УДК 004.056.53+ 519.85

Автентифікація суб'єктів за клавіатурним почерком з використанням диграфів

Автор: *Ю.О. Синиця, Національний університет кораблебудування імені адмірала Макарова, м. Миколаїв*

Науковий керівник: *О.В. Блінцов, к.т.н., доц., Національний університет кораблебудування імені адмірала Макарова, м. Миколаїв*

Вступ. Однією з вимог до функціонування інформаційної системи є забезпечення захисту від несанкціонованого доступу і уразливості компонентів та зв'язків системи. До складової частини комплексу захисту інформаційної системи входить організація ефективного контролю доступу, головним аспектом якого є надійні механізми автентифікації [1].

Постановка проблеми. На сучасному етапі розвитку інформаційних технологій найбільш поширеними є методи автентифікації, що використовують секретні знання суб'єкта (паролі), факт володіння ключем (токеном), або їх комбінацію з формуванням так званої двофакторної автентифікації. Однак подібні методи мають ряд недоліків, що знижують їх надійність: суб'єкт може забути пароль, загубити токен або передати їх сторонній особі внаслідок халатності чи прийомів соціальної інженерії [2].

Для вдосконалення та підвищення надійності автентифікації використовуються методи, що базуються на вимірюваннях психофізіологічних відмінностей суб'єкта – біометрії. Біометричні дані, отримані з вимірювання особливостей поведінки суб'єкта, що складають основу поведінкової біометрії, є складнішими для підроблення та дозволяють виконувати постійну автентифікацію суб'єкта впродовж діяльності. У комбінації з іншими методами автентифікації використання засобів поведінкової біометрії дає змогу знизити ризик несанкціонованого доступу.

Однією з складових поведінкової біометрії є клавіатурний почерк, який характеризує стиль роботи користувача з клавіатурою шляхом аналізу інтервалів між натисканнями клавіш на клавіатурі.

Аналіз останніх досліджень і публікацій. У роботах [3-5] описано методику автентифікації за клавіатурним почерком, що аналізує ритм набору на клавіатурі. Головна ідея методики основана на вимірюванні близькості пред'явленого для автентифікації підпису (паролю або парольного слова) до еталону за допомогою аналізу відстані Хеммінга. Перевагами такого способу автентифікації є відносно невеликі вимоги до оперативної пам'яті

комп’ютера для збереження параметрів та простота використання, а недоліками – залежність процесу формування еталону від тексту, що набирається, необхідність точно повторювати парольну фразу при кожній спробі автентифікації та зависока ймовірність помилок першого і другого роду через невелику довжину парольної фрази [4].

Більшість комп'ютерних систем у якості основного методу автентифікації користувачів використовують первірку логіна (ID) та пароля. У невеликих локальних мережах такий рівень безпеки може бути достатнім, але коли комп'ютери підключені до глобальних мереж, рівень уразливості збільшується.

З метою зниження уразливості та підвищення надійності автентифікації застосовують засоби біометрії. У роботі [6] досліджено використання поведінкової біометрії на основі клавіатурного почерку. Наведено декілька реалізацій засобів аналізу клавіатурного почерку, ефективність яких є залежною від параметрів даних та методики їх отримання. Результати дослідження [6] показали, що рівна ймовірність помилок першого і другого роду (EER) в значній мірі залежить від процесу вибору біометричних параметрів, що аналізуються, і в меншій мірі від вибраного алгоритму автентифікації.

У роботі [7] розглянуті питання біометричної ідентифікації користувачів комп’ютерних систем. Запропоновано модель ідентифікації користувачів розподілених інформаційних систем за клавіатурним почерком. Модель заснована на використанні математичного апарату теорії нечітких множин і теорії ймовірностей. Практична реалізація запропонованої моделі здійснена за допомогою технології програмних агентів.

У роботі [8] розглядається задача ідентифікації користувачів комп'ютерних систем на основі багатозв'язного подання особливостей їх клавіатурного почерку. Запропоновано процедуру ідентифікації, яка дозволяє здійснювати біометричну ідентифікацію в реальному масштабі часу.

Розглянуті роботи присвячені в основному створенню математичних моделей безвідносно до особливостей їх програмної чи апаратної реалізації. Також немає об’єктивної можливості виділення єдиного підходу до створення програмного забезпечення біометричного контролю, яке б було доступне у вільному користуванні.

Дана робота має за мету розроблення практичної методики накопичення біометричних даних для континуальної автентифікації суб’єктів з урахуванням оновлення еталону за клавіатурним почерком з використанням диграфів натиску у якості персональних ознак.

Основний матеріал. Основу методики, що пропонується, складає аналіз n-графів – комбінацій послідовного натиску або відпускання n клавіш. З аналізу характеристик деякої

кількості n-графів можна побудувати еталонну модель клавіатурного почерку суб'єкта автентифікації з можливістю подальшої його ідентифікації [5].

Найбільша доцільність вбачається у використанні диграфів (2-графів). Це пояснюється ефективністю та простотою програмної реалізації методики, пов'язаної з швидкою обробкою та незначною кількістю необхідної пам'яті, у порівнянні з використанням три- та n-графів, а також значним обсягом статистичних даних. Додатковою перевагою використання диграфів є можливість проведення прихованої автентифікації.

Згідно з запропонованою методикою основним біометричним параметром є час виконання двох послідовних подій клавіш (диграфів), окремі комбінації яких зображені на рис. 1. Серед можливих подій клавіші розрізнюється натиснення (press) та відпускання (release).

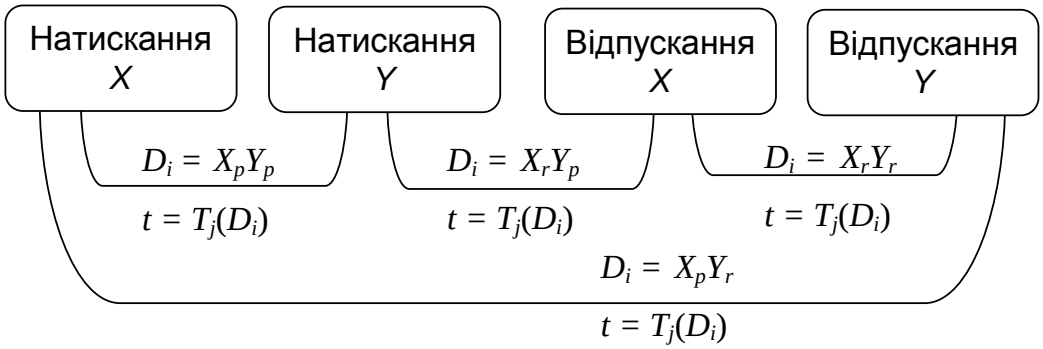

Рис. 1. Приклад диграфів для двох клавіш: D –диграф, t – час між подіями

Під час моніторингу подій клавіатури враховуються як алфавітні символи, так і службові (Shift, Ctrl, Enter, Backspace і т. д.), які разом несуть унікальну характеристику клавіатурного почерку кожного суб'єкта. Дана методика не потребує набору фіксованого зразка тексту, тому що накопичення статистики відбувається з будь-яких подій клавіатури. Відповідно, існує можливість проводити автентифікацію під час набору довільного тексту.

Така особливість методики дозволяє враховувати усі відмінності клавіатурного почерку суб'єкта, включаючи типові помилки і використання навігаційних клавіш. Сформований профіль суб'єкта дозволяє аналізувати не тільки клавіатурний почерк, а й інші характерні ознаки суб'єкта. Наприклад, використання певного програмного забезпечення, що має індивідуальні налаштування гарячих клавіш та впливає на статисти стичні данні сформованого профілю.

Формування профілю клавіатурного набору суб'єкта відбувається шляхом накопичення статисти стичних даних про часові інтервали $T(D_i)$ кожного виділеного диграфу D_i у введеному

суб'єктом тексті. Зазвичай, у процесі набору диграфи повторюються, тому в якості характеристики кожного диграфу доцільно вибрати математичне очікування $M[T(D_i)]$ та дисперсію σ^2 всіх часових інтервалів введення відповідних комбінацій:

$$M\left[T\left(D_i\right)\right] = \mu = \frac{1}{n}\sum_{j=0}^{n} T_j\left(D_i\right),$$
$$\sigma^2 = \frac{1}{n-1}\sum_{j=0}^{n}\left(T_j\left(D_i\right) - M\left[T\left(D_i\right)\right]\right)^2 \quad (1)$$

Таким чином, профіль набору можна представити у вигляді асоціативного масиву (2), в якому елементами є ідентифікатори всіх введених суб'єктом диграфів, а відповідними значеннями – математичне очікування та дисперсія часових інтервалів:

$$D_i = \left(\mu, \sigma^2\right), i \in \left(0, n\right). \quad (2)$$

До профілю заносяться лише диграфи, час введення яких не перевищує граничне значення $T_{\max}$, точне значення якого має бути визначене статистичними дослідженнями. На основі описаного способу накопичення статистики клавіатурного набору суб'єкта можливо побудувати схему біометричної автентифікації. Для формування еталону клавіатурного набору суб'єкт повинен ввести текст, попередньо ознайомившись з ним для зменшення статистичних похибок. З моменту формування еталону система біометричної автентифікації діє автономно, безперервно проводячи автентифікацію суб'єкта під час роботи після накопичення достатнього об'єму статисти стичних даних для перевірки зразків клавіатурного набору.

Автентифікація суб'єкта являє собою аналіз пред'явлених біометричних параметрів на потрапляння у встановлені біометричним еталоном зареєстрованого користувача інтервали:

$$D^s_{\sigma^2} \in \left(D^e_{\sigma^2} - D^e_{\sigma}, D^e_{\sigma^2} + D^e_{\sigma}\right);$$
$$D^s_{\mu} \in \left(D^e_{\mu} - 3\cdot D^e_{\sigma}, D^e_{\mu} + 3\cdot D^e_{\sigma}\right), i \in \left(0, n\right), \quad (3)$$

де D^s_{μ}, $D^s_{\sigma^2}$ – диграфи зразків для автентифікації, D^e_{μ}, $D^e_{\sigma^2}$ – диграфи еталону, e та s – верхні індекси, які означають зразок (sample) та еталон (etalon), σ – стандартне відхилення часових інтервалів введення певного диграфу. З кожною успішною автентифікацією суб'єкта виконується оновлення профілю, що дає змогу врахувати зміни клавіатурного почерку та відповідним чином адаптувати систему автентифікації.

Висновок. Розроблено методику збору поведінкової біометричної статистики для континуальної (безперервної) автентифікації суб'єктів за клавіатурним почерком, використовуючи диграфи у якості біометричних даних та час натискання і відпускання клавіш у якості біометричних параметрів. Розроблена методика дає змогу здійснювати постійний

контроль автентичності суб'єкта, затрачувати менше пам'яті комп'ютера, не залежить від змісту тексту, що вводиться та дає змогу проводити приховану автентифікацію.

Список літератури:

1. Bell E. D., La Padula J. L. Secure computer system: Unified exposition and Multics interpretation. – Bedford, MA: Mitre Corporation. – 1976. – Режим доступу: http://csrc.nist.gov/publications/history/bell76.pdf.

2. BBC News (Technology). LinkedIn passwords leaked by hackers. – 2012. – June. – Режим доступу : http://www.bbc.com/news/technology-18338956.

3. Синиця Ю.О. Біометричний контроль доступу на базі динамічного клавіатурного паролю // Всеукраїнський конкурс студентських наукових робіт, галузь знань «Інформаційна безпека»/ Укл. В. Б. Дудикевич, Ю. Р. Гарасим. – Львів, 2012. – 41 с.

4. Блінцов О.В., Синиця Ю.О. Дослідження системи контролю доступу на основі динамічного клавіатурного паролю. // «Електротехніка і електромеханіка»: Матеріали з Всеукраїнської науково-технічної конференції молодих учених та студентів. – Миколаїв, НУК, 2011. – с. 130-132.

5. Брюхомицкий Ю.А., Казарин М.Н. Исследование биометрических систем динамической аутентификации пользователей ПК по рукописному и клавиатурному почеркам. // Учебно-методическое пособие к циклу лабораторных работ по курсу: «Защита информационных процессов в компьютерных системах». − Таганрог: Изд-во ТРТУ, 2004. – 38 с.

6. A machine learning approach to keystroke dynamics based user authentication / Kenneth Revett, Florin Gorunescu, Marina Gorunescu et al. // Int. J. Electron. Secur. Digit. Forensic. – 2007. – May. – Vol. 1, no. 1. – P. 55–70.

7. Л.Е. Чала. Модель ідентифікації користувачів за клавіатурним почерком. // «Штучний інтелект». – Харків, ХНУРЕ, 2004. – с. 811-817.

8. Л.Е. Чалая. Метод идентификации пользователей информационных систем на основе многосвязного представления клавиатурного почерка. // Системи обробки інформації, – Харків, ХНУРЕ, 2007. – с. 98-101.