УДК 004.491.2

О.С. САВЕНКО
Хмельницький національний університет

ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ МОДЕЛЮВАННЯ ПОШИРЕННЯ ВІРУСНИХ КОДІВ В ГЕТЕРОГЕННИХ МЕРЕЖАХ

Розроблено програмне забезпечення інформаційної технології моделювання поширення вірусних кодів в гетерогенних мережах для прогнозування часу і напрямку розповсюдження вірусних програм з врахуванням топології мережі та системного програмного забезпечення, встановленого на комп'ютерних системах. Застосування розробленого програмного забезпечення надає можливість передбачати ймовірність проникнення вірусів в комп'ютерні системи мережі із врахуванням рівня захищеності комп'ютерних систем та часу їх експлуатації.

Ключові слова: гетерогенна комп'ютерна мережа, комп'ютерний вірус, модель поширення вірусів.

O.S. SAVENKO
Khmelnytsky National University

THE SOFTWARE OF INFORMATION TECHNOLOGY FOR SIMULATION OF VIRAL CODES PROPAGATION IN THE HETEROGENEOUS NETWORKS

The software of information technology for simulation of viral codes propagation in the heterogeneous networks was developed. It allows to predict the time and the direction of spread of malwares and take in to account the network topology and system software. The developed software is based on the using of the most common models of the spread of viruses in the networks such as SI, SIS, SIR, PSIDR, SEIQR and Markov chains. Usage of the developed software makes it possible to predict the possibility of viruses penetrating in the network computer system with accounting the level of security of computer systems and the time of their operation.

Keywords: heterogeneous computer network, computer virus, model of viruses propagation.

Вступ

Прискорений розвиток інформаційних та мережних технологій створює сприятливі передумови для кіберзлочинності. Одним із широко застосовуваних засобів для здійснення кіберзлочинів в мережі Інтернет є комп'ютерні віруси, здатні створювати власні копії та поширюватись по каналам зв'язку з метою здійснення деструктивних дій: зниження продуктивності або пошкодження комп'ютерних систем (КС) та мережних пристроїв, перехоплення або пошкодження інформації на комп'ютерних системах тощо. Зростаючі масштаби кіберзлочинності підтверджують необхідність створення нових систем інформаційної безпеки.

Постановка задачі

З метою підвищення надійності роботи та інформаційної безпеки комп'ютерних мереж, що перебувають в експлуатації, а також на етапі їх проектування, важливим питанням є можливість прогнозування напрямків поширення вірусних кодів в комп'ютерних мережах. Тому, постає задача розробки програмного забезпечення (ПЗ) інформаційної технології моделювання поширення вірусних кодів в гетерогенних мережах, що надасть можливість враховувати характеристики окремих комп'ютерних систем мережі та топологію мережі.

Моделі розповсюдження вірусів в гетерогенних комп'ютерних мережах

Найбільш поширеними моделями розповсюдження вірусів в гетерогенних комп'ютерних мережах є наступні: SI (Suspected-Infected), SIS (Suspected-Infected-Suspected), SIR (Suspected-Infected-Recovered), PSIDR (Progressive Suspected-Infected-Detected-Recovered), SEIQR (Suspected-Exposed-Infected-Quarantined-Recovered) [1-10].

Модель SI не враховує топології мережі та може бути подана узагальненим виразом: $N = S(t) + I(t)$, де N – загальна кількість елементів в мережі; $S(t)$ – кількість вразливих об'єктів; $I(t)$ – кількість інфікованих об'єктів. За моделлю SI, поширення інфекції в мережі не може бути припинене, оскільки вона не враховує роботу антивірусного ПЗ. Внесення функції зв'язності з метою врахування топологічних характеристик мережі трансформує модель SI в SIT (SI-Topology).

Відмінністю моделі SIS є врахування можливості повернення вузла комп'ютерної мережі із інфікованого у вразливий стан, що уповільнює поширення вірусної інфекції в мережі.

Структура комп'ютерної мережі на основі моделі SIR в узагальненому вигляді може бути представлена наступним чином: $N = S(t) + I(t) + R(t)$, де $R(t)$ – кількість вилікованих об'єктів, які володіють імунітетом. За моделлю SIR вірусна епідемія можлива за умови, якщо частота інфікування більша за частоту виліковування. Врахування фактору лікування інфікованих вузлів дозволяє підвищити точність моделювання за наявності антивірусного ПЗ в мережі. Проте ця модель не враховує часових витрат, необхідних для ідентифікації та локалізації шкідливого програмного забезпечення (ШПЗ) з метою подальшого лікування вузла. Внесення функції зв'язності з метою врахування топологічних характеристик

мережі трансформує модель SIR в SIRT (SIR-Topology).

Модель PSIDR в узагальненому вигляді може бути представлена наступним виразом: $N = S(t) + I(t) + D(t) + R(t)$, де $S(t)$ – кількість не інфікованих об'єктів; $I(t)$ – кількість інфікованих об'єктів; $R(t)$ – кількість об'єктів, що володіють імунітетом; $D(t)$ – кількість знайдених інфікованих об'єктів. В моделі PSIDR виділяються два етапи, що уможливлюють незалежність аналізу процесів інфікування та лікування: інфікування об'єктів (ідентичний моделі SI) та лікування (за якого вилікувані вузли не інфікуються повторно). Врахування часової затримки між початками етапів інфікування та лікування, а також часових витрат на ідентифікацію, локалізацію та лікування інфікованих об’єктів дозволяє усунути відповідний недолік моделі SIRT та робить модель PSIDR найбільш придатною для моделювання поширення вірусів мережею. Внесення функції зв’язності з метою врахування топологічних характеристик мережі трансформує модель PSIDR в PSIDRT (PSIDR-Topology).

Відмінністю моделі SEIQR є введення латентного стану, за якого вузол інфікований, але ще не здатен інфікувати інші вузли, і з якого він може перейти до інфікованого стану.

Згідно моделі на основі ланцюгів Маркова, загальний стан мережі в момент часу *t* є сукупністю станів всіх вузлів мережі, та може бути описаний вектором з *N* елементів, де значення *i*-го елемента відповідає стану *i*-го вузла: *I* (infected, якщо вузол інфікований), і *S* (suspected, якщо вузол не інфікований). Стан мережі в наступний момент часу залежить від поточного стану мережі і не залежить від попередніх станів. Перевагою цієї моделі є досить потужне математичне забезпечення, яке дозволяє описати ймовірнісні процеси, що відбуваються в комп’ютерних мережах під час поширення вірусних кодів, і тому надає широкі можливості для аналізу та прогнозування вірусних епідемій в мережах. Основними недоліками моделі на основі ланцюгів Маркова є потреба у значних обчислювальних ресурсах, а також невизначеність певної частини ймовірностей переходів системи.

Програмне забезпечення ІТ моделювання поширення вірусних кодів в гетерогенних мережах

З метою прогнозування поширення вірусів в мережах розроблено інформаційну технологію моделювання поширення вірусних кодів в гетерогенних мережах, яка ґрунтується на використанні вищеописаних моделей поширення вірусів в мережі та ланцюгів Маркова [1–10]. Інформаційна технологія реалізована у вигляді програмного забезпечення, що надає можливість моделювати поширення атак на КС мережі, а саме: прогнозувати напрямки поширення атак; визначати оптимальний шлях від інфікованої КС до КС, на яку спрямовано загрозу; визначати часові характеристики поширення ШПЗ мережею. ПЗ надає можливість здійснювати налаштування характеристик комп’ютерної мережі, зокрема її топології. Процес моделювання відбувається дискретно з покроковим відображенням результатів, допоки не буде досягнуто мети, або до завершення заданої кількості кроків.

UML діаграма варіантів використання розробленої ІТ надана на рис. 1.

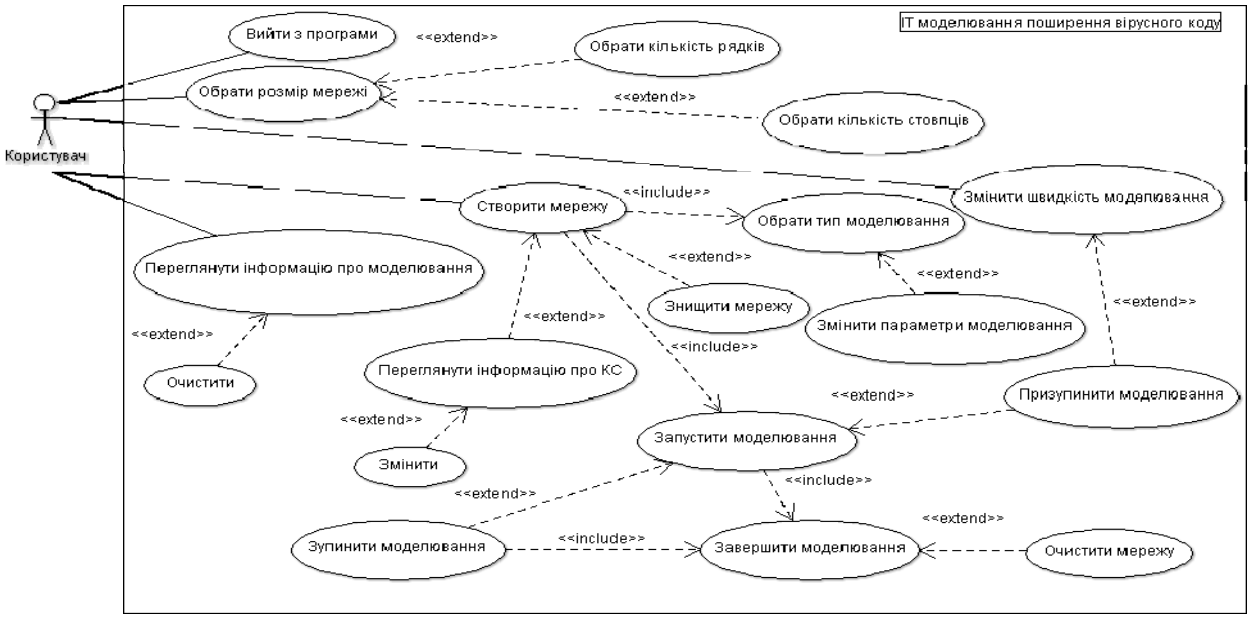


Рис. 1. Діаграма варіантів використання ІТ моделювання поширення вірусних кодів в гетерогенних мережах

Функціонування комп’ютерної мережі розміром n×m моделюється за допомогою взаємодії множини спеціальних класів. Об’єкт класу Comp характеризується властивостями, що містять параметри КС мережі: операційна система (OS), антивірусне програмне забезпечення (AntiVir), фаєрвол (Firewall), вразливості (Vulnerability). Здатність комп’ютерної системи бути інфікованою іншими КС мережі визначається врахуванням її зв’язків з іншими вузлами мережі (у вигляді чотирьох булевих змінних) та прив’язкою до відповідних масивів параметрів комп’ютерних систем мережі: OS_list, Firewall_list, Antivirus_list, Software_list.

Клас Comp пов'язаний з ієрархією класів, що представлена поліморфізмом абстрактного класу Virus.

За опис розміру та структури мережі відповідає клас PC_Matrix. На початку моделювання створюється масив V_Matrix розміром n×m, який заповнюється об'єктами одного з наслідуваних класів відповідно до обраної моделі поширення вірусів: SI , SIS, SIR, PSIDR, SEIQR.

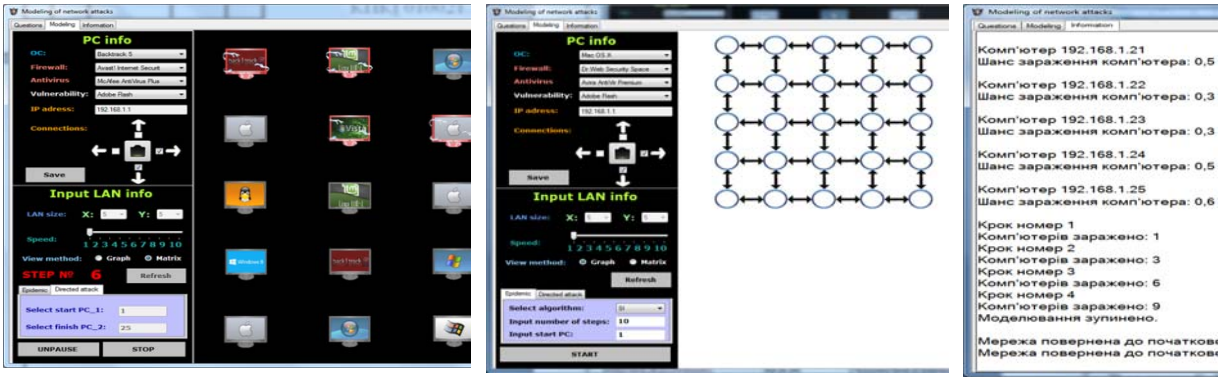
Візуалізація модельованої мережі здійснюється у двох представленнях – у вигляді матриці та графа (рис. 2, а, б) за допомогою двовимірних динамічних масивів об'єктів класу PictureBox та класу ImageList, в яких зберігаються необхідні зображення для відтворення зміни станів масивів PC_Matrix та V_Matrix.

При запуску програмного забезпечення розробленої ІТ користувач має можливість встановити початкові дані на панелі управління моделюванням, а саме: обрати метод моделювання, задати характеристики КС та розмір модельованої мережі. Для відображення інформації стосовно вузла мережі з метою її перегляду або зміни потрібно обрати необхідний елемент мережі на її візуальному представленні. Інформація стосовно обраного вузла буде відображена у блоці PC Info. Зміна прапорців Connections дозволяє змінити мережні з'єднання між КС мережі. Кількість кроків моделювання та швидкість моделювання (від 200 до 2000 мс за 1 крок) також є опціональними параметрами та обираються користувачем.

Вкладка Epidemic (епідемія) надає можливість змоделювати неконтрольоване вільне поширення ШПЗ за обраним алгоритмом з врахуванням номера КС, яка є початковим джерелом поширення інфекції. В режимі «епідемія» від початкової КС, яка позначається інфікованою, для кожної КС мережі перевіряється здатність інфікувати з'єднані з нею КС. Для алгоритму SEIQR віруси в інкубаційній стадії здатні переходити в активну стадію і вже на наступному кроці можуть інфікувати інші КС. В залежності від алгоритму, спроба інфікування може бути успішною за виконання наступних умов: (1) є зв'язок з іншою КС; (2) КС не є інфікованою; (3) КС не є імунною від вірусу; (4) спроба інфікування закінчилась успішно. Для інфікованої КС здійснюється спроба виявити вірус або вилікуватись. Ймовірність інфікування КС визначається з врахуванням встановленого на ній програмного забезпечення та її зв'язків з іншими КС, які визначають ймовірності вразливостей.

Вкладка Directed attack (напрямлена атака) слугує для пошуку найкоротшого шляху поширення ШПЗ між заданими КС мережі. Пошук оптимального шляху відбувається за алгоритмом рекурсивного пошуку. При моделюванні режиму Directed attack масив V_Matrix заповнюється об'єктами наслідуваного класу SI, в якому реалізовано алгоритми здійснення та пошуку оптимального шляху атаки від обраної стартової КС до цільової КС мережі. Реалізація цього алгоритму використовує клас Attack_trace, властивості якого містять послідовність ідентифікаторів комп'ютерних систем мережі та поле, що зберігає добутки ймовірностей їх інфікування. Для пошуку оптимального шляху здійснюється рекурсивний виклик функції від стартової КС, яка позначається інфікованою, у всіх напрямках по мережі до цільової КС зі збереженням ідентифікаторів КС у списку. Подальша побудова шляху можлива, якщо існує зв'язок з іншою КС, і її ідентифікатор відсутній у сформованому списку. Далі перевіряється умова, чи знайдена КС здатна інфікувати наступну, допоки не буде інфіковано цільову КС. Умовою виходу з рекурсії є знаходження шуканої КС або відсутність шляху далі. В останньому випадку шлях відкидається. Після проходження всіх можливих напрямків обирається найбільш успішний шлях, який є найкоротшим.

Стан процесу моделювання відображається на вкладці Information (рис. 2, в)



а) **б)** **в)**

Рис. 2. а) Представлення мережі у вигляді матриці; б) представлення мережі у вигляді графу; в) інформація про результати моделювання

Інтерфейс програмного забезпечення розробленої ІТ дозволяє змінювати параметри об'єктів класу Comp до початку або під час моделювання, що призводить до зміни параметрів відповідного об'єкту класу Virus.

Візуалізація станів КС мережі під час моделювання представлена різнокольоровими позначеннями. В залежності від обраної моделі поширення вірусів, стани КС мережі можуть бути наступними: (1) вразлива до інфікування КС; (2) інфікована КС; (3) імунна до інфікування КС; (4) інфікована КС, на якій виявлено вірус (модель PSIDR); (5) інфікована КС, на якій вірус перебуває в латентній стадії (модель SEIQR).

Експерименти

Для проведення експериментів було обрано множину найбільш поширеного програмного забезпечення (операційні системи, фаєрволи, антивірусне ПЗ) та враховано властиві йому вразливості.

З метою демонстрації роботи розробленого ПЗ інформаційної технології моделювання поширення вірусних кодів в гетерогенних мережах було проведено моделювання із застосуванням різних алгоритмів для мережі 10×10 КС (табл. 1, рис. 3).

Таблиця 1

Результати моделювання для мережі 10×10 КС

Крок	Номер досліду					Середня кількість інфікованих КС
	1	2	3	4	5	
Модель SI						
5	10	8	7	10	10	9
10	40	38	33	32	32	35
15	72	73	74	71	64	70,8
Модель SIS						
5	4	1	10	7	6	5,6
10	19	10	27	25	17	19,6
15	48	29	39	43	47	41,2
Модель SIR						
5	3	2	4	6	4	3,8
10	10	9	10	10	6	9
15	16	10	7	13	3	9,8
Модель PSIDR						
5	8	10	10	5	8	8,2
10	20	20	21	9	15	17
15	31	33	29	21	30	28,8
Модель SEIQR						
5	4	4	2	3	6	3,8
10	5	4	2	3	4	3,6
15	4	1	1	3	2	2,2

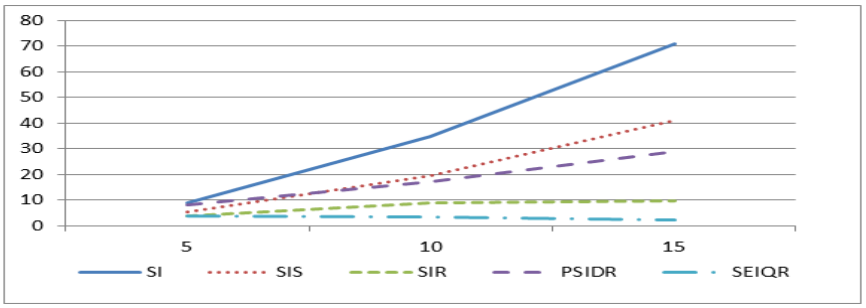


Рис. 3. Залежність кількості інфікованих КС мережі від кількості кроків для різних моделей

Як видно з одержаних результатів моделювання, при використанні моделі SI поширення вірусного коду відбувається без перешкод. При використанні моделі SIS КС мережі можуть бути вилікувані, проте це не впливає на їх захист від інфікування у майбутньому. Для обох моделей через певний проміжок часу мережа скінченних розмірів буде повністю інфікована. Моделі PSIDR та SIR. враховують здатність набування імунітету стосовно вірусу, проте для моделі PSIDR для цього потрібно вдвічі більше часу. Для мережі скінченного розміру через певний проміжок часу всі КС стануть імунними. Модель SEIQR показала найменшу швидкість поширення вірусів, що пов'язане з тим, що в латентному стані вірус не має змоги інфікувати інші КС, хоча і не може бути знищеним. За цією моделлю вірус є вразливим та може бути знищеним в проміжку часу між переходом в активну стадію та початком інфікування сусідніх КС, після чого КС стає імунною.

Результати пошуку оптимального шляху від КС з номером 1 до КС з номером 25 в мережі 5x5 подано на рис. 4, де інфіковані КС позначені колами. Пошук оптимального шляху відбувся за 12 кроків, тобто 3 спроби інфікування були неуспішними.

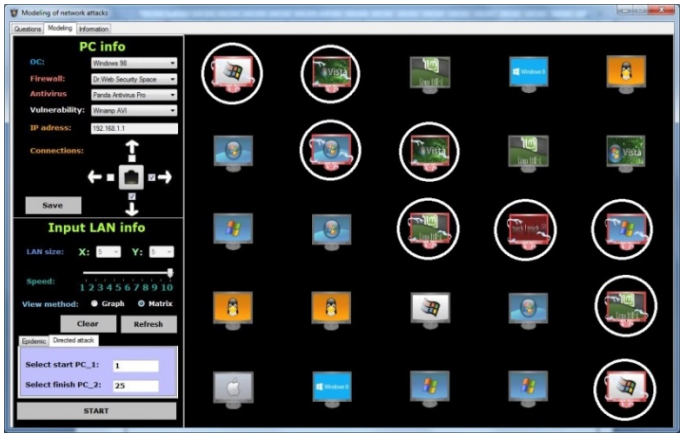


Рис. 4. Візуалізація результатів пошуку оптимального шляху

Висновки

Розроблено програмне забезпечення інформаційної технології моделювання поширення вірусних кодів в гетерогенних мережах. Наведено результати застосування розробленого програмного забезпечення, що надає можливість прогнозувати час і напрямок поширення вірусних програм в мережі з врахуванням топології мережі та системного програмного забезпечення, встановленого на КС мережі. Застосування розробленого ПЗ дозволить вживати заходів стосовно підвищення надійності роботи та інформаційної безпеки комп'ютерних мереж, що перебувають в експлуатації, а також на етапі їх проектування, та здійснювати оцінку ефективності застосування таких заходів.

Література

1. Белим С.В. Нахождение времени заражения локальной сети вирусами на основе сети формальных нейронов / С.В. Белим, С.Ю. Белим //Математические структуры и моделирование. – 2006. – № 1 (16). – С. 84–87.

2. Воронцов В.В. Аналитические модели распространения сетевых червей / В.В. Воронцов, И.В. Котенко // Труды СПИИРАН. – СПб : Наука. – 2007. – № 4. – С. 208–224.

3. Гусаров А.Н. Описание динамики распространения компьютерных угроз в информационно-вычислительных сетях с запаздыванием действия антивирусов / А.Н. Гусаров, Д.О. Жуков, А.В. Косарева // Вестник Московского государственного технического университета им. Н.Э. Баумана. Серия «Приборостроение». – 2010. – № 1. – С. 112–120.

4. Далингер Я.М. Математические модели распространения вирусов в компьютерных сетях различной структуры / Я.М. Далингер, Д.В. Бабанин, С.М. Бурков // Моделирование систем. – 2011. – № 4 (30). – С. 3–11.

5. Новиков С.В. Базовые принципы компьютерных вирусов. Математическая модель распространения вирусных эпидемий / С.В. Новиков // Труды научно-практической конференции «Информационные технологии в науке и образовании» (21-22 марта 2005 года). – Харьков. – С. 102–104.

6. Новиков С.В. Эпидемиологические модели прогнозирования вирусных атак / С.В. Новиков // X международная научно-практическая конференция «Теория и технологии программирования и защиты информации». Санкт-Петербург, 18 мая 2006. – С. 35–36.

7. Семенов С.Г. Математическая модель распространения компьютерных вирусов в гетерогенных компьютерных сетях автоматизированных систем управления технологическим процессом / С.Г. Семенов, В.В. Давыдов // Вестник Национального технического университета Харьковский политехнический институт. Серия: Информатика и моделирование. – 2012. – № 38. – С. 163–171.

8. Семыкина Н.А. Математическая модель защиты компьютерной сети от вируса с последствием / Н.А. Семыкина // Фундаментальные исследования. – 2014. – № 9. – С. 1982–1987.

9. Семыкина Н.А. Оптимальное управление защитой компьютерной сети от вредоносного кода / Н.А. Семыкина // Фундаментальные исследования. – 2015. – № 7-3. – С. 562–567.

10. Jeffrey O. Kephart, Steve R. White. Directed-Graph Epidemiological Models of Computer Viruses // IEEE Symposium on Security and Privacy. – 1991. – P. 343.

Рецензія/Peer review : 26.1.2017 р. Надрукована/Printed :6.2.2017 р.
Рецензент: д.т.н., доц. Мартинюк В.В.