УДК 004.052.042

МАРКОВСЬКИЙ О.П.,
МАЗУР Р.Ф.,
САЇД РЕЗА МАХМАЛІ

ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ІДЕНТИФІКАЦІЇ АБОНЕНТІВ З ВИКОРИСТАННЯМ МОДЕЛІ ПЕРСЕПТРОНУ

У статті пропонується вдосконалення використання нейромережних технологій для ідентифікації віддалених абонентів багатокористувацьких систем. Застосування моделі персептрону в якості перетворювача з неоднозначною зворотною функцією дозволяє прискорити ідентифікацію на основі концепції "нульових знань". Виявлено нові можливості прискорення обчислювальної реалізації моделі персептрону при її застосуванні для ідентифікації абонентів. Запропоновано нову схему обчислень, проведено аналітичне та експериментальне дослідження її ефективності для підвищення продуктивності ідентифікації.

The paper deals with offers how to improve the using of neuronet technologies for remote multi-user systems subscribers identification. Usage of the perceptron model as the ambiguous reversible functional transformation allows to accelerate the identification process based on "zero knowledge" concepts. New ways of speeding up the perceptron model are exposed. What the article brings out is the new calculation scheme. Its efficiency is discovered both analytically and experimentally in order to grade up the identification performance.

Вступ

Динамічний розвиток та поглиблення інформаційної інтеграції викликає зростання числа багатокористувацьких комп'ютерних систем. Такі системи, що пов'язані з абонентами через локальні та глобальні комп'ютерні мережі, дозволяють реалізувати колективний доступ до інтегрованих інформаційних ресурсів. При цьому важливим фактором ефективності багатокористувацьких систем є унеможливлення несанкціонованого доступу до їх інформаційних ресурсів, а також контроль за дотримання прав доступу з боку легальних абонентів. Окрім того, тенденцією останніх років стало не тільки зростання числа таких систем і розширення сфери їх застосування, але й збільшення кількості їх абонентів.

Указані фактори визначають необхідність удосконалення засобів ідентифікації абонентів багатокористувацьких систем, особливо коли їх застосування розширюється в таких сферах людської діяльності, неконтрольоване втручання в які може мати тяжкі наслідки. Це стосується, в першу чергу, систем управління повітряним транспортом, залізницею – тобто систем, які відповідають за забезпечення безпеки людської життєдіяльності. Розширення використання багатокористувацьких систем у галузях економіки стимулює зростання економічних злочинів, що реалізуються шляхом доступу до конфіденційної комерційної інформації. Таким чином, розширення сфер використання багатокористуваньких систем вимагає підвищення ефективності систем ідентифікації їх абонентів шляхом зменшення ризиків неконтрольованого доступу до їх інформаційних ресурсів.

Характерне для сьогоднішнього етапу розвитку багатокористувацьких систем зростання кількості їх абонентів вимагає відповідного підвищення продуктивності ідентифікації.

Оскільки підвищення надійності захисту від неконтрольованого доступу вимагає, як правило, збільшення об'єму часових ресурсів для його реалізації, то комплексне вирішення проблеми зменшення ризиків неконтрольованого доступу та підвищення продуктивності контролю за доступом може бути досягнуто лише за рахунок розробки принципово нових методів та засобів ідентифікації абонентів.

Отже, проблема вдосконалення методів та засобів ідентифікації абонентів багатокористувацьких комп'ютерних систем є актуальною і важливою для сучасного етапу розвитку інформаційних технологій.

Аналіз ефективності існуючих технологій ідентифікації абонентів

Сучасні технології ідентифікації поділяються на два класи: з використанням паролів ("слабка" ідентифікація) та на основі концепції "нульових знань" ("строга" ідентифікація). Остання, з точки зору рівня захищеності від

неконтрольованого доступу, значно переважає ідентифікацію з використанням паролів [2].

Сутність концепції "нульових знань" полягає в тому, що для доведення своєї автентичності абонент має неявним чином виявити знання певної інформації, якою система не володіє, але може перевірити її наявність у абонента. У якості механізму перевірки наявності у абонента знань виступає перетворення, зворотне до якого не може бути отримано аналітичним шляхом і є неоднозначним. У більшості протоколів ідентифікації, що реалізують цю концепцію (схеми Fiat-Shamir, Schnorr [2]) в якості вказаних перетворень виступають незворотні задачі теорії чисел. Базовими для таких перетворень є мультиплікативні модульні операції над числами, довжина яких на порядки перевищує розрядність процесора. Значна обчислювальна складність цих операцій має наслідком низьку продуктивність ідентифікації.

Тому доцільно шукати альтернативні способи реалізації базової концепції "нульових знань". Одним із напрямків такого пошуку може бути застосування перетворень, які володіють властивостями неоднозначності, але не забезпечують умову строгої математичної незворотності. Відповідно до базового принципу захисту інформації, який полягає втому, що рівень захисту вибирається таким чином, щоб потенційний зиск її порушення (отримання несанкціонованого доступу) був завжди меншим вартості ресурсів, що витрачаються на подолання захисту, існує досить широке коло багатокористувацьких систем, для яких застосування концепції "нульових знань" на основі математично строгих незворотних перетворень є надлишковим. Для таких систем, потенційна користь несанкціонованого доступу до ресурсів яких відносно невелика, при реалізації концепції "нульових знань" можуть бути використані більш прості неоднозначні перетворення, що не володіють властивістю строгої математичної незворотності. В останні роки запропоновано ряд подібних перетворень для застосування в області концепції "нульових знань". Вони є більш ефективними в обчислювальному плані, ніж мультиплікативні операції модульної арифметики, а їх незворотність основана на NP-повних задачах. У числі таких альтернативних варіантів реалізації концепції "нульових знань" можна назвати схеми ідентифікації об'єктів на основі задачі перестановки ядер (PKP – Permuted Kernels Problem), аналізу симптомів SD (Sendrom Decoding), з використанням обмежених лінійних рівнянь CLE (Constrained Linear Equations) [3].

Одним із можливих альтернативних варіантів реалізації спрощеної версії концепції "нульових знань" є використання нейромережних технологій, основаних на моделі персептрону [1,3].

Таким чином, метою досліджень є вдосконалення технології ідентифікації абонентів на основі концепції "нульових знань" з використанням для її реалізації моделі персептрону.

Протоколи ідентифікації абонентів з використанням моделі персептрону

До теперішнього часу запропоновано ряд математичних моделей базового компонента нейронної мережі, персептрону [1]. Усі ці моделі володіють властивістю неоднозначності. Це означає, що той самий вихідний сигнал персептрону може бути викликаний великим числом варіантів значень вхідних сигналів. Саме ця властивість персептрону відкриває можливості його використання в реалізації спрощеної версії концепції "нульових знань". Для пересептрону завдання знаходження зворотного перетворення, тобто одержання набору значень вхідних сигналів, які викликають заданий вихідний сигнал, є задачею NP-складності [3].

З огляду на обмежено лінійний характер перетворень, реалізованих у моделі персептрону, завдання знаходження зворотного перетворення не є, у строгому математичному формулюванні, аналітично нерозв'язною. Однак при великій кількості вхідних сигналів, її рішення вимагає значних обчислювальних ресурсів, вартість яких для широкого кола багатокористувацьких систем перевищує вартісне вираження користі від несанкціонованого доступу до її ресурсів.

Найпростіший персептрон описується матрицею W, що складається з m рядків та n стовпчиків, при цьому елементи матриці належать множині $\{-1;1\}$. Нехай існує множина R, яка містить m невід'ємних цілих чисел, кожне з яких не перевищує n. Задача знаходження n-компонентного вектора Y, кожен з елементів якого належить множині $\{-1;1\}$, такого, щоб його добуток з матрицею W був рівний векто-

ру $S = \{s_1, s_2, ..., s_m\}$, $\forall j \in \{1,2,...,m\}: s_j \geq 0 : W \cdot Y = S$, і є задачею NP-складності.

У якості однакового для всіх абонентів відкритого ключа системи використовуються матриця *W* та вектор *S*. Обидва ці компоненти генеруються випадковим чином у системі та розсилаються абонентам при їх реєстрації.

Абонент А після одержання від системи матриці *W* і вектора *S*, виконує наступну послідовність дій.

1. Спочатку випадковим чином генерується *n*-компонентний вектор $Y=\{y_1,y_2,\ldots,y_n\}$, $\forall j\in\{1,\ldots,n\}: y_j\in\{-1,1\}$.
2. Виконується множення зазначеного вектора на матрицю *W*. Результатом цього множення є вектор $R=\{r_1,r_2,\ldots,r_m\}$:$R=W\cdot Y$, кожна з *m* компонент $r_1,r_2,\ldots,r_m$ якого є додатним або від'ємним цілим числом.

Отримана матриця *W* перетворюється в модифіковану матрицю *W′*. Для цього послідовно аналізуються компоненти $r_1,r_2,\ldots,r_m$ вектора R: якщо *i*-та ($i\in\{1,\ldots,m\}$) компонента r_i вектора *R* від'ємна ($r_i < 0$), то *i*-тий рядок модифікованої матриці *W′* утворюється як інверсія однойменного рядка матриці *W*: $\forall j\in\{1,\ldots,n\}$: $w_{ij}=-1\cdot w_{ij}$. В іншому випадку, тобто за умови, що *i*-та ($i\in\{1,\ldots,m\}$) компонента r_i вектора *R* невід'ємна ($r_i \geq 0$), то *i*-тий рядок модифікованої матриці *W′* збігається з однойменним рядком матриці *W*: $\forall j\in\{1,\ldots,n\}$: $w_{ij}=w_{ij}$.

3. Обчислюється *m*-компонентний вектор S_A добутку модифікованої матриці *W′* на вектор *Y*: $S_A = W'\cdot Y$.
4. Код *m*-компонентного вектора S_A відсилається системі як відкритий ключ користувача А. Ці коди, на відміну від *W*, різні для кожного з легальних абонентів системи.

Абонент А формує випадковим чином наступні компоненти, що становлять його секретний ключ.

- Вектор P перестановок рядків матриці W, що містить m неповторюваних компонентів P={p1,p2,…,pm}, при цьому, кожна j-та компонента вектора P не перевищує m: ∀j=1,…,m: pj≤m та показує позицію i-того рядка в перетвореній матриці.
- Вектор Q ={q1,q2,…,qn}, що описує перестановки стовпчиків матриці W зі зміною знаків їхніх компонент. Кожна i-та компонента вектора Q лежить в інтервалі від –n до n: ∀i=1,…,n: -n≤pi≤n і показує позицію i-того стовпчика в перетвореній матриці. Якщо qi <0, то додатково до перестановки змінюються на протилежні знаки всіх компонентів i-того стовпчика матриці W.
- Випадковий вектор Z.

Для подальшої роботи формується вектор $Q'=\{q_1',q_2',\ldots,q_n'\}$, зворотний до вектора перестановки стовпців $Q = \{q_1,q_2,\ldots,q_n\}$. Вектор Q' здійснює зворотну перестановку стовпців матриці зі зміною знака: $\forall i\in\{1,\ldots,n\}$: $q'_{q_i} = i \cdot sign(q_i)$.

При виконанні циклу звернення до системи користувач А виконує наступну послідовність дій.

1. Здійснюється перестановка рядків і стовпчиків матриці *W* відповідно з обраними векторами перестановки рядків *P* та стовпчиків *Q*. У результаті такої перестановки рядків і стовпчиків формується матриця *W′*: $W'=Q(P(W))$.
2. Модифікується вектор *Y* шляхом зворотної перестановки його компонентів відповідно до вектора Q^{-1}. У результаті такої зворотної перестановки буде утворено модифікований вектор Y' : $Y'=Q'(Y)$.
3. Вираховується *n*-компонентний вектор *D* як сума вектора *V* і модифікованого вектора Y' : $D= Z + Y'$.
4. Обчислюються хеш-згортки наступних кодів: конкатенації векторів перестановок *P* та *Q*: $h_0 = H(P|Q)$, хеш-згортки векторів *Z* та *D* окремо: $h_1 = H(Z)$ і $h_2 = H(D)$, а також хеш-згортки добутків зазначених *n-компонентних* векторів на модифіковану матрицю W' : $h_3 = H(W'\cdot Z)$ і $h_4 = H(W'\cdot D)$.
5. Коди h_1, h_2, h_3 та h_4 відсилаються до системи.

Система випадковим чином генерує ціле число $c\in\{0,1,2,3\}$ і відправляє це число користувачеві А.

Якщо c=0, то користувач А відправляє системі три коди: *P*, *Q* та *Z*. Система перевіряє правильність передачі кодів *P* й *Q* шляхом порівняння $H(P|Q)$ з раніше отриманим кодом хеш-згортки h_0; обчислює хеш-згортку $H(Z)$ отриманого від користувача вектора *Z* і порівнює її з отриманим раніше h_1; система виконує перестановку рядків матриці *W*: $P(W)$, здійснює перестановку компонентів отриманого від користувача вектора *Z*: $Q(Z)$ і формує хеш-згортку добутку $P(W)\cdot Q(Z)$. Якщо вона збігається з раніше отриманим кодом h_3, тобто

$H(P(W)\cdot Q(Z)) = h_3$, то користувачеві А надаються права доступу.

Якщо c=1, то користувач А також відправляє системі коди трьох векторів: P, Q та D. Система виконує аналогічні обчислення й перевірки з тією різницею, що те, що виконувалося для вектора Z виконується тепер для вектора D: перевіряється правильність передачі кодів P й Q шляхом порівняння $H(P|Q)$ з раніше отриманим кодом хеш-згортки h_0; обчислюється хеш-згортка $H(D)$ отриманого від користувача n-компонентного вектора D і порівнюється з отриманим раніше h_2; виконується перестановка рядків матриці W: $P(W)$, здійснюється перестановка компонентів отриманого від користувача вектора D: $Q(D)$, і формується хеш-згортка добутку $P(W)\cdot Q(D)$. Якщо хеш-згортка вказаного добутку збігається з раніше отриманим кодом h_4, тобто $H(P(W)\cdot Q(D)) = h_4$, то користувачеві А надаються права доступу.

Якщо c=2, абонент А множить модифіковану матрицю W' на вектор Z і відсилає системі отриманий добуток – $W'\cdot Z$, після чого, аналогічним способом множить модифіковану матрицю W' на вектор Y_A', а отриманий добуток $W'\cdot Y_A'$ відсилає системі. Система, при отриманні вказаних кодів, обчислює хеш-згортку добутку $W'\cdot Z$ і порівнює отриманий результат $H(W'\cdot Z)$ з раніше отриманим кодом h_3, потім система обчислює суму отриманих від абонента добутків $W'\cdot Z$ й $W'\cdot Y_A'$, після чого формується хеш-згортка $H(W'\cdot Z+W'\cdot Y_A')$ цієї суми, що порівнюється з отриманим від абонента кодом h_4. Нарешті, система перевіряє, щоб усі компоненти добутку $W'\cdot Y_A'$ містилися в множині S_A. Якщо всі вказані вище умови виконуються, то система надає абонентові А права доступу до своїх ресурсів.

Якщо c=3, то абонент А відсилає системі сформовані вектори Z й Y_A'. Система, при одержанні цих векторів обчислює хеш-згортку $H(Z)$ вектора Z і порівнює її з кодом h_1, обчислює хеш-згортку суми $H(Z+Y_A')$ отриманих від абонента векторів Z й Y_A', порівнює її з отриманим кодом h_4, якщо у всіх випадках виявляється збіг зазначених компонентів, то абонентові А надаються права на доступ до ресурсів системи.

Обчислювальна складність наведеної процедури ідентифікації може бути зменшена за рахунок урахування особливостей виконання операцій при ідентифікації.

Спосіб зменшення обчислювальної складності реалізації моделі персептрону

Найбільш масовою й трудомісткою операцією, що використовується в наведеній вище схемі ідентифікації на основі одношарового персептрону є скалярне множення. Кожен з компонентів результату при такому множенні є сумою попарних добутків елементів векторів або матриць, які належать множині {-1,1}. Очевидно, що згадані попарні добутки також можуть приймати значення із цієї ж множини {-1,1}.

Лема: Якщо значення чисел a та b належать множині {-1,1}: $a,b\in\{-1,1\}$, а самі числа представлені в доповнюючому коді, то й їхній добуток може бути представленим у вигляді: $a\cdot b = a \oplus b \vee 1$.

Доведення: Як відзначалося вище, якщо значення множників a і b належать бінарній множині {-1,1}, то значення добутку також належить цій множині. При цьому, якщо $a=b$, то $a\cdot b$ =1, а якщо $a\neq b$, то $a\cdot b$ = –1. У першому випадку $a\oplus b$=0 і тому $a\oplus b\vee 1$=1. У другому випадку можливо два варіанти. Для визначеності, нехай a=1, b=–1, тоді $a\oplus b$ = 1⊕(–1) = -2. Отже, $a\oplus b\vee 1$ = –2∨1 = –1. Для зворотніх значень a та b результат буде таким же.

Доведена лема дозволяє реалізувати скалярне множення істотно швидше, за рахунок відмови від операції множення компонент векторів. Це має велике практичне значення для прискорення процесу ідентифікації, оскільки, для того, щоб забезпечувати прийнятний рівень захисту від підбору, реальна розмірність матриць, які використовуються, повинна лежати в діапазоні 200-300 [3].

При прямому скалярному множенні n-компонентних векторів та умові, що компоненти векторів зберігаються в неупакованому форматі, множення пари компонент вимагає однієї операції множення, тривалість якої становить t_m. Для одержання результату – скалярного добутку двох векторів затрачається час, рівний $n\cdot t_m+(n-1)\cdot t_a$, де t_a – час виконання операції додавання. Якщо прийняти, що тривалість t_m операції множення в α раз перевищує тривалість t_a операції додавання ($\alpha=t_m/t_a$), то час T_0 скалярного множення матриці, що складається з m рядків й n стовпчиків на n-компонентний вектор визначається формулою:

$$T_0 = m\cdot(n\cdot t_m+(n-1)\cdot t_a)\approx m\cdot n\cdot(\alpha+1)\cdot t_a \qquad (1)$$

При використанні запропонованого способу реалізації скалярного добутку доцільно зберігати матрицю й вектор, що множаться, в ”упакованому” форматі. У такому форматі використовуються для зберігання кожного компонента 2 розряди, відповідно для зберігання n компонент вектора буде потрібно $2 \cdot n/l$ машинних слів (кодів, рівних розрядності процесора – l). При використанні 2-х розрядів для зберігання доповнюючого коду кожного компонента a векторів, що множаться, значення 1 відповідає коду пари розрядів 01, а значення a = -1 відповідає коду 11. При використанні описаного формату одночасно можна виконувати операцію множення над $l/2$ парами компонентів. Сама операція множення при цьому зводиться до послідовного виконання двох логічних операцій – порозрядної суми по модулю 2 й операції АБО з константою. Для додавання отриманих у результаті цих операцій 2-бітових полів можна використати або табличний суматор, або операції зсуву й додавання. З урахуванням цього, і беручи до уваги те, що тривалість логічних операцій не перевищує тривалості операції додавання, час одержання скалярного добутку двох n-компонентних векторів складе $6 \cdot n \cdot t_a/l$. Тоді час T_1 скалярного множення матриці, що складається з m рядків й n стовпчиків на n-компонентний вектор при використанні запропонованої організації виконання операцій визначається формулою:

$$T_1 = \frac{6 \cdot m \cdot n \cdot t_a}{l} \qquad (2)$$

Порівняння часу скалярного множення матриці на вектор для двох описаних варіантів показує, що запропонований спосіб виконання цієї операції дозволяє зменшити час її виконання в β разів, при цьому чисельне значення β визначається наступним виразом:

$$\beta = \frac{T_0}{T_1} = \frac{l \cdot (\alpha + 1)}{6} \qquad (3)$$

Зокрема, для 32-розрядного процесора (l=32) і значення α =3, час виконання базової операції схеми ідентифікації на основі моделі персептрону, скалярного множення матриці на вектор, при використанні запропонованого способу її виконання скорочується в 21 раз (β=21.3) у порівнянні з традиційним способом її реалізації. Ще однією перевагою запропонованого способу виконання базової операції є скорочення необхідного для зберігання операндів об'єму пам'яті приблизно в $l/2$ разів. У таке ж число раз скорочується час виконання підсумовування векторів і виконання іншої операції – обчислення хеш-згортки.

На жаль, у самому протоколі ідентифікації неможливо отримати вказане прискорення через наявність операцій транспонування матриць у процедурах верифікації. Транспонування матриць з даними, що представлені в упакованому форматі, є досить складною операцією, а тому реальне прискорення в процесах ідентифікації, що було виявлено в ході поставлених експериментів, суттєво відрізняється від номінального в операції знаходження скалярного добутку та складає близько 30% – 50%.

Експериментальні дослідження запропонованої схеми обчислень показали результати, зображені на рисунках 1 та 2.

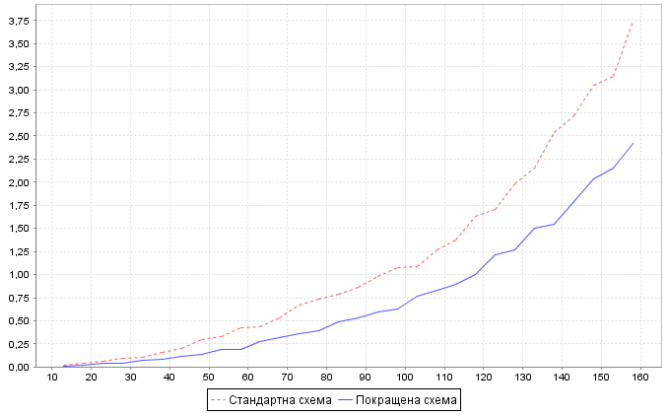

Рис. 1. Швидкість реєстрації

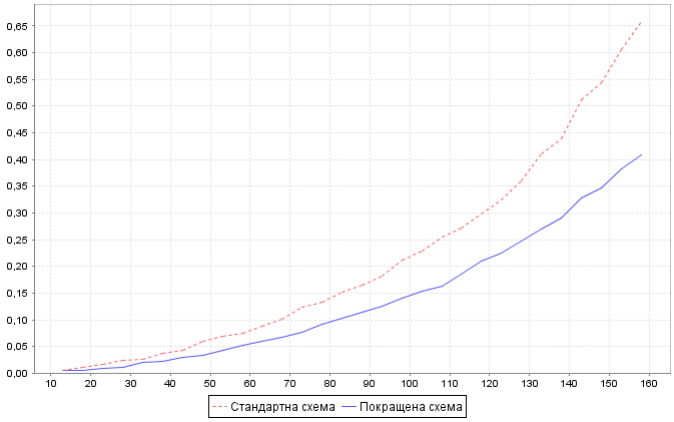

Рис. 2. Швидкість ідентифікації

Графіки показують залежність часу, який втрачається на реєстрацію та ідентифікацію абонентів у залежності від розмірностей матриць та векторів, що використовуються у протоколах відповідних процесів. Початкова розмірність була задана параметрами m=11, n=13.

Для процедури реєстрації було проведено тести при 30 різних розмірностях, нарощуючи її на кожному кроці на 5. Результати для кожної з розмірностей є середнім значенням 10 тис. ітерацій. Дані тестування швидкості ідентифікації були визначені також для 30 розмірностей матриць та векторів з кроком рівним 5. Для кожної з розмірностей було виконано 100 тис. ітерацій.

З наведених графіків випливає, що використання запропонованої організації обчислень дозволяє на 25-35% прискорити ідентифікацію.

Проведені дослідження залежності від розрядності параметрів m та n стійкості такої схеми ідентифікації абонентам до спроб підбору кодів h, V_1 й V_2 для несанкціонованого доступу до ресурсів системи, показали, що для практичного використання значення розрядностей указаних вище параметрів повинно бути не менше 200.

Основні переваги використання моделі персептрону для реалізації ідентифікації абонентів на основі спрощеної концепції “нульових знань” наведено нижче.

1. Ідентифікація використовує відкритий ключ системи – матрицю W великого об'єму однакову для всіх легальних абонентів. Відкриті індивідуальні ключі абонентів – вектори S мають малий об'єм. Це забезпечує істотно менший обсяг пам'яті, необхідний для зберігання в системі інформації, яка використовується для ідентифікації абонентів.

2. Генерація відкритого ключа абонента здійснюється набагато простіше й фактично не вимагає великих обчислювальних ресурсів.

3. Число сеансів звертання практично не обмежено.

Для практичного використання моделі персептрону при реалізації ідентифікації абонентів на основі спрощеної концепції “нульових знань” розроблено програмний продукт.

Висновки

В результаті проведених досліджень показано, що для спрощеної ідентифікації абонентів на основі теоретично строгої концепції нульових знань може бути застосована модель персептрону, яка має властивості неоднозначності зворотного перетворення. Вдосконалено спосіб використання моделі персептрону для ідентифікації абонентів за рахунок урахування особливостей моделі для вказанного застосування. Виявлені особливості дозволили спростити обчислення, повязані з реалізацією моделі персептрону і тим самим підвищити на 25-35% продуктивність ідентифікації. Доведено, що використання моделі персептрону дозволяє спростити генерацію ключів та зменшити об’єм інформації, що використовується в системі для ідентифікації.

Отримані результати можуть бути використані для підвищення ефективності багатокористувацьких систем зберігання інформації.

Перелік посилань

1. Ияд Мохд Маджид Ахмад Шахрури, Магрело В.Д. Потапенко М.М. Ідентифікація віддалених абонентів на основі технології нейронних мереж // Матеріали X Міжнародної науково-технічної конференції ”Системний аналіз та інформаційні технології”.– К.:НТУУ ”КПІ”.–2008.– С.357.
2. Fiat A., Shamir A. How to prove yourself: practical solutions to identification and signature problems // Proceeding of Workshop ”Crypto’86”.– Santa Barbara, USA, August 11-15, 1986.– Berlin: Springer-Verlag.– LNCS-263 – 1986.– P.186-194.
3. Pointcheval D. A new identification schemes based on perceptrons problem// Proceeding of International Conference “Eurocrypt’95”.– Saint-Malo, France, May 21-25, 1995.– Berlin: Springer-Verlag.– LNCS-921 – 1995.– P.319-328.