

**В.Г. ІВАНОВ**, д-р. техн. наук, проф., зав. каф. НЮАУ ім. Я. Мудрого (м. Харків),

**Н.І. МАЗНИЧЕНКО**, ст. викладач НЮАУ ім. Я. Мудрого (м. Харків)

## **ІДЕНТИФІКАЦІЯ КОРИСТУВАЧА ЕОМ НА ОСНОВІ ІНФОРМАЦІЙНОГО ПОЧЕРКУ**

У статті розглянуті основні принципи побудови, області вживання, особливості використання автоматизованих систем ідентифікації користувачів комп'ютерних систем та мереж за інформаційним почерком. Досліджені деякі підходи до ідентифікації користувача ЕОМ за допомогою двох складових інформаційного почерку: клавіатурний почерк та динаміка роботи з манипулятором "миша", розглянутих в комплексі. Проаналізовані можливі сфери вживання, запропоновані перспективні області застосування.

**Ключові слова:** інформаційний почерк, ідентифікація користувача ЕОМ.

**Постановка проблеми.** У час загальної інформатизації особливу важливість і значення набувають задачі захисту інформації. Тому постійно розробляються нові методи захисту, які дозволяють збільшувати надійність і стійкість систем, призначених для вирішення такого роду задач [1].

Основною частиною забезпечення безпеки інформаційної системи є завдання обмеження кола осіб, які мають доступ до конкретної інформації. Ідентифікація користувача ЕОМ – завдання, вирішення якого дозволяє організувати весь процес управління правами доступу, а також реалізувати ряд інших допоміжних завдань, що мають самостійне прикладне значення. З погляду використання прихованого моніторингу роботи за персональним комп'ютером представляє інтерес класифікація психофізичних параметрів користувача, до яких відносяться: клавіатурний почерк, підпис мишею, реакція на події, що відбуваються на екрані [2]. Термін, що з'явився в зв'язку з цим, інформаційний почерк користувача (далі ІПК) відображає стиль роботи користувача з деякими пристроями введення.

**Аналіз літератури.** Останнім часом все більша увага звертається на методи ідентифікації особи користувача і його фізичного стану по динаміці підсвідомих рухів. Мова йде про відпрацьовані рухові навички людини і можливості їх моніторингу по відповідних ознаках. Спостереження за ІПК під час роботи користувача з персональним комп'ютером дозволяє відстежувати його працездатність і характер виконуваної роботи.

У літературі наводяться результати аналізу написання букв [3], спостереження клавіатурного почерку [4, 5] і "інформаційного почерку" [6, 7]. Відомі роботи по розпізнаванню клавіатурного почерку використовують лише наступні параметри:  $V_n$  – середня швидкість натиснення клавіші,  $t_{уд}$  – середній час утримання клавіші,  $t_n$  – середній час між натисненнями. Такий

клавіатурний почерк розпізнається засобами нейронних мереж з надійністю 0,99. [4] При цьому здійснюється не моніторинг роботи користувача, а формування ІПК у момент введення паролльної фрази і розпізнавання її образу при вході в систему. Але в жодній роботі не розглядається безперервний режим моніторингу роботи користувача.

Існуючі дослідження моніторингу маніпулятора миша при роботі користувача в системі показують надійність розпізнавання 0,8 – 0,9 [8, 9]. Але жодне дослідження не розглядає моніторинг всього процесу еволюції системи "користувач-миша" впродовж тривалого (потенційно необмеженого) інтервалу часу спостереження за користувачем.

**Мета статті.** На сьогоднішній день багато питань ідентифікації користувачів за ІПК не вивчено. Актуальна розробка нових методів, алгоритмів і їх програмно-апаратних реалізацій, що дозволить значно підвищити ефективність автоматизованих систем, призначених для цієї мети. Також представляється цікавим розгляд сфер вживання ідентифікації користувача за ІПК, яким до останнього часу призначалося мало уваги, наприклад, моніторинг роботи користувача на протязі тривалого часу з метою відстеження його фізичного стану, що вважається дуже важливим для деяких сфер вживання, де можливі помилки під час роботи можуть мати не тільки важливе значення, але і бути катастрофічними.

**Виклад основного матеріалу.** Людська індивідуальність простежується при роботі користувачів за комп'ютером. Для вирішення цієї задачі активно використовуються методи біометрії – ідентифікації конкретної людини за унікальними, притаманними лише їй біологічними ознаками. Системи доступу і захисту інформації, засновані на таких технологіях є не тільки найнадійніми, але і найзручнішими для користувачів на сьогоднішній день. Біометричні технології активно застосовуються в багатьох областях, пов'язаних із забезпеченням безпеки доступу до інформації і матеріальних об'єктів.

Серед біометричних механізмів ідентифікації можна виділити такі [1]:

1) по статичних ознаках – ті, що практично не міняються згодом, починаючи з народження людини (по відбитку пальця, по малюнку райдужної оболонки ока або сітківки ока, по геометрії грона руки, за формою долоні, будові лица, формі черепа, по формі лица, ДНК і т.д.);

2) по динамічних ознаках – поведінкові характеристики, тобто ті, які засновані на особливостях, властивих для підсвідомих рухів в процесі відтворення якої-небудь дії (по голосу, по почерку, по клавіатурному почерку, по ході, по "мишиному" почерку і т.д.). Динамічні ознаки можуть змінюватися з часом, але не різко, стрибком, а поступово.

У біометричних системах ідентифікації, які засновані на динамічних ознаках, найчастіше використовуються голос, почерк і клавіатурний почерк. Привабливою рисою систем динамічної ідентифікації є те, що вони здатні паралельно з ідентифікацією оцінювати поточний психічний стан особи. З цієї

причини динамічні методи розглядаються як психологічні.

Новим напрямом моніторингу особи користувача є відстеження його поведінкових особливостей, що виявляються в характері роботи з різними маніпуляторами: світлове перо, маніпулятор "миша", клавіатура і ін. В результаті відстеження ІПК в комп'ютері може формуватися унікальний для кожного користувача образ, який може використовуватися як засіб ідентифікації його стану, рівня комп'ютерної письменності, специфіки діяльності і так далі.

Під інформаційним почерком конкретного користувача будемо розуміти область простору параметрів, що відображають особливості роботи користувача з різними пристроями введення (клавіатура, комп'ютерна миша, графічний планшет, джойстик і т. п.). Очікується, що всі користувачі мають ІПК, що розрізняються, в деякій метриці цього простору. Слід сказати, що поняття "інформаційний почерк" зустрінете тільки в роботах [2, 6, 7].

У даній роботі передбачається, що випадковий процес взаємодії користувача з пристроями введення комп'ютера – мишею і клавіатурою – стаціонарний на тимчасовому інтервалі аналізу.

При виборі параметрів для ідентифікації інформаційного почерку дуже важливо врахувати саме ті, які мають важливе значення для надійної ідентифікації користувача.

Стосовно маніпулювання комп'ютерною мишею у [8] визначена множина параметрів траєкторії курсору. Запропоновано кожен траєкторію відображати вектором п'яти параметрів:

$$m = (T, L, U_p, \delta, \alpha), \quad (1)$$

де  $T$  – час руху маніпулятора до його зупинки;  $L$  – довжина траєкторії;  $U_p$  – швидкість розгону;  $\delta$  – час між зупинкою покажчика й підтверджуючим натисненням кнопки маніпулятора;  $\alpha$  – кут напряму початкового руху. Більшість з цих параметрів визначає швидкісні характеристики (динаміку) роботи користувача з комп'ютерною мишею.

В результаті вимірювань множини траєкторій може бути сформована статистична вибірка з  $N$  векторів, представлена у вигляді матриці експерименту розмірністю  $N \times 5$ . Ця матриця є базою для різноманітної обробки з метою формування унікального комп'ютерного образу користувача. Наприклад: вектор середніх (простий випадок); 10 координат вершин гіперпаралелепіеда, що формуються мінімальними і максимальними значеннями в стовпцях матриці експерименту; 5-мірна гістограма – аналог імовірнісної міри з 5-мірною множиною інтервалів (осередків простору параметрів). Проблема нестационарного почерку може бути вирішена, зокрема, регулярним/нерегулярним повторенням акту "навчання" системи, що формує і ідентифікує ІПК.

Стосовно клавіатурного почерку дослідження показали [4, 5], що клавіатурний почерк конкретного користувача володіє стабільністю. Це

дозволяє з великою достовірністю ідентифікувати користувача, що працює з клавіатурою. У якості часових параметрів роботи користувача за клавіатурою автори вказаних робіт пропонують середні значення тимчасових інтервалів  $\tau_{Hi}$  між натисненням клавіш і тривалості часу  $\tau_{уд}$  їх утримання. При цьому тимчасові інтервали між натисненням клавіш характеризують темп роботи, а час утримання клавіш – стиль роботи з клавіатурою: різкий удар або плавне натиснення [4].

До цих двох характеристик додають середню швидкість набору тексту, зміряну в різних епізодах роботи з клавіатурою [4]:

$$U_n = \frac{\sum_{i=1}^n U_i}{n}, \quad (2)$$

де  $U_i$  – середня швидкість набору в  $i$ -му епізоді.

При цьому за деякий час вимірювання з'являється  $M$  значень 3-мірного вектора, які характеризують  $M$  епізодів взаємодії користувача з клавіатурою, – матриця експерименту  $M \times 3$ . Її обробка також може бути зведена до трьох вищеперелічених варіантів, якщо рахувати процес роботи користувача з клавіатурою стаціонарним.

Таким чином, для ідентифікації інформаційного почерку користувача необхідно відстежувати значення як мінімум восьвимірного вектора:

$$P = (T, L, U_p, \delta, \alpha, U_n, t_{y\delta}, t_n). \quad (3)$$

Дослідження взаємозв'язків між параметрами інформаційного почерку на вибірці об'ємом більше 10 000 у декількох користувачів показує, що статистичний зв'язок параметрів клавіатурного почерку і параметрів траєкторій комп'ютерної миші слабкий. Це говорить про самоцінності кожного з компонентів вектора (3) параметрів ІПК. З кореляційної матриці слідують також гіпотеза про відносну автономію кожного з параметрів і відсутності надмірності в сукупності вибраних характеристик ІПК.

У основі більшості алгоритмів розпізнавання лежить гіпотеза компактності, яка полягає в тому, що реалізації одного образу (матриця  $N \times 8$ ) в 8-мірному геометричному просторі утворюють "компактні згустки". Процес ідентифікації ІПК складається з двох основних етапів: навчання (відображення) ІПК кожного користувача і ухвалення рішення про приналежність спостережуваного процесу взаємодії користувача з ПК одному з еталонів (відображень) ІПК.

Процедура формування образу ІПК в реальному часі на етапі "навчання" може бути розбита на три стадії, що повторюються:

1) вимірювання значень параметрів подій і процесів, пов'язаних з вказаними вище пристроями під час роботи користувача з комп'ютером. Формування матриці експерименту. При цьому необхідне нормування значень параметрів;

- 2) формування образу ІПК одним з перерахованих способів;
- 3) шифрування результату.

Поповнення повчальної вибірки може відбуватися, наприклад, через кожних 5 хвилин. Програмне забезпечення для здійснення цих вимірювань створене і апробоване [7]. Висока продуктивність сучасних комп'ютерів робить цей процес абсолютно непомітним для користувача.

На цьому, першому етапі вирішальною є стадія 2. Враховуючи, що вектор  $P$  випадковий, необхідні методи математичної статистики для формування відображень ІПК. При цьому немає проблеми репрезентативності вибірки, оскільки отримання кількості вимірювань в декілька тисяч не складає труднощів.

Перша процедура заснована на відображенні ІПК крапкою в 8-мірному просторі –  $ІПК_c = (p_1, \dots, p_8)$ , де  $p_i$  ( $i = 1, 2, \dots, 8$ ) – середні значення стовпців матриці експерименту.

Друга процедура зводиться до редукції матриці експерименту в матрицю  $2 \times 8$ , першим рядком якої є мінімальні значення, другий – максимальні значення відповідних стовпців матриці експерименту. В цьому випадку образ ІПК – це багатовимірний паралелепіпед ( $ІПК_k$ ):

$$\Delta P = (p_{11} \leq p_1 \leq p_{12}, p_{21} \leq p_2 \leq p_{22}, \dots, p_{81} \leq p_8 \leq p_{82}). \quad (4)$$

Правила формування  $ІПК_c$  і  $ІПК_k$  визначають процедури ідентифікації ІПК користувача, що працює за комп'ютером. Нагадаю, що розглядаємо поки тільки стаціонарний процес взаємодії користувача з комп'ютером.

Можна запропонувати наступні алгоритми ідентифікації користувача.

Алгоритм 1. Доречно розглянути простий варіант зіставлення евклидових відстаней між поточним чином ІПК (поточною крапкою) і кожним із створених образів, допущених до системи користувачів. Істотним обмеженням цього методу є неможливість абсолютної ідентифікації користувача. Правда, цю проблему можна вирішити в процесі тривалого використання методу і створення  $\delta$ -областей  $ІПК_c$ . Для цього необхідно відстежувати флуктуації положення конкретної крапки і визначати цю область у вигляді гіперпаралелепіпеда, об'єм якого

$$V_n = (p_{1, \max} - p_{1, \min}) \times (p_{2, \max} - p_{2, \min}) \times \dots \times (p_{n, \max} - p_{n, \min}).$$

Межі гіперпаралелепіпеда розглядаються у вигляді мінімальних і максимальних значень координат або в статистичних довірчих інтервалах для координат. Останнє зажадає як мінімум перевірки гіпотези на нормальність розподілів координат.

Алгоритм 2 базується на представленні ІПК у вигляді гіперпаралелепіпеда –  $ІПК_k$ . Зокрема, можна використовувати метод "еталонів, що дробляться", – розпізнавання з використанням покриття повчальних вибірок кожного образу простими фігурами, що ускладнюються в міру необхідності [10]. Ухвалення рішення про схожість або відмінність двох  $ІПК_k$

при перетині гіперпаралелепіпедів здійснюється таким чином. При перетині двох образів область перетину ділиться на додаткові образи наступного покоління, які у свою чергу при перетині також можуть бути розбиті. Використання гіперпаралелепіпедів, як фігур що покривають, дозволить спростити процес обчислення об'ємів можливих перетинів різних образів –  $П\text{К}_k$  різних користувачів. Цей метод може дати економію часу розрахунків 30 % [10]. Рішення про схожість/відмінність  $П\text{К}_k$  ухвалюється за наслідками порівняння об'ємів всіх перетинів  $П\text{К}_k$ .

Для абсолютної ідентифікації користувача потрібна істотно велика статистика про  $П\text{К}_k$  конкретного користувача. Далі визначається мінімальне значення об'єму перетину поточного і еталонного  $П\text{К}_k$  і ухвалюється рішення про їх ідентичність.

**Висновки.** Розробки ідентифікації користувача на основі його інформаційного почерку можуть застосовуватися для підвищення надійності систем розмежування доступу, для посилення контролю над користувачами (явного або негласного) в системах, що працюють з критичними даними. На сьогоднішній день рішення питання підвищення ефективності ідентифікації користувачів ЕОМ бачиться через створення комплексних систем, що використовують для прийняття рішень декілька біометричних характеристик користувача: особливості клавіатурного почерку, динаміки при роботі з мишею, тремора (користувачу пропонується обвести за допомогою графічного планшета складну фігуру).

Виконані дослідження дозволили зробити висновки про недостатній рівень розглядання можливостей у використанні саме інформаційного почерку при встановленні особи користувача ЕОМ. Всі представлені існуючі системи взагалі пропонують дуже невелику кількість параметрів для визначення  $П\text{К}_k$ , що автоматично призводить до досить низької надійності систем, побудованих на ідентифікації за цими динамічними ознаками людини. Автоматизованих систем ідентифікації з використанням спостереження за динамікою роботи користувача з маніпулятором "миша" майже не пропонується, а комплексних систем, які б використовували декілька складових  $П\text{К}_k$  разом на сьогоднішній день взагалі не існує. Але ідентифікація на базі  $П\text{К}_k$ , на наш погляд, має певне коло застосувань, тим більше, що це єдині біометричні ознаки, які можливо спостерігати досить тривалий час (під час всієї роботи користувача на ЕОМ), що дозволить значно розширити коло задач, які вирішуються. Наприклад, ідентифікація не тільки під час входу в систему, але й протягом всієї роботи на ЕОМ; відстеження психофізичного стану, його зміни під час роботи за ПК. Ця функція буде дуже важливою в сферах діяльності, де стан користувача і можливі помилки в роботі можуть мати катастрофічні наслідки. Тобто, цей напрямок в системі захисту інформації та обмеження доступу до інформаційних систем та мереж, на наш погляд, дуже прогресивний і актуальний і потребує більш досконалого вивчення та розробки нових можливостей.

**Список литературы:** 1. *Иванов А.И.* Биометрические и нейросетевые механизмы связи с криптографическими механизмами информационной безопасности / Сборник трудов науч.-техн. конф. "Безопасность информационных технологий". – Пенза, ПНИЭИ, 2003. – Том 4. – С. 3 – 6. 2. *Бушуев С.И., Авраменко В.С.* Аутентификация пользователей в автоматизированных системах на основе информационного почерка // Проблемы современной геополитики / Сборник трудов 1-й междунауч. конф. "Проблемы современной геополитики. Продление НАТО на Восток – проблемы безопасности России и стран СНГ". – СПб: Балтийский гос. техн. ун-т "ВОЕНМЕХ". – 1999. – С. 53 – 59. 3. *Епифанцев Б.Н., Ложников П.С.* Идентификация человека по динамике написания слов в компьютерных системах / Материалы 10 междунауч. конф. студентов, аспирантов и молодых ученых "Современные методы и технологии". – Томск: Томский политехн. ун-т, 2004. – С. 163 – 164. 4. *Гузик В.Ф., Галуев Г.А., Десятерик М. Н.* Биометрическая нейросетевая система идентификации пользователя по особенностям клавиатурного почерка // Нейрокомпьютеры. Разработка, применение. – 2001. – № 7 – 8. – С. 104 – 118. 5. *Шарунов Р.Р.* Идентификация и аутентификация пользователей по клавиатурному почерку // Электронное приборостроение. Научн. практ. сб. – Казань: ЗАО "Новое знание", 2005. – Вып. 3 (44). – С. 50 – 54. 6. *Власов А.Н.* Способ представления координатной составляющей информационного почерка пользователя / Материалы междунауч. конф. по мягким вычислениям. – СПб: Изд-во Политехнического ун-та, 2003. – Т. 1. – С. 116 – 119. 7. *Диденко С.М., Шапцев В.А.* Методика отображения информационного почерка пользователя // Вестник кибернетики. – Тюмень: Изд-во ИППОС СО РАН, 2005. – С.74 – 79. 8. *Диденко С.М., Шапцев В.А.* Исследование динамики работы пользователя с манипулятором мышь // Математическое и информационное моделирование. – Тюмень: Изд-во Тюм. ун-та, 2004. – С. 295 – 304. 9. *Диденко С.М.* Разработка и исследование компьютерной модели динамики системы "пользователь – мышь". Автореферат дис. к.т.н. – Тюмень: Изд-во Тюм. гос. ун-та. – 2007. – 19 с. 10. *Загоруйко Н.Г.* Прикладные методы анализа данных и знаний. – Новосибирск: Изд-во Ин-та математики, 1999. – 270 с.

**УДК 519.711.3:343.98**

**Идентификация пользователя ЭВМ на основе информационного почерка / Иванов В.Г., Мазниченко Н.И.** // Вестник НТУ "ХПИ". Тематический выпуск: Информатика и моделирование. – Харьков: НТУ "ХПИ". – 2009. – № 13. – С. 77 – 83.

В статье рассмотрены основные принципы построения, области применения, особенности использования автоматизированных систем идентификации пользователей компьютерных систем и сетей на основе информационного почерка. Исследованы некоторые подходы к идентификации пользователя ЭВМ при помощи двух составляющих информационного почерка: клавиатурного почерка и динамики работы с манипулятором "мышь", рассмотренных в комплексе. Проанализированы возможные сферы использования, предложены перспективные области применения. Библиогр.: 10 назв.

**Ключевые слова:** информационный почерк, идентификация пользователя ЭВМ.

**UDC 519.711.3:343.98**

**Computer user identification on the basis of informative handwriting / Ivanov V.G., Maznichenko N.I.** // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2009. – № 13. – P. 77 – 83.

There are the considered basic principles of construction in the article, application, feature of the use of the automated systems of authentication of users of the computer systems and networks domains on the basis of informative handwriting. Some approaches to computer user identification through two constituents of informative handwriting are explored: keyboard handwriting and dynamics of work with the manipulator «mouse», considered in a complex. The possible spheres of the use are analyses, perspective application domains are offered. Refs: 10 titles.

**Key words:** information handwriting, computer user identification.

*Поступила в редакцию 08.04.2009*