УДК 004.735

Кучеренко О. А., магістрант
(Національний технічний університет України «КПІ». +380 (93) 773 33 91, alexa.kucherenko@gmail.com)

ШЛЯХИ ВДОСКОНАЛЕННЯ АНТИВІРУСНИХ СИСТЕМ

Кучеренко О. А. Шляхи вдосконалення антивірусних систем. В роботі досліджуються перспективні шляхи вдосконалення сучасних антивірусних систем із застосуванням нейронних мереж для розпізнавання вірусної активності в поведінкових аналізаторах. Розглянуті та проаналізовані шляхи розповсюдження та способи розпізнавання веб-орієнтованих скриптових вірусів і троянів. Описанні технології проактивного захисту на основі поведінкових аналізаторів із застосування автоматизованої системи розпізнавання вірусів шляхом аналізу послідовності дій підконтрольного програмного забезпечення. Приведена коротка характеристика основних методів розпізнавання антивірусних систем: шаблонів функціонування, віртуальних машин, довірених додатків, продукційних правил, еврістичних правил. Описані принципи створення нейромережевого аналізатора для розпізнавання комп´ютерних вірусів. Показана доцільність використання нейронної мережі типу двошарового персептрону, наведено методику розрахунку його параметрів.

Ключові слова: антивірусна система, скриптовий вірус, проактивний захист, нейромережевий аналізатор, розпізнавання вірусів, персептрон

Кучеренко А. А. Пути усовершенствования антивирусных систем. В работе исследуются перспективные пути совершенствования современных антивирусных систем с применением нейронных сетей для распознавания вирусной активности в поведенческих анализаторах. Рассмотрены и проанализированы пути распространения и способы распознавания веб-ориентованных скриптових вирусов и троянов. Описании технологии проактивной защиты на основе поведенческих анализаторов с применением автоматизированной системы распознавания вирусов путем анализа последовательности действий подконтрольного программного обеспечения. Приведена короткая характеристика основных методов распознавания антивирусных систем: шаблонов функционирования, виртуальных машин, доверенных дополнений, продукционных правил, эвристических правил. Описаны принципы создания нейросетевого анализатора для распознавания комп?ютерних вирусов. Показана целесообразность использования нейронной сети типа двухслойного персептрона, приведена методика расчета его параметров.

Ключевые слова: антивирусная система, скриптовий вирус, проактивний защита, нейросетевой анализатор, распознавание вирусов, персептрон

Kucherenko O. A. Ways of enhancement of the antvirus systems. . This article investigates the prospective ways of enhancement of the current antivirus systems using the neural networks to detect a virus activity in the behavioural analysing devices. It reviews and analyses the paths of propagation and tools of recognition of web-based script viruses and Trojans. The article gives a description of proactive protection technologies based on the behaviour analysing device using a computer-aided system to recognize the viruses by analysing the sequence of actions of the under-control software. It includes a short description of the following basic methods applicable to recognize the antivirus systems: the performance patterns method, virtual machines method, trusted applications method, production rules method, heuristic rules method. The article specifies the principles of developing a neural network analyser to detect the computer viruses. It shows the applicability of the neural network type of two-layer perceptron and gives the procedure for its parameter estimation.

Keywords: antivirus system, script virus, proactive protection, neural network analyser , recognition of viruses, perceptron

Вступ і постановка задачі. Основною ідеєю розпізнавання шкідливого програмного коду та модулів програм, що можуть зашкодити роботі комп'ютера та окремих його складових, є аналіз та відстеження поведінки таких шкідників. Основою такого підходу є використання поведінкових аналізаторів. Із кожним днем невпинно зростає кількість нових типів шкідливого програмного забезпечення (ПЗ) і антивірусні системи не встигають відстежити всі нові загрози й оновити свої вірусні бази. Проактивні технології – це сукупність технологій і методів, що використовуються в антивірусному програмному забезпеченні, основною метою яких, на відміну від сигнатурних технологій, є запобігання зараженню системи користувача, а не пошук вже відомого шкідливого програмного забезпечення в системі. При цьому проактивний захист намагається блокувати потенційно небезпечну активність програми тільки в тому випадку, якщо ця активність становить реальну загрозу. Серйозний недолік проактивного захисту – блокування легітимних програм.

Подібні причини визначають загальну проблематику даної статті, яка полягає у вдосконаленні поведінкових аналізаторів систем антивірусного захисту.

Метою даної роботи є дослідження перспективних шляхів вдосконалення сучасних антивірусних систем із застосуванням нейронних мереж для розпізнавання вірусної активності в поведінкових аналізаторах.

Середовище існування загроз. Шляхи розповсюдження багато в чому залежать від засобів створення шкідливого ПЗ. Зазначимо, що в даній роботі акцент ставиться на розпізнаванні веб-орієнтованих скриптових вірусів і троянів. В більшості випадків означене таке шкідливе ПЗ створюється за допомогою скриптових мов програмування VBScript, JavaScript та ActionScript, які дозволяють працювати з файловою системою, встановлювати мережеві з'єднання, маніпулювати процесами і потоками, здійснювати виклики функцій API операційної системи та запускати зовнішні програми. Шляхи розповсюдження вказаних вірусів та троянів не обмежені програмним середовищем поштового клієнту або браузеру.

Скриптові віруси, що розповсюджуються за допомогою HTML коду зазвичай складаються з двох основних частин:

- сам код вірусу, написаний на мові програмування JavaScript;
- web-сторінка, розроблена за допомогою мови розмітки тексту HTML, в яку інтегрується небезпечний код.

В основі такого типу вірусів лежить гіпертекстова інформаційна система. Ця система складається з множин інформаційних вузлів і гіпертекстових зв'язків, визначених на цих вузлах і інструментах маніпулювання вузлами і зв'язками. Технологія World Wide Web – це технологія ведення гіпертекстових розподілених систем в Internet, а тому вона повинна відповідати загальним визначенням таких систем. Це означає, що всі перераховані вище компоненти гіпертекстової системи повинні бути і в Web.

Web, як гіпертекстову систему, можна розглядати з двох точок зору. *По-перше,* як сукупність відображуваних сторінок, пов'язаних гіпертекстовими переходами (посиланнями – контейнер ANCHOR). *По-друге,* як множину елементарних інформаційних об'єктів, що становлять відображені сторінки (текст, графіка, мобільний код і т.п.). В останньому випадку множина гіпертекстових переходів сторінки – це такий же інформаційний фрагмент, як і вбудована в текст картинка.

При другому підході гіпертекстова мережа визначається на множині елементарних інформаційних об'єктів самими HTML-сторінками, які і відіграють роль гіпертекстових зв'язків. Цей підхід більш продуктивний з точки зору побудови відображуваних сторінок "на льоту" з готових компонентів.

При генерації сторінок в Web виникає дилема, пов'язана з архітектурою "клієнт-сервер". Сторінки можна генерувати як на стороні клієнта, так і на стороні сервера. У 1995 році фахівці компанії Netscape створили механізм управління сторінками на клієнтській стороні, розробивши мову програмування JavaScript.

Аналіз останніх досліджень та постановка проблеми.

Проактивний захист за допомогою поведінкових аналізаторів – технологія, в якій рішення про небезпечність ПЗ характер об'єкта, що перевіряється, приймається на основі аналізу виконуваних ним операцій. Перші спроби використання поведінкових аналізаторів відомі ще з середини 90-х років. В таких аналізаторах рішення про заборону або дозвіл виконання програмою потенційно небезпечної дії визначалось користувачем. Це стало основною завадою їх широкому застосуванню, адже "підозрілі" дії характерні і великій кількості звичайних програми. Інтерес до поведінкових аналізаторів відновився після того, як стала зрозумілою неминучість зменшення достовірності сигнатурних методів розпізнавання комп´ютерних вірусів. Підвищити ефективність проактивного захисту пропонувалось шляхом застосування автоматизованої системи розпізнавання вірусів на основі аналізу послідовності дій підконтрольного ПЗ. Для цього практично у всіх

антивірусних системах застосовано особливий програмний агент, інтегрований в операційну систему, що аналізує поведінку програм, виявляючи в ній ознаки вірусної поведінки – запис до реєстру, відкриття великої кількості мережевих з'єднань, запис на диск і модифікація важливих файлів, самовільний запуск додатків, блокування роботи тих чи інших утиліт і т. д. Визначені ознаки передаються в блок автоматизованого розпізнавання, методика роботи якого суттєво відрізняється в різних антивірусних системах. При цьому результати досліджень відомих антивірусних систем дозволяють стверджувати, що в основному використовуються наступні методи розпізнавання: шаблонів функціонування, віртуальних машин, довірених додатків, продукційних правил, еврістичних правил. Наведемо коротку характеристику означених методів.

Метод шаблонів функціонування, базується на співставленні дерева функціонування підконтрольного ПЗ з шаблонами поведінки звичайних програм та з шаблонами поведінки вірусів. Недоліки методу пов´язані як з складністю створення означених шаблонів, так і з складністю процесу співвіднесення. Даний метод в основному використовується для аналізу поведінки скриптів і макросів, оскільки відповідні віруси практично завжди виконують ряд однотипних дій.

Метод віртуальних машин, реалізований наприклад в деяких версіях антивірусу Eset Nod32, передбачає запуск програми в обмеженому середовищі віртуальної машини, яка функціонує на підзахисному комп'ютері з наступним аналізом результатів роботи цієї програми. Якщо результати вказують на небезпеку, то приймається рішення про наявність вірусу (шкідливого ПЗ). В якості недоліків методу вказують його високу ресурсоємність, неможливість виявлення вірусів в реальному масштабі часу та можливість обходу системи розпізнавання. Також відомі факти обходу вірусом обмежень віртуальної машини.

Метод довірених додатків, що використовується наприклад у системі DefenseWall HIPS, поділяє всі додатки на довірені і недовірені. Недовірені додатки запускаються з обмеженими правами на модифікацію критичних системних параметрів у спеціально відведеній для них віртуальній зоні, що відокремлює їх від довірених процесів. Спроби виходу із віртуальної зони розцінюються як порушення. Ще одним прикладом подібного рішення може служити технологія російської компанії Protection Technology, під назвою інтелектуальне управління активністю. По своїй суті це спеціальний монітор призначений для контролю взаємодії між прикладним ПЗ і операційною системою. Даний монітор вбудовується в модулі операційної системи, фіксує всі системні виклики, що проходять через них, і в разі небезпеки блокує їх виконання. Однак для такого методу проактивного захисту важливо визначити не правила блокування, а виключення, щоб дати можливість коректно працювати тим програмам, які в процесі виконання свого виконання повинні звертатись до системних викликів. Тому, крім загальної політики контролю всіх програм, потрібні додаткові профілі для кожної програми окремо.

Метод продукційних правил базується на представлені знань про поведінку вірусу у вигляді конструкції "якщо – то". Разом з простотою та ефективністю даного методу слід відзначити неможливість формування відповідних правил для великої кількості шкідливого ПЗ.

Методи еврістичних правил, що використовуються в більшості антивірусів, практично не документовані. Однак практичний досвід дозволяє стверджувати, що в їх основі лежить набір окремих досить розрізнених правил, заданих експертами в галузі антивірусного захисту. При цьому навіть по рекламним заявкам, достовірність розпізнавання описаних методу не перевищує 60-70%.

Відповідно [1, 2], підвищити ефективність розпізнавання можливо за рахунок використання штучних нейронних мереж, які вже довели свою ефективність при вирішення задач розпізнавання в різноманітних галузях. Цим і визначається актуальність досліджень в галузі створення нейромережевих методів розпізнавання комп´ютерних вірусів поведінковими аналізаторами. Слід зазначити, що у відкритих джерелах інформації детального опису використання нейронної мережі в поведінковому аналізаторі не знайдено.

Потенційно небезпечні функції управління розділами **Табл. 1**

Ім´я функції	Призначення функції
DeleteVolumeMountPoint	Розмонтовує розділ від вказаної точки монтування розділу
lpszVolumeMountPoint	Адреса рядка, який вказує точку розмонтування розділу
FindFirstVolume	Повертає ім'я розділу на комп'ютері
FindFirstVolumeMountPoint	Повертає ім'я точки монтування розділу на зазначеному комп'ютері
FindFirstVolumeMountPoint	Відкриває дескриптор пошуку точок монтування та повертає інформацію про першу знайдену точку монтування на зазначеному розділі
FindNextVolume	Продовжує пошук розділів, розпочатий викликом функції FindFirstVolume
FindNextVolumeMountPoint	Продовжує пошук точок монтування розділу, розпочатий викликом функції FindFirstVolumeMountPoint
GetDriveType	Визначає тип дискового пристрою
GetLogicalDrives	Отримує бітову маску, що представляє доступні на поточний момент дискові пристрої
GetLogicalDriveStrings	Заповнює буфер рядками, які визначають дійсні пристрою в системі

За оціночними підрахункам для Win32 в номенклатурі вхідних параметрів слід врахувати приблизно 200-300 потенційно небезпечних функцій. Перехопити виклики цих функцій можливо методом зміни точки входу в таблицю імпорту, або методом зміни початкових байт самої функції. Очевидно, що вхідні параметри нейронної мережі, які відповідають реалізації виклику потенційно небезпечної функції будуть бінарними. Якщо функція викликається, то на відповідний вхід мережі подається 1 і 0 в протилежному випадку. Крім того, в номенклатурі вхідних параметрів доцільно відобразити послідовність викликів небезпечних функцій. Однак ця пропозиція потребує доопрацювання.

Виходячи з позицій практичного застосування та мінімізації структури нейронної мережі можна обмежитись одним виходом, величина якого буде сигналізувати про ймовірність розпізнавання вірусу [6, 7].

Формування номенклатури вхідних та вихідних параметрів дозволяє перейти до визначення архітектури нейронної мережі. В якості основного застосуємо критерій максимізації обчислювальних потужностей мережі. Зазначимо, що на практиці обчислювальні потужності визначаються максимальною кількістю прикладів, яку може запам'ятати мережа для досягнення необхідної достовірності прийняття рішення. Також врахуємо наступні обмеження: мережа може бути попередньо навчена в лабораторних умовах на протязі тривалого терміну, в навчальних прикладах може бути відображений очікуваний вихід мережі, вхідні параметри в навчальних прикладах можуть бути зашумлені, кількість вхідних та вихідних параметрів принципово обмежена, вхідні та вихідні параметри мають числовий характер, розпізнавання повинно відбуватись в реальному масштабі часу, за рахунок оновлення вагових коефіцієнтів синаптичних зв´язків можна відмовитись від донавчання мережі в процесі експлуатації, поставлена задача розпізнавання вірусів відноситься до задач класифікації образів. Відповідно до результатів [8, 9] серед класичних нейромережевих архітектур найбільш повно основному та обмежуючим критеріям вибору відповідає двошаровий персептрон, структура якого показана на Рис.1.

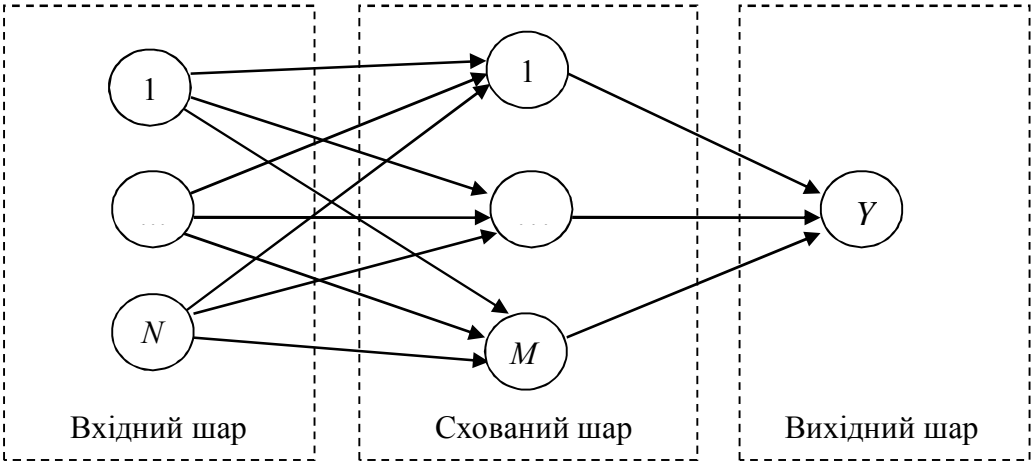

Рис. 1. Структура двошарового персептрону
N – кількість нейронів у вхідному шарі; *M* – кількість нейронів у схованому шарі

Використовуючи критерій мінімізації помилки розпізнавання [2, 3] в першому наближені кількість нейронів в схованому шарі можна визначити так:

$$M \approx (P/N)0.5,$$

де *P* – кількість навчальних прикладів.

Визначену кількість схованих нейронів слід уточнити, аналізуючи результати розпізнавання навчальної та тестової вибірок, сформованих з використанням баз даних антивірусних комплексів.

Висновки. Показано, що одним із найбільш перспективних шляхів вдосконалення сучасних антивірусних систем є застосування нейронних мереж для розпізнавання вірусної активності в поведінкових аналізаторах. Розроблений метод визначення вхідних параметрів нейронної мережі, який базується на застосуванні викликів потенційно небезпечних функцій операційної системи. Доведена доцільність використання нейронної мережі типу двошарового персептрону. Наведено методику розрахунку його параметрів.

Література

1. Петров А. А. Определение оперативно-технических характеристик систем активной защиты информации / А. А. Петров // Захист інформації. – 2009. – № 1 – С. 73–75.

2. Терейковський І. Нейронні мережі в засобах захисту комп'ютерної інформації / І. Терейковський. – К. : ПоліграфКонсалтинг. – 2007. – 209 с.

3. Каллан Р. Основные концепции нейронных сетей / Каллан Р.; пер. с англ. А. Г. Сивака. – Москва : Вильямс, 2003. – 288 с.

4. Корченко О. Г. Шкідливі програми та їх класифікація / О. Г. Корченко, К. П. Ануфрієнко // Защита информации : сб. науч. трудов. – К.: НАУ, 2007. – С.26–32.

5. Гордон Ян. Компьютерные вирусы без секретов / Я. Гордон – Москва : : ИПРЖР, 2010. – 240 с.

6. Огарок А. Виртуальные войны. Искусственный интеллект на защите от вирусов и программных закладок / А. Огарок, Д. Комашинский, Д. Школьников // Конфидент. – 2003. – №2 (50). – С. 64-69.

7. Вилков А. С. Информационная безопасность персональных ЭВМ и мониторинг компьютерных сетей / А. С. Вилков. – Москва : МИНИТ ФСБ России, 2005. – 210 с.

8. Искусственная нейронная сеть [Електронний ресурс] // – Режим доступу : http://ru.wikipedia.org/wiki/Искусственная_нейронная_сеть /

9. Нейронные сети [Електронний ресурс] // – Режим доступу : http://neurones.ru/