УДК 004.056.57

ДОСЛІДЖЕННЯ МОДЕЛЕЙ ПОШИРЕННЯ ШКІДЛИВОГО КОДУ В КОМП'ЮТЕРНО-ІНТЕГРОВАНИХ СИСТЕМАХ УПРАВЛІННЯ

В.Ю. Шадхін, Д.Г. Дель, Я.Ю. Піддубна, І.Ю. Рогальський

Київський національний університет технологій та дизайну

У статті розглядаються основні моделі розповсюдження комп'ютерних вірусів в комп'ютерно-інтегрованих системах управління.

Ключові слова: *модель розповсюдження комп'ютерних вірусів, комп'ютерно-інтегровані системи управління, SIR модель, детермінована модель, стохастична модель.*

Одним із основних напрямків розвитку сучасних систем автоматизації є створення інтегрованих систем управління виробництвом, які вирішують задачу інтеграції традиційних АСУТП і АСУП. Серед основних проблем створення інтегрованої системи управління на підприємстві є забезпечення безпеки, в тому числі захист від шкідливого коду.

Однак для дієвого захисту необхідно дослідити моделі поширення шкідливого коду в комп'ютерно-інтегрованих системах управління.

Постановка завдання

Протидіяти поширенню шкідливих кодів досить складна задача, яка має безліч аспектів. Одним з таких аспектів є моделювання розповсюдження шкідливих епідемій. Математичне моделювання є важливим інструментом для моделювання, отримання результатів, розуміння поширення шкідливого коду.

За допомогою створення математичних моделей розповсюдження шкідливих кодів і подальшого моделювання на базі цих моделей еволюції шкідливих програм можна оцінити масштаби можливої загрози, вивчити динаміку зміни числа заражених комп'ютерів і т.д. Крім того, результати чисельного моделювання можна використовувати для оцінки ефективності різних заходів протидії поширенню шкідливих програм.

Об'єкт та методи досліджень

Моделі розповсюдження комп'ютерних вірусів в комп'ютерно-інтегрованих системах управління. Статистичні дані про поширення шкідливих програм, математичні епідеміологічні моделі і результати, отримані з їх допомогою.

Результати досліджень та їх обговорення

Існує декілька моделей дослідження поширення шкідливого коду в комп'ютерно-інтегрованих системах управління, а саме: детермінована модель та стохастична модель. Детермінована модель (протилежна моделі стохастичній) — математична модель, параметри і змінні якої залежать один від одного функціонально, тобто не підпорядковані випадковим коливанням процесу, в зв'язку з чим характер системи в будь-який час повністю визначається умовами обраними спочатку.

Стохастичні моделі використовуються, коли ймовірність коливання параметрів або врахування різнорідності є важливими як у малій, так і в ізольованій мережі. Стохастичні моделі дають змогу перевірити кожен комп'ютер в мережі на ймовірнісній основі. Але вони можуть бути дуже трудомісткими й потребують багато симуляції для того, щоб отримати корисні прогнози. Тим не менше, врахування можливості зміни у процесах передачі вірусу забезпечує певний діапазон можливих результатів. Ці моделі можуть бути математично дуже складними і не надають пояснення поширення шкідливого коду. Перш ніж переходити до розгляду детермінованих моделей, важливо розуміти, яким чином виникає епідемія шкідливого забезпечення в комп'ютерних мережах. Для комп'ютерних вірусів кількість вразливих комп'ютерів зменшується із часом. Перед першим спалахом поширення шкідливого коду частка вразливих у мережі є 100%. А отже, частка заражених та тих, яких вилучили, є 0%. Коли починається епідемія, то кількість вразливих зменшується, а кількість заражених збільшується.

У стохастичної моделі, швидкість збільшення кількості антивірусів пов'язана з числом вірусів, вже існуючих в даний момент в мережі. Для SIR моделі швидкість появи антивірусів не залежить від числа наявних вірусів, зміна їх кількості має постійну швидкість. Крім того, стохастична модель, на відміну від SIR моделі дозволяє пророкувати результат, при якому в даній мережі може виникнути новий спалах епідемії, так як частина комп'ютерів залишається незахищеною антивірусом, що дозволяє отримати більш точний результат при моделюванні процесу поширення шкідливого коду.

Розглянемо спочатку епідемію найпростішого виду, тобто випадок, коли шкідливе програмне забезпечення розповсюджується серед групи сприйнятливих комп'ютерів, але видалення їх із мережі за рахунок виходу з ладу, виліковування або ізоляції не відбувається. Припустимо, що маємо n комп'ютерів, сприйнятливих до даного шкідливого програмного забезпечення і те, що в момент часу $t=0$ в групу попадає одне джерело вірусу.

Природно починати дослідження з детерміністської моделі, хоча можна вважати, що вона не дуже підійде для епідемії, що розпочинаються при невеликій кількості джерел вірусів, так як статистичні коливання в групі заражених комп'ютерів будуть відчутними навіть в тому випадку, коли значення *n* достатньо велике.

Кількість заражених та незаражених комп'ютерів розраховується за наступними формулами (1) та (2) відповідно. Графік зміни розподілу ймовірності зараження та незараження комп'ютерів та відповідні асимптотичні значення побудовано на основі формул (1) та (2) та зображено на рисунку 1.

$$p_0(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}, \quad (1)$$

$$p_1(t) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}, \quad (2)$$

де p_0 – кількість заражених ПК;

p_1 – кількість незаражених ПК;

λ – інтенсивність взаємодії між ПК;

μ – інтенсивність виліковування ПК.

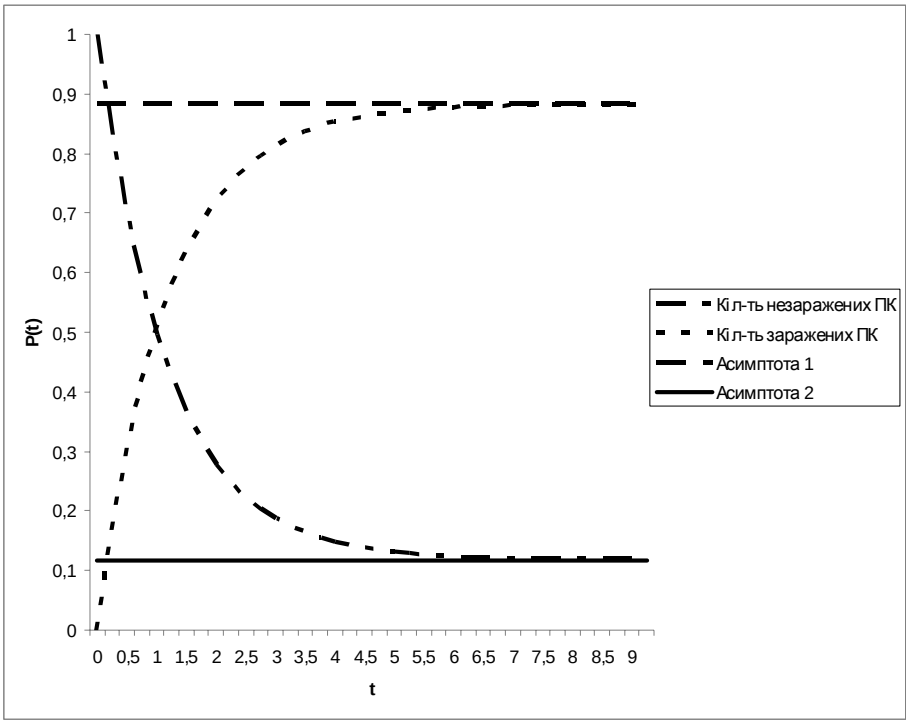

Рис. 1. **Графік зміни розподілу ймовірності зараження та незараження комп'ютерів та відповідні асимптотичні значення**

Розглянемо загальний випадок детермінованої моделі [1, 2]. Припустимо, що є група із n комп'ютерів, в якій в момент часу t є x сприйнятливих комп'ютерів, y джерел вірусу та z вилучених комп'ютерів. Таким чином, $x+y+z=n$. Припустимо, що частота контактів дорівнює β, таким чином середнє число нових випадків зараження в інтервалі Δt дорівнює $\beta xy\Delta t$. Слід також врахувати частоту випадків вилучення γ, тобто кількість комп'ютерів, які в інтервалі $\gamma y\Delta t$ покидають заражену групу.

Рівняння руху для цього процесу мають вигляд:

$$\left.\begin{aligned} \frac{dx}{dt} &= -\beta xy \\ \frac{dy}{dt} &= \beta xy - \gamma y \\ \frac{dz}{dt} &= \gamma y \end{aligned}\right\} \qquad (3)$$

При початковій умові $(x, y, z)=(x_0, y_0, 0)$ в момент $t=0$. Якщо вихідне число джерел вірусу, тобто y_0, дуже мале, то можна вважати, що $x_0 \approx n$. Із другого рівняння системи безпосередньо слідує, що епідемія не може починатися, якщо не виконується умова $x_0>\gamma/\beta$. Назвемо величину $\rho=\gamma/\beta$ відносною частотою вилучення заражених комп'ютерів з мережі; тоді пороговим значенням цієї величини буде значення $\rho= x_0 \approx n$. При щільності сприйнятливих комп'ютерів, що лежить нижче цього значення, початкові випадки зараження шкідливим кодом зникнуть раніше, чим вірус передасться іншим комп'ютерам. Але якщо ця щільність вище порогової, то епідемія виникає навіть в тому випадку, якщо початкове число джерел вірусу дуже невелике.

Хоча для системи рівнянь можна знайти точний розв'язок, в даному випадку можна скористуватися наближеними методами обчислень. Невідоме y можна виключити, розділивши перше рівняння на третє. В результаті маємо: $dx/dz=-x/\rho$, що після інтегрування дає:

$$x=x_0 e^{-z/\rho} \qquad (4)$$

Підставляючи в рівняння (3) системи значення $y=n-x-z$ та взявши x з формули, маємо

$$\frac{dz}{dt} = \gamma(n-x-z) = \gamma\left(n - z - x_0 e^{-\frac{z}{\rho}}\right) \qquad (5)$$

Розв'язуючи рівняння отримуємо $z(t)$, що розраховується за формулою (6).

$$z(t) = c\gamma t^2 + \frac{\gamma t(b\gamma^3 - 2c) + b\gamma^3 + 2c - a\gamma^2}{\gamma^3} + const\, e^{\gamma t} \quad (6)$$

де $a = \gamma n$, $b = x_0 \gamma \rho$ і $c = \frac{x_0 \gamma}{2}$ – константи.

Криву *z(t)*, зображену на рисунку 2, зручно розглядати як епідемічну криву для даної моделі, так як ізолюють тільки ті комп'ютери, у яких проявляються ознаки враження.

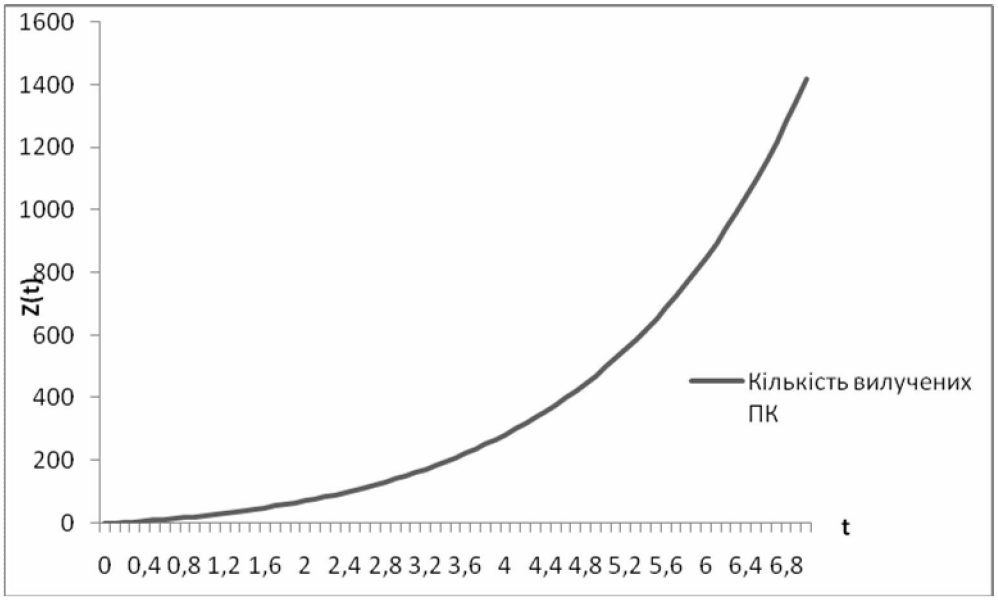

Рис. 2. **Кількість вилучених ПК**

Висновки

В ході дослідження поширення шкідливого коду в комп'ютерно-інтегрованих системах управління було зроблено висновок, що при використанні детермінованої моделі розповсюдження комп'ютерних вірусів, ми маємо стрімке зростання заражених до певної точки максимуму, після якого відбувається знижується до нуля. Стохастична модель не дає чітких значень, так як у неї є тільки один змінний параметр - це час, за який будуть заражені всі комп'ютери. Звідси слідує, що характер протікання епідемії піддається різким коливанням, обумовленим випадковими причинами, і в тих характерних випадках, коли епідемія розповсюджується дуже повільно або навпаки занадто швидко. Отже, детермінована модель дає більш точний опис поширення шкідливого коду, що дозволить вчасно застосувати заходи протидії її поширенню в комп'ютерно-інтегрованих системах управління.

ЛІТЕРАТУРА

1. Н.Т.Дж. Бейли; Математическая теория эпидемий; / Н.Т.Дж. Бейли – Хафнер 1957.
2. Шадхін В.Ю. Детерміністська модель розповсюдження комп'ютерних вірусів / Шадхін В.Ю., Копотій А.В., Бєлов К.Е., Костомаров О.В., Попов О.В. // Тези доповідей І Міжнародна науково-технічна конференція «Обчислювальний інтелект – 2011 (результати, проблеми, перспективи)». – Черкаси, 2011. – С. 495.

В.Ю. Шадхин, Д.Г. Дель, Я.Ю. Поддубная, И.Ю. Рогальский

Исследование моделей распространения вредоносного кода в компьютерно-интегрированных системах управления

В статье рассматриваются основные модели распространения компьютерных вирусов в компьютерно-интегрированных системах управления.

Ключевые слова: *модель распространения компьютерных вирусов, компьютерно-интегрированные системы управления, SIR модель, детерминированная модель, стохастическая модель.*

V. Y. Shadhin, D.H. Del, Y.Y. Piddubna, I.J. Rogalsky

The Investigation models the spread of malicious code in computer-integrated control systems

The article reviews the main models of the spread of computer viruses in computer-integrated control systems.

Keywords*: model the spread of computer viruses, computer integrated control system, SIR model, deterministic model, stochastic model.*

