Використання Темпоральних Правил в Задачах Захисту Інформації в Комп'ютерних Системах

Віктор Левикін
кафедра інформаційних управляючих систем
Харківський національний університет
радіоелектроніки
Харків, Україна
levykinvictor@gmail.com

Оксана Чала
кафедра інформаційних управляючих систем
Харківський національний університет
радіоелектроніки
Харків, Україна
oksana.chala@nure.ua

Using Temporal Rules to Information Security Tasks in Computer Systems

Victor Levykin
Department of Informatiom Control Systems
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
levykinvictor@gmail.com

Oksana Chala
Department of Informatiom Control Systems
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
oksana.chala@nure.ua

***Анотація*—Розглянуто проблему виявлення вторгнень в роботу комп'ютерної системи. Показано, що комбінація існуючих підходів виявлення аномалій та виявлення зловживань дає можливість підвищити ефективність виявлення атак. Запропоновано підхід до виявлення вторгнень в роботу комп'ютерної системи з використанням логічних фактів та темпоральних правил, що дозволяє комбінувати виявлення аномальних станів та переходів до цих станів, які свідчать про втручання. Кожен зважений факт відображає стан одного із процесів в комп'ютерній системі. Факт задається на основі вектору змінних, які містять значення атрибутів об'єктів предметної області. Темпоральні правила задають типові зв'язки між станами процесу. Виявлення аномалій виконується шляхом порівняння ваги поточного стану з вагами нормальних та аномальних станів для аналогічних векторів змінних. Виявлення втручань реалізовано шляхом порівняння ваг темпоральних правил, що поєднують нормальний та аномальний факти.**

Ключові слова—вторгнення; аномальний стан; темпоральне правило; компьютерна система

Abstract—The problem of detecting intrusions in the computer system is considered. It has been shown that the combination of existing approaches to detecting abnormalities and detecting abuses enables to increase the effectiveness of detecting attacks. The approach to detecting intrusions into the work of a computer system with the use of logical facts and temporal rules is proposed, which allows combining the detection of abnormal states and transitions to these states that indicate intervention. Each weighted fact reflects the state of one of the processes in the computer system. The fact is set on the basis of a vector of variables that contain the values of the attributes of objects in the subject area. Temporal rules specify typical relationships between process states. Detection of anomalies is performed by comparing the weight of the current state with the weights of normal and abnormal states for similar vectors of variables. Detection of interventions is realized by comparing the weight of the temporal rules, which combine normal and abnormal facts.

Keywords—-intrusion; abnormal condition; temporal rule; computer system.

I. Вступ

Захист інформації в комп'ютерній системі здійснюється шляхом аналізу дій процесів, що в ній виконуються. Кожен процес є завантаженою у пам'ять програмою, яка має свій ідентифікаційний код та доступ до комп'ютерних ресурсів – процесору, пам'яті та зовнішніх приладів. Доступ до зовнішніх приладів реалізується через стандартизований інтерфейс, що дозволяє відобразити поточний стан цих приладів, дані, та команди які їм передаються. Це дає можливість зафіксувати доступ до даних та зовнішніх приладів у стандартизованій формі, а потім провести аналіз поведінки процесів в комп'ютерній системі на основі аналізу цих записів.

При проведенні аналізу дій процесу розглядається виконання вимог щодо доступності, конфіденційності та цілісності даних [1] та інших ресурсів комп'ютерної системи. Відповідно до цих вимог дані повинні бути доступними тільки для авторизованих користувачів і змінюватися тільки авторизованим способом.

Вторгнення – це спроба неавторизованого доступу до ресурсів комп'ютерної системи [1]. Для того, щоб виявити вторгнення, аналізується інформація щодо операцій, що виконуються в цій системі. Така інформація зазвичай представлена у вигляді послідовності станів процесів у комп'ютерній системі з зазначенням часу виникнення кожного стану.

Кожен стан характеризується вектором змінних, які містять інформацію як про поточний процес, так і про стан об'єктів, з якими цей процес взаємодіє. Останній будемо розглядати як контекст процесу. Слід відзначити, що ми розглядаємо контекст у широкому сенсі, з точки зору користувача, на відміну від традиційного регістрового контексту для процесу в комп'ютерній системі.

Контекст в широкому сенсі задається поточними значеннями властивостей об'єктів предметної області, з якими цей процес взаємодіє. Наприклад, поточний стан процесу може бути охарактеризований даними про мережеві пакети, IP-адреси, системні виклики тощо. Набір змінних, що містить інформацію про вторгнення в комп'ютерну систему, відрізняється від вектору змінних для типових станів, пов'язаних з її нормальною роботою. Такі стани містять нетипові значення атрибутів, наприклад IP-адресу, з якої здійснюється атака та яка не належить до поточної комп'ютерної мережі.

Виявлення аномального стану одного з процесів свідчить про можливе вторгнення в комп'ютерну систему. Для виявлення аномального стану виконується порівняльний аналіз послідовностей станів процесів і виявляються характерні ознаки або патерни цих станів. Використання таких патернів дозволяє виявити атаку. Після знаходження атаки виконується її класифікація та формується попередження із зазначенням типу атаки.

На сьогодні при виявленні вторгнень в комп'ютерну систему застосовуються два взаємодоповнюючі підходи: Виявлення аномалій (Anomaly Detection) і виявлення зловживань (Misuse Detection). Ілюстрацію відмінностей між цими підходами представлено на рис.1.

Відповідно до першого підходу при виявленні вторгнень виконується пошук відхилень від апріорно відомої нормальної поведінки процесів в комп'ютерній системі. Попередньо на основі масиву відомих послідовностей станів формуються патерни «нормальної» поведінки процесів. Тоді розпізнавання вторгнень реалізується шляхом порівняння вектору змінних для поточного (або підмножини останніх станів) з патерном нормальної поведінки та виявлення аномалій функціонування комп'ютерної системи.

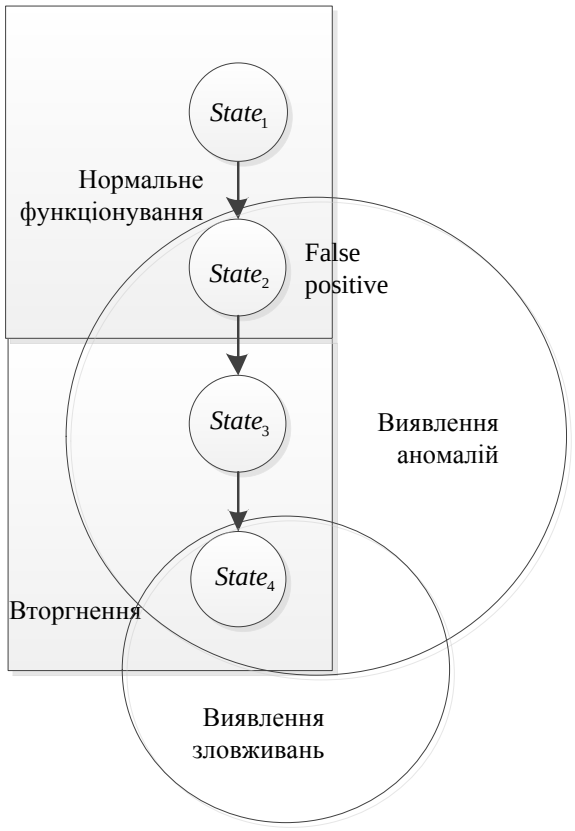

Рис. 1. Порівняння Підходів Виявлення Аномалій та Виявлення Зловживань для Послідовності Станів

Перевага підходу розпізнавання вторгнень на основі виділення аномальної поведінки процесів в комп'ютерній системі полягає в можливості виявити нові, невідомі раніше атаки. Однак дана парадигма має істотний недолік. Вторгнення може бути розпізнане помилково в разі зміни поточних завдань і появи нових програм, а також при зміни поведінки поточних процесів. В результаті при використанні даного підходу для динамічних процесів в комп'ютерній системі потрібно виконувати постійну адаптацію патернів нормальної поведінки. На рис. 1 проілюстровано приклад помилкового виявлення вторгнення на основі інформації про $State_2$, в той час як ця подія відповідає нормальній поведінці процесу в комп'ютерній системі.

Відповідно до другого підходу розробляються патерни вторгнень на основі відомих послідовностей станів. Тому такі системи розпізнають відомі атаки з невеликою кількістю помилок. Недолік цього підходу полягає в тому, що не виявляються невідомі атаки, тобто такі атаки, для яких ще не побудовані патерни. На рис.1 представлений приклад помилки у виявленні вторгнення згідно даного підходу: $State_3$ не виявляється як вторгнення, в силу відсутності відомого патерну для такої атаки. В той же час $State_4$ визначається як вторгнення на основі відомого патерну.

Для того, щоб компенсувати недоліки обох підходів, необхідно одночасно проводити аналіз нових станів з тим, щоб зіставити їх і з аномаліями і відомими атаками. Однак при безпосередньому порівнянні доведеться враховувати значне число різних патернів. Ми можемо вирішити цю проблему на основі порівняння інтегральних характеристик процесів нормальної роботи і процесів, пов'язаних з вторгненням.

В якості інтегральних характеристики можуть бути використані логічні факти [2, 3], які задають стан кожного процесу в комп'ютерній системі, та темпоральні правила, які задають поведінку цього процесу [4, 5]. Темпоральне правило має вагу, яка залежить від ймовірності його виникнення для процесів в комп'ютерній системі [6].

Дана робота містить такі результати. Ми пропонуємо підхід до порівняння і класифікації станів для виявлення вторгнень з використанням логічних фактів та темпоральних правил. Кожен факт визначає стан процесу в комп'ютерній системі на основі вектору змінних, які містять значення атрибутів об'єктів предметної області. Темпоральні правила задають типові зв'язки між станами процесу. В якості вхідних даних для запропонованого підходу використовується журнал, що містить події всіх процесів в комп'ютерній системі. Кожна подія описує поточний стан системи та час виникнення цього стану. Ці події розмічені: вказані нормальні стани та стани, що пов'язані із вторгненням.

Відповідно до запропонованого підходу, журнал розбивається на дві підмножини станів. Перший містить впорядковані за часом події для нормального функціонування комп'ютерної системи, а друга – події, які пов'язані з вторгненням. Для станів на кожній із трас обчислюються ваги як фактів, так і правил. Детектування вторгнень здійснюється на основі порівняння ваг фактів та правил. Запропонований підхід дозволяє об'єднати переваги підходів Anomaly and Misuse Detection.

II. Використання темпоральних правил

A. Темпоральні залежності на послідовностях станів

В якості об'єкта аналізу ми розглядаємо сукупність процесів, що виконуються в комп'ютерній системі. Вказані процеси при роботі взаємодіють з використанням загальних ресурсів і даних цієї системи. Тому при аналізі вторгнень доцільно об'єднати всю сукупність системних процесів в один інтегральний процес. Цей процес характеризується єдиною послідовністю станів, упорядкованих у часі.

Для опису поведінки такого інтегрального процесу можуть бути використані темпоральні правила трьох типів: Next, Future, Until. Правило першого типу задає темпоральний зв'язок між послідовними станами поєднаного процесу. Це означає, що перший стан завжди передуватиме другому. Друге правило задає упорядкованість у часі між парою станів процесу в комп'ютерній системі, причому між цими станами можуть бути проміжні стани. Третє правило задає зв'язок між станами процесу з урахуванням змін значень у векторі змінних, що характеризують ці стани.

В даній роботі розглядаються правила першого типу, оскільки мета полягає в тому, щоб виявити перехід до аномального стану. Правило другого типу об'єднує декілька правил Next. Правило третього типу при вирішенні поставленої в роботі задачі доцільно розглядати як уточнююче відносно правила типу Next.

Упорядкована послідовність станів $S = s_1,...,s_j,s_{j+1}...$ процесів в комп'ютерній системі відображає відповідні події вхідного журналу. Ця послідовність може бути представлена у вигляд послідовності логічних фактів на векторі змінних кожного стану: $ft_1,...,ft_j,ft_{j+1}...$, причому кожен факт ft_j задається як предикат на значенні змінних, що характеризують відповідний стан s_j:

$$ft_j = (a_j^1 = \alpha_j^{1,l}) \wedge ... \wedge (a_j^k = \alpha_j^{k,l}) \wedge ... \wedge (a_j^K = \alpha_j^{K,l}), \quad (1)$$

де a_j^k – змінна, що характеризує стан s_j; $\alpha_j^{k,l}$ – значення змінної a_j^k для поточного стану s_j.

Тоді послідовність станів інтегрального процесу представимо такою послідовністю темпоральних правил типу Next:

$$S = ft_1 X ft_1 X ... X ft_j X ft_{j+1} ... ft_J, \quad (2)$$

де X – оператор темпоральної логіки Next.

У відповідності до виразу (2) правила першого типу задають допустимі послідовності переходів між станами в часі. Послідовність (2) розбивається на окремі правила, що мають вигляд $ft_{j-1} X ft_j$, $ft_j X ft_{j+1}$.

При вирішенні задачі виявлення вторгнень інтегральний процес в комп'ютерній системі розділяється на два підпроцеси:

- нормального функціонування;
- аномальної роботи, яка виникає в результаті зовнішніх впливів.

Тоді задача виявлення вторгнень зводиться до задачі поділу послідовності станів на дві непересічні підмножини. Перша підмножини містить впорядковані стани, які відповідають нормальній роботі комп'ютерної системи. Друга підмножина містить стани, які відповідають аномальній поведінці, в результаті непередбачених впливів, причому для будь-якої пари станів перший стан може відповідати нормальній поведінці, а другий - аномальній:

$$S = S_{norm} \bigcup S_{\text{int}r} \,|\, \exists s_j \in S_{normi}, s_{j+1} \in \Pi_{\text{int}r}, \quad (3)$$

де $S_{norm}, S_{\text{int}r}$ – множини станів, що відповідають типовій та аномальній поведінці, $S_{norm} \bigcap S_{\text{int}r} = \varnothing$.

Основна ідея пропонованого підходу згідно (3) заснована виявленні таких темпоральних правил виду $ft_{norm} X ft_{abnorm}$, що перший факт кожного правила відповідає нормальному стану, а другий – аномальному.

Іншими словами, необхідно знайти такі темпоральні залежності для пар станів, які дозволяють при фіксації поточного стану як другого стану правила віднести його до підмножини S_{norm} або до $S_{\text{int}r}$. Якщо поточний стан віднесено до підмножини $S_{\text{int}r}$, то це означає що ми виявили вторгнення.

B. Підхід до виявлення вторгнень

Кожен процес в комп'ютерній системі включає в себе послідовність дій і станів, що виникають в результаті цих дій. Кожна подія вхідного журналу характеризує стан інтегрального процесу через множину значень змінних, що і встановлюється відповідними фактами ft_j. Тому при порівнянні подій доцільно порівнювати значення відповідних змінних, що відповідають значенням атрибути об'єктів предметної області. В якості атрибутів виступають, наприклад, IP адреса, порт, тощо.

Дії процесу відповідають переходам між станами. Вказані дії представлені темпоральними правилами $ft_j Xft_{j+1}$. Це дозволяє порівнювати пари атрибутів аналогічних об'єктів, що відображають старий і новий стан процесу, наприклад пару IP адрес: (192.168.220.51; 192.168.100.11). Тобто зміна стану пов'язана з тим, що для двох послідовних в часі подій відбулося перемикання на роботу з іншою адресою.

Тоді доцільно припустити, що відмінності між підмножинами значень атрибутів для двох послідовних станів будуть більш значними для різних процесів. Наприклад, при порівнянні атрибутів двох станів для процесу нормальної роботи, у них можуть повторюватися IP-адреси, порти і т.п. Аномальний процес матиме інші підмножини адрес і портів.

Наведене дає можливість сформувати темпоральні правила типу Next для нормальної роботи, для зловживань та для переходу від нормальної до аномальної поведінки.

Запропонований підхід базується на ідеї побудови темпоральної бази знань, представленій в роботі [4], та складається з трьох фаз. На першій фазі розраховуються характеристики логічних фактів, що відображають стани, та правил, що відображають дії. Вони визначаються для відомих підмножин станів з нормальною та аномальною поведінкою. На другій фазі розраховуються характеристики логічного факту та правила для нового (поточного) стану з використанням результатів робіт [4-6]. На третій фазі виконується класифікація нового стану. Тобто визначається, відповідає він нормальній роботі чи аномальній поведінці. Класифікація виконується на основі порівняння правил та фактів.

III. Експериментальні Результати

Запропонований підхід експериментально перевірено шляхом аналізу штучного розміченого журналу комп'ютерної системи [7], який моделює послідовність роботи процесів та втручань в системі. Кожен стан системи характеризується множиною значень таких змінних: IP–адреси джерела та отримувача; порти джерела та отримувача; Інтернет-протокол; мітка втручання.

Антецеденти та консеквенти кожного правила представлені у вигляді логічних фактів – кон'юнкції змінних із зданими значеннями, наприклад Протокол = «ICMP» ∧ Порт_джерела = «441» ∧ … Такі логічні факти визначаються окремо для підмножин S_{norm} та $S_{\text{int}r}$. Потім формуються правила із антецедентом із S_{norm} та консеквентом із $S_{\text{int}r}$. При формуванні цих правил використовуються пари логічних фактів, які є послідовними в часі та відповідають станам з підмножини S_{norm} та $S_{\text{int}r}$. Ваги фактів та правил розраховуються за методом [6]. Альтернативний варіант розрахунку ваг – за частотною характеристикою комбінацій атрибутів та послідовностей фактів, відповідно. Проведений аналіз показав, що ваги фактів відрізняються зазвичай на один – два порядки, а правила – на два й більше порядків для нормальної роботи та для втручань. Така різниця ваг дозволяє класифікувати кожний новий стан комп'ютерної системи. Зазначимо, що для опису нормальної роботи використовувались правила, в яких і антецедент, і консеквент відображають стани із S_{norm}. Однак такі результати пов'язані, на нашу думку, з тим, що вхідний журнал був згенерований штучно.

IV. Висновки

Запропоновано підхід до побудови й використання темпоральних правил для розпізнавання атак в комп'ютерних системах. Підхід передбачає послідовне формування бази зважених темпоральних правил для нормальної та аномальної роботи комп'ютерної системи, а потім класифікацію поточного стану на основі порівняння ваг цих правил. Підхід дозволяє поєднати переваги виявлення аномалій та вторгнень і тим самим підвищити ефективність виявлення втручань у роботу комп'ютерної системи.

Література REFERENCES

[1] R. G. Bace, Intrusion Detection. USA: MacMillan Technical Publishing, 2000.

[2] O. Chala, "Logical-probabilistic representation of casual dependencies between events in business-process management", *Науково-технічний журнал, Сучасні інформаційні системи*, Том 2, № 2, с. 40-44. 2018.

[3] О. В. Чала, "Розробка представлення знань на основі марківських логічних мереж в системі процесного управління", *Вісник НТУ «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*, № 22 (1298), с. 22-26. 2018.

[4] V. Levykin, O. Chala, "Method of automated construction and expansion of the knowledge base of the business process management system", *EUREKA: Physics and Engineering*, Vol. 4. pp. 29-35. 2018. DOI: http://dx.doi.org/ 10.21303/ 2461-4262. 2018.00676.

[5] В. М Левикін, О. В. Чала, "Розробка представлення причинно-наслідкових залежностей для бази знань системи процесного управління", *Вісник НТУ «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*, № 21 (1297). с. 48-53. 2018.

[6] V. Levykin, O. Chala, "Method of determining weights of temporal rules in markov logic network for building knowledge base in information control system", *EUREKA: Physics and Engineering*, 2018. Vol. 5. pp. 3-10. DOI: http://dx.doi.org/10.21303/2461-4262.2018. 00713.

[7] M. Ring, S. Wunderlich, D. Gr udl, D. Landes, A. Hotho, "Flow-based benchmark data sets for intrusion detection" in *Cyber Warfare and Security*: 16th European Conference ECCWS, 2017, pp. 361-369.