

ДОСЛІДЖЕННЯ ДВОМОДАЛЬНОГО МЕТОДУ ІДЕНТИФІКАЦІЇ КОРИСТУВАЧІВ ЕОМ ЗА ІНФОРМАЦІЙНИМ ПОЧЕРКОМ

Маслій Д.В.

Науковий керівник – к.т.н., доц. Горелов Д.Ю.

Харківський національний університет радіоелектроніки
(61166, Харків, пр. Науки, 14, навчально-наукова лабораторія «Систем
технічного захисту інформації (відеоспостереження, охоронні сигналізації
і контроль доступу)», тел. (057) 702-14-78

The objective of the work is explore the possibility of using nonstop
secretive monitoring of keystroke and mouse dynamics in computer networks'
user identification tasks.

У математичній постановці біометричні системи вирішують задачу
класифікації векторів біометричних ознак на класи: «свій» або «чужий». Побудова
вирішальних правил для класифікації заснована на порівнянні векторів користувачів,
які претендують на доступ, з еталонними векторами всіх зареєстрованих користувачів,
що зберігаються в базі даних. Імовірнісний характер векторів біометричних ознак
приводить до необхідності побудови середньостатистичних еталонів на основі серії
зразків, пред'явлених системі в режимі навчання.

Побудова вирішальних правил для класифікації біометричних ознак
зазвичай складна задача по наступних причинах: відсутність відомостей
про закони й числові характеристики розподілів векторів; має місце
статистична нестабільність розподілів векторів; відсутня представницька
вибірка навчальних векторів «чужих» користувачів (що впливає на рівень
помилки FAR); обмежений розмір представницької вибірки навчальних
векторів «своїх» користувачів (що впливає на рівень помилки FRR); велика
розмірність простору вхідних даних (векторів). Окремо взяті методи
біометричної ідентифікації мають певні переваги й недоліки, пов'язані з
імовірністю неправильного спрацьовування або відмови: жоден з нині
існуючих методів не дозволяє забезпечити 100% імовірність коректної
ідентифікації. Застосування комплексного підходу, заснованого на
сукупності декількох методів ідентифікації, дозволяє підвищити точність й
мінімізувати ймовірність некоректного спрацьовування.

Для проведення досліджень в якості бази даних клавіатурних почерків
було використано розташований у вільному доступі Keystroke Dynamics
Benchmark Data Set. Цей датасет містить 31 часовий параметр вводу
парольної фрази «.tie5Roanl», яку набирали 51 користувач по 400 раз
кожен. В якості бази даних курсорних почерків (динаміки миші) було
використано розташований у вільному доступі Mouse Dynamics Challenge
Data Set. Для цієї БД було розраховано 36 інформаційних параметрів
курсорного почерку 10 користувачів. У якості програмного засобу

прототипування процесу класифікації випростовувався Neural Network Toolbox математичного пакету Matlab.

Для перевірки системи ідентифікації за клавіатурним почерком 200 профілів по 31 інформаційному параметру було розділено на три множини: 140 (70 %) біометричних векторів було використано для навчання нейронної мережі, 30 (15 %) векторів було використано для оптимізації структури мережі та 30 (15 %) векторів – для перевірки якості роботи мережі. Точність системи не дуже висока – 72 % – це пояснюється малим об’ємом навчальних даних – лише по 20 профілів для кожного з 10 користувачів. Збільшення кількості профілів призведе на значного росту точності.

Для перевірки системи ідентифікації за динамікою комп’ютерної миші 200 профілів по 36 інформаційних параметрів було розділено на три множини: 140 (70 %) біометричних векторів було використано для навчання нейронної мережі, 30 (15 %) векторів було використано для оптимізації структури мережі та 30 (15 %) векторів – для перевірки якості роботи мережі. Точність системи низька – лише 56 %. Це можна пояснити малим об’ємом навчальних даних.

Для перевірки системи ідентифікації за інформаційним почерком 200 профілів по 67 інформаційні параметри було розділено на три множини: 140 (70 %) біометричних векторів було використано для навчання нейронної мережі, 30 (15 %) векторів було використано для оптимізації структури мережі та 30 (15 %) векторів – для перевірки якості роботи мережі. Точність системи склала 89.5 %, що можна рахувати задовільним результатом.

Отже, використання двомодальної системи, а саме розпізнання користувачів ЕОМ за клавіатурним почерком та динамікою комп’ютерної миші, дозволяє значно підвищити точність ідентифікації у порівнянні з одномодальними системами. Особливо це важливо для випадків малих об’ємів навчальних баз даних.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Keystroke Dynamics Benchmark Data Set. Режим доступу: <http://www.cs.cmu.edu/~keystroke>.
2. Balabit Mouse Dynamics Challenge Data Set. Режим доступу: <https://github.com/balabit/Mouse-Dynamics-Challenge>.
3. Traore, I. Woungang, M. Obaidat, Y. Nakkabi, and I. Lai. Combining mouse and keystroke dynamics biometrics for risk-based identification in web environments. In 4th Int. Conf. on Digital Home, pages 138–145, 2012.
4. Aliksieiev Vasyl, Elena Sharapova, Olena Ivanova, Gorelov Denis, Synytsia Yuliia. Web-Based Application to Collect and Analyze Users Data for Keystroke Biometric Identification. In Proceedings of the First IEEE Ukraine Conference on Electrical and Computer Engineering (UKRCON). Pages 917-922, 2017.