Ідентифікація користувача за його поведінкою в системі

Ігор Рубан[1], Віталій Улітічев[2]

1. Кафедра електронних обчислювальних машин,
Харківський національний університет радіоелектроніки,
УКРАЇНА, г. Харків, пр. Науки, 14,
E-mail: ihor.ruban@nure.ua

2. Кафедра безпеки інформаційних технологій,
Харківський національний університет радіоелектроніки,
УКРАЇНА, г. Харків, пр. Науки, 14,
E-mail: vitalii.ulitichev@nure.ua

Коротка анотація – The widespread information systems that integrate client server technology with the global network have posed numerous problems. It turned out that standard identification methods are already out of date. Particularly the problem is that the universally recognized separation of methods of controlling access to and access to information is no longer effective. To resolve this issue, you need to provide identification methods that will be able to identify users through a set of actions taken by users while working with DIS.

Ключові слова – інформаційні системи, нейронна мережа, автентифікація, ідентифікація.

I. Вступ

Ідентифікація та автентифікація можуть вважатися основними програмними та технічними засобами безпеки, оскільки решта служб орієнтована на підтримку названих об'єктів. Ідентифікація та автентифікація - це перша лінія захисту інформаційної зони організації.

Відповідно до звіту [1] інституту ДАНС «Тенденції витрачання на IT-безпеку», який вказує на відсоток витрат організацій на забезпечення різних технічних засобів інформаційної безпеки. На рис 1 представлена схема вартості кожної технологічної одиниці.

Проаналізувавши діаграму, можна визначити два типи технологій кібербезпеки, які представлені найважливішими, такими як:

– доступ до автентифікації;

– розширити захист від зловмисного програмного забезпечення;

Зараз розповсюджені інформаційні системи стають популярнішими, ніж раніше. Сумнівно, що існує можливість знайти будь-яку програму, яка не використовує компоненти різних постачальників. Чим складніше сучасні програми, тим більше потребує використання компонентів, розподілених віддаленою машиною. Послуги, які використовуються веб-додатками, дуже часто розміщуються не на території інформаційного центру закладу.

Поширені інформаційні системи, що об'єднують технологію сервера клієнта з глобальною мережею Інтернет, викликають численні проблеми. Виявилося, що стандартні методи ідентифікації вже застаріли. Особливо проблема полягає в тому, що загальновизнаний поділ методів контролю за фізичним доступом і контролю над доступом до інформації вже не дієвий. Для вирішення цієї проблеми необхідно застосувати методи ідентифікації, які матимуть можливість ідентифікувати користувачів за допомогою сукупності дій, реалізованих користувачами в процесі роботи з DIS.

II. Основний матеріал

Найпоширенішими сьогодні є методи ідентифікації користувачів, засновані на використанні паролів, які, на жаль, можуть бути втрачені, викрадені або порушені багатьма способами. Таким чином, може бути чудово реалізувати ідею комбінації стандартного пароля з методом ідентифікації користувачів на основі їх поведінки в системі. У цьому випадку, навіть якщо зловмисник отримує доступ до пароля, загальний доступ до комп'ютерної системи може бути відхилений через ідентифікацію поведінки, зображену на. У цьому випадку структура ідентифікації через поведінку користувача в системі є наступне: якщо користувач введе неправильний пароль, йому буде відмовлено у доступі негайно. Якщо пароль є

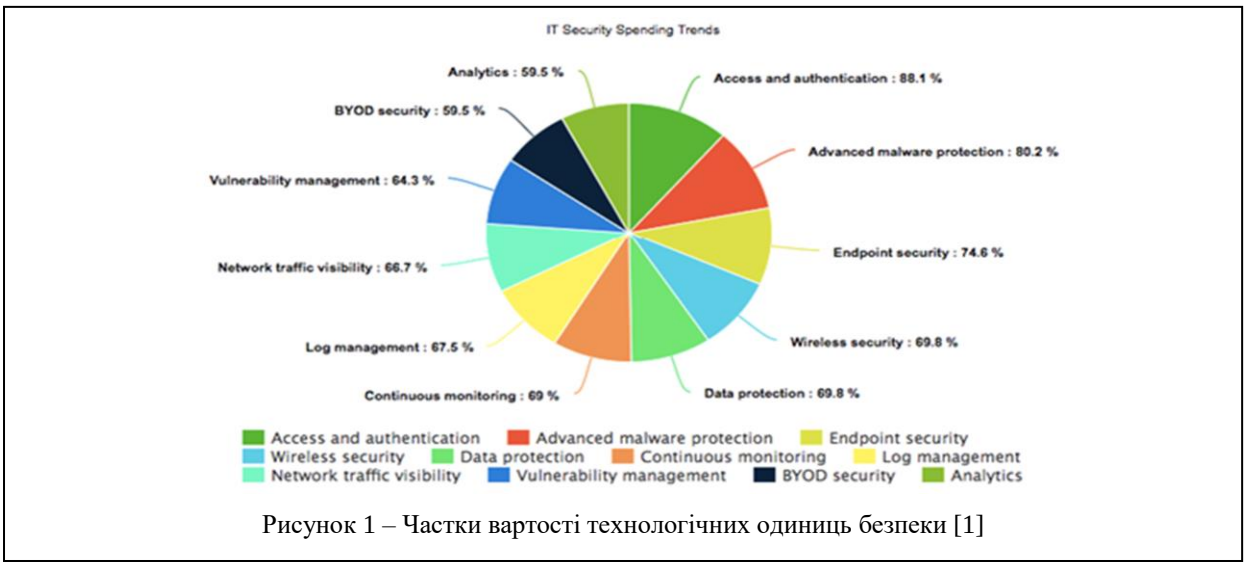

Рисунок 1 – Частки вартості технологічних одиниць безпеки [1]

правильним протягом певного періоду часу, T відбувається процес збору даних про його поведінку, після чого ці дані порівнюються з прикладом зареєстрованого користувача з бази даних, яка вже пройшла через ідентифікацію. Залежно від необхідної точності в процесі даних порівняння доступу може бути дозволено або заборонено.

Завдання ідентифікації можна розділити на кілька етапів, основними з яких є процес викладання та, відповідно, процес розпізнавання. На початку система бере і зберігає в базі даних інформацію про поведінку користувача в системі; на основі значень цих налаштувань існує шаблон або профіль поведінки користувача. Потім відбувається процес порівняння даного шаблону з уже збереженим у базі даних в системі, який є фактично ідентифікаційним.

Ідентифікація відбувається завдяки результатам контролю за поведінкою користувача в процесі роботи в інформаційній системі. Як вихідні дані використовуються матричні індикатори X у доказуванні моніторингу поведінки користувача та стовпець Y, який складається з множинності {0,1}, де 1 виникає, якщо поведінка користувача стосується визначеного користувача та 0, якщо ні.

Матриця X складається з векторів, які мають таку форму:

$$x_i = \{S_i^T, ST_i^T\}, i = 1, \dots, l$$

де S_i^T – множина станів в яких знаходився користувач за період і-ої сесії, а ST_i^T – множина статистичних параметрів побудованих на основі множини S_i^T, наприклад тривалість перебування користувача в кожному стані з множини S_i^T, середній час перебування користувача в кожному стані S_i^T та інші. Сесії користувачів виділені таким чином, що вони не можуть бути довше часу T або містить більш B станів. Тобто сесія вважається закінченою або коли користувач знаходився в B станах, або коли сесія зайняла за часом більш T часу .

Завдання побудови шаблону поведінки користувача буде складатися в побудові функції

$$\alpha: X \to Y, \qquad (2)$$

Яка здатна ідентифікувати довільний об'єкт $x_i \in X$. Для побудови функції шаблону α можна використовувати лінійну модель з вектором параметрів $w=\{w_0, w_{1,\dots,} w_n\}$, де n довжина вектора x_i. Звідси функція шаблон α має такий вигляд:

$$\alpha(x, w) = w_0 + w_1 x_1 + \dots + w_n x_n$$

при цьому завдання ідентифікації користувача можна звести до задачі бінарної класифікації з множиною $Y=\{-1;1\}$. В цьому випадку шаблон користувача буде мати наступний вигляд:

$$\alpha(x, w) = sign \sum_j^n w_j x_j, \ (x_0 = 0)$$

Знайдені параметри і будуть шаблоном поведінки користувача і забезпечувати оптимальне значення функціоналу якості. У цьому завданню мінімізується функціонал помилок - це середня кількість розбіжностей, де $L(\alpha, x_i)$ – це функція втрат.

Функція помилок має такий вигляд:

$$Q(\alpha, X) = \frac{1}{l}\sum_i^l L(a, x_i) = \frac{1}{l}\sum_i^l [a(x_i) - y_i] \to min, \quad (5)$$

Висновки

Підвівши підсумки результатів експерименту можна виділити наступні переваги і недоліки ідентифікації користувача по його поведінці в інформаційній системі.

Переваги.

Простота реалізації і впровадження. Не потрібно спеціального апаратного забезпечення, дані беруться з системи моніторингу стану інформаційної системи, а значить - використання не потрібне придбання ніякого додаткового обладнання. Це найдешевший спосіб ідентифікації по біометричних характеристик суб'єкта доступу.

Не вимагає від користувача ніяких додаткових дій і навичок. не представляється можливим скопіювати профіль поведінки аутентифицированного користувача.

- Можливість прихованої ідентифікації.

Недоліки.

На етапі первинного експлуатування потрібні додаткові витрати по часу, для побудови профілю поведінки.

Сильно залежить від безлічі станів в яких може перебувати конкретний користувач системи.

Список літератури

[1] [SANS 2016м IT Security Spending Strategies Survey // [Online] https://www.sans.org/webcasts/2016-security-spending-strategies-survey-100997

[2] How to Get the Absolute Most from Your Cybersecurity Budget // [Online] https://www.stickman.com.au/how-to-get-the-absolute-most-from-your-cybersecurity-budget/

[3] Kim H., Lee E. A. Authentication and Authorization for the Internet of Things //IT Professional. – 2017. – Т. 19. – №. 5. – С. 27-33.

[4] Ali M. L. et al. Keystroke biometric systems for user authentication //Journal of Signal Processing Systems. – 2017. – Т. 86. – №. 2-3. – С. 175-190.

[5] Wu F. et al. A lightweight and robust two-factor authentication scheme for personalized healthcare systems using wireless medical sensor