1. *Keldysh I.V.* Zh. Eksp. Theor. Fiz. **47** 1515 (1964) (Sov. Phys. JETP **20** 1018 (1964).
2. *Kadanoff L. P. and Baym G.* Quantum Statistical Mechanics (Benjamin, New York, 1962).
3. *Kubo R. J.* Phys. Soc. Japn. **12**, 570 (1957). Martin P. C., Schwinger J. Phys. Rev. **115**, 1342 (1959).
4. *Garzon S., Zuti´c I., Webb R.A.* Phys. Rev. Lett. **94**, 176601 (2005).
5. *Myöhä P., Stan A., Stefanucci G., R. van Leewen,* J. Phys. Europhys. Lett. **84** 67001 (2008).
6. *Myöhä P., Stan A., Stefanucci G., R. van Leewen*, Phys. Rev. B**80**. 115107 (2009).

Поступила 30.08.2010р.

УДК 004.056.5

Б. Я. Корнієнко, к.т.н., НАУ, м. Київ
Г.О. Бойко, НАУ, м. Київ
О.С. Снігур, НАУ, м. Київ

ВИКОРИСТАННЯ ГЕНЕТИЧНИХ АЛГОРИТМІВ ДЛЯ ВИЯВЛЕНЯ ВТОРГНЕННЯ В КОМП'ЮТЕРНІ МЕРЕЖІ

This paper describes how it is possible to apply Genetic Algorithm in Intrusion Detection Systems. A brief overview of the Intrusion Detection System, genetic algorithm, and related detection techniques is presented. Compare with other implementations of the same problem, this implementation considers both temporal and spatial information of network connections in encoding the network connection information into rules in IDS. The main system architecture and diagram of GA are shown.

Вступ

Генетичні алгоритми широко застосовуються в системах виявлення вторгнення. Генетичні алгоритми використовуються для створення правил поведінки системи в разі вторгнення. Мережне з'єднання та його характер можна представити так, щоб у відповідність йому можливо було поставити правило для прийняття рішення про те розцінювати дане з'єднання як вторгнення або ні. Ці правили моделюються як хромосоми певної популяції. Популяція перебуває в розвитку до тих пір поки вона не буде задовольняти встановленим критеріям. Сформовані правили використовуються системою для прийняття рішення про характер з'єднання (вторгнення або ні). Генетичні алгоритм не є самостійною системою безпеки, а являють собою механізм формування правил для використання в системах безпеки [1-3]. Останнім часом найбільш досліджуваною сферою комп'ютерної безпеки є системи виявлення вторгнення. Ця технологія виявлення використовується як

основний засіб для забезпечення цілісності інформації та роботи системи при вторгненні.

Постановка задачі

Метою даної статті є дослідження основних принципів побудови та функціонування генетичного алгоритму в системах виявлення вторгнення.

Принцип роботи генетичних алгоритмів

Коли порушник виконує спробу проникнути в інформаційну систему, або особа виконує дії які їй не дозволені, таку активність можна вважати спробою вторгнення. Зловмисників умовно можливо розділити на дві категорії: зовнішні та внутрішні. Зовнішні – це ті порушники що не мають права доступу до системи та проводять атаку використовуючи різні техніки проникнення. Внутрішні – в свою чергу це ті особи, що мають права доступу але намагаються виконати не авторизовані дії. Техніка вторгнення може використовувати різні недоліки програмного забезпечення, нестабільну роботу системи, визначення паролів, прослуховування незахищеного трафіку, чи помилок в роботі мережних протоколів. Система виявлення вторгнення – це система для виявлення вторгнення, збереження цього факту та оповіщення про них уповноваженим особам чи виконання регламентованих дії.

Існують дві основні категорії виявлення вторгнення: виявлення зловживань та виявлення підозрілої активності. Виявлення зловживань – до цієї категорії відносяться технології, що використовують відомі методи проникнення в систему. Такі методи проникнення бувають двох видів: такі, що відбуваються за схемою та такі, що мають особливий характер. Методи вторгнення, що відбуваються за схемою та такі, що мають особливий характер можуть бути представлені певною дією або послідовністю дій. Реакція системи залежить від виду дії, які спрямовані на проникнення. Виявлення підозрілої активності – технологія, яка має за основу аналіз нормальних показників роботи системи. Параметри, що відрізняються від нормальних розцінюються як спроба вторгнення [4,5].

Також системи виявлення вторгнення можливо розділити на дві групи в залежності від місця застосування: мережні та локальні. Мережні – аналізують трафік, що проходить через мережу та роботу всієї мережі в цілому. Локальні – спрямовані на аналіз роботи лише певної робочої станції. Також існують і нові технології, зокрема – система виявлення та блокування вторгнення (комбінується з локальною СВВ та має змогу модифікувати правила брандмауера).

Генетичні алгоритми - обчислювальна модель, що базуються на принципах еволюції та природного відбору. Ці алгоритми моделюють конкретну поставлену проблему та використовують інформаційну структуру схожу на хромосоми, розвивають ці хромосоми використовуючи оператори селекції, рекомбінації та мутації. Робота генетичного алгоритму зазвичай розпочинається з вибору довільної популяції хромосом. Ці хромосоми відображають проблему що потрібно вирішити. Згідно атрибутів певної

проблеми, різни позиції кожного хромосому інтерпретуються як біти, символи або цифри. Позиції, що відповідають атрибутам задачі інколи називають генами, гени в свою чергу змінюються у випадковому порядку протягом еволюції. Набір хромосом протягом процесу розвитку називається популяцією. В процесі еволюції використовується функція оцінки, яка визначає придатність кожної хромосоми. Еволюція популяції досягається за допомогою двох основних операторів: схрещування та мутації. Переваги в виборі хромосоми в процесі комбінації та селекції надаються хромосомам більшої придатності.

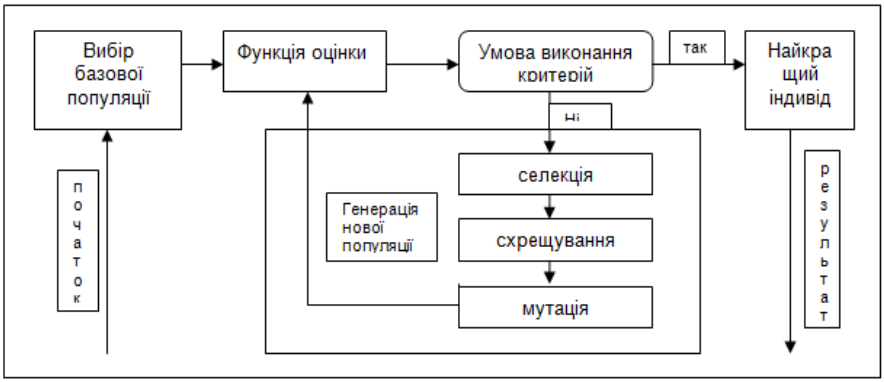

Рис. 1. Структура генетичного алгоритму.

На рис.1 зображено структуру найпростішого генетичного алгоритму. Він починається з випадково вибраної популяції, що проходить свою еволюцію завдяки селекції, схрещуванню та мутації. Наприкінці відбираються найкращі індивіди (хромосоми), що і є оптимальним розв`язком поставленої задачі.

Використання генетичних алгоритмів для виявлення вторгнення

Генетичний алгоритм може використовуватися як інструмент удосконалення простих правил для мережного трафіку [6-8]. Ці правила використовуються для того, щоб відрізняти нормальне з'єднання від підозрілого з'єднання. Підозрілі з'єднання відносяться до подій, що пов'язані з вторгненням. Правила зазвичай зберігаються в основній базі правил в такому вигляді:

if {умова} then {дія}

Для проблеми, що була розглянута вище, умова зазвичай є відношення інформації поточного з'єднання до інформації, що міститься в базі даних правил СВВ - IP адрес джерела, IP адрес призначення, час з'єднання та інші важливі атрибути, що в свою чергу може свідчити про вторгнення. Дія – зазвичай описує дії, що заздалегідь визначені уповноваженою особою. Наприклад, повідомлення адміністратора, відключити з'єднання, зберегти

запис в звіт безпеки. Далі буде наведено правило що використовується для запобігання протиправних дій:

if {якщо з'єднання має наступну інформацію: IP адреса джерела 98.2.33.12; IP адреса призначення 111.99.55.11; порт призначення 21; тривалість з'єднання 10.2 секунди } **then** {припинити з'єднання}

Дане правило можливо розглянути як: Якщо існуюче мережне з'єднання з IP адресою джерела 98.2.33.12, і IP адресою призначення 111.99.55.11 підключено до 21 порту і триває 10.2 секунди, то таке з'єднання необхідно зупинити. Тому що IP адреса 98.2.33.12 знаходиться в забороненому списку СВВ всі процеси пов'язані з даною адресою будуть завершені.

Основною задачею впровадження генетичного алгоритму є формування правил які будуть відповідати лише підозрілим підключенням. Правила перевіряються на з'єднаннях, що вже відбулись (історія з'єднань), а потім використовуються для фільтрації нових з'єднань з метою виявлення серед них небезпечних.

Отже, мережний трафік, що використовується генетичними алгоритмами попередньо класифікований, і розділений на нормальні з'єднання та підозрілі. Цей трафік збирається за допомогою спеціальної програми (мережного сніферу) і класифікується оператором. Потім він використовується для розвитку ГА. Починаючи генетичний алгоритм з малої кількості довільно сформованих правил, можна створити великий набір даних що містить в собі правила для СВВ.

Представлення інформації

Для повного використання підозрілого з'єднання нам необхідно проаналізувати всі значення, що відносяться до підозрілого з'єднання. Для простити використання виділимо основні параметри мережевого з'єднання в таблиці 1.

Таблиця 1. Правила визначення з'єднання та діапазон значень його параметрів.

Атрибут	**Діапазон значень**	**Приклад значення**	**Опис**
IP адреса джерела	0.0.0.0~255.255.255.255	d1.0b.**.** (209.11.??.??)	Під мережа с IP адресою 209.11.0.0 – 209.11.255.255
IP адреса призначення	0.0.0.0~255.255.255.255	82.12.b*.** (130.18.176+?.??)	Під мережа с IP адресою 130.18.176.0 – 130.18.255.255
Порт джерела	0~65535	42335	Порт через котрий було виконано з'єднання

Порт призначення	0~65535	00080	Порт на який біло виконано з’єднання, http сервіс
Тривалість	0~99999999	00000482	Тривалість з’єднання – 482 секунди
Статус	1~20	11	З’єднання припинено
Протокол	1~9	2	Протокол TCP
Кількість байт надіслана відправником	0~9999999999	0000007320	Відправник надіслав 7320 байт
Кількість байт надіслана отримувачем	0~9999999999	0000038891	Отримувач надіслав 38891 байта

На основі таблиці 1 можливо сформувати наступне правило:

If {з’єднання має наступну інформацію*:* IP адреса джерела 209.11.??.??; IP адреса призначення 130.18.176+?.??; порт джерела 42335; порт призначення 80; тривалість 482 с; з’єднання зупинено; використовується протокол TCP; відправник надіслав 7320 байт; отримувач надіслав 38891 байта **}** then { завершити з’єднання}

Дане правило у вигляді хромосоми наведено на рис. 2.

(d,1,0,b,-1,-1,-1,-1,8,2,1,2,b,-1,-1,-1,4,2,3,3,5,0,0,0,8,0,0,0,0,0,0,4,8,2,1,1,2,0,0,0,0,0,0,7,3,2,0,0,0,0,0,0,3,8,8,9,1)

Рис. 2. Хромосомна структура

Загалом в кожному хромосомі буде міститись 57 ген. Якщо має місце мережне з’єднання з IP адресою джерела 209.11.??.??; IP адреса призначення 130.18.176+?.??; порт джерела 42335; порт призначення 80; тривалість 482 с; з’єднання зупинено; використовується протокол TCP; відправник надіслав 7320 байт; отримувач надіслав 38891 байта, то таке з’єднання можна розцінювати як підозріле, а отже можливою спробою вторгнення. Якщо дане правило здатне виявити підозрілу поведінку, перевага буде надана цьому хромосому. Якщо правило розцінить з’єднання, як нормальне, хромосому буде надано негативну оцінку. Очевидно, що неможливо використовувати одне правило, яке зможе розділити всі нормальні та підозрілі з’єднання. Популяція, що складається з хромосом, які представляють собою правила, повинна розвиватися, щоб досягти оптимального розв’язку.

В таблиці 1 використовується символ *, а відповідна йому гена має значення -1. Цей символ використовується для представлення певного діапазону значень.

Генетичний алгоритм розпочинається з довільно вибраної популяції. Популяція може розвиватися за допомогою операторів схрещування та мутацій. Зважаючи на ефективність функції відбору, успішна популяція базується на правилах, що виявляють підозрілі з'єднання. У підсумку алгоритм зупиняється і нові генеровані правила додаються до загальної бази системи виявлення вторгнення.

Параметри генетичного алгоритму

Існує дуже багато параметрів, які слід враховувати при використанні генетичного алгоритму.

Функція оцінки є одним з основних параметрів генетичного алгоритму. Щоб обчислити функцію оцінки необхідно підрахувати кінцевий результат та придатність.

Кінцевий результат обчислюється на основі всіх атрибутів поточного з'єднання, що збігаються з попередньо вибраними даними, і потім множаться на «вагу» кожного атрибуту. Значення збіжності є 1 або 0.

$$кінцевий\ \ результат = \sum_{i=1}^{57} збіжність * вагу(i) \qquad (1)$$

Порядок вагових коефіцієнтів зображений на рис. 3. Цей порядок розбивається на категорії згідно даних про атрибути, що були надані мережним сніфером.

Рис. 3. Порядок вагових коефіцієнтів функції оцінки

Абсолютна різниця між кінцевим результатом хромосому та фактичним рівнем загрози:

$$\Delta = \left| кінцевий\ \ результат - рівень\ \ загрози \right| \qquad (2)$$

Якщо мережеві з'єднання що порівнюються не співпадають хромосому надається штраф:

$$штраф = (\Delta * категорія)/100 \quad (3)$$

Значення категорії залежить від того наскільки важко ідентифікувати порушення. Придатність хромосоми розраховується на основі результату «Штрафу»:

$$придатність = 1 - штраф \quad (4)$$

Значення придатності буде знаходитись в межах від 0 до 1.

Схрещування та мутація

Використовуючи генетичний алгоритм, необхідно знайти локальний максимум, як протилежність глобального максимуму. Використовуючи поняття підпросторів можна знайти декілька локальних максимумів. Ця техніка базується на аналогії з натуральною природою, де в середовищі існують підпростори в яких розвиваються різні типи життя. Проводячи аналогію генетичний алгоритм може містити різноманітність різних популяцій у мультимодальному домені, що підходить для доменів де необхідно знайти декілька оптимальних розв`язків. Використовуються два методи: скупчення та обмін. Метод скупчення використовує майже схожі елементи для потрібних замін, тим самим гальмує розвиток популяції і сходить до єдиної точки в наступних поколіннях. Метод обміну зменшує придатність популяції зі схожими членами та підвищує придатність індивідів, що дозволяє переходити до інших локальних оптимумів малих популяцій. Недолік даного методу полягає в тому, що для його коректної роботи потрібно багато вхідних параметрів.

Операція мутації є вагомою протягом еволюції. Наприклад, кожний сегмент IP адреси не повинен перевищувати значення 255. Мутація використовується для забезпечення відповідності правил таблиці 1.

Також існують параметри які необхідно узгоджувати, наприклад: рівень мутації, рівень схрещування, кількість популяцій і кількість поколінь. Ці параметри залежать від середовища використання генетичного алгоритму.

Рис. 4. Архітектура впровадження генетичного алгоритму

Архітектура системи

Необхідно зібрати необхідну кількість історичної інформації, для цього система виявлення вторгнення виділяє набір даних що передається до сніферу для аналізу, після чого інформація потрапляє до генетичного

алгоритму. В результаті генетичний алгоритм завдяки функції оцінки формую набір правил, які потрапляють в загальну базу даних правил.

Висновок

В статті розглянуто методологію застосування генетичного алгоритму в системах виявлення вторгнення, короткий опис системи виявлення вторгнення та генетичного алгоритму. Наведено архітектуру системи. Розглянуто фактори, що впливають на роботу генетичного алгоритму.

1. *Bezroukov, Nikolai.* “Intrusion Detection (general issues).” Softpanorama: Open Source Software Educational Society. *Nikolai Bezroukov.* URL: http://www.softpanorama.org/Security/intrusion_detection.shtml, 2000.
2. *Bridges, Susan, and Rayford B. Vaughn.* “Intrusion Detection Via Fuzzy Data Mining.” *In Proceedings of 12th Annual Canadian Information Technology Security Symposium*, pp. 109-122. Ottawa, Canada, 2000.
3. *Crosbie, Mark, and Gene Spafford.* “Applying Genetic Programming to Intrusion Detection.” *In Proceedingsof 1995 AAAI Fall Symposium on Genetic Programming*, pp. 1-8. Cambridge, Massachusetts. URL:http://citeseer.nj.nec.com/crosbie95applying.html, 1995.
4. *Graham, Robert.* “FAQ: Network Intrusion Detection Systems.” RobertGraham.com Homepage. *Robert Graham*. URL: http://www.robertgraham.com/pubs/network-intrusion-detection.html, 2001.
5. *Jones, Anita. K. and Robert. S. Sielken*. “Computer System Intrusion Detection: A Survey.” Technical Report. Department of Computer Science, University of Virginia, Charlottesville, Virginia, 2000.
6. *McHugh, John*. “Intrusion and Intrusion Detection.” Technical Report. CERT Coordination Center, Software Engineering Institute, Carnegie Mellon University, 2001.
7. *Paxson, Vern.* “Bro: A System for Detecting Network Intruders in Real-time.” *In Proceedings of 7th USENIX Security Symposium*, pp. 31-51. San Antonio, Texas, 1998.
8. *Pohlheim, Hartmut.* “Genetic and Evolutionary Algorithms: Principles, Methods and Algorithms.” Genetic and Evolutionary Algorithm Toolbox. Hartmut Pohlheim. URL: http://www.geatbx.com/docu/algindex.html, 2003.

Поступила 4.08.2010р.

УДК 519.6

С. Ю. Протасов, ЧГТУ, г. Черкассы

ДИНАМИЧЕСКИЕ ХАРАКТЕРИСТИКИ ЛИНЕЙНЫХ ОБЪЕКТОВ С ПЕРЕМЕННЫМИ ПАРАМЕТРАМИ

В данной статье проанализированы основные виды интегральной зависимости между выходным и входным сигналами линейных динамических объектов с переменными параметрами.

