УДК 004.891:005.3

# ПРОТОТИПІЗАЦІЯ UEBA СИСТЕМИ

Єременко Д.М.
Науковий керівник – к.т.н., доц. Іванова О.О.
Харківський національний університет радіоелектроніки, каф. КРіСТЗІ,
м. Харків, Україна
e-mail: dmytro.ieremenko@nure.ua

The work is devoted to the development of a prototype system for detecting anomalous user behavior based on the analysis of their behavioral biometric characteristics to create new ways to provide analytical data to the analyzing service with a description of why the identified actions are considered anomalous.

Прийняття управлінських рішень керівниками організацій за результатами розслідування інцидентів інформаційної безпеки має ґрунтуватися на основі реальних даних, що збираються з аналізованого об'єкта. Таким об'єктом, у нашому випадку, є співробітник організації, а управлінськими рішеннями – рішення, які робить керівництво з урахуванням отриманих з допомогою мобільного додатку поведінкових даних та їх відхилень від еталонного профілю користувача.

Однак, на основі вибраних з пристрою користувача даних досить складно здійснити прийняття будь-якого управлінського рішення, оскільки дані є різнорідними, а їх обсяги дуже великі. Для вирішення цієї проблеми можна використовувати методи інтелектуального аналізу даних (класифікація, регресія, асоціація, кластеризація, послідовні шаблони, аналіз відхилень). Це дозволить отримувати короткий перелік результуючих параметрів та приймати виважені управлінські рішення. Таким чином, для реалізованої UEBA системи з функціоналом підтримки прийняття рішень (DSS – Decision Support System), заснованої на аналізі поведінкових біометричних характеристик персоналу організації, у зв'язку з великим обсягом вхідних даних, що аналізуються, пропонується використовувати методи інтелектуального аналізу даних.

Для збору вхідних даних використовуються мобільні пристрої співробітників організації з ОС Android. Програмне забезпечення вказує на певні відхилення поведінкових характеристик користувача та пропонує здійснити ряд дій адміністратору. У деяких випадках адміністратор системи ухвалює рішення про блокування користувача. Для поведінкового аналізу пропонується використовувати такі методи:

1) нейронні мережі,

2) метод k найближчих сусідів.

Нейронні мережі застосовуються для аналізу таких даних, як записані дзвінки, записаний звук з диктофона та фотографії. Для знаходження в них відхилень відбувається попереднє навчання мережі.

Для знаходження відхилень від еталонного профілю користувача в таких даних як історія переміщень співробітника (GPS), текст, що набирається, отримуваний текст застосовується метод k найближчих сусідів. Використання цього методу дозволяє зменшити навантаження під час аналізу даних, і скоротити кількість ітерацій під час навчання. У процесі навчання цей метод лише зберігає тренувальні дані. Класифікація здійснюється при отриманні на вході алгоритму нових немаркованих даних. У цьому випадку відбувається перевірка отриманих від користувача даних та пошук їх приналежності до певної групи користувачів або певного користувача.

Порівняння характеристик користувача здійснюється шляхом пошуку Евклідової відстані до всіх записів отриманої вибірки. Потім проводиться відбір k записів, для яких евклідова відстань від поточного запису до нового буде мінімальною. Далі для кожного користувача здійснюється підрахунок суми зворотних квадратів відстаней між записами цього класу та новим записом. Новому запису присуджується клас, у якого сума зворотних квадратів виходить найбільшою.

Якщо ідентифікатор користувача або групи, присвоєний алгоритмом класифікації до аналізованого запису, відповідає ідентифікатору користувача або групи, отриманого під час початкової авторизації в системі, вважається, що отримані характеристики відповідають еталонним і відхилення від еталонного профілю не знайдено.

Якщо ідентифікатор отриманий при початковій авторизації в системі не відповідає ідентифікатору, присвоєному методом до нового сформованого запису, то вважається, що отримані характеристики відрізняються від еталонних або належать іншому користувачеві або групі користувачів.

На основі «просіяних» даних програмне забезпечення пропонує здійснити низку дій адміністратору, вказуючи на певні відхилення користувача від еталонного профілю. У деяких випадках адміністратор системи ухвалює рішення про блокування користувача.

На кожен мобільний пристрій, підключений до системи, встановлюється мобільний додаток – клієнт. Після встановлення програми, на мобільний пристрій співробітника, адміністратор системи призначає користувачеві перелік параметрів, що збираються. Набір аналізованих параметрів, які будуть збиратися на пристрої та аналізуватися на сервері, залежить від користувача / групи користувачів. Формування переліку аналізованих параметрів та груп користувачів здійснюється адміністратором системи.

Мобільний додаток запускається при старті мобільного пристрою як сервіс. Мобільний пристрій запитує список команд із сервера через певний інтервал часу. Після отримання команд іде їх обробка, отримання відповідної інформації та відправлення даних на сервер. Команди мають різні пріоритети виконання. Також команди мають різні статуси, такі як

одноразове виконання та циклічне виконання з таймером. Після надсилання даних на сервер відбувається їх прийом головним сервером, обробка, подальший аналіз, та запис до бази даних вхідних та результуючих параметрів. У разі помилки команда виконується повторно.

Архітектура побудована таким чином, що з боку клієнтських пристроїв неможливо отримати інформацію з бази даних, унеможливлює витік інформації про користувачів. Мобільні пристрої отримують лише перелік команд, які мають виконати та відповісти серверу.

Панель адміністратора підключається безпосередньо до головного сервера і має наступні можливості:

1) керування групою чи певним користувачем;

2) додавання нових команд;

3) створення звітів.

Прямий доступ до сервера забезпечує постійний доступ до управління на випадок атаки ddos атак.

Підводячи підсумки, слід зазначити наступне.

Завдяки застосуванню методів інтелектуального аналізу даних у прототипованій системи можна з впевненістю очікувати високої інформативності даних про переміщення співробітників. Застосування цих методів дозволить знаходити аномалії у переміщеннях кожного користувача системи та приймати виважені управлінські рішення щодо працівника та застосовувати до нього відповідні дисциплінарні санкції.

Список використаних джерел:

1. Cai L., Zhu Y. The challenges of data quality and data quality assessment in the big data era. Data science journal. 2015. 14.

2. Cao J. et al. Big data: A parallel particle swarm optimization-back-propagation neural network algorithm based on MapReduce. PloS one. 2016; 11(6).

3. Chen H., Chiang R. H. L., Storey V. C. Business intelligence and analytics: From big data to big impact .MIS quarterly. 36(4). 2012.

4. Dutt A., Ismail M. A., Herawan T. A systematic review on educational data mining .IEEE Access. 5. 2017.

5. Wang J., Neskovic P., Cooper L. N. Improving nearest neighbor rule with a simple adaptive distance measure .Pattern Recognition Letters. 2007; 28(2): 2007. P. 207-213.

6. Yan Z. et al. Energy-efficient continuous activity recognition on mobile phones: // An activityadaptive approach: 16th international symposium on wearable computers. IEEE. 2012.P.16

7. Грицаненко Я. Ю. UBA-аналіз як засіб підвищення інформаційної безпеки автоматизованих систем / Я. Ю. Грицаненко // Радіоелектроніка та молодь у XXI столітті : тези доповідей 27-го Міжнародного молодіжного форуму, 10–12 травня 2023 р. – Харків : ХНУРЕ, 2023. – Т. 3. – С. 233–234.