Висновки

Розроблено програмне забезпечення, що реалізує методи перевірки однорідності вибірок багатовимірних статистичних даних. Обчислення належних імовірнісних характеристик розподілів, що виконують роль індикаторів перевірки, виконується самою програмою. Як наслідок, звертання до спеціальних статистичних програм або таблиць не потрібне. Дане програмне забезпечення може бути застосовано для розв'язання часто виникаючих на практиці задач порівняння результатів різних технологічних процесів, опитувань різних верств населення, соціальних стандартів.

Література

1. "Statistical Tables for the Social, Biological, and Physical Sciences" by Albert A. Bennett and James O. Berger: Springer, 2021, 632 p.
2. "Tables of F-and Related Distributions with Algorithms" by Mardia K., Zemroch P.: Academic Press, 1978, 254 p.
3. Краснитський С.М. Щербань В.Ю. та ін. Векторні випадкові величини і випадкові процеси — К.: КНУТД, 2008. — 191 с.

ПУТІЄНКО В.Р., КРАСНИТСЬКИЙ С.М.

ЗАСТОСУВАННЯ ІМОВІРНІСНОГО КЛАСИФІКАТОРА БАЙЄСА ДЛЯ ВИЯВЛЕННЯ НЕСАНКЦІОНОВАНИХ ДОСТУПІВ ДО КОМП'ЮТЕРНИХ МЕРЕЖ

PUTIIENKO V.R., KRASNITSKIY S.M.

APPLICATION OF THE BAYESIAN PROBABILITY CLASSIFIER FOR DETECTING UNAUTHORIZED ACCESS TO COMPUTER NETWORKS

With the growing reliance on computer networks for critical information exchange and data storage, the need to ensure their security has become paramount. Unauthorized access to computer networks can lead to significant consequences, including data breaches, privacy violations, and financial losses. To address this issue, this paper explores the application of a Naive Bayesian classifier as a tool for detecting unauthorized access to computer networks. *Keywords: unauthorized access, computer networks, naive Bayesian classifier, intrusion detection*

Вступ

Виявлення несанкціонованого доступу до комп'ютерних мереж є складним завданням через постійну зміну природи кіберзагроз та досвідченості зловмисників. Для вирішення цього виклику звертають увагу на використання алгоритмів машинного навчання, оскільки вони ефективно аналізують великі обсяги даних мережевого трафіку і розпізнають шаблони, що вказують на спроби несанкціонованого доступу.

Основною метою цієї роботи є дослідження застосування імовірнісного класифікатора Байєса як інструменту для виявлення несанкціонованого доступу до комп'ютерних мереж. Наївний байєсівський класифікатор — це ймовірнісний алгоритм машинного навчання, відомий своєю простотою та ефективністю в класифікаційних завданнях. Використовуючи можливості цього класифікатора та аналізуючи дані мережевого трафіку, можна розробити надійну систему виявлення вторгнень, яка зможе точно ідентифікувати спроби несанкціонованого доступу та відрізняти їх від нормальної поведінки мережі.

Основна частина

Виявлення вторгнень є критично важливим компонентом безпеки мережі, який передбачає виявлення несанкціонованих дій або моделей поведінки в комп'ютерній мережі.

Система виявлення атак, відома також як СВА (Intrusion Detection System, IDS), є системою, яка автоматизує процес аналізу подій в інформаційно-комунікаційній системі з метою забезпечення її безпеки. В сучасному світі СВА вважається невід'ємною складовою інфраструктури безпеки, що має вирішальне значення[1]. Ці системи мають за мету виявити будь-які спроби несанкціонованого доступу, зміни конфігурації або шкідливої активності, що можуть вказувати на потенційні загрози безпеці. Вони аналізують мережевий трафік, системні журнали та інші джерела даних, шукаючи аномалії, невідповідності або характеристики, які вказують на вторгнення.

Наївні методи Байєса — це набір контрольованих алгоритмів навчання, заснованих на застосуванні теореми Байєса з «наївним» припущенням про умовну незалежність між кожною парою ознак, заданою значенням змінної класу. Теорема Байєса формулює наступну залежність, припускаючи класову змінну *y* та залежний вектор ознак x_1 через x_n [2]:

$$P(y \mid x_1, \dots, x_n) = \frac{P(y)P(x_1, \dots, x_n \mid y)}{P(x_1, \dots, x_n)}$$

Використовуючи наївне припущення про умовну незалежність, маємо

$$P(x_i \mid y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = P(x_i \mid y)$$

для усіх *i*, це співвідношення спрощено до

$$P(y \mid x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^{n} P(x_i \mid y)}{P(x_1, \dots, x_n)}$$

Оскільки $P(y \mid x_1, \dots, x_n)$ стала з урахуванням вхідних даних, ми можемо використати таке правило класифікації:

$$P(y \mid x_1, \dots, x_n) \propto P(y) \prod_{i=1}^{n} P(x_i \mid y)$$

$$\downarrow$$

$$\hat{y} = \arg\max_y P(y) \prod_{i=1}^{n} P(x_i \mid y)$$

Ми можемо використовувати оцінку Maximum A Posteriori (MAP) для оцінки $P(y)$ та $P(x_i \mid y)$, перша є відносною частотою класу y в навчальному наборі.

Наївний байєсівський класифікатор працює шляхом оцінки розподілу ймовірностей ознак у навчальному наборі даних і використання цієї інформації для класифікації нових екземплярів. У контексті виявлення вторгнень такі функції мережевого трафіку, як IP-адреса джерела, IP-адреса призначення, тип протоколу та номер порту, можна використовувати як вхідні характеристики для класифікатора. Для реалізації запропонованого алгоритму було обрано мову програмування Python та бібліотеку Scikit-learn.

Попередні дослідження[3, 4] показали багатообіцяючі результати у застосуванні наївного байєсовського класифікатора для виявлення вторгнень у комп'ютерні мережі. Навчаючи класифікатор на позначених наборах даних, що містять як нормальний, так і зловмисний мережевий трафік, він може навчитися розрізняти звичайні дії від зловмисних.

Здатність байєсівського класифікатора обробляти багатовимірні дані та адаптуватися до мінливих умов мережі робить його придатним вибором для виявлення несанкціонованого доступу на основі наборів даних мережевого трафіку. Його простота та ефективність роблять його практичним рішенням для систем виявлення вторгнень у режимі реального часу.

Висновок

Успішне застосування наївного байєсовського класифікатора у виявленні несанкціонованого доступу має значний вплив на безпеку комп'ютерних мереж. Шляхом включення цього алгоритму до систем виявлення вторгнень, організації можуть покращити свої заходи безпеки мережі та зміцнити захист від несанкціонованого доступу.

Використання алгоритмів машинного навчання може допомогти автоматизувати процес виявлення та зменшити залежність від ручних підходів на основі правил або сигнатур. Ця автоматизація дозволяє ефективно та своєчасно виявляти потенційні вторгнення, дозволяючи організаціям швидко реагувати та пом'якшувати вплив неавторизованого доступу.

Результати цього дослідження дають цінну інформацію для розробки передових систем виявлення вторгнень. Оптимальні методи вибору функцій і методи попередньої обробки даних, виявлені під час дослідження, сприяють вдосконаленню майбутніх алгоритмів виявлення вторгнень, покращуючи їх продуктивність і точність.

Література

1. М. В. Грайворонський, О. М. Новіков. Безпека інформаційно-комунікаційних систем — 2009. — 608 с
2. Naive Bayes – Scikit learn [Електронний ресурс] – Режим доступу до ресурсу: https://scikit-learn.org/stable/modules/naive_bayes.html
3. Mukherjeea S. and Sharmaa N., "Intrusion Detection using Naive Bayes Classifier with Feature Reduction," Procedia Technology, vol. 4, 2012.
4. Kanagalakshmi R. and Naveenantony V., "Network Intrusion Detection Using Hidden Naive Bayes Multiclass Classifier Model," International Journal of Science, Technology and Management, vol. 3, no. 12, 2014

КРАСНИТСЬКИЙ С.М., СВЕРГУН М.М.

РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ЛІНІЙНОГО ПРОГНОЗУВАННЯ ЧИСЛОВИХ ПОКАЗНИКІВ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ

KRASNYTSKY S.M., SVERHUN M.M.

DEVELOPMENT OF SOFTWARE FOR LINEAR FORECASTING OF NUMERICAL INDICATORS OF TECHNOLOGICAL PROCESSES

Develop software for linear forecasting of values of quantitative characteristics of output and input products of selected stages of technological processes, as well as for correlation analysis of relationships between these characteristics of selected stages.

Keywords: multi-stage technological process, covariance matrix, linear prediction.

Вступ

Прогнозування значень характеристик кінцевих та проміжних продуктів технологічних процесів є важливою задачею їх кількісного аналізу та керування ними. При розв'язанні таких задач виявив себе досить адекватним і був перевіреним на багатьох прикладах реальних ситуацій метод лінійного прогнозування. Оскільки метод потребує доволі