DOI: 10.58254/viti.9.2026.16.199

УДК 004.056:519.876

Паламарчук С. А. ORCID: 0000-0001-7483-9165 (ВІТІ ім. Героїв Крут)
Паламарчук Н. А. ORCID: 0000-0001-8818-7794 (ВІТІ ім. Героїв Крут)
Вороной О. В. ORCID: 0009-0007-1427-4431 (ВІТІ ім. Героїв Крут)
Побережець Т. В. ORCID: 0000-0001-8007-8614 (ВІТІ ім. Героїв Крут)

МОДЕЛЬ ВИЗНАЧЕННЯ СТІЙКОСТІ КОРИСТУВАЧІВ ДО АТАК СОЦІАЛЬНОЇ ІНЖЕНЕРІЇ НА ОСНОВІ ТЕОРІЙ НАВЧАННЯ ТА ДЕГРАДАЦІЇ ЗНАНЬ

У сучасних умовах зростання кількості кібератак, зокрема соціальної інженерії, людський фактор залишається одним із ключових елементів кібербезпеки організацій. Здатність протидіяти таким атакам значною мірою залежить від рівня навченості та обізнаності користувачів (стійкості), що формується в процесі навчання. Актуальність дослідження обґрунтовано з урахуванням посилення ролі людського фактора та поширення соціально-інженерних атак як на основі міжнародних стандартів і рекомендацій, так і на основі статистичних даних. Незважаючи на значну кількість досліджень, існуючі підходи (моделі), як правило, не враховують у явному вигляді рівень навченості та обізнаності користувача (його стійкість) як змінну, що безпосередньо впливає на здатність протидіяти атакам соціальної інженерії. У статті розглянуто питання кількісного оцінювання впливу навчання користувачів на рівень їхньої стійкості до атак соціальної інженерії, зокрема фішингу. Запропоновано модель визначення стійкості користувача під впливом навчання та деградації знань внаслідок забування або появи нових механізмів атак, обґрунтовано параметри моделювання на основі статистичних та емпіричних даних. Окрім цього, визначається залишковий ризик, пов'язаний із людським фактором для системи. Результати моделювання підтверджують адекватність запропонованої моделі реальним процесам навчання та втрати знань користувачем, зокрема через наявність ефекту насичення та деградації знань. Отримані результати можуть бути використані для визначення пріоритетних груп ризику, оптимізації періодичності навчання, а також врахування залишкового ризику, пов'язаного з людським фактором у моделях (методиках) оцінювання ризиків системи.

Ключові слова: *кібербезпека, людський фактор, навчання користувачів, деградація знань, соціальна інженерія, стійкість користувача, оцінювання ризиків, фішинг.*

S. Palamarchuk, N. Palamarchuk, O. Voronoi, T. Poberezhets. Model for determining user resistance to social engineering attacks based on the theories of learning and knowledge degradation

In today's conditions of increasing cyberattacks, in particular social engineering, the human factor remains one of the key elements of cybersecurity of organizations. The ability to counter such attacks largely depends on the level of user training and awareness (resilience) that is formed in the training process. The relevance of the study is justified taking into account the increasing role of the human factor and the spread of social engineering attacks, both on the basis of international standards and recommendations and on the basis of statistical data. Despite a significant number of studies, existing approaches (models), as a rule, do not explicitly take into account the level of user training and awareness (his resilience) as a variable that directly affects the ability to counter social engineering attacks. The article considers the issue of quantitative assessment of the impact of user training on the level of their resistance to social engineering attacks, in particular, phishing. A model for determining user resilience under the influence of learning and knowledge degradation due to forgetting or the emergence of new attack mechanisms is proposed, and modeling parameters are justified based on statistical and empirical data. In addition, the residual risk associated with the human factor for the system is determined. The modeling results confirm the adequacy of the proposed model to real processes of learning and knowledge loss by the user, in particular, due to the presence of the saturation effect and knowledge degradation. The results obtained can be used to determine priority risk groups, optimize the frequency of training, and take into account the residual risk associated with the human factor in models (methods) for assessing system risks.

Keywords: *cybersecurity, human factors, user training, knowledge degradation, social engineering, user resilience, risk assessment, phishing.*

Актуальність та постановка завдання. У сучасних умовах розвитку цифрових технологій та ведення гібридних війн соціальна інженерія перетворилася на один із найбільш ефективних інструментів реалізації кіберзагроз, зокрема й у сфері оборони. Незважаючи на створення авторизованих систем із безпеки та активне впровадження технічних засобів кіберзахисту, систем виявлення і реагування на кіберінциденти (їх інтелектуалізації, здатності до навчань), вплив на людський фактор залишається найбільш вразливою ланкою будь-якої

системи (сервісу). Зловмисники цілеспрямовано експлуатують психологічні особливості людини/користувача, фактор терміновості, страх та "критичність моментної загрози" (*homo intuivus, emotus et fallibilis* – інтуїтивна, емоційна та схильна до помилок людина), що суттєво підвищує результативність методів соціальної інженерії, зокрема фішингу. Вразливість користувача та його стійкість (здатність протидіяти соціальним атакам) є динамічними характеристиками, що визначаються сукупністю когнітивних і поведінкових факторів, а також рівнем його обізнаності та підготовки. Згідно з підходами, визначеними в *NIST SP 800-30* [1], вразливість розглядається як слабкість, яка може бути використана джерелом загрози, причому в контексті кібербезпеки такі слабкості все частіше пов'язані саме з людським фактором. Відповідно, недостатня сформованість навичок безпечної поведінки та відсутність періодичного оновлення знань призводять до зростання ймовірності реалізації атак соціальної інженерії, що обумовлює необхідність розроблення моделі визначення стійкості користувачів до атак соціальної інженерії.

За даними *Microsoft* за 2025 рік та *Institute of Information Security (*далі – *InfoSec)* визначено, що соціальна інженерія, зокрема фішинг, займає провідне місце серед сучасних видів загроз, причому розвиток штучного інтелекту та використання його із зловмисними намірами суттєво розширює можливості атак, фішингові кампанії на основі штучного інтелекту мають на 42 % вищий рівень успіху, ніж звичайні [2]. Згідно зі статистикою Національної команди реагування на кіберінциденти *CERT-UA* [3–6], у 2025 році опрацьовано 5 927 кіберінцидентів, що на 37,4 % більше порівняно з попереднім роком. Серед найпоширеніших типів атак – фішинг, 1 727 випадків (це майже вдвічі більше ніж у 2024 році, і складає ~ 30 % від усіх опрацьованих кіберінцидентів), розповсюдження та зараження шкідливим програмним забезпеченням та компрометація облікових записів. Значна частка інцидентів припадає на сектор безпеки й оборони, органи державної влади та критичну інфраструктуру, що свідчить про цілеспрямований характер та системність атак.

При цьому, дослідження, що проведені в Україні у 2021 та 2023 роках за підтримки *CRDF Global* [7; 8] показують, що серед значної частини користувачів (понад 41 %), які вже зіткнулися з кіберінцидентами, важливим поведінковим фактором ризику є переконання у власній "нецікавості" для зловмисників, таку позицію поділяє понад половина опитаних. Тому, результати досліджень демонструють наявність розриву між рівнем обізнаності користувачів у сфері кібербезпеки та їхньою реальною поведінкою, зокрема значна частина користувачів ігнорує базові практики кібергігієни. У більшості досліджень і звітів з кібербезпеки зазначається, що значна частина інцидентів починається саме з фішингу або маніпуляції користувачем, що актуалізує необхідність аналізу, розроблення та впровадження моделей визначення стійкості користувачів до атак соціальної інженерії.

Аналіз останніх досліджень і публікацій. Незважаючи на значну кількість досліджень, існуючі підходи, як правило, не враховують у явному вигляді рівень навченості та обізнаності користувача (його стійкість) як змінну, що безпосередньо впливає на здатність протидіяти атакам соціальної інженерії. Оцінювання впливу навчання користувачів на їхню здатність протидіяти атакам соціальної інженерії здійснюється з використанням різних підходів, зокрема нормативних, поведінкових і статистичних. Міжнародні публікації, такі як рекомендації *NIST* та *SANS Institute* [9; 10], визначають необхідність періодичного навчання та підвищення обізнаності користувачів. У цих підходах ефект забування враховується опосередковано через вимогу періодичності навчання (4–6 місяців), однак відсутня його формалізація у вигляді математичних залежностей, що обмежує можливість кількісного оцінювання з метою оптимізації навчання.

В [11] розглянуто узагальнену модель поведінки користувача, яка базується на концепції співіснування кількох критеріїв прийняття рішень при отриманні фішингового повідомлення. Автори виділяють такі поведінкові стани користувача (стани "розуму"), як рутинний, імпульсивний, поведінковий та раціональний (обдуманий), які визначають вибір дій

користувача. У таких моделях увага приділяється процесу прийняття рішень, проте процеси навчання та забування не мають явного математичного опису.

В [12] розроблено методику прогнозування фішингових атак на основі статистичних методів. Для аналізу трендів застосовано часові ряди, сформовані на основі відкритих статистичних даних, що дозволило побудувати модель динаміки кількості фішингових атак за визначений період та сформувати прогноз їх розвитку на майбутнє. Отримані результати підтверджують ефективність статистичних підходів для оцінювання тенденцій розвитку загроз.

В [13] досліджено вплив безперервного навчання на довгострокову стійкість користувача. Показники успішності фішингу (із врахуванням контекстної інформації та емоційних тригерів) знизилися майже вдвічі протягом перших шести місяців і стабілізувалися близько до галузевих показників, що підкреслює довгострокову ефективність ініціатив щодо сталого підвищення обізнаності (протягом 12 місяців).

Водночас аналіз існуючих досліджень показав, що в підходах до оцінювання ризиків соціальної інженерії рівень стійкості користувача переважно не враховується як формалізований показник, а процеси навчання та деградації знань не мають узагальненого математичного опису. Це обумовлює необхідність розроблення моделі визначення стійкості користувачів до атак соціальної інженерії на основі теорій навчання та деградації знань.

Мета статті. Розробка моделі визначення стійкості користувачів до атак соціальної інженерії на основі теорій навчання та деградації знань**.**

Виклад основного матеріалу дослідження. Вразливість (*vulnerability*) – це слабкість активу (системи, сервісу, програмного забезпечення тощо) або контролю, яка може бути використана однією або кількома загрозами [1; 14]. Після виявлення вразливості удосконалюються (модернізуються) як самі технологічні середовища (системи, сервіси, програмне забезпечення тощо), так і засоби виявлення та реагування (зокрема системи виявлення вторгнень (*IDS/IPS*), управління подіями безпеки *(SIEM)*, захисту кінцевих точок *(EDR/XDR)*, засоби автентифікації та контролю доступу, а також спеціалізовані рішення протидії фішингу тощо). Водночас при соціально-інженерних атаках об'єктом впливу є людина – користувач сервісу (системи або пристрою). У концептуальних документах і стандартах термін "вразливість користувача" (дослівно "вразливість людського фактору", *human factor vulnerability*) прямо не формалізований, однак його зміст визначається контекстно через загальне поняття вразливості як слабкості, що може бути використана при реалізації загрози. Зокрема, відповідно до *ISO/IEC 27001* [14] вразливість трактується як властивість активу або групи активів, яка може бути використана загрозою, причому до таких активів належить і персонал. Аналогічно, у *NIST SP 800-30* [1] під вразливістю розуміється слабкість інформаційної системи, процедур безпеки або внутрішніх контролів, що може бути використана джерелом загрози, включаючи недоліки, пов'язані з людським фактором. У документах Агенства Європейського союзу з питань кібербезпеки (*European Union Agency for Cybersecurity*, далі – *ENISA*) [15] також підкреслюється, що сучасні кіберзагрози базуються на експлуатації поведінкових характеристик користувача та людських помилок (*exploiting human error and behaviour*). У цьому контексті "вразливість користувача" доцільно розглядати як сукупність когнітивних, поведінкових та організаційних характеристик, які зумовлюють можливість успішного впливу соціально-інженерних атак та прийняття користувачем небезпечних рішень. Таким чином, під "вразливістю користувача" пропонується розуміти слабкі сторони знань, сприйняття, мислення та поведінки користувача, які можуть бути використані для реалізації загроз соціальної інженерії, що узгоджується із загальним визначенням вразливості, яке надане в *ISO/IEC 27001* та *NIST SP 800-30*, а також із підходами *ENISA* щодо ролі людського фактору.

Стійкість (*resilience*) – це здатність організації, системи або інфраструктури протистояти кіберінцидентам, адаптуватися до них і швидко відновлюватися після них, зберігаючи

критично важливі функції [14]. У контексті атак соціальної інженерії "стійкість користувача" – це здатність користувача розпізнавати, протидіяти та не піддаватися впливу соціально-інженерних атак, зберігаючи безпечну поведінку навіть в умовах невизначеності та маніпулятивного впливу. Відповідно, поняття "вразливість користувача" – "стійкість користувача" є взаємопов'язаними характеристиками, що мають обернено пропорційний характер, підвищення рівня навченості та обізнаності користувача призводить до зменшення його вразливості та, відповідно, до зростання стійкості до соціально-інженерних атак.

Міжнародні стандарти, *NIS2 Directive (EU) 2022/2555* (далі – *NIS2*) [16] та *NIST SP 800-53* [9], містять вимоги (контролі), що виокремлюють і формують вимоги, які в сукупності призведуть до зменшення вразливості користувача та підвищення рівня його стійкості через впровадження заходів навчання, обізнаності, безперервності діяльності та управління ризиками. Згідно з *NIS2* (стаття 21. *Cybersecurity risk-management measures*) [16] вимоги такі:

1. Навчання та обізнаність персоналу *(Art. 21(2)(g). Cybersecurity training and awareness)*:
   періодичне навчання персоналу щодо кіберзагроз;
   формування навичок розпізнавання фішингу, соціальної інженерії.
2. Управління інцидентами *(Art. 21(2)(b). Incident handling)*: виявлення та реагування на інциденти, включає інциденти, спричинені соціальною інженерією.
3. Безперервність діяльності *(Art. 21(2)(c). Business continuity and crisis management)*: підготовка до інцидентів, у тому числі людського фактору.
4. Управління ризиками *(Art. 21(1). Risk management framework)*: системне оцінювання кіберризиків, включаючи поведінкові фактори, інтегрує соціальну інженерію у загальну модель ризику.

NIST SP 800-53 [9] розглядає соціальну інженерію як один із основних сценаріїв реалізації загроз (експлуатуючи поведінкові вразливості користувачів), і відповідно, вимагає впровадження багаторівневих організаційних і технічних заходів для її запобігання, виявлення та мінімізації наслідків.

При цьому, навчання та обізнаність користувачів (персоналу) досягається через оцінку ефективності навчання (доцільно з визначенням початкового рівня стійкості користувача), періодичні тренінги, симуляцію атак (phishing campaigns) та обов'язкове проведення тестування. За зазначеними стандартами, обізнаність користувача (*awareness training*) зменшується з часом за відсутності оновлення, і навчання має бути періодичним для підтримки ефекту. Тому, моделюючи атаки, соціальну інженерію можливо розглядати як інструмент перевірки стійкості користувача.

Розробка моделі визначення стійкості користувачів до атак соціальної інженерії на основі теорій навчання та забування.

Загальна постановка задачі. В будь-якій інформаційній системі (далі – системі) розглядається користувач, здатність якого протидіяти атакам соціальної інженерії змінюється у часі під впливом процесів навчання, практично набутого досвіду, природного забування отриманих знань, а також від еволюції та оновлення механізмів соціоінженерних атак. Формально вводиться поняття рівня стійкості користувача, який відображає поточну здатність користувача виявляти та протидіяти атакам соціальної інженерії. Вказаний рівень є змінною величиною і залежить від початкового рівня стійкості та обізнаності користувача (навчання або забування інформації). При цьому, підвищення рівня обізнаності користувачів у сфері кібербезпеки має властивість до зменшення граничного ефекту (за класичними дослідженнями у сфері теорії навчання та когнітивної психології) – початкові етапи навчання дають найбільший приріст знань, тоді як подальше навчання супроводжується поступовим насиченням і зменшенням темпів засвоєння. Така поведінка узгоджується з класичними результатами досліджень, зокрема кривою забування, запропонованою *Hermann Ebbinghaus*, яка описує швидке початкове зниження рівня запам'ятовування з подальшим його уповільненням [17].

Модель має визначати рівень стійкості користувача у часі як функцію попереднього стану знань (враховуючи вхідні дані – обґрунтовані параметри моделі та індикатор атаки (наявність або відсутність атаки)) та забезпечувати покрокове оновлення рівня стійкості користувача за результатами навчання (забування).

За результатами моделювання можливо визначити, як змінюється стійкість різних користувачів, обґрунтовувати оптимальну періодичність проведення навчань і тренінгів персоналу, а також використати отримані результати для оцінювання залишкового ризику, пов’язаного із людським фактором для системи. Схему моделі визначення рівня стійкості користувача до атак соціальної інженерії наведено на рисунку 1.

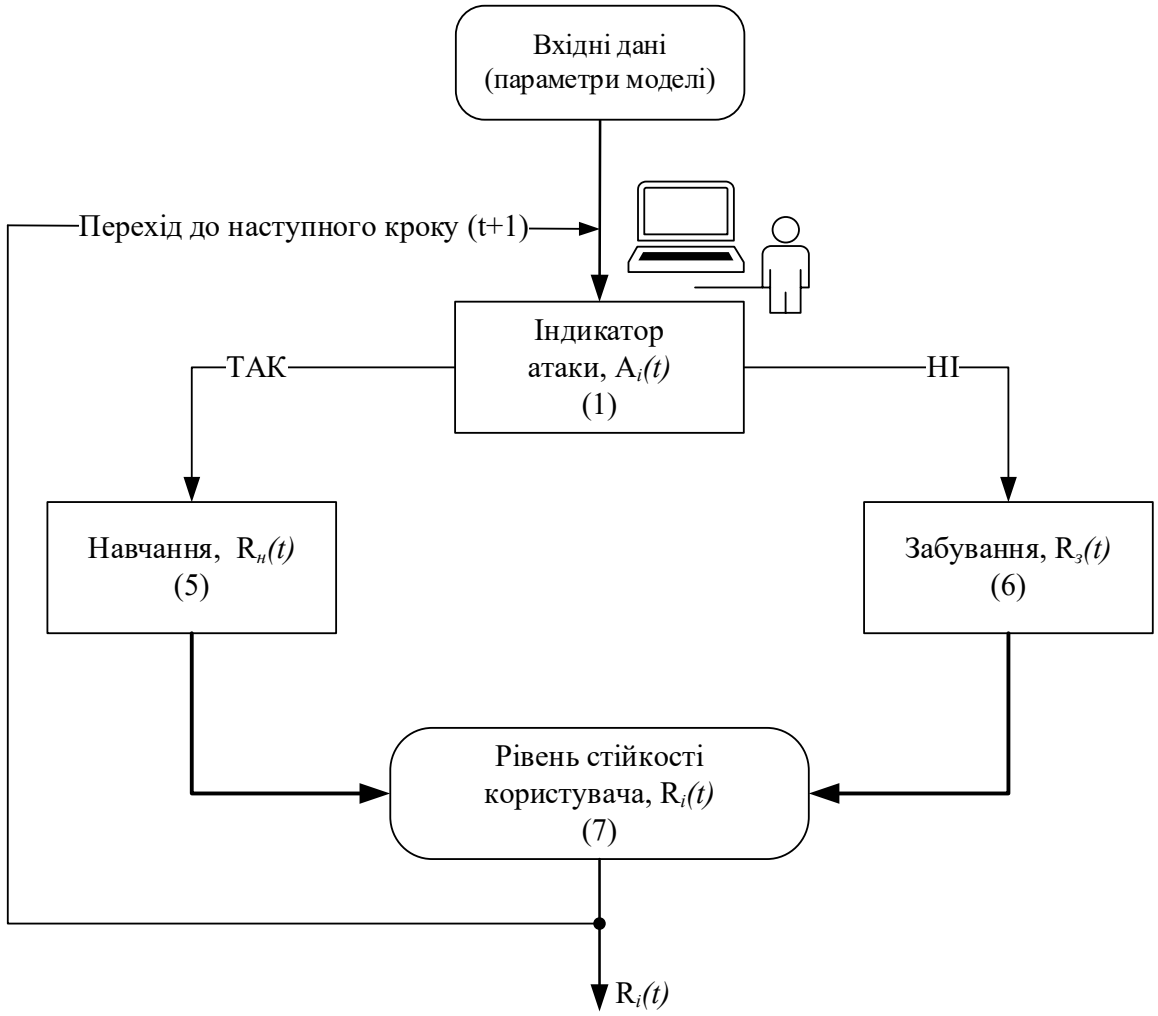

Рис. 1. Схема моделі визначення рівня стійкості користувача до атак соціальної інженерії

Вхідні дані, припущення та обмеження. Згідно з [18] наведено статистичні показники про користувача-жертву фішингових атак, а саме:

нетренований користувач має схильність піддатись фішингу в 33 % атак;

після 90 днів тренування ризик зменшується на 41 %;

протягом року навчань ризик зменшується на 82 %.

Також, згідно зі статистикою соціальної інженерії за 2026 рік [19], 1,5 % співробітників все ще натискають небезпечні посилання в симуляції фішингу, навіть після неодноразового навчання. Окрім зазначеного, для розуміння глибини/шкоди від реалізації атак за цією статистикою [18]:

65 % нападів соціальної інженерії були фішинговими;

79 % захоплень облікових записів почалися з фішингових електронних листів.

Найбільш поширеними результатами атак соціальної інженерії були викрадення облікових даних (29 %), крадіжка даних (18 %) і вимагання (13 %).

Припущення: Кожна атака є незалежною подією. Навчання підвищує рівень стійкості користувача, забування – знижує рівень стійкості користувача.

Обмеження: Моделювання проведено із врахуванням середньостатистичного рівня початкової стійкості користувача (R_0), що тотожне початковому рівню знань користувача (варіативність початкового рівня стійкості користувача буде враховано в подальших дослідженнях).

Формалізація моделі. Для формалізації процесу визначення рівня стійкості користувача в момент часу t до атак соціальної інженерії введемо змінну $R_i(t) \in [0; 1]$ – нормований інтегральний показник стійкості користувача, який інтерпретується як здатність протидіяти атаці.

Для врахування факту наявності атаки (або симуляції) в момент часу t вводиться індикаторна функція $\mathrm{A}_i(t)$ – індикатор атаки ("1" якщо вона відбулася, або "0" якщо атака відсутня), яка має вигляд (1):

$$\mathrm{A}_i(t)=\begin{cases}1, & \text{якщо атака відбулася}\\ 0, & \text{якщо інше}\end{cases}. \quad (1)$$

Значення індикатора атаки визначає, який із двох процесів домінує у поточний момент часу, навчання або забування. Індивідуальні процеси навчання та забування доцільно описати експоненційною функцією з насиченням (2), що відображено в [20; 21]:

$$R(t)= R_0 e^{-\lambda t}, \quad (2)$$

де $R(t)$ – рівень знань через час t;

R_0 – початковий рівень знань (початковий рівень стійкості користувача);

λ – параметр процесу, який залежить від складності матеріалу, індивідуальних особливостей та умов навчання/забування.

Експоненційні моделі ефективно описують поведінку пам'яті у більшості навчальних ситуацій, хоча конкретні значення λ можуть варіюватись залежно від типу знань і когнітивних навичок особи, яка навчається (обґрунтування параметрів та їхніх значень для моделювання наведено в таблиці 1).

Для уникнення нефізичних значень доцільно використовувати обмежену форму моделі, тоді ефект навчання користувача із врахуванням початкового рівня знань матиме вигляд (3):

$$R(t)=R_{max}-(R_{max}-R_0)e^{-kt}, \quad (3)$$

де R_{max} – максимальний рівень знань (максимальний рівень стійкості користувача), який нормується [0; 1];

k – коєфіцієнт навчання.

Аналогічно, ефект забування користувача із врахуванням мінімального рівня знань описується як залежність рівня знань в момент часу t від початкового рівня знань (4):

$$R_з(t)=R_{min}+(R_0-R_{min})e^{-dt}, \quad (4)$$

де R_{min} – мінімальний рівень знань (мінімальний рівень стійкості користувача);

d – коєфіцієнт деградації знань (внаслідок забування або появи нових механізмів атак).

Для практичного застосування неперервні моделі (3)–(4) апроксимуються дискретними рекурсивними формами (5)–(6). Так як процес навчання та забування змінюють один одного, то більш доцільно описати залежності не від моменту часу *t*, а від попереднього моменту часу.

Ефект навчання (3) набуває вигляду (5):

$$R_{\text{н}}(t)=R_{max}-(R_{max}-R_{t-1})*k, \tag{5}$$

де $R_{t\text{-}1}$ – рівень знань (стійкості користувача) в попередній момент часу.

Ефект забування (4) набуває вигляду (6):

$$R_{\text{з}}(t)=R_{min}+(R_{min}-R_{t-1})*d. \tag{6}$$

Тоді, рівень стійкості користувача, $R_i(t)$ (із врахуванням індикатору атаки (1), функцій навчання (5) та забування (6)) визначається згідно з (7):

$$R_i(t)=\text{A}_i(t)R_{\text{н}}(t)+(1-\text{A}_i(t))\,R_{\text{з}}(t). \tag{7}$$

Параметри моделі. Для реалізації моделі проведено обґрунтування її параметрів на основі статистичних та емпіричних даних (табл. 1).

Таблиця 1

Обґрунтування параметрів моделі на основі статистичних та емпіричних даних

№ з/п	Параметр	Значення	Опис припущення	Що відображає	Обґрунтування
1	Початкова стійкість користувача (R_0)	0,67	На основі статистичних даних або результатів тестування користувача	Базовий рівень здатності протидії фішингу	Виходячи зі статистики ~ 33 % користувачів піддаються фішингу: стійкість ≈ 1−0,33 = 0,67
2	Стандартне відхилення R_0	± 0,10	Варіативність користувачів	Індивідуальні відмінності	Враховує різний рівень обізнаності та досвіду
3	Максимальне значення стійкості (R_{max})	0,985	Обмеження моделі (граничний рівень)	Верхня межа ефективності навчання	Навіть після навчання частина користувачів (≈ 1,5 %) допускає помилки
4	Мінімальне значення стійкості (R_{min})	0,67	Обмеження моделі (граничний рівень)	Нижня межа	Відповідає $R_{min} \le R_0$
5	Коєфіцієнт навчання (експонен-ційний)	0,41 (за 90 днів)	Емпіричний параметр	Приріст знань	Визначено шляхом апроксимації експоненційної моделі згідно з виразом (4) (зниження ризику на 41 % за 90 днів), $k \approx 0{,}005$
6	Коєфіцієнт деградації знань (експонен-ційний)	[0;1]	Емпіричний параметр	Втрата знань	Нормується в межах [0;1] та характеризує інтенсивність втрати знань (внаслідок забування або появи нових механізмів атак)
7	Функція складності $f(t)$	[0;1]	Рівень складності атаки	Вплив атаки	Враховує різні типи складності атак

Визначення початкової стійкості користувача (п. 1 табл. 1) здійснено на основі статистичних даних про фішингові атаки, тоді як для реальних систем (умов) можливий варіант визначення стійкості користувача шляхом проведення його тестування за різними ступенями складності тесту. Кожен параметр може мати як математичну інтерпретацію, так і емпіричне обґрунтування, що забезпечує узгодженість моделі з реальними статистичними даними та дозволяє адекватно відобразити нелінійний ефект навчання, ефект насичення та деградацію знань (забування).

Апробація моделі. Для апробації запропонованої моделі здійснено її програмування мовою *Python*, з використанням бібліотеки *numPy.* Визначення рівня стійкості користувача здійснюється згідно з виразом (7), час в моделі інтерпретується як кількість ітерацій навчання, одна ітерація моделі відповідає умовному часовому інтервалу (t). Динаміку зміни рівня стійкості користувачів залежно від кількості ітерацій навчання (на прикладі 100 ітерацій) наведено на рисунку 2.

Початковий рівень стійкості користувача для досліджуваних категорій прийнято усередненим (0,67 згідно з п. 1 табл. 1), що дає змогу порівнювати динаміку зміни рівня стійкості за однакових початкових умов. Подальша зміна рівня стійкості визначається здатністю користувача до навчання, яка в моделі описується коефіцієнтами навчання та деградації знань, що характеризує поступову втрату набутих навичок у часі. Так, слабкий користувач характеризується коєфіцієнтом навчання $k = 0{,}001$ (найнижчий рівень засвоєння знань), середній користувач – коефіцієнтом навчання, який визначено емпіричним шляхом $k = 0{,}005$ (п. 5. табл. 1), сильний користувач характеризується коєфіцієнтом навчання $k = 0{,}01$.

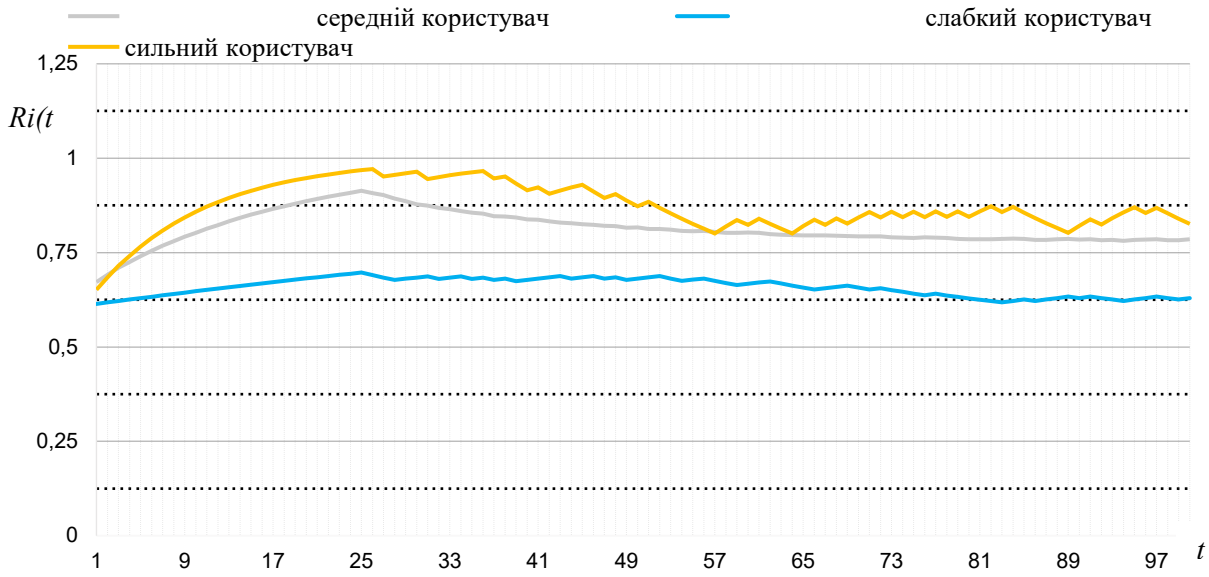

Рис. 2. Динаміка зміни рівня стійкості користувача залежно від кількості ітерацій навчання

Максимальний ефект навчання (рівень стійкості користувача 0,9) досягається приблизно на 20–25-й ітерації навчання, після чого спостерігається поступова деградація знань. Така динаміка узгоджується з класичними уявленнями про криві навчання та забування. Із часом, у користувачів спостерігається уповільнення темпів зростання показника стійкості, що пояснюється ефектом насичення знань та наближенням до граничного рівня засвоєння.

На прикладі системи з чисельністю 1000 користувачів (N) проведено розрахунок рівня їхньої стійкості (7). Користувачів розподілено та згруповано за інтервалами значень показника стійкості (рис. 3).

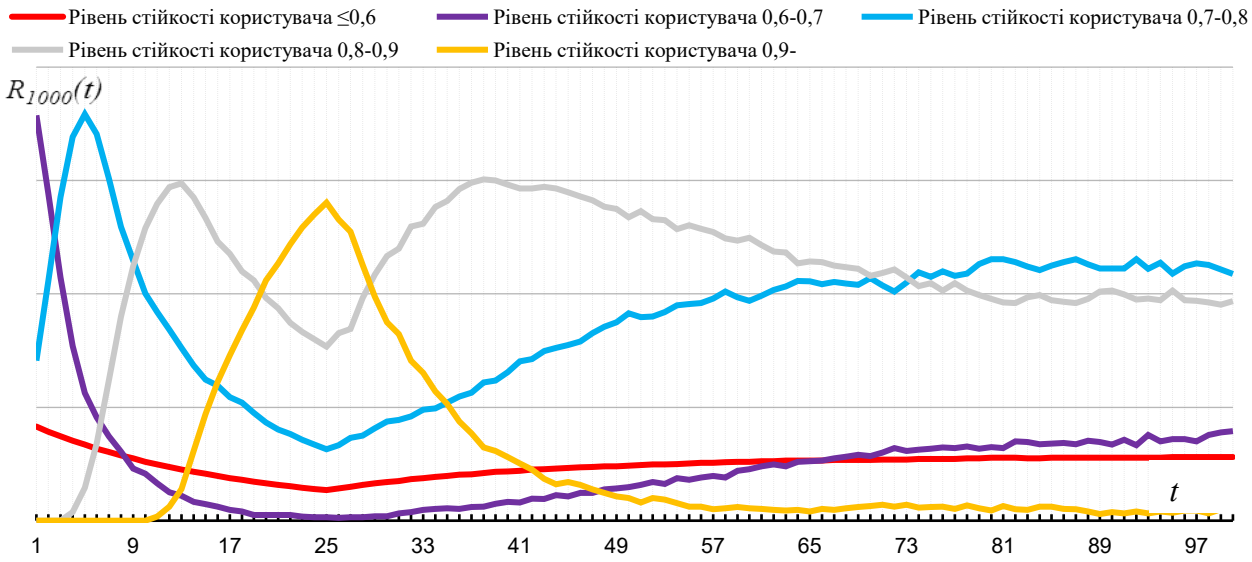

Рис. 3. Розподіл кількості користувачів за рівнем їхньої стійкості залежно від ітерацій навчання

Аналіз графіка дозволяє оцінити співвідношення між зазначеними групами користувачів у системі та виявити пріоритетні групи ризику для формування підходів до їх навчання (тренування). Зменшення частки користувачів із рівнем стійкості 0,6–0,7 та зростання груп із рівнями 0,7–0,8, 0,8–0,9 і понад 0,9 свідчить про позитивний вплив навчання на загальний рівень кіберстійкості організації.

Також, за результатами навчання користувачів доцільно визначити залишковий ризик, пов'язаний із людським фактором R_l для системи із N користувачів (як практична складова визначення стійкості користувачів) з метою його врахування при оцінці загального ризику системи.

За класичним визначенням [22] залишковий ризик людського фактору для всієї системи (сервісу) після навчання користувачів оцінюється (8):

$$R_l = \frac{1}{N}\sum_{i=1}^{n}\prod_{t=1}^{t_{end}}\left(A_i(t)*(1-R_i(t))*f(t)\right), \tag{8}$$

де $1-R_i(t)$ – вразливість користувача i в момент часу t;

$f(t)$ – функція складності атаки в момент часу t.

Окремо на графіку (див. рис. 3) червоним кольором виділено категорію найбільш вразливих користувачів, із рівнем стійкості нижчим за поріг $1-R_i(t)=0{,}6$. Вразливість впливає на залишковий ризик системи, пов'язаний із людським фактором. Отримані результати демонструють доцільність диференційованого підходу до підготовки користувачів з урахуванням їхніх індивідуальних особливостей до навчання та деградації знань.

Таким чином, розробка моделі надає можливість кількісного визначення стійкості користувачів до атак соціальної інженерії, а саме, визначення рівня стійкості користувача, $R_i(t)$ згідно з виразом (7) та оцінювання залишкового ризику людського фактору R_l для системи згідно з виразом (8) після навчання користувачів.

Висновок та перспективи подальших досліджень. У статті запропоновано модель визначення стійкості користувачів до атак соціальної інженерії, яка, на відміну від існуючих підходів, у явному вигляді враховує рівень знань користувача, ефект насичення, а також деградації знань. Модель інтегрує процеси навчання та втрати знань в єдиній аналітичній формі та дозволяє оцінювати залишковий ризик, пов'язаний із людським фактором для системи після навчання користувачів. Результати моделювання підтверджують адекватність запропонованої моделі, оскільки вона коректно відтворює основні закономірності процесів

навчання та втрати знань і узгоджується із теоріями навчання та деградації знань. Практична цінність моделі полягає у можливості її використання для визначення пріоритетних груп ризику, визначення оптимальної періодичності навчання та інтеграції показників людського фактору у моделі (методики) оцінювання ризиків для системи.

Подальші дослідження спрямовані на емпіричну валідацію моделі, уточнення та розширення її параметрів, врахування варіативності користувачів, а також оцінювання залишкового ризику, пов’язаного із людським фактором.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. NIST SP 800-30 Rev. 1, Guide for Conducting Risk Assessments. CSRC. NIST Computer Security Resource Center / CSRC. URL: https://csrc.nist.gov/pubs/sp/800/30/r1/final (дата звернення: 04.04.2026).

2. Microsoft Digital Defense Report 2025. Lighting the path to a secure future. A Microsoft Threat Intelligence report. URL: https://spacelift.io/blog/social-engineering-statistics (дата звернення: 14.04.2026).

3. Кібератаки UAC-0001 на сектор безпеки та оборони із застосуванням програмного засобу LAMEHUG, що використовує LLM (велику мовну модель). URL: https://cert.gov.ua/article/6284730 (дата звернення: 14.04.2026).

4. Кібератака UAC-0125 з використанням тематики “Армія+”. URL: https://cert.gov.ua/article/6281701 (дата звернення: 14.04.2026).

5. Російські кібероперації аналітика за ІІ півріччя 2024 року // Державна служба спеціального зв’язку та захисту інформації України. URL: https://cip.gov.ua/ua/statics/analitichni-materiali-derzhspeczv-yazku (дата звернення: 12.03.2026).

6. Російські кібероперації аналітика за І півріччя 2025 року // Державна служба спеціального зв’язку та захисту інформації України. URL: https://cip.gov.ua/ua/statics/analitichni-materiali-derzhspeczv-yazku (дата звернення: 12.03.2026).

7. Звіт про базове дослідження щодо інформованості цільових аудиторій про основні аспекти кібербезпеки. Для CRDF Global. Short URL service Surli – FREE Short Links. URL: https://surl.li/lzxqpx (дата звернення: 12.03.2026).

8. Звіт про проміжне дослідження щодо інформованості цільових аудиторій про основні аспекти кібербезпеки. Для CRDF Global. 2023. 72 с. URL: https://surl.li/icvklc (дата звернення: 12.03.2026).

9. NIST SP 800-53 rev. 5 Security and Privacy Controls for Information Systems and Organizations, 2020. URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r5.pdf (дата звернення: 04.04.2026).

10. SANS Cybersecurity Training Overview // SANS Institute. URL: https://www.sans.org/cyber-security-training-overview (дата звернення: 14.04.2026).

11. Embrey I., Kaivanto K. Many phish in the C: A coexisting-choice-criteria model of security behavior // Risk Analysis. April 2023. Vol. 43, Issue 4. P. 645–866. URL: https://doi.org/10.1111/risa.13947 (дата звернення: 04.04.2026).

12. Dobryshyn Y. STATISTICAL METHODS FOR PREDICTING PHISHING ATTACKS // Cybersecurity: Education, Science, Technique. 2024. Vol. 3, no. 23. P. 56–70. URL: https://doi.org/10.28925/2663-4023.2024.23.5670 (дата звернення: 04.04.2026).

13. Toth R. et al. Sustaining Cyber Awareness: The Long-Term Impact of Continuous Phishing Training and Emotional Triggers // arXiv preprint arXiv:2510.27298. 2025. URL: https://arxiv.org/html/2510.27298v1 (дата звернення: 18.04.2026).

14. ДСТУ EN ISO/IEC 27001:2022 Інформаційні технології. Методи захисту. Системи управління інформаційною безпекою. Вимоги (EN ISO/IEC 27001:2017, IDT; ISO/IEC 27001:2013 including Cor 1:2014 and Cor 2:2015, IDT). URL: https://www.iso.org/ru/standard/27001 (дата звернення: 04.04.2026).

15. Cybersecurity Culture Guidelines: Behavioural Aspects of Cybersecurity ENISA. Home ENISA. URL: https://www.enisa.europa.eu/publications/cybersecurity-culture-guidelines-behavioural-aspects-of-cybersecurity (дата звернення: 04.04.2026).

16. DIRECTIVE (EU) 2022/2555 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive). EUR-Lex – Access to European Union law – choose your language. URL: https://eur-

lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022L2555&qid=1708526 (дата звернення: 14.04.2026).

17. Крива забування Еббінгауза – Вікіпедія. URL: uk.wikipedia.org/wiki/Крива_забування_Еббінгауза (дата звернення: 14.04.2026).

18. 2025 Phishing By Industry Benchmark Report | KnowBe4. Beyond Security Awareness Training // KnowBe4 Human Risk Mgmt Platform. URL: https://www.knowbe4.com/resources/reports/phishing-by-industry-benchmarking-report (дата звернення: 14.04.2026).

19. Michalowski M. 70 Social Engineering Statistics for 2026 // Spacelift. URL: https://spacelift.io/blog/social-engineering-statistics (дата звернення: 14.04.2026).

20. Vianna, Leonardo Silva, Alexandre Leopoldo Gonçalves, and João Artur Souza. Analysis of learning curves in predictive modeling using exponential curve fitting with an asymptotic approach // Plos one. 2024. No. 4 : e0299811. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0299811. (дата звернення: 14.04.2026).

21. Murre J. M. J., Chessa A. G. Power laws from individual differences in learning and forgetting: mathematical analyses // Psychonomic Bulletin & Review. 2011. Vol. 18, no. 3. P. 592–597. URL: https://doi.org/10.3758/s13423-011-0076-y (дата звернення: 14.04.2026).

22. ISO/IEC 27005:2022 Інформаційна безпека, кібербезпека та захист конфіденційності – Керівництво з управління ризиками інформаційної безпеки. URL: https://www.iso.org/standard/80585.html (дата звернення: 22.04.2026).

Надійшла до редколегії 11.05.2026.
Схвалена до друку 22.05.2026.
Дата публікації 29.05.2026.