быть представлена с помощью выражения:

$$N = S(t) + E(t) + I(t) + R(t) + Q(t), \qquad (7)$$

где S(t) – количество уязвимых объектов; E(t) – количество объектов, зараженных вирусом, но не распространяющих инфицирование (латентный период вируса); I(t) – количество зараженных объектов; R(t) – количество вылеченных объектов, обладающие иммунитетом; Q(t) – количество объектов в карантине.

Динамическое изменение характеристик данной модели описывается с помощью системы:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\dfrac{\beta I(t)}{N} S(t); \\ \dfrac{dE(t)}{dt} = \dfrac{\beta I(t)}{N} S(t) - (\alpha + k) E(t); \\ \dfrac{dI(t)}{dt} = \alpha E(t) - (\gamma + \delta) I(t); \\ \dfrac{dQ(t)}{dt} = \delta I(t) - \gamma Q(t); \\ \dfrac{dR(t)}{dt} = kE(t) + \gamma (Q(t) + I(t)), \end{cases} \qquad (8)$$

где α – коэффициент перехода вируса в латентное состояние; k, γ – коэффициенты лечения; δ – коэффициент карантирования; β – коэффициент заражения.

Схема переходов различных объектов системы представлена на рис. 5.

Кривые зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SEIQR представлены на рис. 6.

Проведенные исследования модели SEIQR показали, что введение дополнительных типов объекта управления и учет возможного фактора введения в карантин, позволило повысить точность конечного результата в условиях наличия обновляемого антивирусного ПО.

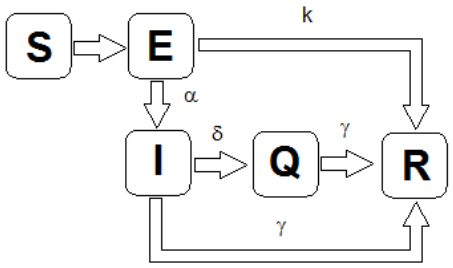

Рис. 5. Схематическая диаграмма переходов объектов системы SEIQR

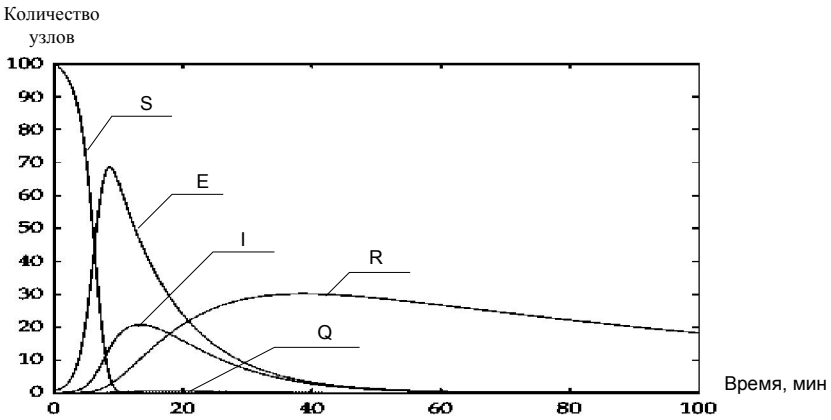

Рис. 6. Графики кривых зависимости изменения количества зараженных, незараженных, вылеченных, карантинных, подверженных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели SEIQR

Однако проблемы, связанные с отсутствием учета топологических особенностей компьютерных сетей в данной модели не устранены. Кроме того, в реальных условиях систем АСУТП введение карантина неприменимо, т.к. требует наличия «холодного» резерва, что влечет за собой дополнительные финансовые затраты.

Данный фактор в модели SEIQR не учитывается, что также снижает область применения данной модели, исключая ее использование на системах управления конвейерными технологическими процессами.

4. Модель PSIDR. Модель PSIDR [5] характеризуется наличием 4-х типов объектов управления: зараженные (I), не зараженные (S), вылеченные объекты, обладающие иммунитетом (R) и найденные зараженные объекты (D).

Поведение системы в условиях воздействия злоумышленного ПО, описанное с помощью модели PSIDR, предполагает 2 этапа:

1) (аналог SI модели) заражение объектов (с частотой β). Изначально злоумышленное ПО инфицирует один объект в сети. В течении определенного времени злоумышленное ПО распространяется незаметно;

2) (по истечении времени t_d) добавляется лечение объектов (с частотой δ).

3) По истечении времени t_d, вирус обнаруживается. Осуществляется выделение его сигнатур и внесение их в базы антивирусного ПО. Неинфицированные узлы на данный момент становятся невосприимчивы к вирусу, а инфицированные – излечиваются.

Динамическое изменение характеристик данной модели описывается с помощью системы:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\beta S(t)I(t) - \mu S(t); \\ \dfrac{dI(t)}{dt} = \beta S(t)I(t) - \mu I(t); \\ \dfrac{dR(t)}{dt} = \delta D(t) + \mu S(t); \\ \dfrac{dD(t)}{dt} = \mu I(t) - \delta D(t); \\ \dfrac{dS(t)}{dt} + \dfrac{dI(t)}{dt} + \dfrac{dR(t)}{dt} + \dfrac{dD(t)}{dt} = 0, \end{cases} \quad (9)$$

где β - частота заражения; δ – частота лечения; μ – вероятность вылечивания (переход в состояние R); S(t) – количество уязвимых объектов; I(t) – количество зараженных объектов; R(t) – количество вылеченных (с иммунитетом) объектов (на первой стадии = 0); D(t) – количество обнаруженных зараженных объектов (на первой стадии = 0).

Кривые зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях угасания эпидемии (второй этап) по модели PSIDR представлена на рис. 7.

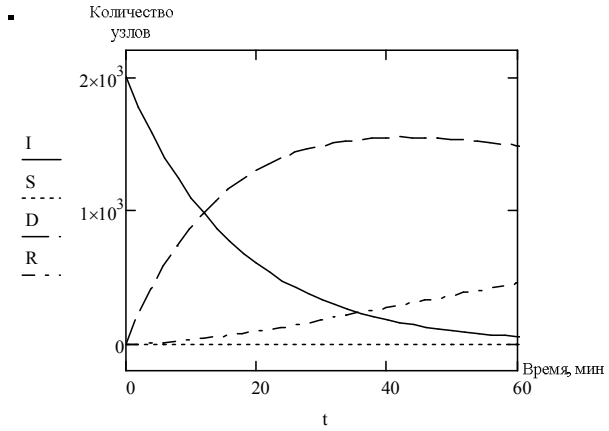

Рис. 7. Графики зависимости изменения количества зараженных, незараженных, вылеченных узлов от времени функционирования компьютерной системы в условиях угасания эпидемии (второй этап) по модели PSIDR

На рис. 8 представлена кривые зависимости изменения количества зараженных узлов от времени функционирования компьютерной системы в условиях как распространения так и угасания эпидемии (оба этапа) по модели PSIDR .

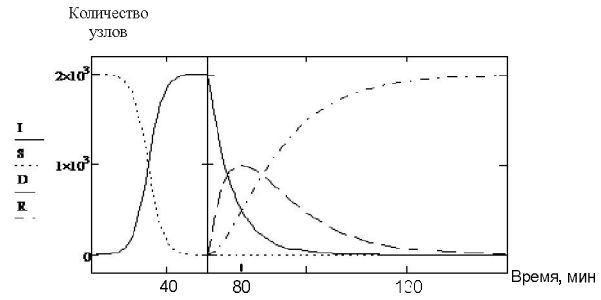

Рис. 8. Графики зависимости изменения количества зараженных, незараженных, вылеченных узлов от времени функционирования компьютерной системы в условиях распространения эпидемии по модели PSIDR

Проведенный анализ модели PSIDR показал, что разбиение модели распространения компьютерных угроз на два этапа дает возможность независимого анализа процесса заражения и лечения. Введение задержки между началами этих двух этапов, идентификации, локализация и лечение злоумышленного ПО, позволило устранить один из недостатков модели SIR.

Выводы

Анализ исследуемых моделей показал, что все модели имеют схожий недостаток – отсутствия учета связности и топологии сети. В дальнейшем предполагается усовершенствовать модель PSIDR, добавив в нее фактор связности сети, так как она наиболее адекватно описывает процесс распространения эпидемии в автоматизированных системах управления технологическими процессами: выделяет отдельно процесс распространения и процесс лечения с задержкой во времени.

Список литературы

1.Котенко И.В. Аналитические модели распространения сетевых червей / И.В. Котенко, В.В. Воронцов // Труды СПИИРАН. – СПб.: Наука, 2007. – Вып. 4.

2.Rohloff K. Stochastic Behavior of Random Constant Scanning Worms / K. Rohloff , T. Basar // Computer Communications and Networks, 2005. ICCCN 2005. Proceedings. 14th International Conference on 17-19 Oct. 2005. – P. 339-344.

3.Cohen F. Computer viruses, theory and experiments / F. Cohen // Computers & Security. – 1987. – Vol. 6. – P. 22-35.

4.SEIQR-SIS Epidemic Network Model and its stability / W. Jumpen, S. Orankitjaroen, P. Boonkrong, B. Wiwatanapataphe // International journal of mathematics and computers in simulation. – 2011. – Issue 4, Vol. 5. – P. 326-333.

5.Williamson M. Epidemiological model of virus spread and cleanup [Электронный ресурс] / M. Williamson, J. Leveille // HP Laboratories Bristol (February 27th, 2003).

Поступила в редколлегию 1.03.2012

Рецензент: канд. физ.-мат. наук, с.н.с. А. А. Можаев, Национальный технический университет «ХПИ», Харьков.

ПОРІВНЯЛЬНИЙ АНАЛІЗ МОДЕЛЕЙ РОЗПОВСЮДЖЕННЯ КОМП'ЮТЕРНИХ ВІРУСІВ В АВТОМАТИЗОВАНИХ СИСТЕМАХ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМ ПРОЦЕСОМ

В.В. Давидов

Проведено порівняльне дослідження математичних моделей розповсюдження комп'ютерних вірусів, виявлені їх основні особливості, достоїнства і недоліки. Побудовані порівняльні графіки залежності кількості заражених вузлів від часу функціонування комп'ютерної системи при розповсюдженні епідемії. Зроблений висновок про доцільність використання моделі PSIDR як основу нових розробок і удосконалення моделі розповсюдження комп'ютерних вірусів в автоматизованих системах управління технологічними процесами з урахуванням зв'язності і топології мережі.

Ключові слова: *автоматизована система управління технологічним процесом, комп'ютерні віруси, модель PSIDR.*

COMPARATIVE ANALYSIS OF MODELS OF DISTRIBUTION OF COMPUTER VIRUSES IS IN THE AUTOMATED TECHNOLOGICAL PROCESS CONTROL SYSTEMS

V.V. Davydov

Comparative research of mathematical models of distribution of computer viruses is conducted, their basic features, dignities and failings, are exposed. The comparative graphs of dependence of amount of the infected knots are built от time of functioning of the компьютерной system at distribution of epidemic. A conclusion is done about expedience of the use of model of PSIDR as basis of new developments and improvement of model of distribution of computer viruses in the automated control of technological processes systems taking into account coherentness and topology of network.

Keywords: *automated technological process control system, computer viruses, model of PSIDR.*