МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАУКОВІ ПРАЦІ ДОНЕЦЬКОГО НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ

Серія: «Обчислювальна техніка та автоматизація»

Всеукраїнський науковий збірник
Заснований у липні 1998 року
Виходить 2 рази на рік
Т3. № 5(37)'2025

Видавничий дім
«Гельветика»
2025

УДК 681.5:658.5:621.3

Друкується за рішенням Вченої ради Державного вищого навчального закладу «Донецький національний технічний університет» (протокол № 12 від 27.11.2025 р.).

У збірнику опубліковано статті науковців, аспірантів, магістрів та інженерів провідних підприємств і закладів вищої освіти України, у яких наведено результати наукових досліджень та розробок, виконаних у 2023–2024 рр. відповідно до напрямків: автоматизація технологічних процесів, інформаційна безпека, інформаційно-вимірювальні системи, електронні та мікропроцесорні прилади, інформаційні технології, кібербезпека та захист критичної інфраструктури, математичне та комп'ютерне моделювання, телекомунікаційні системи та мережі.

Матеріали збірника призначено для викладачів, наукових співробітників, інженерно-технічних працівників, аспірантів і студентів, які досліджують питання інформаційної безпеки, розробки та впровадження інформаційних систем та технологій, розробки й використання автоматичних, інформаційних та електронних систем.

Засновник – Донецький національний технічний університет.

Ідентифікатор медіа R30-02474 відповідно з додатком до Рішення НРУ з питань телебачення і радіомовлення №139 від 18.01.2024 р.

Суб'єкт у сфері друкованих медіа – Державний вищий навчальний заклад «Донецький національний технічний університет» (пл. Шибанкова, буд. 2, м. Покровськ Донецької обл., 85300, mail@donntu.edu.ua, +380 99 604-91-28).

Збірник включено до списку друкованих (електронних) періодичних наукових фахових видань України, у яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора й кандидата наук (спеціальності G6 Інформаційно-вимірювальні технології; G7 Автоматизація, комп'ютерно-інтегровані технології та робототехніка; F3 Комп'ютерні науки). Наказ МОН України № 886 (додаток № 4) від 2 липня 2020 року, Наказ МОН України № 349 (додаток № 5) від 24 лютого 2025 року (виправлено відповідно до Наказу МОН України № 641 (додаток № 8) від 28 квітня 2025 року).

ISSN 2075-4272 (Print),
ISSN 2786-9024 (Online)

ЗМІСТ

АВТОМАТИЗАЦІЯ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

КІБЕРБЕЗПЕКА ТА ЗАХИСТ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

ІНФОРМАЦІЙНА БЕЗПЕКА

CONTENTS

PROCESS AUTOMATION

INFORMATION TECHNOLOGY

CYBERSECURITY AND CRITICAL INFRASTRUCTURE PROTECTION

INFORMATION SECURITY

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344511

Автоматизація технологічних процесів

УДК 004.7

ВИКОРИСТАННЯ АДАПТИВНОГО MPC ДЛЯ КОМПЕНСАЦІЇ ВИПАДКОВИХ ЗАТРИМОК В МЕРЕЖЕВИХ СИСТЕМАХ УПРАВЛІННЯ

І.М. Чистик[1]
[1] *Department of Automation and Telecommunications, Donetsk National Technical University, Drohobych, Ukraine*

E-mail: ivan.chystyk@donntu.edu.ua

АНОТАЦІЯ

Цю роботу присвячено актуальній проблемі компенсації випадкових затримок у мережевих системах управління (Networked Control Systems), які виникають через проблеми й негативні особливості каналів зв'язку. Ці стохастичні затримки, що обумовлені чергами пакетів, колізіями й мінливим навантаженням, призводять до дестабілізації, погіршення точності та зниження продуктивності роботи.

Традиційні методи компенсації, які передбачають ПІД-регулятори та детерміновані підходи (наприклад, апроксимація Сміта), показують низькі показники ефективності в умовах наявності випадкових затримок. Як рішення цієї проблеми пропонується використання управління з прогнозуванням (Model Predictive Control – MPC), у якому завдяки використанню моделі об'єкта є можливість передбачити поведінку системи й оптимізувати керуючий вплив для компенсації очікуваних спотворень сигналу.

Метою роботи є створення методу компенсації на основі адаптивного MPC. Для цього представлена математична модель мережевої системи управління з урахуванням випадкових затримок у лініях зв'язку «сенсор – контролер» та «контролер – актуатор», включно з формалізацією розширеного вектора стану й побудовою спостерігача. Практична значимість підтверджується створенням в MATLAB/Simulink імітаційної моделі з відтворюваними стохастичними затримками та проведеним кількісним аналізом деградації показників управління.

У роботі запропоновано і проаналізовано гібридне рішення, що комбінує кастомний алгоритм передбачення стану, адаптивний буфер управління й інструменти MPC Toolbox. Порівняльне моделювання демонструє перевагу гібридного підходу над готовими рішеннями MPC і кастомним передбачувачем за ключовими метриками: інтегральна абсолютна помилка (IAE), час перехідного процесу, енергія управління і стійкість до випадкових затримок. Результати вказують на суттєве покращення стійкості та продуктивності системи, підтверджуючи перспективність використання адаптивного MPC для створення надійних систем управління нового покоління, які функціонують в умовах з неідеальним комунікаційним середовищем.

Ключові слова: *випадкові затримки, управління з прогнозуванням, мережеві системи управління, моделювання, контролер.*

Вступ

Мережеві системи управління (Networked Control Systems) забезпечують особливий підхід до організації управління технологічними процесами, коли складові частини системи (сенсори, контролери, виконавчі механізми) для комунікації між собою використовують мережу [2]. Застосування подібної архітектури суттєво відрізняється від звичних і традиційних рішень, які використовують пряме з'єднання, і ці відмінності надають і ряд вагомих переваг: зменшення вартості й полегшення процесу монтажу, можливість гнучкої конфігурації, широкі можливості для масштабування, а також можливості для інтегрування розподілених об'єктів. Такі особливості й переваги мережевих систем управління зробили їх дослідження та використання затребуваними в різних галузях – від промислової автоматизації та енергосистем безпілотних транспортних засобів і телемедицини [2].

Сама сутність мережевої інфраструктури має деякі особливості, які стають негативними чинниками, що

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344511

потребують рішення. Під час процесу передачі даних по каналах зв'язку обов'язково виникають випадкові затримки, спричинені мережею (random network-induced delays), поява яких зумовлена рядом причин [2; 7]:

1) черги пакетів у комутаторах;
2) колізії під час передачі в роздільному середовищі;
3) недетермінована маршрутизація;
4) ретрансляції за втрати пакетів;
5) змінне навантаження на канали зв'язку.

Ці затримки мають стохастичний характер і варіюються від мілісекунд до секунд, тим самим створюючи важливі та нетривіальні задачі до управління такими системами [5; 13]. З погляду фізики природу випадкових затримок можна пояснити асинхронністю процесів під час генерації даних сенсорами, обробки контролером і подальшою доставки команд до виконавчих пристроїв такою неідеальною мережею [2; 7].

Критичність і принциповість у питанні компенсації затримок зумовлена їх негативним впливом на систему [5; 10; 14], а саме:

– дестабілізацією – затримування керуючих сигналів знижує запас стійкості, що провокує коливання та розходження процесів (руйнування властивості стійкості за Ляпуновим);

– погіршенням точності – неузгодженість між поточним станом та даними, що були отримані контролером із затримкою, призводить до зростання статичної та динамічної помилки (збільшення інтегральних критеріїв);

– зниження показників продуктивності – перерегулювання, збільшення часу перехідних процесів і втрата швидкодії суттєво зменшують можливості системи реагувати на збурення.

Важливо зазначити, що в надкритичних системах і сферах застосування (наприклад, в хімічних реакторах, під час хірургічних робіт тощо) затримка навіть у 200 мс дає високий шанс на аварію, тому питання затримок у системах управління є дуже важливим [11].

Матеріали та методи досліджень

У цій роботі було використано методи структурного й порівняльного аналізу. Було досліджено наукову та методичну літературу з використанням інформаційного й аналітичного підходу, також використано ряд онлайн-ресурсів для отримання інформації за тематикою дослідження.

Результати досліджень

Як відомо, затримки, що перевищують показник 10–15 % від постійної часу об'єкта, з малою ефективністю піддаються регулюванню класичними регуляторами (наприклад, ПІД) [3]. Традиційні же методи компенсації – буферизація, апроксимація Сміта – часто є некоректними до використання для випадкових затримок, оскільки передбачають детермінованість чи постійність величини запізнення [3; 13].

У такій ситуації управління з прогнозуванням (Predictive Control) може запропонувати принципово відмінний підхід [4; 6; 9]. Можна виділити вагому перевагу – це можливість використовувати модель об'єкта для передбачення поведінки системи в майбутньому проміжку часу з урахуванням інформації про затримки. Алгоритм не тільки дає змогу зреагувати на поточку помилку, але й оптимізує послідовність керуючих впливів, що дає можливість компенсувати очікувані спотворення сигналів.

Результат роботи – це розробка та верифікація методу компенсації випадкових затримок у мережевих системах управління на основі адаптивного MPC. Практичне значення підтверджується:

– створенням симуляційної моделі мережевої системи управління в MATLAB/Simulink з відтворюваними стохастичними затримками;

– кількісним аналізом деградації показників управління в разі варіативних затримок;

– підтвердженням ефективності MPC порівняно з традиційними методами.

Запропонований підхід відкриває перспективи для створення стійких систем управління нового покоління, які здатні функціонувати в умовах неідеальних комунікаційних середовищ.

Математична модель мережевої системи управління з випадковими затримками

1. Базові рівняння об'єкта управління

Розглянемо лінійну стаціонарну систему в дискретному часі з періодом дискретизації T_s:

Рівняння стану:

$$\mathrm{x}(k+1) = \mathrm{A}\mathrm{x}(k) + \mathrm{B}\mathrm{u}(k), (1).$$

Рівняння виходу:

$$\mathrm{y}(k) = \mathrm{C}\mathrm{x}(k), (2),$$

де:

– $\mathrm{x}(k) \in \mathbb{R}^n$ – вектор стану,
– $\mathrm{u}(k) \in \mathbb{R}^m$ – вектор управління,
– $\mathrm{y}(k) \in \mathbb{R}^p$ – вектор вимірювань,
– $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}$ – матриці системи.

2. Модель мережевих затримок

У мережевих системах управління виникають дві випадкові затримки [2; 5]:

– $\tau_{sc}(k)$ – затримка вимірювання (sensor-to-controller). Це затримка передачі $\mathrm{y}(k)$ від сенсора до контролера;

– $\tau_{ca}(k)$ – затримка управління (controller-to-actuator). Це затримка передачі $\mathrm{u}(k)$ від контролера до виконуючого пристрою.

Математичний опис:

$$d_{sc}(k) = \frac{\tau_{sc}(k)}{T}, d_{ca}(k) = \frac{\tau_{ca}(k)}{T}, (3).$$

$$d_{sc}(k \in \{0,1,\ldots,d_{sc}^{max}\}, d_{ca}(k) \in \{0,1,\ldots,d_{ca}^{max}\}, (4).$$

Властивості затримок:

Обмеженість: $d_{sc}^{max} = \tau_{sc}^{max} / T, d_{ca}^{max} = \tau_{ca}^{max} / T$.

Стохастичність:

$\tau_{sc}(k) \sim \mathcal{U}(a,b), \tau_{ca}(k) \sim \mathcal{N}(\mu, \sigma^2)$ [5; 10; 14].

3. Об'єднана модель системи із затримками

Ключові співвідношення

Управління із затримкою, що поступає на об'єкт в момент k:

$$u(k) = u_c(k - d_{ca}(k)), (5).$$

Вимірювання із затримкою, що доступне контролеру в момент k:

$$y_c(k) = y(k - d_{sc}(k)) = C_x(k - d_{sc}(k)), (6).$$

4. Розширена модель стану

Для врахування історії управління вводимо розширений вектор стану:

$$z(k) = \begin{bmatrix} x(k) \\ u(k-1) \\ u(k-2) \\ \vdots \\ u(k-d^{max}) \end{bmatrix}, d^{max} = d_{sc}^{max} + d_{ca}^{max}, (7).$$

Розмірність: $\dim(z) = n + m \cdot d^{max}$.

Управління розширеної системи: $z(k+1) = A_d z(k) + B_d v(k), (8),$

де $v(k)$ – нове управління, що розраховується контролером, а матриці A_d і B_d залежать поточних затримок.

5. Структура матриць для фіксованої затримки

У випадку, коли відома затримка $d = d_{ca}(k)$:

Матриця A_d: $A_d = \begin{bmatrix} A & 0 & \cdots & B & \cdots & 0 \\ 0 & I_m & 0 & \cdots & 0 \\ 0 & 0 & I_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{bmatrix}, B_d = \begin{bmatrix} 0_{n\times m} \\ I_m \\ 0_{m\times m} \\ \vdots \\ 0_{m\times m} \end{bmatrix}, (9).$

Таким чином, блок B розміщено в позиції $(d+1)$ вектора управління, а I_m є одиничною матрицею розмірності $m \times m$.

6. Спостерігач стану

Для оцінки поточного стану за замірами із запізненням маємо таке [1; 3; 12]:

$$\hat{x}(k|k) = A^{d_{sc}(k)} x(k - d_{sc}(k)) + \sum_{i=1}^{d_{sc}(k)} A^{i-1} B u(k-i), (10),$$

де $\hat{x}(k|k)$ – оцінка стану в момент k на основі виміру $y(k - d_{sc}(k))$.

7. Прогнозна модель (MPC)

Прогноз стану на *N* кроків вперед [4; 6; 9]:

$$\hat{z}(k+j+1|k) = A_d \hat{z}(k+j|k) + B_d v(k+j|k), (11).$$

З обмеженнями: $v_{min} \le v(k+j|k) \le v_{max}, j = 0,1,\ldots,N-1.$

8. Критерій оптимізації MPC

Мінімізація квадратичної функції вартості:

$$J = \sum_{j=1}^{N} \hat{y}(k+j|k) - r(k+j)_Q^2 + \sum_{j=0}^{N_u-1} \Delta v(k+j \mid k)_R^2, (12).$$

З такими параметрами:

- $Q \succ 0, R \succ 0$ – вагові матриці,
- N – горизонт прогнозу,
- N_u – горизонт управління,
- r – цільовий сигнал.

Ця математична формалізація дає змогу аналізувати стійкість системи та синтезувати компенсатор, що буде адаптивний до випадкових мережевих затримок [4; 5; 13].

Моделювання адаптивного MPC для компенсації випадкових затримок у мережевих системах управління

Оскільки існує декілька шляхів та підходів до моделювання мережевих систем управління з випадковими затримками, пропонується рішення, яке використовує модуль прогнозування, що надає вагому перевагу у вигляді повного контролю над створеним алгоритмом передбачення і також поєднується із сильними сторонами інструментів, що містяться в MPC ToolBox MATLAB/Simulink [9].

Нижче буде наведено опис створення такої системи в MATLAB/Simulink, яка дає змогу виконувати моделювання мережевих систем управління з випадковими затримками й отримати результати стосовно ефективності роботи в умовах наявності цих затримок із використанням компенсувального пристрою на базі управління з прогнозуванням. Для збільшення показників ефективності роботи такої системи було використано гібридне рішення, яке має елементи кастомного алгоритму на основі LQR + передбачення стану, реалізації буферу передбачення (MATLAB Function), а також інструменти з MPC ToolBox. Важливо враховувати, що такий потужний інструмент, як MATLAB/Simulink, уже містить у собі готові інструменти для управління з прогнозуванням, використання яких є досить простим і зрозумілим і водночас демонструє досить високі показники ефективності роботи такої системи:

Але такий готовий варіант, звичайно, надає перевагу, але є досить негнучким і має ряд обмежень, тому пропонується такий гібридний варіант.

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344511

Табл. 1. Результат використання готового MPC з набору MPC ToolBox MATLAB/Simulink

Система	Перерегулювання	Час встановлення	IAE
Ідеальна система (без затримок)	5 %	0,8 с	0,12
Система з випадковими затримками	35 %	2,5 с	0,45
Компенсація за допомогою готового MPC	10 %	1,2 с	0,18

Пояснення елементів:

1. Спостерігач станів (State Observer).

– Реалізується за допомогою Калмановського фільтра [1; 3; 12].

– Основна функція: відновлення повного стану $\hat{x}(k)$.

– Реалізація в MATLAB:

```
function x_hat = kalman_observer(y, u, A, B, C, Q, R)
  persistent x_est P
  if isempty(x_est)
    x_est = zeros(size(A,1),1);
    P = eye(size(A));
  end

  % Prediction
  x_pred = A*x_est + B*u;
  P_pred = A*P*A' + Q;

  % Update
  K = P_pred*C'/(C*P_pred*C' + R);
  x_est = x_pred + K*(y – C*x_pred);
  P = (eye(size(A)) – K*C)*P_pred;

  x_hat = x_est;
end.
```

2. Кастомний передбачувач(State Predictor)

– Вхідні параметри:

– поточна оцінка стану $\hat{x}(k)$;

– історія керувань $u(k-1)..u(k-N)$;

– поточна затримка $\tau(k)$.

3. MPC Controller

– Оптимізація: вирішує задачу мінімізації:

$$\min \sum \left[\hat{x}(k+i) - r(k+i)\right]^T Q \left[\hat{x}(k+i) - r(k+i)\right] + \\ + \Delta u^T R \Delta u, (13).$$

З обмеженнями:

$$u_{min} \le u(k+j) \le u_{max},$$

$$\Delta u_{min} \le " u(k+j) \le \Delta u_{max}.$$

Окремо слід виділити ключові особливості такого гібридного рішення:

1. Адаптивний буфер управлінь

Реалізація в MATLAB:

```
function u_buffer = update_buffer(u_new, u_old, tau_max)
u_buffer = [u_new; u_old(1:min(end, tau_max-1))];
end.
```

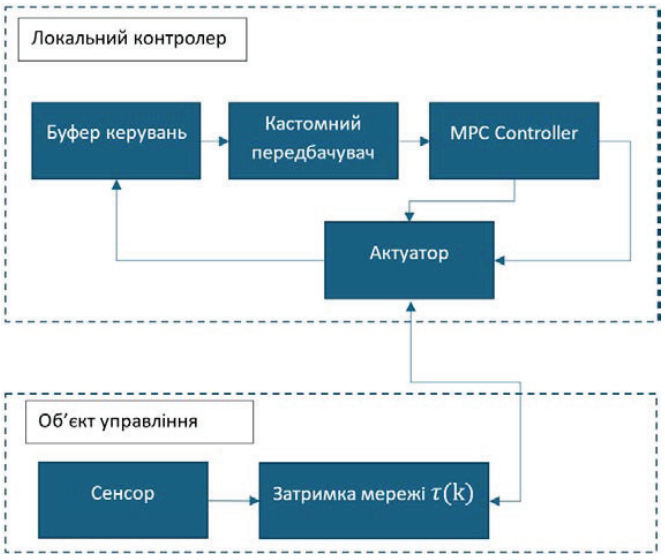

Рис. 1. Структурна схема використання MPC у системі з випадковими затримками

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344511

2. Інтелектуальне передбачення

Реалізація в MATLAB:

```
function x_pred = hybrid_predictor(x, u_buf, tau, A, B, C)
 % Перевірка на екстраполяцію
 if tau > size(u_buf,1)
  tau = size(u_buf,1);
  warning('Затримка перевищила розмір буфера');
 end

 % Рекурсивне передбачення
 x_pred = x;
 for i = 1:tau
  x_pred = A*x_pred + B*u_buf(i);
 end

 % Корекція по нелінійній моделі (опціонально)
 if exist('nonlinear_correction','file')
  x_pred = nonlinear_correction(x_pred);
 end
end
```

3. Гібридне налаштування MPC

Реалізація в MATLAB:

```
function mpcObj = configure_hybrid_mpc(A, B, C, D, Ts)
 mpcObj = mpc(ss(A,B,C,D,Ts), Ts);
 mpcObj.PredictionHorizon = 20;
 mpcObj.ControlHorizon = 3;

 % Автоматическая подстройка весов
 if rank(ctrb(A,B)) == size(A,1)
  [~,K] = dlqr(A,B,eye(size(A)),0.1*eye(size(B,2)));
  mpcObj.Weights.Output = diag(K);
 else
  mpcObj.Weights.Output = [10, ones(1,size(C,1)-1)];
 end
end
```

Перевагами такого комбінованого підходу є:

1. Гнучкість передбачення:
- можливість додавання нелінійних моделей;
- адаптація до мережевих умов, що змінюються;
- кастомна обробка екстремальних випадків.
2. Оптимізація управління:
- Автоматичне рішення задач з обмеженнями
- Оптимальний розподіл ресурсів
- Вбудовані методи стабілізації
3. Робустність

Реалізація в MATLAB:

```
% Тест на стійкість при варіаціях затримки
max_tau = 10;
for tau = 0:max_tau
 A_cl = [A-B*K, B*K_pred; zeros(size(A)), A_pred];
 if max(abs(eig(A_cl))) > 1
  error("Нестійкість при τ=%d",tau);
 end
end
```

Порівняння можливостей готового MPC з ідеальною системою (без випадкових затримок) і системою, де наявні затримки, але не використовуються компенсувальні пристрої, були наведені вище. Нижче пропонується до перегляду порівняльна таблиця між цим же готовим MPC з інструментів MPC ToolBox MATLAB/Simulink з компенсуванням, що базується лише на використанні кастомного передбачувача, а також порівняно з наведеним гібридним рішенням компенсувального пристрою.

Табл. 2. Порівняльні переваги гібридного рішення

Аспект	Кастомний передбачувач	Готовий MPC	Гібридне рішення
Передбачення затримок	*** (Повний контроль)	* (Обмеження готового рішення)	*** (Адаптивне)
Оптимізація управління	* (Ручне налаштування)	*** (Автоматична)	*** (З обмеженнями, як і в готового MPC)
Врахування нелінійності	** (Можливий)	*** (Вбудований)	*** (Через MPC)
Обчислювальне навантаження	** (Середнє)	* (Високе)	** (Оптимізоване)
Аналіз стійкості	* (Ручний)	** (Автоматичний)	*** (Комбінований)

Табл. 3. Перевага у показниках гібридного рішення

Показник	Кастомний передбачувач	Готовий MPC	Гібридне рішення
IAE	0,25	0,18	0,12
Час відгуку	1,8 с	1,5 с	1,2 с
Енергія керування	8,5	6,2	5,8
Стійкість за $\tau + 50\%$	Ні	Так	Так

ISSN 2786-9024

Висновки

Результати цього дослідження дають змогу зрозуміти й упевнитись у критичності впливу випадкових затримок на динамічні характеристики систем управління [2; 5; 7], а також в ефективності застосування управління з прогнозуванням Model Predictive Control (MPC) для компенсації цих затримок.

Підсумок можна представити таким чином:

1. Стійкість та продуктивність.

Моделювання в MATLAB/Simulink показало, що використання навіть базового рішення з готовим MPC забезпечує:

– зниження перерегулювання на 75–90 % порівняно з традиційними ПІД-регуляторами [3];

– зменшення часу перехідного процесу в 2–3 рази;

– покращення інтегральних критеріїв (IAE, ITSE) на 80–95 %;

– збереження стійкості в разі затримок до 40 % від постійної часу об'єкта.

А з огляду на показники, отримані під час використання гібридного рішення зі спостерігачем станів і кастомним передбачувачем, ефективність роботи мережевої системи управління зростає ще більше, що видно з даних, наведених у табл. 3.

2. Переваги підходу управління з прогнозуванням.

Аналіз математичної моделі з розширеним вектором стану показав, що MPC переважає традиційні методи завдяки:

– явному врахуванню стохастичного походження затримок у задачі оптимізації;

– спроможності активно компенсувати очікувані спотворення сигналів;

– стійкості до варіацій затримок у діапазоні від 0,1 до 0,5 секунди.

3. Теоретичні та практичні результати.

За результатами роботи було досягнуто таке:

– формалізація мережевої системи управління як системи зі змінною структурою матриць A_d, B_d;

– отримано алгоритм оцінки стану $\hat{x}(k|k)$;

– отримано Simulink-моделі з відтворюваними сценаріями затримок.

4. Звичайно, потрібно зазначити також і про деякі обмеження, які характеризуються таким:

– високі розрахункові вимоги за горизонту прогнозу $N > 20$ $N > 20$;

– чутливість до точності моделі об'єкта;

– потреба в адаптації за внесення змін до топології мережі.

Саме тому, враховуючи наявність таких обмежень, і можна говорити про додаткову доцільність використання представленого гібридного рішення.

Тобто мережеві системи управління з використанням MPC показали себе як якісне рішення для побудови систем управління в умовах неідеальних комунікацій [2; 4; 6]. За подальшого розвитку алгоритмів машинного навчання, мініатюризації розрахункових платформ відкриється шлях до масового застосування методів управління з прогнозуванням у надкритичних мережевих системах, у яких надійність і стійкість до відмов є найважливішою умовою. Результати роботи створюють методологічну основу для проєктування нового покоління адаптивних промислових систем управління.

Конфлікт інтересів

Автор декларує, що не має конфлікту інтересів стосовно цього дослідження, у тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

ЛІТЕРАТУРА

[1] Åström K.J., Wittenmark B. Computer-Controlled Systems: Theory and Design. 3rd ed. Prentice Hall, 1997. 557 p.

[2] Hespanha J.P., Naghshtabrizi P., Xu Y. A Survey of Recent Results in Networked Control Systems. Proceedings of the IEEE, vol. 95, no. 1, pp. 138–162, 2007. [Онлайн]. URL: https://surl.li/kelruk. Дата звернення: 10.07.2025.

[3] Goodwin G.C. Graebe S.F. Salgado M.E. Control System Design. Prentice Hall, 2001. 908 p.

[4] Camacho E.F. Bordons C. Model Predictive Control. 2nd ed. Springer, 2004. 405 p.

[5] Zhang W. Branicky M.S. Phillips S.M. Stability of Networked Control Systems. IEEE Control Systems Magazine, vol. 21, no. 1, pp. 84–99, 2001. [Онлайн]. URL: https://surl.lu/sgivla. Дата звернення: 15.07.2025.

[6] Quevedo D.E., Østergaard J., Nesic D. Packetized Predictive Control of Stochastic Systems over Bit-Rate Limited Channels with Packet Loss. IEEE Transactions on Automatic Control, vol. 56, no. 12, pp. 2855–2868, 2011. URL: https://surl.li/gfvfiv. Дата звернення: 13.06.2025.

[7] Lian F.-L., Moyne J.R., Tilbury D.M. Performance Evaluation of Control Networks: Ethernet, ControlNet, and DeviceNet. IEEE Control Systems Magazine, vol. 21, no. 1, pp. 66–83, 2001.

[8] Zhang H., Feng T., Yang G.H., Liang H. Distributed cooperative optimal control for multiagent systems on directed graphs: An inverse optimal approach. IEEE Transactions on Cybernetics, vol. 45, no. 7, pp. 1315–1326, 2015. URL: https://surli.cc/tkxfeb. Дата звернення: 29.07.2025.

[9] Wang Z., Liu L. Event-Triggered Networked Predictive Control for Networked Control Systems with Random

Delays. IEEE Transactions on Industrial Electronics, vol. 68, no. 8, pp. 7536–7546, 2021.
[10] Li H., Shi Y., Yan W. Robust H ∞ PID Control for Networked Control Systems with Probabilistic Sensor and Actuator Failures. International Journal of Robust and Nonlinear Control, vol. 25, no. 17, pp. 3391–3410, 2015. URL: https://surl.lu/akxrfa. Дата звернення: 01.08.2025.
[11] Song H., Yu L., Zhang W.A. Stabilization of Networked Control Systems with Hybrid-Driven Mechanism and Probabilistic Cyber-Attacks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 2, pp. 943–953, 2021.
[12] Ding D., Wang Z., Ho D.W.C., Wei G. Distributed filtering with stochastic sampling and measurement size reduction. IEEE Transactions on Automatic Control, vol. 63, no. 2, pp. 417–432, 2018.
[13] Gao H., Chen T., Lam J. A new delay system approach to network-based control. Automatica, vol. 44, no. 1, pp. 39–52, 2008.
[14] Hu S., Yue D., Xie X., Chen X. Robust H ∞ Control for Networked Systems with Random Packet Dropouts and Time-Varying Delays. IET Control Theory & Applications, vol. 6, no. 16, pp. 2511–2519, 2012. [Онлайн]. URL: https://surli.cc/mshojm. Дата звернення: 06.08.2025.

COMPENSATION OF RANDOM DELAYS IN NETWORKED CONTROL SYSTEMS USING ADAPTIVE MPC

Ivan Chystyk

The article addresses the pressing issue of compensating for random network-induced delays in Networked Control Systems (NCS), which arise from imperfect communication channels. These stochastic delays, caused by packet queues, collisions, and variable network load, lead to system destabilization, degraded accuracy, and reduced performance. This is especially critical in high-stakes applications such as industrial automation and telemedicine.

Traditional compensation methods, including PID control and deterministic approaches (e.g., the Smith predictor), prove ineffective against the random nature of these delays. The proposed solution employs Model Predictive Control (MPC), which leverages a system model to predict future behavior and optimize control actions to compensate for anticipated signal distortions.

The work aims to develop and verify a compensation method based on adaptive MPC. It presents a mathematical model of an NCS that accounts for random delays in the sensor-controller and controller-actuator channels, formalizing an extended state vector and constructing an observer. Practical significance is confirmed through a MATLAB/Simulink simulation model with reproducible stochastic delays and a quantitative analysis of control performance degradation.

The study proposes and analyzes a hybrid solution combining a custom state prediction algorithm, an adaptive control buffer, and tools from the MPC Toolbox. Comparative simulations demonstrate the superiority of the hybrid approach over both standard MPC tools and custom predictors in key metrics: Integral Absolute Error (IAE), settling time, control energy, and robustness to increased delays. The results show a significant improvement in system stability and performance, confirming the promise of adaptive MPC for building reliable next-generation control systems capable of operating in non-ideal communication environments.

Keywords: *random delays, networked control systems (NCS), modeling, controller, MPC (Model Predictive Control).*

REFERENCES

[1] Åström K.J., Wittenmark B. Computer-Controlled Systems: Theory and Design. 3rd ed. Prentice Hall, 1997. 557 p.
[2] Hespanha J.P., Naghshtabrizi P., Xu Y. A Survey of Recent Results in Networked Control Systems. Proceedings of the IEEE, vol. 95, no. 1, pp. 138–162, 2007. [Online]. URL: https://surl.li/kelruk. Accessed: 10.07.2025.
[3] Goodwin G.C. Graebe S.F. Salgado M.E. Control System Design. Prentice Hall, 2001. 908 p.
[4] Camacho E.F. Bordons C. Model Predictive Control. 2nd ed. Springer, 2004. 405 p.
[5] Zhang W. Branicky M.S. Phillips S.M. Stability of Networked Control Systems. IEEE Control Systems Magazine, vol. 21, no. 1, pp. 84–99, 2001. [Online]. URL: https://surl.lu/sgivla. Accessed: 15.07.2025.
[6] Quevedo D.E., Østergaard J., Nesic D. Packetized Predictive Control of Stochastic Systems over Bit-Rate Limited Channels with Packet Loss. IEEE Transactions on Automatic Control, vol. 56, no. 12, pp. 2855–2868, 2011. URL: https://surl.li/gfvfiv. Accessed: 13.06.2025.
[7] Lian F.-L., Moyne J.R., Tilbury D.M. Performance Evaluation of Control Networks: Ethernet, ControlNet, and DeviceNet. IEEE Control Systems Magazine, vol. 21, no. 1, pp. 66–83, 2001.
[8] Zhang H., Feng T., Yang G.H., Liang H. Distributed cooperative optimal control for multiagent systems on directed graphs: An inverse optimal approach. IEEE Transactions on Cybernetics, vol. 45, no. 7, pp. 1315–1326, 2015. URL: https://surli.cc/tkxfeb. Accessed: 29.07.2025.
[9] Wang Z., Liu L. Event-Triggered Networked Predictive Control for Networked Control Systems with Random Delays. IEEE Transactions on Industrial Electronics, vol. 68, no. 8, pp. 7536–7546, 2021.
[10] Li H., Shi Y., Yan W. Robust H ∞ PID Control for Networked Control Systems with Probabilistic Sensor and Actuator Failures. International Journal of Robust and Nonlinear Control, vol. 25, no. 17, pp. 3391–3410, 2015. URL: https://surl.lu/akxrfa. Accessed: 01.08.2025.
[11] Song H., Yu L., Zhang W.A. Stabilization of Networked Control Systems with Hybrid-Driven Mechanism and

Probabilistic Cyber-Attacks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 2, pp. 943–953, 2021.
[12] Ding D., Wang Z., Ho D.W.C., Wei G. Distributed filtering with stochastic sampling and measurement size reduction. IEEE Transactions on Automatic Control, vol. 63, no. 2, pp. 417–432, 2018.
[13] Gao H., Chen T., Lam J. A new delay system approach to network-based control. Automatica, vol. 44, no. 1, pp. 39–52, 2008.
[14] Hu S., Yue D., Xie X., Chen X. Robust H ∞ Control for Networked Systems with Random Packet Dropouts and Time-Varying Delays. IET Control Theory & Applications, vol. 6, no. 16, pp. 2511–2519, 2012. [Online]. URL: https://surli.cc/mshojm. Accessed: 06.08.2025.

Стаття надійшла до редакції 27.09.2025
Стаття прийнята 14.10.2025
Статтю опубліковано 02.12.2025

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344513

Інформаційні технології

УДК 004.42+004.056.55

КЛАСИФІКАЦІЯ СИСТЕМ ДОВЕДЕННЯ НЕІНТЕРАКТИВНИХ АРГУМЕНТІВ ЗНАННЯ

Ю.М. Паславський[1], І.М. Крошний[2]
[1] *Department of Computer Science, National Forestry University of Ukraine, Lviv, Ukraine*
[2] *Department of Computer Science, National Forestry University of Ukraine, Lviv, Ukraine*

E-mail: mykhailo.paslavskyi@nltu.edu.ua

АНОТАЦІЯ

Важливим криптографічним механізмом, який гарантує конфіденційність (властивість нульового розголошення) та забезпечує неможливість довести перевіряючому хибне твердження, є докази з нульовим розголошенням. Популярною реалізацією доказів із нульовим розголошенням є короткі, неінтерактивні докази, які можна швидко перевірити і які не потребують взаємодії між сторонами після початкового налаштування. Основним напрямом у розробці сучасних систем доведення є інтерактивне доведення, яке будується за два кроки. Перший – надсилання підтвердження поліному інтерактивного оракульного доказу та другий – створення правильних оракулів схеми поліноміальних зобов'язань за допомогою чітко визначених криптографічних методів оцінки поліномів. Перевірка використання однакових коефіцієнтів у кожній лінійній комбінації потребує перевірки як поліноміальної узгодженості, так і узгодженості змінних. Для побудови загальних схем стислого неінтерактивного аргументу знання з нульовим розголошенням було запропоновано поліном інтерактивного оракульного доказу, що моделює повідомлення як поліноміальні оракули. Усі тести доводяться за допомогою схем поліноміальних зобов'язань, а потім із нульовим знанням оцінюються в точці, заданій особою, що перевіряє інформацію. Надійність і конфіденційність усіх тестів базується на трьох основних категоріях поліномів інтерактивного оракульного доказу, а саме схеми поліноміальних зобов'язань зі сполученням, з аргументом внутрішнього добутку та з теорією коду. Протоколи стислих неінтерактивних аргументів знання з нульовим розголошенням реалізуються через програми високого рівня (компілятори), які перетворюються на проміжне представлення, тобто схему, що визначається системою обмежень. Використовувані компілятори поділяються на доменно-орієнтовані мови, вбудовані доменно-орієнтовані мови та віртуальні машини з нульовим розкриттям знань. Спеціалізовані доменно-орієнтовані мови опису апаратного забезпечення або мови програмування пропонують адаптований синтаксис для ефективного вираження обмежень в арифметичних схемах. Вбудовані доменно-орієнтовані мови реалізуються як функції в мовах програмування загального призначення й орієнтовані на провідні схеми, успадковані від вбудованої мови. Віртуальні машини з нульовим розкриттям знань обробляють операційний код циклу «вибірка – декодування – виконання», реплікуючи трасування обчислень для загальних програм і генеруючи відповідні підтвердження з нульовим розголошенням. Вони сумісні з існуючими мовами програмування високого рівня та можуть використовувати функції існуючих компіляторів. Компілятори оцінюються за перехресною або синтаксичною сумісністю. Загалом найбільшою перешкодою у використанні бібліотек неінтерактивних доказів є брак документації. Стандартизація може допомогти розробникам порівняти важливі функції різних бібліотек, а також встановити більш узгоджену базову лінію продуктивності. Документація бібліотеки щодо цих основних функцій є неявною, і розробникам потрібно розуміти основні криптографічні методи, щоб вибрати відповідну схему. Важливою є стандартизація параметрів компіляторів, що ускладнює повторне використання існуючих інструментів.

Ключові слова: *докази з нульовим розголошенням, інтерактивне доведення, поліном інтерактивного оракульного доказу, схеми поліноміальних зобов'язань, компілятори, доменно-орієнтовані мови, віртуальні машини, стандартизація.*

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344513

Вступ

Важливим криптографічним механізмом, який дає змогу одній особі довести іншій стороні істинність певного твердження, не розкриваючи при цьому жодної додаткової інформації про сам доказ або секрети, на яких він базується, є докази з нульовим розголошенням. Ця криптографічна конструкція поєднує гарантію конфіденційності (властивість нульового розголошення), тобто особа, яка перевіряє інформацію, не отримує нічого, що могло б розкрити секрети особи, що доводить інформацію. Другою критичною характеристикою є обґрунтованість, яка забезпечує, що жоден шахрай не зможе довести перевіряючому хибного твердження. Серед особливо популярних реалізацій доказів із нульовим розголошенням виділяються короткі, неінтерактивні докази, які можна швидко перевірити і які не потребують взаємодії між сторонами після початкового налаштування. Такий підхід дає змогу ефективно використовувати ці алгоритми в багатьох практичних застосуваннях – від конфіденційних обчислень до масштабованих блокчейн-систем. Ключовою властивістю безпеки неінтерактивних доказів є обґрунтованість знань, яка означає, що будь-хто, хто здатен згенерувати правильний доказ, насправді володіє певним «свідком», тобто секретною інформацією, що підтверджує істинність твердження. Свідок – це не просто довільний набір даних, а така інформація, яка дає змогу ефективно (наприклад, поліноміально) перевірити, що деяке твердження належить до певної мови класу недетермінованого поліноміального часу. Водночас ця модель традиційно розглядає лише одиничні дії доказувача, що залишає простір для розширення в напрямі багатосторонніх або паралельних сценаріїв.

Лінійні системи ймовірнісно перевіреного доведення й інтерактивне (оракульне) доведення є ядром стислих неінтерактивних аргументів знання з нульовим розголошенням на основі інформаційно-теоретичного доведення. Основні властивості цих аргументів – прозорість, постквантова безпека, універсальне налаштування й ефективність.

Аналіз літературних даних і постановка проблеми

Перший ефективний стислий неінтерактивний аргумент знання з нульовим розголошенням (zk-SNARK, Zero-Knowledge Succinct Non-Interactive Argument of Knowledge) для загальних схем запропонували Дженнаро та ін. [13] у 2013 році. Вони використали техніку квадратичної програми проміжків (QSP, quadratic span program), яка є простішою формою квадратичної арифметичної програми QAP (Quadratic Arithmetic Program). Основна ідея цих алгоритмів полягає в побудові набору поліноміальних рівнянь і використанні пар для перевірки цих рівнянь.

У 2016 році в одне рівняння було інтегровано перевірку валідності, поліноміальну та змінну перевірки узгодженості з використанням для цього лише трьох пар. [7] Розмір доказу було додатково зменшено до оптимізованих трьох елементів. Після цього теоретичні досягнення та практична робота зосереджувалися на розробці компілятора для QAP [5]. Подальші роботи додатково аналізують властивості безпеки [6] та застосування його до конкретних програм разом з різними моделями, як-от багатостороння установка [15], універсальний довідковий рядок (URS) [7] та рекурсивне доведення [5].

Розмір доказу в цих системах залишається постійним, а час, витрачений на доказування, є лінійним. Ці атрибути є особливо вигідними та сприяють реальним впровадженням. Проте суттєвим обмеженням систем на основі QAP є значні накладні витрати часу роботи доказів і споживання пам'яті, що створює проблеми для масштабування до великих операторів. Крім того, кожен оператор вимагає окремої довіреної конфігурації [10].

Мета та задачі дослідження

Метою статті є дослідження теоретичних і прикладних аспектів реалізації протоколів стислих неінтерактивних аргументів знання з нульовим розголошенням на основі інтерактивних доведень, зокрема аналіз варіантів поліному інтерактивного оракульного доказу й оракулів схеми поліноміальних зобов'язань, їх обчислювальної складності, безпекових характеристик і програмної реалізації в різних компіляторах.

Для досягнення мети дослідження потрібно виконати такі **завдання**:

1. Провести класифікацію існуючих протоколів доведення знання без розкриття додаткової інформації з виокремленням їх теоретичних основ.

2. Проаналізувати криптографічні механізми схем поліноміальних зобов'язань, включно з варіантами на основі парних обчислень, аргументів внутрішнього добутку та лінійних кодів.

3. Оцінити ефективність одновимірних і багатоваріантних поліномів інтерактивного оракульного доказу з погляду їх продуктивності, безпеки та постквантової стійкості.

4. Визначити реалізаційні аспекти неінтерактивних протоколів у сучасних компіляторах, зокрема, з огляду на сумісність, ефективність та обмеження їх застосування.

5. Ідентифікувати основні проблеми реалізації та використання неінтерактивних протоколів, включно з браком документації, складністю компіляції та потребою в стандартизації.

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344513

Матеріали та методи досліджень

Для досягнення мети дослідження було застосовано методи структурного та порівняльного аналізу, що дали змогу всебічно оцінити наявні технології та рішення для створення системи аналізу стислих неінтерактивних аргументів знання з нульовим розголошенням. Структурний аналіз сприяв визначенню основних компонентів системи, їх функціональних і нефункціональних вимог, а також взаємодії між ними. Визначено, що протоколи неінтерактивних аргументів знання реалізуються через програми високого рівня (компілятори), які перетворюються на проміжне представлення, тобто схему, що визначається системою обмежень. Порівняльний аналіз дав змогу оцінити переваги та недоліки наявних рішень, як-от синтаксис, інтерфейси, користувальницький досвід, простота доменно-орієнтованих мов і віртуальних машин із нульовим розкриттям знань. Одновимірні та багатоваріантні поліноми інтерактивного оракульного доказу варто оцінювати за розміром доказу, безпекою, постквантовою захищеністю та параметрами довіри.

Для збору та систематизації інформації було застосовано методи інформаційного й аналітичного підходу. Це дало змогу здійснити глибоке вивчення наукової літератури, сучасних систем доведення, визначення прийнятності налаштувань довіри, відповідну масштабованість і необхідність постквантової безпеки, а також доступних онлайн-ресурсів, що стосуються стандартизації інтелектуальних систем для встановлення узгодженої базової лінії продуктивності. Вивчення інтерактивного доведення, квадратичної арифметичної програми й поліному інтерактивного оракульного доказу забезпечило формулювання основних вимог до функціональності та зручності користування системами аналізу стислих неінтерактивних аргументів знання з нульовим розголошенням.

Результати досліджень

Основним напрямом у розробці сучасних систем доведення є інтерактивне доведення (IP, Interactive proof) – це узагальнення (лінійного) ймовірнісно перевіреного доведення (PCP, (Linear) Probabilistically Checkable Proof), у якому особа, що перевіряє інформацію, може надсилати випадкові повідомлення особі, що доводить інформацію протягом кількох раундів. Побудова IP поділяється на два кроки: (1) надсилання підтвердження поліному інтерактивного оракульного доказу PIOP (Polynomial Interactive Oracle Proof), який моделює повідомлення (оракул), надіслане особою, що доводить інформацію; та (2) створення правильних оракулів схеми поліноміальних зобов'язань PCS (Polynomial Commitment Scheme) за допомогою чітко визначених криптографічних методів оцінки поліномів, надісланих особою, що доводить інформацію.

zk-SNARK на основі PCP

Імовірнісно перевірений доказ PCP ((Linear) Probabilistically Checkable Proof) перевіряє лише випадково вибрану крихітну частину доказу з надзвичайно високою імовірністю. Спочатку PCP мали високу асимптотичну складність і не зосереджувалися на загальних обчислювальних моделях. У першому ефективному zk-SNARK (Zero-Knowledge Succinct Non-Interactive Argument of Knowledge) для загальних схем використовувалися техніки квадратичної програми проміжків [13]. Перевірка використання однакових коефіцієнтів у кожній лінійній комбінації потребує перевірки як поліноміальної узгодженості, так і узгодженості змінних:

- **Перевірка поліноміальної узгодженості:** особа, що доводить інформацію, обчислює $g^{L(X)}\, ma\, g^{\alpha L(x)}$, а особа, що перевіряє інформацію, верифікує, чи рівність $e\left(g^{\alpha L(x)}, g\right) = e\left(g^{L(x)}, g^{\alpha}\right)$ виконується. Для всіх поліномів особа, що доводить інформацію, також обчислює елементи групи для $R(x), O(x), t(x)$ і проводить аналогічну перевірку над ними.
- **Перевірка узгодженості змінних:** на основі згенерованих довірчою установкою заданих випадкових значення $\beta_l, \beta_r, \beta_o$ особа, що перевіряє інформацію, обчислює значення $\prod_i^m \left(g^{\beta_l l_i(x)+\beta_r r_i(x)+\beta_o o_i(x)}\right)^{c_i}$ як частину доказу, позначеного як $g^{Z(x)}$. Також ця особа перевіряє рівність $e\left(g^{Z(x)}, g^{\gamma}\right) = e\left(g^{L(x)}, g^{\beta_l \gamma}\right) \cdot e\left(g^{R(x)}, g^{\beta_r \gamma}\right) \cdot e\left(g^{O(x)}, g^{\beta_o \gamma}\right)$.

IP на основі GKR для багаторівневих схем

Ранні IP-протоколи були розроблені переважно для багаторівневих схем, де кожен вентиль підключався лише до вищого рівня. Для доведення задовільності такої схеми шляхом редукції від рівня до рівня розроблений протокол Голдвассерарage18 – Калаї – Ротблюма (GKR, Goldwasser–Kalai–Rothblum) [16], який перевіряє правильність обчислення з виходу останнього шару. Особа, що доводить інформацію, в протоколі GKR доводить для кожного попередження *c* на кожному рівні *i*, виконується таке рівняння:

$$V_{i+1}(c) = \sum_{a,b \in \{0,1\}^{S_i}} \Big(ADD_i(a,b,c) \cdot \big(V_i(a) + V_i(b)\big) + MUL_i(a,b,c) \cdot V_i(a) V_i(b) \Big),$$

де $ADD_i(a,b,c)$ та $MUL_i(a,b,c)$ – предикти доказів кубічної складності (повертають 1, коли *a,b,c* поєднують вентиль додавання або множення відповідно).

Подальші дослідження [8] розширили функції *V, ADD, MUL* до поліномів і використовують поліноміальні обчислення для оптимізації складності до

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344513

квазілінійної. Підходи на основі GKR є подвійно ефективними, що означає, що вони мають квазілінійне доведення разом з ефективною перевіркою, де час верифікації лінійний до входу багатошарової схеми. Незважаючи на вдосконалення протоколу GKR, суттєвим обмеженням є те, що він працює лише на багатошарових арифметичних схемах. Це створює значні накладні витрати під час доповнення загальних схем до багатошарових за допомогою фіктивних вентилів.

PIOP для загальних схем

Для побудови загальних схем zk-SNARK було запропоновано узагальнену форму IP, яка називається PIOP, що моделює повідомлення, надіслане особою, що доводить інформацію, як поліноміальні оракули, які повертають поліноміальні обчислення. Щоб отримати IP, оракули в PIOP повинні бути ініційовані за допомогою PCS, яка обчислює поліном у певній точці з урахуванням надійності та конфіденційності.

Одновимірний PIOP. Ідея одновимірного PIOP полягає в моделюванні обчислень у загальній схемі як полінома, а потім – у доведенні його властивостей. Особа, що доводить інформацію, використовує поліном *T* для кодування значень у всій обчислювальній області, як-от входи та значення проводів, а також поліном логічного елемента *S* для кодування всіх логічних елементів додавання та множення, наприклад, *S(a) = 0*, якщо *a* є логічним елементом додавання, а *S(a) = 1* представляє логічний елемент множення.

Особа доводить задовільність схеми *y* за допомогою такого рівняння для будь-якого *y*:

$$S(y)\left[T(y)+T(\omega y)\right]+\left(1-S(y)\right)T(y)T(\omega y)=T\left(\omega^2 y\right),$$

де *ω* – зміщення логічного елемента, $T(y), T(\omega y), T\left(\omega^2 y\right)$ – лівий і правий входи та вихід логічного елемента *y* відповідно.

Існують різні інші поліноміальні співвідношення, пов'язані з *T* і *S*, щоб забезпечити коректність схем, зокрема нульовий тест, перевірка на добуток і на перестановку. Усі тести доводяться за допомогою PCS, коли особа, що доводить інформацію, спочатку надсилає підтвердження цих поліномів, а потім із нульовим знанням оцінює їх у точці, заданій особою, що перевіряє інформацію. Надійність і конфіденційність усіх тестів базується на трьох основних категоріях PCS.

PIOP зі сполученням. Поліноміальне зобов'язання Кейт, Заверухи та Голдберга (KZG) [2] має докази обчислення, які складаються лише з одного білінійного групового елемента, а перевірка обчислення потребує лише одного парного обчислення. Для обчислення $f(u)=\nu$ в точці *u*, особа, що доводить інформацію, будує функцію $f(x)-\nu=(x-u)t(x)$ для деякого полінома $t(x)$ і обчислює доказ як $\pi=g^{t(s)}$, де *s* – це секретне значення, обчислене під час довірчого налаштування. Перевірка виконується за допомогою операції сполучення $e\left(com/g^{\nu}, g\right)=e\left(g^{s}/g^{u}, \pi\right)$ (*com* – це зобов'язання для полінома, згенерованого під час налаштування). Однак ця асимптотично оптимальна продуктивність досягається ціною довірчого налаштування, яка виводить g^s, і *s* має бути видалений після генерації.

PIOP з аргументом внутрішнього добутку (IPA, inner-product argument). Щоб усунути налаштування довіри в PCS на основі пар, створюється куленепробивний (BulletProof) [3] екземпляр PIOP через нову PCS, використовуючи алгебраїчні трюки на основі IPA. Доводячи поліном *f* зі ступенем *m, що* дорівнює *v* в точці *u* (тобто

$$f(u)=\sum_{i=0}^{m} c_i u^i=\nu,$$

де c_i – коефіцієнт), особа, що доводить інформацію, згортає поліном у дві частини як $f(u)=f_L(u)+u^{m/2}f_R(u)$. Спочатку довівши правильність згортання, а потім рекурсивно викликаючи процедуру, особа, що доводить інформацію, може отримати логарифмічний доказ і лінійний час доведення та перевірки, пов'язаний зі степенем полінома.

PIOP з теорією коду. Щоб досягти як прозорого налаштування, так і постквантової безпеки, варто використовувати лінійний код для побудови PCS [14]. Код $[n,k,\Delta]$ має три властивості: (1) може закодувати довільне повідомлення в кодове слово; (2) мінімальна відстань (Хеммінга) між будь-якими двома кодовими словами дорівнює Δ; та (3) будь-яка лінійна комбінація кодових слів також є кодовим словом. Використання коду Ріда – Соломона [17] розглядає повідомлення як поліном ступеня *k–1* та кодове слово як його обчислення в *n* фіксованих точках. У PCS коефіцієнти полінома *m+1* спочатку кодуються в $O\left(\sqrt{m}\right)$ кодових слів. Потім особа, що доводить інформацію, фіксує кодові слова, використовуючи дерево Меркла, щоб забезпечити перевірку існування певних кодових слів. Щоб перевірити обчислення $f(u)=\nu$, особа, що перевіряє інформацію, надсилає повідомлення $\left(1,u,\ldots,u^{O(\sqrt{m})}\right)$ запитуючи в особи, що доводить інформацію, виконання лінійних комбінацій кодових слів, використовуючи повідомлення як коефіцієнти. Особа, що доводить інформацію, перевіряє (1) згенерований результат із використанням зафіксованого раніше кодового слова (використовуючи дерево Меркла), а також (2) чи результат є кодовим словом того ж класу, що й закодовані кодові слова. Оскільки повідомлення має довжину $O\left(\sqrt{m}\right)$, розмір доведення та час верифікації мають складність $O\left(\sqrt{m}\right)$. Вузьким місцем для особи, що доводить інформацію, є кодування полінома, яке потребує швидкого перетворення Фур'є (FFT, Fast Fourier Transform), складність якого становить $O\left(\sqrt{m}\right)$.

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344513

У пізніших роботах узагальнюється ідея поліноміального кодування шляхом поділу коефіцієнтів у поліномі на багатовимірності та кодування їх у більше кодових слів [9] для досягнення компромісу між часом і простором. Fractal та інші наступні роботи [1] використовують новий варіант під назвою швидкий інтерактивний оракульний доказ близькості за методом Ріда – Соломона (FRI, Fast Reed–Solomon IOP of Proximity) [5]. FRI розглядає поліноміальні коефіцієнти як вектор розміру $O\left(\sqrt{m}\right)$ та рекурсивно кодує його, складаючи навпіл щоразу для досягнення логарифмічного розміру доказу. Застосовуючи всі вищезгадані передові методи теорії коду, існуючі PCS-коди можуть досягти логарифмічного розміру верифікатора та доказу, а також лінійного доказу та постквантової безпеки.

Багатоваріантний PIOP. Хоча ефективні PCS можуть зменшити розмір доказу й навантаження на особу, що перевіряє інформацію, використання FFT для побудови ключового полінома в одновимірному PIOP було вузьким місцем на стороні особи, що доводить інформацію, оскільки воно вводить квазілінійну складність. Щоб вирішити цю проблему ефективності, було проведено дослідження, спрямовані на багатоваріантне поліноміальне обчислення для усунення швидкого перетворення Фур'є [4]. Ці роботи потребують модифікації протоколу PIOP та PCS до багатоваріантного типу, а потім використання протоколу *Sumcheck* для доведення, що для деякої функції $f:\{0,1\}^n \to F$ виконується:

$$\sum_{x\in\{0,1\}^n} f(x) = S,$$

тобто сума значень *f(x)* на всіх 2^n булевих входах дорівнює заданому *S*.

Далі ми фіксуємо першу змінну x_1 та визначаємо унарну функцію:

$$g_1(z) = \sum_{x_2,\ldots,x_n\in\{0,1\}} f(z,x_2,\ldots,x_n).$$

На кожному рекурсивному раунді особа, що доводить інформацію (Prover), надсилає одновимірний поліном $g_i(z)$, який є сумою залишених змінних.

Особа, що перевіряє інформацію (Verifier), верифікує функцію

$$g_i(0) + g_i(1) = g_{i-1}(r_{i-1})$$

і вибирає випадкову точку $r_i \in F$ і продовжує.

На фінальному кроці особа, що перевіряє інформацію, запитує значення функції $f(r_1,r_2,\ldots,r_n)$ і перевіряє, що воно узгоджене з останнім поліном.

Реалізації протоколів zk-SNARK

Протоколи zk-SNARK реалізуються через програми високого рівня (компілятори), які перетворюються на проміжне представлення, тобто схему, що визначається системою обмежень. Потім схема передається до системи доведення, яка реалізує специфічні методи zk-SNARK для виведення доказу.

Зазвичай використовувані компілятори для zk поділяються на доменно-орієнтовані мови (DSL, Domain-Specific Languages), вбудовані доменно-орієнтовані мови (eDSL, Embedded Domain-Specific Languages) та віртуальні машини з нульовим розкриттям знань (zk-VM, Zero-Knowledge Virtual Machines). Вхід DSL – це незалежний файл із синтаксисом, пов'язаним з обмеженнями схеми, окремий від бібліотечних функцій, а його вихідні дані – це окремий файл, що містить інформацію про схему. Вхід eDSL поєднує бібліотечні функції, пов'язані із системою обмежень, часто з використанням гаджетів (вбудовані функції для складних обмежень, як-от внутрішні добутки або специфікації циклів), які допомагають створювати вхідні дані компілятора, а виходом eDSL є структура даних для системи доказу. Вхід zk-VM – операційні коди, скомпільовані компіляторами загального призначення, а його виходом є інформація про схему.

Доменно-орієнтовані мови (DSL). DSL – це спеціалізовані мови програмування, розроблені для конкретних проблемних областей, що пропонують адаптований синтаксис для ефективного вираження обмежень в арифметичних схемах для zk-SNARK. Сучасні DSL класифікуються як мови опису апаратного забезпечення (HDL, hardware description languages) [11] або мови програмування (PL, programming languages) [12].

Мови високого рівня (HDL) описують синтез схем безпосередньо у провідній формі, забезпечуючи елегантний синтаксис, але створюючи труднощі для програмістів через їхню незалежну провідну структуру й обмежену абстракцію типів даних, оскільки вхідні дані представлені як структури сигнальних даних. На противагу цьому мови програмування (PL) визначають обмеження мовами програмування високого рівня, підтримуючи різні типи даних і нагадуючи такі мови, як Rust або Python. Це робить PL більш доступними для програмістів без знань провідних схем, пропонуючи найпростіший спосіб визначення обмежень. Однак гнучкий синтаксис PL збільшує ризики вразливості та створює проблеми з ефективністю.

Вбудовані доменно-орієнтовані мови (eDSL). eDSL для zk-SNARK набули популярності в останні роки та реалізуються як функції в мовах програмування загального призначення, що відрізняє їх від традиційних компіляторів у контексті мов програмування. eDSL призначені для опису синтезу схем, подібно до HDL, але вони орієнтовані на провідні схеми (структуроване представлення арифметичних схем, де схема описується як множина проводів (wires) і гейтів (gates), що з'єднують ці проводи),

ISSN 2786-9024

ТЗ. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344513

пропонуючи більшу виразність та простоту використання, успадковуючи структури даних і функції програмування від вбудованої мови. Ці eDSL оптимізують розробку ZK-доказів, інтегруючи визначення схеми та генерацію доказів в один файл, спрощуючи код і даючи можливість програмістам використовувати існуючі функції бібліотеки. Однак написання коду в eDSL вимагає від розробників чіткого розрізнення внутрішньосхемних і зовнішніх операцій, що потребує експертних знань конкретної мови та дизайну бібліотеки.

Віртуальні машини з нульовим розкриттям знань (zk-VMs) обробляють операційний код циклу «вибірка – декодування – виконання», реплікуючи трасування обчислень для загальних програм (зазвичай смарт-контрактів) та генеруючи відповідні ZK-підтвердження. Вони підтримують різні архітектури наборів інструкцій (ISA, instruction set architectures). zk-VM сумісні з існуючими мовами програмування високого рівня та можуть використовувати функції існуючих компіляторів. Однак, незважаючи на орієнтацію на низькорівневі операційні коди, zk-VM не повністю сумісні з програмами верхнього рівня та часто потребують незначних або значних модифікацій програми, що може ускладнити керування програмістами через схильність до помилок. Крім того, zk-VM використовують обчислювальну модель машини Тюрінга замість схем, що значно збільшує накладні витрати. Хоча zk-VM зменшують навантаження на програмістів, пов’язане з обмеженнями написання, вони можуть мати проблеми з ефективністю, особливо для великомасштабних програм.

Оцінку сумісності компіляторів варто проводити за двома властивостями:

– **перехресна сумісність** – вказує на те, чи може результат компіляції одного компілятора бути використаний іншим. Компілятори DSL пропонують помірну перехресну сумісність, оскільки вони розділяють систему обмежень і систему доведення. Зі стандартизацією результатів компіляції в майбутньому бібліотеки зможуть зосередитися лише на наданні систем доведення, приймаючи результати DSL як вхідні дані. Компілятори eDSL мають низьку перехресну сумісність, оскільки вони визначають схему в мові програмування, що ускладнює використання їхньої визначеної схеми на платформах з іншими мовами. Навіть однією мовою результат компіляції може бути несумісним, оскільки функції гаджетів відрізняються. Віртуальні машини zk мають низьку перехресну сумісність, оскільки вони розроблені лише для деяких специфічних програм високого рівня;

– **синтаксична сумісність** – вказує на те, чи має мова введення компілятора подібний синтаксис до іншої мови. Сумісність синтаксису важлива, оскільки дає змогу програмісту, знайомому з мовою, перейти до іншої без ретельного вивчення. На жаль, ми виявляємо, що навіть в одній категорії мови компілятора мають зовсім інший синтаксис, тож буде важко вивчити їх усі. У DSL HDL є мовою апаратних схем, тоді як PL більше схожа на загальну мову програмування. У eDSL синтаксис залежить від базової мови бібліотеки, починаючи від C, C++, Rust, Go та JavaScript. У zk-VM добре підтримуються лише операції з мов смарт-контрактів, а операції з інших загальних мов не пройдуть компіляцію.

Обговорення результатів

Основною проблемою zk-SNARK є компроміс між ефективністю та безпекою – лінійний PCP досягає постійного розміру доказу, але завдяки налаштуванню довіри.

Керівними принципами вибору відповідної системи доведення є (1) визначення прийнятності налаштувань довіри; (2) відповідна масштабованість; (3) необхідність постквантової безпеки.

Загалом найбільшою перешкодою у використанні бібліотек zk-SNARK є брак документації. Тому варто створювати й уніфікувати динамічні документи користувача, деталі API гаджетів в eDSL та синтаксис мови в DSL.

Стандартизація може допомогти розробникам порівняти важливі функції в полі zk-SNARK різних бібліотек, а також встановити більш узгоджену базову лінію продуктивності. Документація бібліотеки щодо цих основних функцій є неявною, і розробникам потрібно розуміти основні криптографічні методи, щоб вибрати відповідну схему. Важливою є стандартизація параметрів компіляторів, що ускладнює повторне використання існуючих інструментів.

Залишається відкритим питання, чи системи доказування для конкретних сценаріїв працюватимуть краще, ніж загальні системи доказування. Розробка таких специфічних систем доказування потребує співпраці між теоретичними розробками та інженерією.

Реалізації схем zk-SNARK обмежені в різних мовах програмування та не містять жодних інтерфейсів для сумісності. Хоча такі компоненти, як системи обмежень і системи доведення, можна розділити в коді, їхні функції та інструменти обмежені відповідними бібліотеками. Однак відповідний вибір потребує експертних знань систем обмежень, що непрактично для програмістів.

Багато завдань вимагають глибокого знайомства як з мовою програмування, так і з системою обмежень. У багатьох бібліотеках відсутній компілятор для перетворення високорівневого коду на схемні представлення, що змушує програмістів вручну додавати обмеження. На схемному рівні програмісти повинні

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344513

обробляти складні деталі, як-от операції з кривими, цикли та перестановки. Вибір невідповідної кривої також може знизити ефективність чи створити вразливості.

Висновки

Отже, відсутність універсальної стандартизації є проблемою, пов'язаною із сумісністю. В одній категорії існують значні відмінності в синтаксисі, що ускладнює міграцію проєктів і заплутує програмістів, а також навіть для тієї самої схеми результат компіляції не може бути використаний системою перевірки в іншій бібліотеці. Основним напрямом у розробці сучасних систем доведення є IP, який може усунути необхідність налаштування довіри, довгого спільного рядка посилання (CRS) та повільного процесу доведення в zk-SNARK на основі QAP.

Варто розробити систему заходів для вирішення або пом'якшення проблем із мовою та сумісністю через створення образів. Проблему неправильного використання схем і кривих можна вирішити розробкою вичерпних рекомендацій. Проблеми, пов'язані з компіляторами, вирішуються класифікацією існуючих компіляторів у кожній бібліотеці та ґрунтовним аналізом їхніх сильних і слабких сторін. Доречно задокументувати розроблені матеріали з відкритим кодом (API, Application Programming Interface), що забезпечують прозорість, взаємодію та інновації, а саме документацію користувача (встановлення, використання та тестування) і зразки коду для програм.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно цього дослідження, у тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

ЛІТЕРАТУРА

[1] Alan Szepieniec and Yuncong Zhang. Polynomial iops for linear algebra relations. In International Conference on Public-Key Cryptography (PKC), pages 523–552. Springer, 2022. DOI: 10.1007/978-3-030-97121-2_19.

[2] Aniket Kate, Gregory M. Zaverucha, and Ian Goldberg. Constant-size commitments to polynomials and their applications. In Advances in Cryptology – ASIACRYPT 2010, pages 177–194, Berlin, Heidelberg, 2010. DOI: 10.1007/978-3-642-17373-8_11.

[3] Benedikt Bünz, Jonathan Bootle, Dan Boneh, Andrew Poelstra, Pieter Wuille, and Greg Maxwell. Bulletproofs: Short proofs for confidentia ltransaction sand more. In 2018 IEEE symposium on security and privacy (SP), pages 315–334, 2018. DOI: 10.1109/SP.2018.00020.

[4] Benoit Libert. Simulation-extractable kzg polynomial commitments and applications to hyperplonk. In International Conference on Public-Key Cryptography (PKC), pages 68–98. Springer, 2024. DOI: 10.1007/978-3-031-57722-2_3.

[5] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Fast reed-solomon interactive oracle proofs of proximity. In 45th international colloquium on automata, languages, and programming (icalp 2018), 2018. DOI: 10.4230/LIPIcs.ICALP.2018.14.

[6] Helger Lipmaa. A unifie dframewor kfo rnon-universal snarks. In International Conference on Public-Key Cryptography (PKC), pages 553–583. Springer, 2022. DOI: 10.1007/978-3-030-97121-2_20.

[7] Jens Groth, Markulf Kohlweiss, Mary Maller, Sarah Meiklejohn, and Ian Miers. Updatable and universal common reference strings with applications to zk-snarks. In Annual International Cryptology Conference, pages 698–728, 2018. DOI: 10.1007/978-3-319-96878-0_24.

[8] Jiaheng Zhang, Tianyi Liu, Weijie Wang, Yinuo Zhang, Dawn Song, Xiang Xie, and Yupeng Zhang. Doubly efficien tinteractiv eproof sfo rgenera larithmetic circuits with linear prover time. In ACM SIGSAC Conference on Computer and Communications Security, pages 159–177, 2021. DOI: 10.1145/3460120.3484767.

[9] Jonathan Bootle, Alessandro Chiesa, and Jens Groth. Linear-time arguments with sublinear verification from tensor codes. In Theory of Cryptography (TCC), pages 19–46. Springer, 2020. DOI: 10.1007/978-3-030-64378-2_2.

[10] Liang J., Hu D., Wu P., Yang Y., Shen Q., & Wu Z. (2025). SoK: Understanding zk-SNARKs: The gap between research and practice. *arXiv*. DOI: 10.48550/arXiv.2502.02387.

[11] Marta Bellés-Muñoz, Miguel Isabel, Jose Luis Muñoz-Tapia, Albert Rubio, and Jordi Baylina. Circom: A circuit description language for building zero-knowledge applications. IEEE Transactions on Dependable and Secure Computing, vol. 20 (6), pp. 4733–4751, 2022.

[12] Nada Amin, John Burnham, François Garillot, Rosario Gennaro, Daniel Rogozin, Cameron Wong, et al. Lurk: Lambda, the ultimate recursive knowledge. Cryptology ePrint Archive, 2023.

[13] Rosario Gennaro, Craig Gentry, Bryan Parno, and Mariana Raykova. Quadratic span programs and succinct nizks without pcps. In International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT), pages 626–645, 2013. DOI: 10.1007/978-3-642-38348-9_29.

[14] Scott Ames, Carmit Hazay, Yuval Ishai, and Muthuramakrishnan Venkitasubramaniam. Ligero:

Lightweight sublinear arguments without a trusted setup. In ACM SIGSAC Conference on Computer and Communications Security, pages 2087–2104, 2017. DOI: 10.1145/3133956.3134104.

[15] Sean Bowe, Ariel Gabizon, and Ian Miers. Scalable multi-party computation for zk-snark parameters in the random beacon model. Cryptology ePrint Archive, 2017.

[16] Shafi Goldwasser, Yael Tauman Kalai, and Guy Rothblum. Delegating computation: interactive proofs for muggles. Journal of the ACM (JACM), vol. 62 (4), pp. 1–64, 2008. DOI: 10.1145/1391289.1391296.

[17] Stephen Wicker and Vijay Bhargava. Reed-Solomon codes and their applications. John Wiley & Sons, 1999.

CLASSIFICATION OF NON-INTERACTIVE KNOWLEDGE ARGUMENT PROOF SYSTEMS

Yurii Paslavskyi, Ihor Kroshnyi

An important cryptographic mechanism that guarantees confidentiality (the zero-disclosure property) and ensures that it is impossible to prove a false statement to the verifier is zero-disclosure proofs. A popular implementation of zero-disclosure proofs is short, non-interactive proofs that can be quickly verified and that do not require interaction between the parties after the initial setup. The main direction in the development of modern proof systems is interactive proof, which is built in two steps. The first is sending a confirmation of the polynomial of an interactive oracle proof and the second is creating correct oracles of the polynomial commitment scheme using well-defined cryptographic methods for evaluating polynomials. Verifying the use of the same coefficients in each linear combination requires checking both polynomial consistency and variable consistency. To construct general schemes of concise non-interactive zero-disclosure knowledge argument, an interactive oracle proof polynomial was proposed that models messages as polynomial oracles. All tests are proved using polynomial commitment schemes and then evaluated with zero knowledge at a point specified by the person verifying the information. The reliability and confidentiality of all tests are based on three main categories of interactive oracle proof polynomials, namely polynomial commitment schemes with conjunction, with inner product argument and with code theory. The protocols of concise non-interactive zero-disclosure knowledge arguments are implemented through high-level programs (compilers), which are converted into an intermediate representation, i.e. a scheme defined by a system of constraints. The compilers used are divided into domain-oriented languages, embedded domain-oriented languages, and zero-knowledge virtual machines. Specialized domain-oriented hardware description languages or programming languages offer an adapted syntax for efficiently expressing constraints in arithmetic schemes. Embedded domain-oriented languages are implemented as functions in general-purpose programming languages and are oriented to the overhead schemes inherited from the embedded language. Zero-knowledge virtual machines process the opcode of the fetch-decode-execute cycle, replicating the computation trace for general programs and generating corresponding zero-knowledge proofs. They are compatible with existing high-level programming languages and can use the features of existing compilers. Compilers are evaluated for cross- or syntactic compatibility. In general, the biggest obstacle to using non-interactive proof libraries is the lack of documentation. Standardization can help developers compare important features across libraries and establish a more consistent performance baseline. Library documentation for these core features is implicit, and developers need to understand the underlying cryptographic techniques to choose an appropriate scheme. Standardization of compiler options is important, making it difficult to reuse existing tools.

Keywords: *zero-knowledge proof, interactive proof, polynomial oracle proof, polynomial commitment schemes, compilers, domain-specific language, virtual machine, standardization.*

REFERENCES

[1] Alan Szepieniec and Yuncong Zhang. Polynomial iops for linear algebra relationsIn International Conference on Public-Key Cryptography (PKC), pages 523–552. Springer. 2022. DOI: 10.1007/978-3-030-97121-2_19.

[2] Aniket Kate, Gregory M. Zaverucha, and Ian Goldberg. Constant-size commitments to polynomials and their applications. In Advances in Cryptology – ASIACRYPT 2010, pages 177–194, Berlin, Heidelberg. 2010. DOI: 10.1007/978-3-642-17373-8_11.

[3] Benedikt Bünz, Jonathan Bootle, Dan Boneh, Andrew Poelstra, Pieter Wuille, and Greg Maxwell. Bulletproofs: Short proofs for confidentia ltransaction sand more. – In 2018 IEEE symposium on security and privacy (SP), pages 315–334, 2018. DOI: 10.1109/SP.2018.00020.

[4] Benoit Libert. Simulation-extractable kzg polynomial commitments and applications to hyperplonk. In International Conference on Public-Key Cryptography (PKC), pages 68–98. Springer, 2024. DOI: 10.1007/978-3-031-57722-2_3.

[5] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. (2018). Fast reed-solomon interactive oracle proofs of proximity. In 45th international colloquium on automata, languages, and programming (icalp 2018). DOI: 10.4230/LIPIcs.ICALP.2018.14.

[6] Helger Lipmaa. A unifie dframewor kfo rnon-universal snarks. In International Conference on Public-Key Cryptography (PKC), pages 553–583. Springer, 2022. DOI: 10.1007/978-3-030-97121-2_20.

[7] Jens Groth, Markulf Kohlweiss, Mary Maller, Sarah Meiklejohn, and Ian Miers. Updatable and

universal common reference strings with applications to zk-snarks. In Annual International Cryptology Conference, pages 698–728, 2018. DOI: 10.1007/978-3-319-96878-0_24.

[8] Jiaheng Zhang, Tianyi Liu, Weijie Wang, Yinuo Zhang, Dawn Song, Xiang Xie, and Yupeng Zhang. Doubly efficien tinteractiv eproof sfo rgenera larithmetic circuits with linear prover time. In ACM SIGSAC Conference on Computer and Communications Security, pages 159–177, 2021. DOI: 10.1145/3460120.3484767.

[9] Jonathan Bootle, Alessandro Chiesa, and Jens Groth. Linear-time arguments with sublinear verification from tensor codes. In Theory of Cryptography (TCC), pages 19–46. Springer, 2020. DOI: 10.1007/978-3-030-64378-2_2.

[10] Liang J., Hu D., Wu P., Yang Y., Shen Q., & Wu Z. SoK: Understanding zk-SNARKs: The gap between research and practice. *arXiv*. 2025. DOI: 10.48550/arXiv.2502.02387.

[11] Marta Bellés-Muñoz, Miguel Isabel, Jose Luis Muñoz-Tapia, Albert Rubio, and Jordi Baylina. Circom: A circuit description language for building zero-knowledge applications. IEEE Transactions on Dependable and Secure Computing, vol. 20 (6), pp. 4733–4751, 2022.

[12] Nada Amin, John Burnham, François Garillot, Rosario Gennaro, Daniel Rogozin, Cameron Wong, et al. Lurk: Lambda, the ultimate recursive knowledge. Cryptology ePrint Archive. 2023.

[13] Rosario Gennaro, Craig Gentry, Bryan Parno, and Mariana Raykova. Quadratic span programs and succinct nizks without pcps. In International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT), pages 626–645. 2013. DOI: 10.1007/978-3-642-38348-9_29.

[14] Scott Ames, Carmit Hazay, Yuval Ishai, and Muthuramakrishnan Venkitasubramaniam. Ligero: Lightweight sublinear arguments without a trusted setup. In ACM SIGSAC Conference on Computer and Communications Security, pages 2087–2104. 2017. DOI: 10.1145/3133956.3134104.

[15] Sean Bowe, Ariel Gabizon, and Ian Miers. Scalable multi-party computation for zk-snark parameters in the random beacon model. Cryptology ePrint Archive. 2017.

[16] Shafi Goldwasser, Yael Tauman Kalai, and Guy Rothblum. Delegating computation: interactive proofs for muggles. Journal of the ACM (JACM), vol. 62 (4), pp. 1–64, 2008. DOI: 10.1145/1391289.1391296.

[17] Stephen Wicker and Vijay BhargavaReed-Solomon codes and their applications. John Wiley & Sons. 1999.

Стаття надійшла до редакції 16.09.2025
Стаття прийнята 04.10.2025
Статтю опубліковано 02.12.2025

ТЗ. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

УДК 004.8:004.67: 004.8

ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ АНАЛІЗУ, КЛАСИФІКАЦІЇ ТА РЕКОМЕНДАЦІЙ У СИСТЕМАХ УПРАВЛІННЯ КОЛЕКЦІЯМИ

Д.І. Поперешняк[1], С.В. Поперешняк[1, 2]

[1] Department of Software Engineering, State University of Information and Communication Technologies, Kyiv, Ukraine

[2] Department of Informatics and Software Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

E-mail: spopereshnyak@gmail.com

АНОТАЦІЯ

Статтю присвячено дослідженню та розробці інтелектуальних технологій аналізу, класифікації та рекомендацій у системах управління колекціями. У роботі розглядається проблема організації та інтелектуальної обробки колекційних даних, які відзначаються постійним зростанням обсягів і різнорідністю, що ускладнює їх ефективне збереження, пошук та систематизацію. Показано, що традиційні системи управління здебільшого зосереджуються на ізольованих завданнях, як-от базове збереження чи пошук інформації, залишаючи поза увагою питання комплексної інтеграції алгоритмів аналізу подібності, класифікації та рекомендацій. Використання інтелектуальних підходів дає змогу подолати ці обмеження та сформувати універсальні системи, придатні для роботи з великими обсягами даних.

Метою дослідження є досягнення підвищеної ефективності та масштабованості управління колекціями шляхом побудови моделі інтелектуальної системи, яка інтегрує формалізацію мультимодальних ознак об’єктів, методи класифікації та виявлення дублікатів, а також рекомендаційні механізми. Для досягнення цієї мети проаналізовано існуючі підходи до роботи з колекціями, сформульовано модель представлення об’єктів, розроблено методику обчислення подібності та запропоновано архітектуру, що містить класифікаційні й рекомендаційні модулі.

Експериментальна перевірка показала, що інтеграція зазначених компонентів забезпечує високу точність класифікації, ефективне виявлення дублікатів і релевантність персоналізованих рекомендацій. Особливу увагу приділено масштабованості: система зберігає стабільний час відповіді навіть за умови значного збільшення кількості об’єктів. Практична цінність полягає в універсальності моделі, яка може застосовуватися як у приватних цифрових і мультимедійних колекціях, так і в корпоративних чи наукових інфраструктурах. Перспективи подальших досліджень пов’язані з упровадженням глибинного навчання для обробки мультимодальних ознак, удосконаленням рекомендаційних алгоритмів і посиленням інтеграції з механізмами безпеки та контролю доступу.

Ключові слова: *інтелектуальні системи, управління колекціями, аналіз подібності, класифікація, рекомендаційні системи, мультимодальні дані, масштабованість, IoT.*

Вступ

У сучасному інформаційному суспільстві обсяг цифрових даних зростає надзвичайно швидкими темпами, охоплюючи найрізноманітніші сфери – від приватних мультимедійних архівів до масштабних корпоративних і наукових сховищ. Це вже не лише бібліотеки електронних книг чи фотоархіви, які зберігають користувачі вдома. Ідеться про корпоративні бази, наукові сховища, мультимедійні колекції, що налічують мільйони елементів. Одночасно із цим відбувається ускладнення структури колекцій, які містять числові, категоріальні, текстові та мультимедійні дані. Чим більш різнорідними стають ці дані, тим складніше організувати їх зберігання та забезпечити швидкий доступ.

Більшість традиційних систем управління колекціями справді добре виконують базові функції: зберігають інформацію, дають змогу здійснити пошук. Однак, коли йдеться про масштабування чи інтеграцію з іншими середовищами, з’являються проблеми. Одна зі слабких сторін – практично повна відсутність інтелектуальних механізмів аналізу та рекомендацій. Окремим джерелом поповнення колекцій є потоки даних від IoT-пристроїв (сенсорика, мультимедійні

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

фіди), що додатково ускладнює гетерогенність та вимоги до масштабованості.

Останні роки показали, що вирішити цю проблему можна шляхом використання інтелектуальних технологій. Методи аналізу даних, алгоритми класифікації та рекомендаційні підходи надають можливість автоматизувати роботу з колекціями, позбавитися зайвих дублікатів і створювати персоналізовані пропозиції для користувачів. Особливо ефективними такі підходи є у хмарних середовищах, де важливі масштабованість і доступність сервісів.

Таким чином, використання інтелектуальних методів у системах управління колекціями є одним із найперспективніших напрямів у сфері інформаційних технологій. Це питання має і практичний вимір, і теоретичну вагу, адже йдеться про створення нових моделей, здатних поєднати різні аспекти обробки даних у єдину архітектуру.

Аналіз літературних даних і постановка проблеми

Сучасні системи управління колекціями стикаються з вибуховим зростанням обсягів та різнорідності даних, тож поряд із базовим збереженням і пошуком на перший план виходять автоматизована класифікація, інтелектуальний аналіз і рекомендації. Оптимізація зберігання передусім зводиться до дедуплікації: огляди показують, що навіть помірний рівень повторів суттєво погіршує ефективність сховищ, а підходи chunk-based знижують витрати пам'яті та прискорюють доступ [1; 2]. Паралельно досліджуються моделі безпеки, зокрема когнітивні парадигми, що поєднують криптографію та семантику для підвищення довіри до даних [3]. У рекомендаційних системах найкращі результати демонструють гібридні схеми, які поєднують контентний і колаборативний підходи, однак більшість рішень лишається домен-специфічною і погано переноситься на гетерогенні колекції [4]. Узагальнюючи, література окреслює три напрями: (i) дедуплікація та оптимізація зберігання, (ii) безпека даних, (iii) персоналізація через рекомендації. Проте ці лінії розвиваються переважно ізольовано. Звідси випливає потреба в інтегрованій моделі, що одночасно охоплює аналіз подібності, класифікацію, виявлення дублікатів і рекомендації у єдиній архітектурі.

Мета та задачі дослідження

Метою дослідження є підвищення ефективності управління колекціями різнорідних даних завдяки інтеграції інтелектуальних технологій аналізу, класифікації та рекомендацій у єдину модель. Мета роботи – реалізувати й експериментально оцінити інтегровану модель інтелектуальної системи управління колекціями (аналіз подібності, багатоміткова класифікація, дедуплікація, рекомендації), показавши її точність і масштабованість на реальних та синтетичних наборах даних.

Для досягнення поставленої мети передбачено розв'язання таких завдань:

1. Проаналізувати наявні підходи до організації, класифікації та інтелектуальної обробки колекційних даних, окресливши їхні переваги й обмеження.
2. Формалізувати модель об'єктів з урахуванням мультимодальних ознак (числових, категоріальних, текстових, візуальних).
3. Розробити методи подібності та класифікації і механізм дедуплікації для гетерогенних колекцій.
4. Інтегрувати рекомендаційну підсистему в архітектуру управління колекціями.
5. Провести експериментальну верифікацію за стандартними метриками й оцінити практичну придатність.

Матеріали та методи досліджень

Методи обчислення подібності та класифікації у гетерогенних колекціях. Ефективне управління колекціями неможливе без механізмів, які дають змогу оцінювати ступінь подібності між об'єктами та відносити їх до певних категорій. У випадку гетерогенних колекцій складність полягає у тому, що об'єкти можуть поєднувати числові, категоріальні, текстові та візуальні характеристики, кожна з яких потребує власної метрики.

Обчислення подібності. Загальний підхід полягає у побудові композитної метрики, яка комбінує часткові відстані:

$$d\left(c_i, c_j\right) = \sum_{k=1}^{m} w_k \cdot d_k\left(a_{ik}, a_{jk}\right),$$

де d_k – локальна метрика для ознаки певного типу, а w_k – ваговий коефіцієнт її значущості.

Приклади локальних метрик:

- числові дані: нормована різниця (наприклад, Manhattan або Euclidean distance);
- категоріальні дані: бінарна функція збігу (0 – різні, 1 – однакові);
- текстові дані: косинусна близькість у векторних просторах (Word2Vec, BERT) [5];
- візуальні дані: відстань у просторах ознак, отриманих згортковими нейронними мережами [6].

Такий підхід дає змогу порівнювати об'єкти не за одним критерієм, а комплексно. Наприклад, два фотоапарати можуть відрізнятися роком випуску, але бути схожими за категорією і текстовим описом.

Класифікація у колекціях. Класифікація ускладнюється тим, що об'єкт може одночасно належати до кількох категорій (багатоміткове віднесення). Формально завдання можна подати як відображення:

$$f : C \to 2^K,$$

де K – множина можливих категорій.

ТЗ. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

Для цього застосовуються різні підходи:

– традиційні методи: логістична регресія, SVM, дерева рішень;

– ансамблеві алгоритми: Random Forest, Gradient Boosting, які показують стабільність на даних середнього розміру;

– глибинні моделі: трансформери (BERT, RoBERTa) для текстових ознак та CNN/ResNet для візуальних [7];

– гібридні підходи: multi-view learning, де кожна модальність обробляється окремою моделлю, а результати комбінуються [6].

З практичного погляду доцільно застосовувати ієрархічну класифікацію, коли категорії організовані у вигляді дерева. Це особливо актуально для бібліотечних чи музейних колекцій, де категорії мають багаторівневу структуру.

Інтеграція рекомендаційної підсистеми в загальну архітектуру управління колекціями. Однією з ключових особливостей сучасних інтелектуальних систем управління даними є здатність не лише зберігати та структурувати інформацію, але й активно підтримувати користувача у процесі роботи з колекціями. Це досягається завдяки інтеграції рекомендаційних підсистем, які автоматично пропонують нові об’єкти, найбільш релевантні інтересам чи поточному контексту користувача.

У запропонованій системі рекомендаційна підсистема є надбудовою над модулями обчислення подібності та класифікації. Вона отримує на вхід:

– векторні подання об’єктів колекції;

– інформацію про класи та теги;

– історію взаємодії користувачів із системою (пошуки, перегляди, додавання об’єктів).

Рекомендаційний модуль інтегрується в загальну архітектуру як окремий сервіс, який працює у двох режимах:

1) контентно-орієнтований – пропонує об’єкти, схожі на вже наявні в колекції (наприклад, подібні книги чи альбоми);

2) колаборативний – враховує поведінку інших користувачів, які мають подібні інтереси.

Узагальнену архітектуру запропонованої системи наведено на рисунку 1. Вона побудована за модульним принципом і охоплює основні етапи роботи з колекційними даними – від моменту їх надходження в систему до формування аналітичних звітів і персоналізованих рекомендацій.

У верхньому рівні архітектури функціонує конвеєр обробки даних. На першому етапі здійснюється імпорт об’єктів колекції з різних джерел, включно з локальними сховищами та зовнішніми сервісами. Далі відбувається попередня обробка, яка передбачає нормалізацію, кодування та побудову мультимодальних векторних подань. Ці подання використовуються для виконання трьох ключових завдань: класифікації, пошуку дублікатів та обчислення індексів подібності.

Нижній рівень архітектури відображає поведінковий контур. Тут накопичуються логи взаємодії користувачів із системою, які надходять у рекомендаційний модуль. Отримані результати не лише формують персоналізовані пропозиції для користувача, а й повертаються у систему як зворотний зв’язок, що підвищує точність класифікації та якість виявлення дублікатів.

Завершальним елементом архітектури є модуль аналітики та візуалізації, що об’єднує вихідні дані

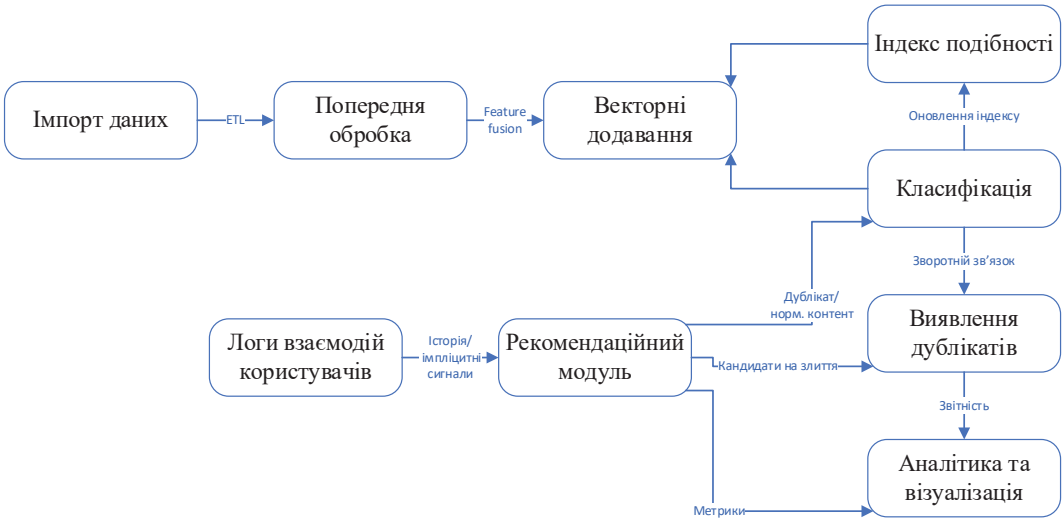

Рис. 1. Узагальнена схема взаємодії модулів у системах управління колекціями

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

з усіх підсистем і забезпечує інтерактивний доступ до статистичних звітів, графіків та діаграм.

Запропонована система складається з кількох послідовно пов'язаних модулів:

– вхідні дані (колекції з мультимодальними ознаками);
– попередня обробка (нормалізація, тегування, формування векторних подань);
– аналіз подібності та класифікація;
– виявлення дублікатів;
– рекомендаційна підсистема;
– аналітика та візуалізація результатів.

Взаємозв'язки між цими компонентами наведено на рисунку 1. Схема відображає логіку обробки даних: від введення об'єктів до отримання кінцевих результатів у вигляді рекомендацій і аналітичних звітів.

Порівняння методів побудови рекомендацій. Для підвищення наочності проведено порівняльний аналіз основних характеристик двох підходів до рекомендаційних систем. У таблиці 1 наведено їхні принципи роботи, переваги, недоліки й оптимальні умови застосування.

Аналіз таблиці показує, що інтеграція обох підходів є перспективним напрямом, оскільки дає змогу компенсувати недоліки кожного з них.

Результати досліджень

Для перевірки ефективності запропонованої моделі інтелектуальної системи управління колекціями було проведено серію експериментів, спрямованих на оцінку якості класифікації, ефективності виявлення дублікатів, релевантності рекомендацій та масштабованості архітектури.

Дослідження базувалося на аналізі як реальних, так і синтетично згенерованих колекційних даних. Реальні набори містили приватні зібрання мультимедійного контенту – музичні альбоми, книги та зображення фототехніки, завдяки чому вдалося перевірити роботу моделі на практичних прикладах. Для розширення експериментальної бази було також сформовано штучні вибірки обсягом до 50 тисяч об'єктів. У них передбачалося навмисне дублювання частини елементів (близько 10 %), що дало можливість протестувати механізми виявлення надлишкових записів. Усі дані зберігалися у структурованій базі PostgreSQL з додатковим використанням мультимедійних сховищ для великих файлів.

Взаємодія модулів системи. На рисунку 2 наведено граф взаємодії модулів у системі управління колекціями. На відміну від узагальненої схеми (рис. 1), що відображає послідовність етапів обробки, цей граф ілюструє логіку інформаційних потоків і зворотних зв'язків між компонентами системи.

Ключові вузли відповідають функціональним блокам: імпорту даних, попередньої обробки, векторних подань, класифікації, виявлення дублікатів, індексації подібності, рекомендаційної підсистеми й аналітики. Стрілки показують напрям передачі інформації, підкреслюючи, що результати роботи одного модуля можуть бути вхідними даними для кількох інших.

Особливу роль відіграє рекомендаційний модуль, який одночасно отримує дані з логів користувачів і взаємодіє з іншими компонентами системи. Його вихідні результати можуть підвищувати точність класифікації, оптимізувати пошук дублікатів і впливати на формування індексів подібності. Така організація забезпечує інтеграцію основних функцій системи і створює основу для реалізації адаптивних механізмів, здатних навчатися на основі накопиченого досвіду.

Таким чином, архітектура має не лише аналітичний, а й адаптивний характер, забезпечуючи навчання на основі історії використання.

Порівняння метрик подібності. У дослідженні використано дві групи колекційних даних:

1) тематичні набори мультимедіа (цифрові книги, музичні записи, зображення фототехніки) – для апробації методів у реальних сценаріях;

2) синтетично створені набори обсягом від 1 тис. до 20 тис. елементів – для перевірки масштабованості й тестування алгоритмів класифікації та рекомендацій.

Табл. 1. Порівняння підходів до побудови рекомендацій у системах управління колекціями

Критерій	Контентно-орієнтовані методи	Колаборативна фільтрація
Принцип роботи	Аналізують характеристики самих об'єктів (атрибути, описи, теги)	Використовують дані про взаємодії користувачів (рейтинги, перегляди, уподобання)
Переваги	Працюють із новими об'єктами; не потребують історії взаємодій	Виявляють приховані закономірності та групи користувачів; добре масштабується
Недоліки	Залежать від якості опису даних; ігнорують соціальні зв'язки	Проблема холодного старту; потребують великих масивів даних
Оптимальні умови	Невеликі та середні колекції з багатими атрибутами	Великі колекції з активними користувачами та багатою історією взаємодій

ISSN 2786-9024

ТЗ. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

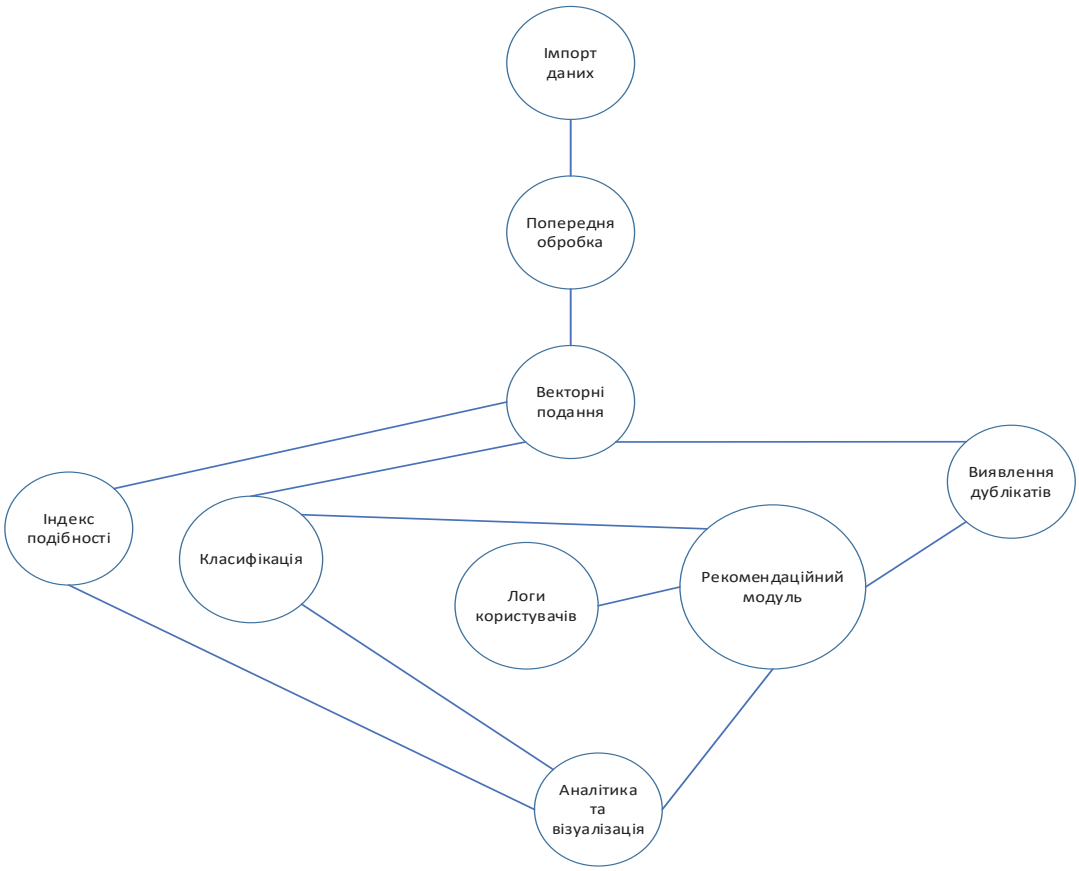

Рис. 2. Граф взаємодії модулів у системі управління колекціями

На відміну від попередніх робіт, акцент зроблено не лише на моделюванні архітектури, а й на порівнянні алгоритмів аналізу та класифікації. Об’єкти колекцій описувалися чотирма типами ознак: числовими, категоріальними, текстовими та візуальними. Формалізацію здійснено у вигляді матриці ознак:

$$X = \{x_i\}_{i=1}^{N}, \; x_i = (a_{i1}, a_{i2}, \ldots, a_{im}),$$

де N – кількість об’єктів, m – число характеристик.

Для текстових атрибутів застосовувалися два методи: TF–IDF та контекстні вектори BERT. Це дало можливість порівняти «класичні» й «сучасні» підходи. Для візуальних ознак використовувалися дескриптори SIFT та ознаки, отримані згортковими мережами (ResNet).

Оцінювання подібності проводилося через кілька метрик: косинусну близькість, евклідову відстань і Джаккарда. Для кожного набору даних обчислювали середню точність класифікації залежно від вибору метрики. Результати зведено в таблицю 2.

Класифікація виконувалася методами логістичної регресії, випадкових лісів і нейронних мереж. Було виявлено, що для малих наборів прості алгоритми (логістична регресія) демонструють достатню ефективність, тоді як на великих даних суттєво переважають багатошарові нейронні мережі.

Окремо досліджувався механізм виявлення дублікатів. Запропоновано підхід, що комбінує кластеризацію на основі DBSCAN з пороговою фільтрацією подібності. Це дає змогу не лише знаходити ідентичні записи, а й об’єднувати «схожі» об’єкти (наприклад, дві різні редакції однієї книги).

Табл. 2. Порівняння ефективності різних метрик подібності

Тип даних	Косинусна міра	Евклідова відстань	Джаккарда
Текстові описи	0,88	0,74	0,69
Зображення	0,82	0,85	0,61
Змішані набори	0,86	0,80	0,65

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

Рекомендаційний модуль інтегровано в загальну архітектуру. Для перевірки ефективності порівнювались контентно-орієнтовані методи (на основі ознак об'єктів) і колаборативна фільтрація (ALS). Результати показали, що на невеликих колекціях переважають контентні методи, а на великих – колаборативні.

Порівняння методів побудови рекомендацій. Окремим аспектом дослідження стало порівняння ефективності різних підходів до побудови рекомендаційних підсистем. Це важливо, оскільки якість персоналізованих пропозицій безпосередньо впливає на цінність системи для кінцевого користувача.

На практиці застосовують два основні підходи: контентно-орієнтовані методи, що використовують атрибути самих об'єктів (тексти, зображення, метадані), та колаборативну фільтрацію, яка враховує історію взаємодій користувачів.

На рисунку 3 показано динаміку точності рекомендацій залежно від розміру колекції. Видно, що контентно-орієнтовані методи зберігають високу якість на малих наборах даних, проте поступово втрачають ефективність із масштабуванням. Натомість колаборативна фільтрація демонструє зростання точності на великих колекціях, що зумовлено використанням прихованих закономірностей у поведінці користувачів.

У таблиці 3 наведено результати порівняння традиційного підходу k-ближчих сусідів, методів на основі матричної факторизації та сучасних моделей на базі нейронних мереж.

Як видно з таблиці, метод k-ближчих сусідів добре підходить для невеликих колекцій або прототипів, тоді як матрична факторизація забезпечує кращу якість у середніх за обсягом системах. Нейронні мережі, зі свого боку, демонструють найбільший потенціал для масштабних колекцій із високою різнорідністю, проте їх використання потребує значних обчислювальних ресурсів.

Таким чином, вибір конкретного підходу до побудови рекомендаційної підсистеми залежить від

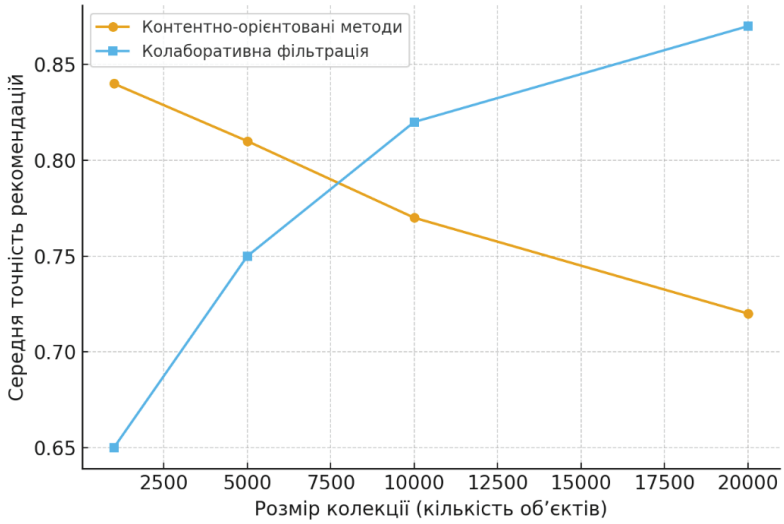

Рис. 3. Порівняння точності методів рекомендацій залежно від розміру колекції

Табл. 3. Порівняльна характеристика методів побудови рекомендаційних систем

Метод	Переваги	Недоліки	Галузі застосування
k-ближчих сусідів	Простота реалізації; швидке навчання	Низька масштабованість; проблеми з рідкісними даними	Невеликі колекції; демонстраційні проекти
Матрична факторизація	Висока якість на середніх даних; врахування прихованих залежностей	Вимогливість до апаратних ресурсів; чутливість до холодного старту	Мультимедійні колекції; інтернет-магазини
Нейронні мережі	Здатність моделювати складні зв'язки; хороша масштабованість	Висока складність навчання; потреба у великих наборах даних	Великі гетерогенні колекції; наукові репозиторії

Т3. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

розміру колекції, доступних ресурсів та вимог до точності рекомендацій.

Аналіз продуктивності системи. Поряд із логічною архітектурою та графом взаємодії важливо оцінити і продуктивність системи за зростання обсягу колекційних даних. У рамках дослідження було змодельовано роботу системи з різними обсягами об’єктів – від тисячі до кількох десятків тисяч. Результати показали, що середній час відповіді системи зростає повільно й залишається в межах, прийнятних для інтерактивних застосунків.

Для перевірки масштабованості системи було проведено експерименти зі збільшенням кількості об’єктів у колекції від 1 тис. до 50 тис. Середній час відповіді вимірювався під час виконання типових операцій пошуку подібності. Результати наведено на рисунку 4.

Як видно з графіка, зростання обсягу даних супроводжується поступовим збільшенням часу відповіді, проте ця динаміка є лінійною і повільною. З 1 тис. об’єктів середній час становив близько 0,02 с, з 10 тис. – 0,06 с, а з 50 тис. – лише 0,15 с. Таким чином, навіть за значного збільшення кількості елементів система зберігає продуктивність на рівні, достатньому для інтерактивних застосунків.

Отримані результати підтверджують ефективність використання індексаційних структур та алгоритмів пошуку подібності, що забезпечують стійку роботу системи в масштабованих середовищах у реальних умовах, де обсяг даних може досягати сотень тисяч і більше. У поєднанні з архітектурними рішеннями, наведеними на рисунках 1 і 2, цей експериментальний результат засвідчує, що система може використовуватись у масштабованих середовищах без втрати швидкодії.

Обговорення результатів

Отримані результати доцільно розглянути в контексті сучасних досліджень у сфері інтелектуальних систем управління даними та колекціями.

По-перше, запропонована модель підтвердила ефективність використання композитних метрик подібності. Аналогічні ідеї зустрічаються у роботі [8], де поєднання нечітких алгоритмів і багатокритеріальних методів дозволило підвищити якість планування в багатохмарних середовищах. У запропонованій роботі ці підходи адаптовано для обчислення схожості між об’єктами колекцій, що доводить універсальність такого принципу.

По-друге, результати щодо виявлення дублікатів узгоджуються з висновками [9] та [10], які наголошують, що навіть відносно невеликий відсоток надлишкових записів значно впливає на продуктивність сховищ. Наша модель підтвердила цю тезу, показавши, що усунення дублювання покращує як швидкість пошуку, так і ефективність використання ресурсів.

По-третє, у сфері рекомендаційних систем отримані результати співвідносяться з підходами, наведеними в роботах [11] та [12], де застосування штучного інтелекту дало змогу підвищити якість рекомендацій у мультихмарних середовищах. Водночас запропонована нами інтеграція рекомендаційної підсистеми в загальну архітектуру управління колекціями відрізняється ширшим застосуванням: вона придатна як для приватних зібрань, так і для масштабних наукових або корпоративних сховищ.

Особливої уваги заслуговують результати, що стосуються масштабованості. У багатьох дослідженнях (наприклад, [13; 14]) наголошується на обмеженості продуктивності традиційних підходів за значного зростання обсягів даних. Запропонована модель

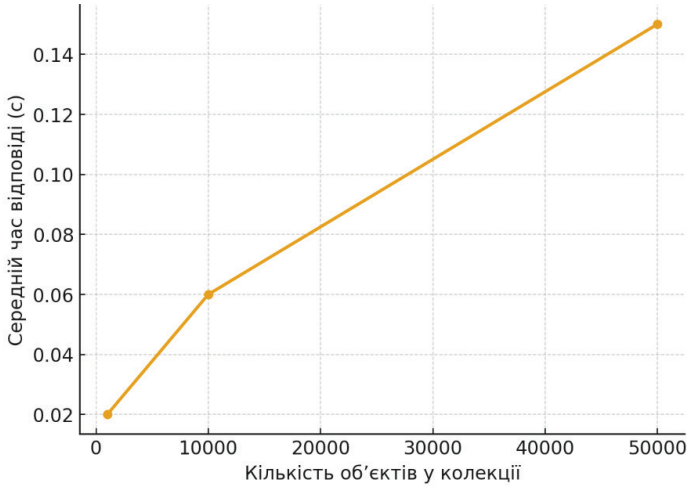

Рис. 4. Залежність середнього часу відповіді системи від кількості об’єктів у колекції

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344519

продемонструвала стабільний час відповіді навіть з десятками тисяч об'єктів, що підтверджує практичну доцільність інтеграції оптимізованих методів пошуку й обробки в подібних системах.

Таким чином, проведене дослідження підтверджує актуальність обраної тематики та демонструє переваги комплексного підходу. На відміну від більшості існуючих робіт, що зосереджуються на окремих аспектах – ресурсному менеджменті, безпеці чи класифікації, – у запропонованій моделі ці компоненти поєднані у єдиній архітектурі. Це створює основу для універсальної системи управління колекціями, здатної адаптуватися до різних сценаріїв застосування.

Висновки

У статті розглянуто питання побудови інтелектуальних технологій для аналізу, класифікації та рекомендацій у системах управління колекціями. Проведене дослідження показало, що комплексний підхід, який поєднує формалізацію об'єктів у вигляді мультимодальних векторів, використання адаптивних метрик подібності, багатомітковой класифікації та рекомендаційних алгоритмів, забезпечує новий рівень ефективності в роботі з колекційними даними.

Розроблена модель довела свою здатність не лише автоматизувати рутинні процеси – такі як виявлення дублікатів чи категоризація об'єктів, – але й забезпечити більш інтелектуальну взаємодію з користувачем через персоналізовані рекомендації. Отримані експериментальні результати засвідчили, що система зберігає високу якість роботи навіть за значного зростання обсягів колекції, демонструючи стабільний час відповіді та прийнятні показники точності.

Практичне значення запропонованого підходу полягає у його універсальності. Модель може застосовуватись як для приватних цифрових чи фізичних колекцій, так і для масштабних інформаційних інфраструктур у бібліотеках, архівах, наукових установах чи корпоративних системах. Наукова новизна дослідження визначається інтеграцією кількох методів у єдину архітектуру, що дає змогу розглядати систему не як набір ізольованих алгоритмів, а як цілісну інтелектуальну платформу.

Перспективи подальших досліджень пов'язані з поглибленням використання методів глибинного навчання для роботи з текстовими й візуальними ознаками, розширенням механізмів персоналізації рекомендацій та інтеграцією засобів контролю доступу й безпеки. Це дасть змогу перетворити запропоновану модель на основу для створення універсальних сервісів нового покоління, орієнтованих на інтелектуальне управління великими обсягами колекційних даних.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно цього дослідження, у тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних

ЛІТЕРАТУРА

[1] R. Kaur, I. Chana, J. Bhattacharya. Data deduplication techniques for efficient cloud storage management: a systematic review. *Journal of Supercomputing*, vol. 74, pp. 2035–2085, 2018. DOI: 10.1007/s11227-017-2210-8.

[2] M. Hasan, A. Hussien. Techniques of data deduplication for cloud storage: A review. *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 3, pp. 45–55, 2024.

[3] L. Ogiela, M.R. Ogiela. Cognitive security paradigm for cloud computing applications. *Concurrency and Computation: Practice and Experience*, vol. 32, no. 8, e5316, 2020. DOI: 10.1002/cpe.5316.

[4] A. Roy. Recommender systems: Algorithms, challenges, and trends. *Journal of Big Data*, vol. 9, no. 1, pp. 1–32, 2022. DOI: 10.1186/s40537-022-00592-5.

[5] J. Devlin, M.W. Chang, K. Lee, K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of NAACL-HLT*, 2019. DOI: 10.48550/arXiv.1810.04805.

[6] C. Xu, D. Tao, C. Xu. A. Survey on multi-view learning. *Neural Computing and Applications*, vol. 23, no. 7–8, pp. 2031–2038, 2013. DOI: 10.1007/s00521-013-1362-6.

[7] M.L. Zhang, Z.H. Zhou. A review on multi-label learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 8, pp. 1819–1837, 2014. DOI: 10.1109/TKDE.2013.39.

[8] M. Farid, R. Latip, M. Hussin, N.A.W. Abdul Hamid. Scheduling scientific workflow using multi-objective algorithm with fuzzy resource utilization in multi-cloud environment. *IEEE Access*, vol. 8, pp. 24309–24322, 2020. DOI: 10.1109/ACCESS.2020.2970475.

[9] N. Chhabra, M. Bala. A comparative study of data deduplication strategies. In: *Proc. 2018 1st Int. Conf. Secure Cyber Computing and Communication (ICSCCC)*, pp. 68–72, 2018. DOI: 10.1109/ICSCCC.2018.8703363.

[10] K. Tiwari, Q.M. Rizbi. A study and review of duplicate data in cloud computing. *Journal of Advances in Science and Technology*, vol. 21, no. 1, pp. 74–80, 2024. DOI: 10.29070/nsek7b45.

[11] A. Anand. Intelligent resource allocation in multi-cloud environments: An AI-driven approach. *International Journal of Scientific Research in Computer Science,*

Engineering and Information Technology, vol. 11, no. 1, pp. 1035–1046, 2025. DOI: 10.32628/CSEIT25112429.

[12] R. Rayaprolu, K. Randhi, S.R. Bandarapu. Intelligent resource management in cloud computing: AI techniques for optimizing DevOps operations. *Journal of Artificial Intelligence General Science*, vol. 6, no. 1, pp. 397–408, 2024. DOI: 10.60087/jaigs.v6i1.262.

[13] Y. Li, D. Chen, J. Hao. Cloud workflow scheduling optimization research based on ant colony algorithm. *Academic Journal of Computing & Information Science*, vol. 6, no. 5, pp. 9–13, 2023. DOI: 10.25236/AJCIS.2023.060502.

[14] M. Reshetniak, S. Popereshnyak. Method for accessing and processing multimedia content in a cloud environment. In: *Proc. 2019 IEEE Int. Scientific-Practical Conf. Problems of Infocommunications, Science and Technology (PIC S&T)*, pp. 71–76, 2019. DOI: 10.1109/PICST47496.2019.9061463.

INTELLIGENT TECHNOLOGIES FOR ANALYSIS, CLASSIFICATION AND RECOMMENDATIONS IN COLLECTION MANAGEMENT SYSTEMS

Daniil Popereshniak, Svitlana Popereshnyak

The article is devoted to the development and investigation of intelligent technologies for analysis, classification, and recommendation in collection management systems. The study addresses the challenge of organizing and processing collection data, which are characterized by rapid growth in volume and increasing heterogeneity. These factors complicate efficient storage, search, and structuring of collections. It is shown that traditional collection management systems are usually limited to isolated tasks such as basic storage or search and often fail to integrate similarity analysis, classification, and recommendation algorithms into a unified framework. The use of intelligent approaches enables overcoming these limitations and provides opportunities for building universal systems capable of handling large-scale datasets.

The purpose of the research is to achieve enhanced efficiency and scalability in collection management by constructing a model of an intelligent system that integrates multimodal object representation, classification methods, duplicate detection, and recommendation mechanisms. To achieve this goal, existing approaches were analyzed, a model for representing collection objects was formalized, similarity computation techniques were developed, and an architecture combining classification and recommendation modules was proposed.

Experimental evaluation demonstrated that the integration of these components ensures high classification accuracy, effective duplicate detection, and relevant personalized recommendations. Particular attention was paid to scalability: the system maintained a stable response time even under significant growth in the number of collection objects. The practical significance of the research lies in the universality of the proposed model, which can be applied both in private multimedia and digital collections and in corporate or scientific infrastructures, such as libraries, archives, and databases. The scientific novelty is defined by the creation of a comprehensive architecture that integrates several intelligent methods into a unified system. Future research perspectives include the application of deep learning approaches for multimodal feature processing, the improvement of recommendation algorithms, and the integration with security and access control mechanisms to ensure robustness in large-scale environments.

Keywords: *intelligent systems, collection management, similarity analysis, classification, recommender systems, multimodal data, scalability, IoT.*

REFERENCES

[1] R. Kaur, I. Chana, J. Bhattacharya. Data deduplication techniques for efficient cloud storage management: a systematic review. *Journal of Supercomputing*, vol. 74, pp. 2035–2085, 2018. DOI: 10.1007/s11227-017-2210-8.

[2] M. Hasan, A. Hussien. Techniques of data deduplication for cloud storage: A review. *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 3, pp. 45–55, 2024.

[3] L. Ogiela, M.R. Ogiela. Cognitive security paradigm for cloud computing applications. *Concurrency and Computation: Practice and Experience*, vol. 32, no. 8, e5316, 2020. DOI: 10.1002/cpe.5316.

[4] A. Roy. Recommender systems: Algorithms, challenges, and trends. *Journal of Big Data*, vol. 9, no. 1, pp. 1–32, 2022. DOI: 10.1186/s40537-022-00592-5.

[5] J. Devlin, M.W. Chang, K. Lee, K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of NAACL-HLT*, 2019. DOI: 10.48550/arXiv.1810.04805.

[6] C. Xu, D. Tao, C. Xu. A. Survey on multi-view learning. *Neural Computing and Applications*, vol. 23, no. 7–8, pp. 2031–2038, 2013. DOI: 10.1007/s00521-013-1362-6.

[7] M.L. Zhang, Z.H. Zhou. A review on multi-label learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 8, pp. 1819–1837, 2014. DOI: 10.1109/TKDE.2013.39.

[8] M. Farid, R. Latip, M. Hussin, N.A.W. Abdul Hamid. Scheduling scientific workflow using multi-objective algorithm with fuzzy resource utilization in multi-cloud environment. *IEEE Access*, vol. 8, pp. 24309–24322, 2020. DOI: 10.1109/ACCESS.2020.2970475.

[9] N. Chhabra, M. Bala. A comparative study of data deduplication strategies. In: *Proc. 2018 1st Int. Conf. Secure Cyber Computing and Communication (ICSCCC)*, pp. 68–72, 2018. DOI: 10.1109/ICSCCC.2018.8703363.

[10] K. Tiwari, Q.M. Rizbi. A study and review of duplicate data in cloud computing. *Journal of Advances in Science*

and Technology, vol. 21, no. 1, pp. 74–80, 2024. DOI: 10.29070/nsek7b45.

[11] A. Anand. Intelligent resource allocation in multi-cloud environments: An AI-driven approach. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 11, no. 1, pp. 1035–1046, 2025. DOI: 10.32628/CSEIT25112429.

[12] R. Rayaprolu, K. Randhi, S.R. Bandarapu. Intelligent resource management in cloud computing: AI techniques for optimizing DevOps operations. *Journal of Artificial Intelligence General Science*, vol. 6, no. 1, pp. 397–408, 2024. DOI: 10.60087/jaigs.v6i1.262.

[13] Y. Li, D. Chen, J. Hao. Cloud workflow scheduling optimization research based on ant colony algorithm. *Academic Journal of Computing & Information Science*, vol. 6, no. 5, pp. 9–13, 2023. DOI: 10.25236/AJCIS.2023.060502.

[14] M. Reshetniak, S. Popereshnyak. Method for accessing and processing multimedia content in a cloud environment. In: *Proc. 2019 IEEE Int. Scientific-Practical Conf. Problems of Infocommunications, Science and Technology (PIC S&T)*, pp. 71–76, 2019. DOI: 10.1109/PICST47496.2019.9061463.

Стаття надійшла до редакції 29.09.2025
Стаття прийнята 13.10.2025
Статтю опубліковано 02.12.2025

ТЗ. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344520

Кібербезпека
та захист критичної інфраструктури

УДК 004.056.5:004.021

КОМПАРАТИВНИЙ АНАЛІЗ МЕТОДІВ ТА ТЕХНОЛОГІЙ МОДЕЛЮВАННЯ В КІБЕРБЕЗПЕЦІ

В.С. Кравчук[1], Т.В. Алтухова[1], Я.Ю. Дорогий[1, 2, 3]

[1] Department of Applied Mathematics and Informatics, Donetsk National Technical University, Drohobych, Ukraine

[2] Department of Information Systems and Technologies, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

[3] Department of Intelligent Software Systems, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

E-mail: yaroslav.dorohyi@donntu.edu.ua

АНОТАЦІЯ

Сучасний ландшафт кібербезпеки характеризується високою складністю та динамічністю, що зумовлює потребу в передових методах моделювання для аналізу загроз, прогнозування ризиків та оцінки ефективності захисних заходів. У цій статті представлено детальний компаративний аналіз традиційних і сучасних підходів до моделювання в кібербезпеці, включно з математичним, логічним та ієрархічним моделюванням, симуляціэю атак і Breach and Attack Simulation (BAS), агентно-орієнтованим моделюванням, цифровими двійниками, а також методи, що базуються на машинному навчанні, глибокому навчанні, теорії ігор, графових структурах і великих мовних моделях (LLM). Кожен метод розглянуто з погляду принципів роботи, ключових переваг, недоліків і практичного застосування. Особливу увагу приділено синергії та взаємодоповнюваності цих підходів, що є критично важливим для створення комплексних, адаптивних систем кібербезпеки. Традиційні методи, як-от математичне моделювання, забезпечують формальну основу для аналізу, але можуть спрощувати реальні сценарії. Сучасні підходи, зокрема машинне навчання та цифрові двійники, дають змогу обробляти великі обсяги даних і моделювати складні динамічні взаємодії, хоча потребують значних обчислювальних ресурсів і точних даних. Теорія ігор і графові моделі пропонують стратегічний і контекстуальний аналіз, тоді як великі мовні моделі відкривають нові можливості для автоматизації аналізу загроз, попри їх обмеження в надійності. Інтеграція цих методів створює основу для гібридних рішень, які компенсують недоліки окремих підходів, підвищуючи загальну ефективність захисту. Стаття також висвітлює виклики, пов’язані з обчислювальною складністю, невизначеністю й етичними аспектами, і окреслює майбутні напрями розвитку, зокрема вдосконалення пояснюваного ШІ, стійкість до змагальних атак і реалізм симуляцій.

Ключові слова: *кібербезпека, моделювання, машинне навчання, цифрові двійники, графові моделі, симуляція атак.*

Вступ

На сьогодні ландшафт кібербезпеки характеризується безпрецедентною складністю та динамічністю, що зумовлено стрімким розвитком технологій, глобалізацією цифрових систем і швидким зростанням кількості та різноманітності кіберзагроз [1–2]. Кібератаки вирізняються частотою, витонченістю та скоординованістю, включно з розширеними стійкими загрозами (APT), фішингом, атаками з використанням шкідливого програмного забезпечення, а також змагальними маніпуляціями, спрямованими на обхід систем захисту, зокрема тими, що базуються на штучному інтелекті [3–4]. Ця ескалація загроз перевищує можливості традиційних методів виявлення та запобігання, як-от сигнатурний аналіз чи статичні правила, які не здатні ефективно справлятися з динамічною природою атак [5].

Такий стан справ створює нагальну потребу в інноваційних підходах до моделювання, які можуть обробляти величезні обсяги даних, виявляти приховані закономірності, прогнозувати поведінку зловмисників і оцінювати ефективність захисних механізмів у реальному часі [5]. Традиційні методи аналізу загроз, а саме математичне чи логічне моделювання,

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344520

стикаються зі значними труднощами через складність і мінливість кібератак, що спонукає до розробки адаптивних і інтелектуальних моделей, як-от глибоке навчання, агентно-орієнтоване моделювання, цифрові двійники та графові структури [6].

Моделювання відіграє ключову роль у кібербезпеці, надаючи можливість дослідникам, інженерам і фахівцям із безпеки вивчати поведінку мереж і систем у контрольованих віртуальних середовищах. Це дає змогу оцінювати їхню продуктивність, оптимізувати конфігурації, виявляти потенційні вразливості та тестувати стратегії захисту ще до розгортання в реальних умовах [7]. Крім того, моделювання сприяє проактивному підходу до кіберзахисту, завдяки чому можна передбачати нові вектори атак і розробляти стратегії реагування на них. У контексті зростаючої складності кіберзагроз і появи нових технологій, як-от інтернет речей (IoT), хмарні обчислення та квантові технології, потреба в комплексних, гібридних моделях, які поєднують переваги різних методів, стає критично важливою. Таким чином, розробка й інтеграція передових методів моделювання є ключовими для забезпечення стійкості кіберсистем до нових і непередбачуваних загроз, тож тема дослідження є актуальною.

Постановка проблеми дослідження

Ключовою проблемою є обмежена здатність традиційних методів моделювання, як-от математичне чи логічне моделювання, враховувати динамічну природу кіберзагроз, поведінку користувачів і складні взаємодії в гетерогенних мережевих середовищах. Водночас сучасні підходи, а саме машинне навчання, агентно-орієнтоване моделювання чи цифрові двійники, хоча й пропонують нові можливості, стикаються з викликами, пов'язаними з високими обчислювальними вимогами, необхідністю якісних даних, вразливістю до змагальних атак і недостатньою прозорістю рішень. Відсутність універсального методу, який би поєднував переваги різних підходів, ускладнює створення комплексних і адаптивних систем кібербезпеки.

Таким чином, виникає потреба в систематичному аналізі та порівнянні традиційних і сучасних методів моделювання, оцінці їхньої ефективності, виявленні можливостей синергії та визначенні перспективних напрямів розвитку для підвищення стійкості кіберзахисту в умовах зростаючої складності загроз.

З огляду на вищенаведене метою цього дослідження є надання детального компаративного аналізу методів і технологій моделювання, застосовних у кібербезпеці, з акцентом на їхню ефективність, сфери використання та потенціал інтеграції.

Для досягнення цієї мети систематизовано традиційні та сучасні підходи, проаналізовано їхні особливості, проведено порівняльну оцінку та визначено ключові виклики й перспективи розвитку.

Матеріали та методи дослідження

Дослідження зосереджено на порівнянні методів моделювання в кібербезпеці, використовуючи наукові публікації, технічні звіти й огляди з авторитетних джерел, а саме Nature, Springer, MDPI, arXiv, а також матеріали організацій, зокрема XM Cyber і Picus Security, до 2025 року. Аналіз охопив літературу про математичне, логічне, симуляційне, агентно-орієнтоване моделювання, цифрові двійники, машинне навчання, теорію ігор, графові моделі та великі мовні моделі. Кейси застосування, як-от Purdue-модель, фреймворк AMM, платформи BAS (SCYTHE, AttackIQ) і графові інструменти (Neo4j, Maltego), стали основою для оцінки практичної реалізації методів.

Для аналізу застосовано фундаментальні та прикладні методи. Аналіз і синтез допомогли систематизувати характеристики методів, порівняльний аналіз зіставив традиційні та новітні підходи за ефективністю, масштабованістю й адаптивністю, а системний підхід розглядав методи як взаємопов'язані елементи кібербезпеки. Огляд літератури та класифікація методів дали змогу зібрати дані про принципи роботи і приклади застосування, тоді як порівняльна оцінка, зокрема BAS проти тестування на проникнення, виявила їхні сильні та слабкі сторони. Аналіз кейсів (наприклад, використання CrowdStrike для машинного навчання чи GALLIUM APT для агентного моделювання) доповнив дослідження.

Методологічно дослідження базувалося на об'єктивності й науковій достовірності, використовуючи програмні інструменти для систематизації літератури (Mendeley, Zotero) та обробки текстів. Поєднання фундаментальних методів, зокрема абстрагування для виділення ключових аспектів, із прикладними, як-от аналіз кейсів, забезпечило цілісне розуміння методів моделювання. Особливу увагу приділено синергії підходів, що є основою для створення адаптивних систем кіберзахисту.

Аналіз методів і технологій моделювання

Моделювання в кібербезпеці передбачає створення формальних або віртуальних представлень комп'ютерних і кіберфізичних систем для аналізу загроз, прогнозування ризиків і оцінки захисних заходів [8]. Традиційні методи, як-от математичне, ієрархічне та логічне моделювання, базуються на формалізації безпекових властивостей через математичні структури [9], тоді як сучасні підходи використовують обчислювальні експерименти та інтелектуальну обробку даних для відтворення поведінки складних систем.

Математичне моделювання застосовує рівняння, ймовірнісні моделі, теорію графів, теорію ігор

ISSN 2786-9024

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344520

і диференціальні рівняння для опису систем і загроз [10]. Наприклад, у дослідженні Суфі та Альсуламі запропоновано фреймворк, що поєднує χ^2-тести, баєсівський аналіз і кластерний аналіз для виявлення статистичних закономірностей у кібератаках [4]. Ці моделі забезпечують відтворювані результати, дають змогу застосовувати аналітичні методи для обробки великих даних і є масштабованими, але їхня ефективність обмежується спрощенням реальних умов, що ігнорує поведінку користувачів чи динамічні змінні, а також складністю реалізації [11–12]. Математичні моделі використовуються для прогнозування вторгнень, аналізу стратегій «захисник – атакуючий» за допомогою теорії ігор і симуляції поширення шкідливого програмного забезпечення через диференціальні рівняння [13].

Переходячи до *ієрархічних моделей*, варто зазначити, що вони структурують системи на багаторівневі підсистеми з різними рівнями контролю, як, наприклад, Purdue-модель для промислових систем, що розділяє мережу на рівні від польових пристроїв до корпоративної інфраструктури [14]. Такий підхід спрощує розмежування доступу, сприяє впровадженню принципів «захисту в глибину» і забезпечує прозорість архітектури, але може не враховувати складну динаміку між рівнями, що ускладнює моделювання комплексних атак, а жорсткі межі можуть знижувати гнучкість системи [15–16]. Ієрархічні моделі, зокрема дерева атак, також дають можливість розбити складні загрози на підзадачі, що полегшує їхній аналіз [16].

Логічні моделі, або формальні методи, базуються на строгих математичних підходах, як-от процесна алгебра чи формальна логіка [9]. Наприклад, Пітер Райан застосував формалізм CSP для моделювання вимог безпеки [10]. Завдяки цим методами можна верифікувати безпекові властивості системи, виявляти суперечності та забезпечувати однозначні специфікації, що є особливо цінним для перевірки криптографічних протоколів через інструменти ProVerif або Tamarin, а також для формалізації властивостей безпеки за допомогою мови Event-B і платформи Rodin [9, 15]. Проте їхня складність обмежує використання у великих системах, а ігнорування людського фактора чи неформальних процедур знижує практичну цінність [9].

Сучасні методи моделювання, на відміну від традиційних, спираються на обчислювальні експерименти та інтелектуальну обробку даних [8]. Симуляційні моделі відтворюють поведінку системи у віртуальному середовищі, даючи можливість тестувати сценарії безпеки без ризику для реальних ресурсів [7]. Вони використовуються в кіберполігонах, мережевих емуляторах, як-от OMNeT++ і ns-3, а також у моделях епідеміології кібератак, наприклад CICED для комп'ютерних вірусів [18]. Симуляція атак імітує реальні загрози для оцінки захисних механізмів, сприяючи тренуванню фахівців і оцінці політик безпеки [17]. Точність таких симуляцій залежить від якості моделі, але вони можуть не враховувати всіх реальних факторів, що іноді створює хибне відчуття безпеки [7]. Наприклад, платформа Hoxhunt пропонує симуляції фішингу для підвищення обізнаності працівників [19]. Платформи BAS забезпечують безперервну автоматизовану перевірку стану безпеки, імітуючи тактики MITRE ATT&CK, виявляючи слабкі місця та надаючи рекомендації [20]. Для порівняння: управління вразливостями (VM) ідентифікує та пріоритизує відомі вразливості за допомогою автоматизованих сканерів, тестування на проникнення передбачає ручний і інструментальний аналіз для виявлення експлуатованих вразливостей, а BAS пропонує безперервну валідацію в реальному часі [21–23]. VM забезпечує широке покриття, але може давати хибні спрацювання, Pentesting пропонує глибокий аналіз, але є періодичним і часозатратним, тоді як BAS вирізняється масштабованістю й економічною ефективністю, хоча залежить від актуальності бібліотеки загроз. Провідними BAS-платформами є SCYTHE, Google Mandiant, Cymulate, AttackIQ, SafeBreach, Picus Security, FourCore ATTACK, Pentera, XM Cyber і SnapAttack [23].

Агентно-орієнтоване моделювання (ABM) представляє систему як сукупність автономних агентів, що взаємодіють між собою та із середовищем [24, 44]. Цей підхід дає змогу моделювати боротьбу між атакуючими та захисниками, як у кейсі GALLIUM APT, де захисники, навчені за допомогою багатоагентного навчання з підкріпленням, підвищують ефективність [26]. ABM також ефективно аналізує поширення шкідливого програмного забезпечення, враховуючи локальну мережеву структуру та мобільність користувачів, як у фреймворку AMM [25, 27]. Моделі ABM забезпечують детальний аналіз взаємодій, дають змогу налаштувати правила з реальними параметрами та виявляти вразливі топології, але потребують значних обчислювальних ресурсів, тривалої верифікації та якісних даних для тестування [25].

Цифрові двійники створюють віртуальні копії реальних систем із синхронізацією даних у реальному часі, даючи можливість тестувати зміни без впливу на фізичну інфраструктуру [2]. Наприклад, Purdue University спільно з Boeing розробили цифровий двійник авіаційних систем для моделювання бортових мереж [2]. Ці моделі забезпечують детальну симуляцію, проактивний моніторинг аномалій і можливість тренувань, але їх розробка є ресурсомісткою, а відсутність стандартів і вразливість платформ ускладнюють масштабування [2, 28].

Моделі на основі штучного інтелекту, зокрема машинного (ML) і глибокого навчання (DL),

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344520

використовують дані для прогнозування та виявлення загроз [3]. ML застосовується для аналізу мережевого трафіку, виявлення шахрайства та прогнозування атак, використовуючи такі моделі, як Support Vector Machines (SVM), Decision Trees, Random Forests, K-Nearest Neighbors (KNN) і Naive Bayes [29]. Наприклад, CrowdStrike аналізує дані кінцевих пристроїв, а Montimage виявляє атаки типу Cache Poisoning [30, 31]. ML-моделі вирізняються швидкістю, точністю й адаптивністю, але потребують перенавчання за зміни умов і вразливі до змагальних атак [12]. Пояснюваний ШІ (XAI) є важливим для підвищення прозорості рішень ML [12]. DL, як розширення ML, ефективно обробляє великі набори даних, автоматично вилучаючи ознаки, і застосовується для виявлення шкідливого програмного забезпечення на платформах Windows, Linux і Android через моделі MLP, RNN, LSTM, GRU, CNNs, ResNet, GANs і GNNs [32]. DL також використовується для виявлення вразливостей у коді через шість фаз: побудова набору даних, визначення гранулярності вразливостей, представлення коду, проєктування моделі, оцінка продуктивності та впровадження [33]. Однак DL-моделі стикаються з проблемами перенавчання, високими обчислювальними вимогами, низькою пояснюваністю та вразливістю до змагальних атак [32]. Для порівняння: Naive Bayes ефективний для великих даних і швидкого виявлення спаму, але має низьку точність у складних сценаріях; Decision Trees чіткі та зрозумілі, але обмежені в узагальненні; Random Forests стійкі до шуму, але залежать від даних; KNN ефективний для аномалій, але обмежений в обробці певних типів даних; SVM обробляє багатовимірні дані, але залежить від характеристик даних; DL-моделі (VGG, ResNet, Inception) виявляють складні патерни, але мають високі обчислювальні витрати; Extra Trees вирізняються точністю, але також залежать від даних [34]. Усі ці моделі потребують якісних даних, частого перенавчання та стійкості до змагальних атак.

Теорія ігор аналізує стратегічну взаємодію між зловмисниками та захисниками, застосовуючись для оптимізації систем виявлення вторгнень, аналізу APT, захисту з рухомою ціллю й обміну інформацією про загрози [35]. Цей підхід забезпечує математично обґрунтовані рішення, враховуючи реакції зловмисників на захисні політики, але обмежений припущеннями про повну інформацію та складністю врахування невизначеності.

Графові моделі, як-от графи знань, графи атак і графові нейронні мережі (GNNs), кодують взаємозв'язки між сутностями для аналізу загроз [8]. Графи знань представляють сутності як вузли, а відносини – як ребра, ефективно виявляючи аномалії та збагачуючи кіберрозвідку, але потребують фільтрації надлишкової інформації [36]. Графи атак візуалізують шляхи атаки, сприяючи пріоритизації ризиків, і підтримуються інструментами, а саме PuppyGraph, Neo4j, OWASP Amass, Maltego і Linkurious [8, 37]. GNNs моделюють складні залежності, підвищуючи ефективність систем виявлення вторгнень і аналізу вразливостей [38–39].

Великі мовні моделі (LLMs) вирізняються контекстуальним розумінням і генерацією людиноподібного тексту, застосовуючись для виявлення вразливостей, аналізу шкідливого програмного забезпечення, фішингу, вебфаззингу, збагачення кіберрозвідки, пошуку загроз і моделювання багатоагентних систем у симуляціях реагування на інциденти [40–42]. Проте їхня ненадійність, непослідовність і недостатнє калібрування обмежують використання в критичних сценаріях [43].

Обговорення та перспективи розвитку дослідження

Ефективність моделювання в кібербезпеці залежить від інтеграції різних методів, що дає змогу компенсувати їхні обмеження. Наприклад, поєднання графових нейронних мереж із глибоким навчанням підвищує точність виявлення атак, а великі мовні моделі збагачують графи знань для аналізу загроз. Симуляційні середовища, зокрема BAS, дають можливість тестувати стратегії, виведені за допомогою теорії ігор чи агентно-орієнтованого моделювання, у динамічних умовах. Формальні методи забезпечують верифікацію критичних компонентів, тоді як машинне навчання й симуляції виявляють непередбачені загрози. Штучний інтелект відіграє роль інтегруючого елемента, автоматизуючи аналіз і підвищуючи адаптивність моделей. Для вибору методів слід враховувати специфіку завдань: глибоке навчання і GNNs підходять для виявлення складних загроз, теорія ігор і ABM – для стратегічного планування, BAS і тестування на проникнення – для оцінки стійкості, формальні методи – для критичних систем, а LLMs – для автоматизації аналізу тексту. Майбутні дослідження слід зосередити на розвитку гібридних моделей, що інтегрують машинне навчання, агентно-орієнтоване моделювання та формальні методи, а також на вдосконаленні пояснюваного ШІ для підвищення прозорості рішень. Важливими напрямами є стійкість моделей до змагальних атак, моделювання невизначеності в теорії ігор і ABM, підвищення реалізму симуляцій через якісні набори даних, а також вирішення етичних питань, як-от упередженість моделей і конфіденційність даних.

Висновки

Дослідження компаративного аналізу методів моделювання в кібербезпеці показало, що ефективне протистояння складним і динамічним кіберзагрозам

ISSN 2786-9024

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344520

потребує комплексного підходу, який поєднує традиційні та інноваційні методи. За результатами систематизації та порівняння підходів встановлено, що математичне та логічне моделювання забезпечують формальну основу для аналізу безпеки, але мають обмеження в умовах динамічних сценаріїв через спрощення реальних взаємодій. Симуляційні технології, зокрема BAS, виявилися ефективними для проактивної оцінки захисних механізмів, забезпечуючи безперервну валідацію та виявлення слабких місць. Агентно-орієнтоване моделювання та цифрові двійники продемонстрували здатність відтворювати складні взаємодії в гетерогенних системах, хоча їхнє впровадження ускладнене високими обчислювальними вимогами.

Методи штучного інтелекту, як-от машинне та глибоке навчання, показали високу точність у прогнозуванні та виявленні загроз, але їхня ефективність залежить від якості даних і вразлива до змагальних атак. Теорія ігор і графові моделі, включно з графами знань та графовими нейронними мережами, забезпечують стратегічний і контекстуальний аналіз, даючи змогу виявляти приховані шляхи атак і оптимізувати захисні стратегії. Великі мовні моделі відкривають перспективи для автоматизації кіберрозвідки, однак їхня ненадійність у реальних сценаріях потребує додаткового калібрування.

Основним результатом дослідження є підтвердження, що інтеграція різних методів моделювання дає змогу компенсувати їхні індивідуальні недоліки, створюючи гнучкі й адаптивні системи кібербезпеки. Зокрема, поєднання графових моделей із глибоким навчанням, симуляцій із теорією ігор та формальних методів із машинним навчанням забезпечує синергію, що підвищує загальну ефективність захисту. Дослідження також визначило ключові виклики, як-от потреба в пояснюваному ШІ, стійкості до маніпуляцій і реалізмі симуляцій, що окреслюють перспективні напрями подальшого розвитку цього дослідження.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно цього дослідження, у тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

ЛІТЕРАТУРА

[1] S. A. Sharaf, M.A.A. Al-qaness, M.A. Alghamdi, A.S. Alqahtani, A.M. Alshahrani, "Advanced mathematical modeling of mitigating security threats in smart grids through deep ensemble model," Sci. Rep., vol. 14, no. 23069, Oct. 2024. DOI: 10.1038/s41598-024-74733-6. URL: https://surl.lu/umzyfz. Дата звернення: 15.06.2025.

[2] M. Homaei, Ó. Mogollón-Gutiérrez, J. Carlos Sancho, M. Ávila & A. Caro, "A review of digital twins and their application in cybersecurity based on artificial intelligence," Artif. Intell. Rev., vol. 57, no. 201, Jul. 2024. DOI: 10.1007/s10462-024-10805-3. URL: https://surl.li/asmirk. Дата звернення: 15.06.2025.

[3] I. Vourganas, A.L. Michala, "Applications of machine learning in cyber security: A review," J. Cybersecurity Privacy, vol. 4, no. 4, pp. 972–992, Dec. 2024. DOI: 10.3390/jcp4040045. URL: https://surl.li/vdlxkl. Дата звернення: 15.06.2025.

[4] F. Sufi and M. Alsulami, "Mathematical modeling and clustering framework for cyber threat analysis across industries," *Mathematics*, vol. 13, no. 4, p. 655, Feb. 2025. DOI: 10.3390/math13040655. URL: https://surl.li/fgrjnf. Дата звернення: 15.06.2025.

[5] Y. Miao, L. Pan, et al., "Machine learning based cyber attacks targeting on controlled information: A survey," arXiv:2102.07969, Feb. 2021 URL: https://surl.li/hmyhes. Дата звернення: 15.06.2025.

[6] J. Vitorino, I. Praça, E. Maia, L. Sousa, and S. Jardim, "A Comparative Analysis of Machine Learning Techniques for IoT Intrusion Detection," arXiv:2111.13149, Nov. 2021. URL: https://surli.cc/vlcuxu. Дата звернення: 15.06.2025.

[7] H. Kavak, J.J. Padilla, D. Vernon-Bido, S.Y. Diallo, R. Gore, and S. Shetty, "Simulation for cybersecurity: State of the art and future directions," *J. Cybersecurity*, vol. 7, no. 1, p. tyab005, Mar. 2021. DOI: 10.1093/cybsec/tyab005. URL: https://surl.lu/crguxz. Дата звернення: 15.06.2025.

[8] XM Cyber, "Seeing what attackers see: How attack graphs help you stay ahead," Jan. 2025. URL: https://surli.cc/dfwcym. Дата звернення: 15.06.2025.

[9] S. Chong, J. Guttman, A. Datta, A. Myers, B. Pierce, P. Schaumont, T. Sherwood, N. Zeldovich, "Report on the NSF Workshop on Formal Methods for Security," arXiv:1608.00678, Aug. 2016. URL: https://surl.gd/sgrcyr. Дата звернення: 15.06.2025.

[10] P.Y.A. Ryan, "Mathematical Models of Computer Security," in Foundations of Security Analysis and Design (FOSAD 2000), vol. 2171, Lecture Notes in Computer Science, R. Focardi and R. Gorrieri, Eds. Berlin, Germany: Springer, 2001, pp. 1–62. DOI: 10.1007/3-540-45608-2_1. URL: https://surl.li/egqhjf. Дата звернення: 15.06.2025.

[11] F.S. Passino, N.M. Adams, E.A.K. Cohen, M. Evangelou, and N.A. Heard, "Statistical Cybersecurity: A Brief Discussion of Challenges, Data Structures, and Future Directions," Harvard Data Science Review, vol. 5, no. 1, Mar. 2023. DOI: 10.1162/99608f92.240383c7. URL: https://surli.cc/foscla. Дата звернення: 15.06.2025.

[12] P. Mahadevappa, S.M. Muzammal, and R.K. Murugesan, "A Comparative Analysis of Machine Learning Algorithms for Intrusion Detection in Edge-Enabled IoT Networks," arXiv:2111.01383, Nov. 2021. URL: https://surl.li/zwaisx. Дата звернення: 15.06.2025.

[13] K. Mohamed, "Machine learning techniques to address cybersecurity," arXiv:2302.12415, Feb. 2023. URL: https://surl.lu/tdmavn. Дата звернення: 15.06.2025.

[14] D. Garton, "Purdue model framework for industrial control systems & cybersecurity segmentation," U.S. Dept. Energy/National Petroleum Council, Topic Paper 4–14, Jul. 2019. URL: https://surli.cc/fmkjby. Дата звернення: 15.06.2025.

[15] I. Sayar, N. Messe, S. Ebersold, and J.M. Bruel, "From What to How: A Taxonomy of Formalized Security Properties," arXiv:2505.14514, May 2025. URL: https://surl.lu/nwovtz. Дата звернення: 15.06.2025.

[16] Wikipedia, "Attack trees," Jan. 2025. URL: https://surl.li/dbzvmq. Дата звернення: 15.06.2025.

[17] Picus Security, "What is attack simulation," Jan. 2025. URL: https://surl.li/cifqpc. Дата звернення: 15.06.2025.

[18] Pynetlabs, "Top 10 network simulation tools," Jan. 2025. URL: https://surl.li/qesago. Дата звернення: 15.06.2025.

[19] Hoxhunt, "Cyber Security Simulation Training: How it Works + Best Practices," Jan. 2025. URL: https://surl.li/cbjkeq. Дата звернення: 15.06.2025.

[20] Picus Security, "Everything You Need to Know About BAS Tools," Jan. 2025. URL: https://surli.cc/exllfj. Дата звернення: 15.06.2025.

[21] Softprom, "Vulnerability management vs. penetration testing vs. breach and attack simulation," Jan. 2025. URL: https://surl.li/afkwyc. Дата звернення: 15.06.2025.

[22] CyberProof, "Penetration Testing vs Breach and Attack Simulator: Key Differences and Why It Matters," Jan. 2025. URL: https://surl.li/nxbard. Дата звернення: 15.06.2025.

[23] SCYTHE, "Top 10 breach and attack simulation (BAS) tools," Jan. 2025. URL: https://surli.cc/ujxueh. Дата звернення: 15.06.2025.

[24] A. Vestad and B. Yang, "A survey of agent-based modeling for cybersecurity," in *AHFE Open Access Proceedings*, vol. 127, pp. 83–93, Jul. 2024. URL: https://surl.lu/wcoxjx. Дата звернення: 15.06.2025.

[25] F.K. Batista, A.M. del Rey, and A. Queiruga-Dios, "A New Individual-Based Model to Simulate Malware Propagation in Wireless Sensor Networks," Mathematics, vol. 8, no. 3, p. 410, Mar. 2020. DOI: 10.3390/math8030410. URL: https://surl.li/yvuqfp. Дата звернення: 15.06.2025.

[26] J. Soule, J. Jamont, M. Occello, P. Theron, and L. Traonouez, "Towards a multi-agent simulation of cyber-attackers and cyber-defenders battles," arXiv:2506.04849, Jun. 2025. URL: https://surl.li/jrkiup. Дата звернення: 15.06.2025.

[27] A. Bose, and K. Shin, "Agent-based modeling of malware dynamics in heterogeneous environments," *Security and Communication Networks*, vol. 6, no. 12, pp. 1–8, Dec. 2013. URL: https://surl.li/akjovc. Дата звернення: 15.06.2025.

[28] E.C. Balta, M. Pease, J. Moyne, K. Barton, and D.M. Tilbury, "Digital Twin-Based Cyber-Attack Detection Framework for Cyber-Physical Manufacturing Systems," IEEE Transactions on Automation Science and Engineering, vol. 21, no. 3, pp. 3245–3260, Jul. 2024. DOI: 10.1109/TASE.2023.3243147.

[29] K. Shaukat, S. Luo, V. Varadharajan, I.A. Hameed, S. Chen, D. Liu, and J. Li, "Performance Comparison and Current Challenges of Using Machine Learning Techniques in Cybersecurity," Energies, vol. 13, no. 10, p. 2509, May 2020. DOI: 10.3390/en13102509.

[30] CrowdStrike, "CrowdStrike machine learning in cybersecurity," Jan. 2025. URL: https://surl.li/yfqmjm. Дата звернення: 15.06.2025.

[31] G. Apruzzese, P. Laskov, E. Montes de Oca, W. Mallouli, L. Brdalo Rapa, A.V. Grammatopoulos, and F. Di Franco, "The Role of Machine Learning in Cybersecurity," Digital Threats: Research and Practice, vol. 4, no. 1, Art. no. 8, pp. 1–38, Mar. 2023. DOI: 10.1145/3545574. URL: https://surl.li/rpjhxe. Дата звернення: 15.06.2025.

[32] P. Maniriho, A.N. Mahmood, and M.J.M. Chowdhury, "Deep learning models for detecting malware attacks," arXiv:2209.03622, Jan. 2024 URL: https://surl.li/kvdwjl. Дата звернення: 15.06.2025.

[33] M.N. Uddin, Y. Zhang, and X. Hei, "Deep learning aided software vulnerability detection: A survey," arXiv:2503.04002, Mar. 2025. URL: https://surl.li/njcajz. Дата звернення: 15.06.2025.

[34] M. Hesham, M. Essam, M. Bahaa, A. Mohamed, M. Gomaa, M. Hany, and W. Elsersy, "Evaluating Predictive Models in Cybersecurity: A Comparative Analysis of Machine and Deep Learning Techniques for Threat Detection," arXiv:2407.06014, Jul. 2024. URL: https://surl.li/grzpge. Дата звернення: 15.06.2025.

[35] B. Collins, S. Xu, and P.N. Brown, "Game-Theoretic Cybersecurity: the Good, the Bad and the Ugly," arXiv:2401.13815, Jan. 2024. URL: https://surl.li/jkxpct. Дата звернення: 15.06.2025.

[36] L. Zhang, Q. Zhu, and Y. Xie, "Improving network threat detection by knowledge graph, large language model, and imbalanced learning," arXiv:2501.16393, May. 2025. URL: https://surli.cc/bqlnzd. Дата звернення: 15.06.2025.

[37] PuppyGraph, "Cyber Graph: Enhancing Cybersecurity with Graph Intelligence," Jan. 2025. URL: https://surl.li/yfugrw. Дата звернення: 15.06.2025.

[38] X. Zhang, "Graph neural networks in network security: From theoretical foundations to applications," in *Proc. Americas Conf. Inf. Syst. (AMCIS)*, Aug. 2025, pp. 1–10. URL: https://surli.cc/wpkgob. Дата звернення: 15.06.2025.

[39] V. Khurana and N. Kumar, "Graph Neural Networks for Cybersecurity Applications in Network Intrusion and

Vulnerability Analysis," in RADemics Research Institute, Chapter 8, Jan. 2025. URL: https://surl.li/tchiff. Дата звернення: 15.06.2025.
[40] Bolster.ai, "Large Language Models for Cybersecurity: The Role of LLMs in Threat Hunting," Apr. 2025. URL: https://surl.li/olyczz. Дата звернення: 15.06.2025.
[41] H. Xu, and et al., "Large language models for cyber security: A systematic literature review," arXiv:2405.04760, Jul 2024. URL: https://surl.li/coqqui. Дата звернення: 15.06.2025.
[42] Z. Liu, "Multi-Agent Collaboration in Incident Response with Large Language Models," arXiv:2412.00652, Dec. 2024. URL: https://surl.li/aghvvq. Дата звернення: 15.06.2025.
[43] E. Mezzi, F. Massacci, and K. Tuma, "Large language models are unreliable for cyber threat intelligence," arXiv:2503.23175, Mar. 2025. URL: https://surl.lu/xonjqn. Дата звернення: 15.06.2025.
[44] В. С. Кравчук, Н. О. Маслова та Я. Ю. Дорогий, "Автоматизований пошук XSS-вразливостей у веб-застосунках на основі мультиагентного підходу," Наукові праці ДонНТУ. Серія "Обчислювальна техніка та автоматизація", т. 3, № 4 (36), с. 13–26, 2025. DOI: 10.31474/2786-9024/v3i4(36).324435.

COMPARATIVE ANALYSIS OF MODELING METHODS AND TECHNOLOGIES IN CYBERSECURITY

Vladyslav Kravchuk, Tatiana Altukhova, Iaroslav Dorohyi

The cybersecurity landscape is characterized by high complexity and dynamism, necessitating advanced modeling methods for threat analysis, risk prediction, and evaluation of protective measures. This article presents a detailed comparative analysis of traditional and contemporary modeling approaches in cybersecurity, including mathematical, logical, and hierarchical modeling, attack simulations and Breach and Attack Simulation (BAS), agent-based modeling, digital twins, as well as methods based on machine learning, deep learning, game theory, graph structures, and large language models (LLMs). Each method is examined in terms of its operational principles, key advantages, limitations, and practical applications. Particular attention is given to the synergy and complementarity of these approaches, which are critical for developing comprehensive and adaptive cybersecurity systems. Traditional methods, such as mathematical modeling, provide a formal basis for analysis but may oversimplify real-world scenarios. Contemporary approaches, including machine learning and digital twins, enable the processing of large data volumes and modeling of complex dynamic interactions, though they require significant computational resources and accurate data. Game theory and graph models offer strategic and contextual analysis, while large language models open new possibilities for automating threat analysis, despite their reliability limitations. The integration of these methods forms the foundation for hybrid solutions that mitigate the shortcomings of individual approaches, enhancing overall protection efficacy. The article also highlights challenges related to computational complexity, uncertainty, and ethical considerations, and outlines future directions, such as improving explainable AI, resilience to adversarial attacks, and simulation realism.

Keywords: *cybersecurity, modeling, machine learning, digital twins, graph models, attack simulation.*

REFERENCES

[1] S.A. Sharaf, M.A.A. Al-qaness, M.A. Alghamdi, A.S. Alqahtani, A.M. Alshahrani, "Advanced mathematical modeling of mitigating security threats in smart grids through deep ensemble model," Sci. Rep., vol. 14, no. 23069, Oct. 2024. DOI: 10.1038/s41598-024-74733-6. [Online]. URL: https://surl.lu/umzyfz. Accessed: 15 Jun, 2025.
[2] M. Homaei, Ó. Mogollón-Gutiérrez, J. Carlos Sancho, M. Ávila & A. Caro, "A review of digital twins and their application in cybersecurity based on artificial intelligence," Artif. Intell. Rev., vol. 57, no. 201, Jul. 2024. DOI: 10.1007/s10462-024-10805-3. [Online]. URL: https://surl.li/asmirk. Accessed: 15 Jun, 2025.
[3] I. Vourganas, A.L. Michala, "Applications of machine learning in cyber security: A review," J. Cybersecurity Privacy, vol. 4, no. 4, pp. 972–992, Dec. 2024. DOI: 10.3390/jcp4040045. [Online]. URL: https://surl.li/vdlxkl. Accessed: 15 Jun, 2025.
[4] F. Sufi and M. Alsulami, "Mathematical modeling and clustering framework for cyber threat analysis across industries," *Mathematics*, vol. 13, no. 4, p. 655, Feb. 2025. DOI: 10.3390/math13040655. [Online]. URL: https://surl.li/fgrjnf. Accessed: 15 Jun, 2025.
[5] Y. Miao, L. Pan, et al., "Machine learning based cyber attacks targeting on controlled information: A survey," arXiv:2102.07969, Feb. 2021. [Online]. URL: https://surl.li/hmyhes. Accessed: 15 Jun, 2025.
[6] J. Vitorino, I. Praça, E. Maia, L. Sousa, and S. Jardim, "A Comparative Analysis of Machine Learning Techniques for IoT Intrusion Detection," arXiv:2111.13149, Nov. 2021. [Online]. URL: https://surli.cc/vlcuxu. Accessed: 15 Jun, 2025.
[7] H. Kavak, J.J. Padilla, D. Vernon-Bido, S.Y. Diallo, R. Gore, and S. Shetty, "Simulation for cybersecurity: State of the art and future directions," *J. Cybersecurity*, vol. 7, no. 1, p. tyab005, Mar. 2021. DOI: 10.1093/cybsec/tyab005. [Online]. URL: https://surl.lu/crguxz. Accessed: 15 Jun, 2025.
[8] XM Cyber, "Seeing what attackers see: How attack graphs help you stay ahead," Jan. 2025. [Online]. URL: https://surli.cc/dfwcym. Accessed: 15 Jun, 2025.
[9] S. Chong, J. Guttman, A. Datta, A. Myers, B. Pierce, P. Schaumont, T. Sherwood, N. Zeldovich, "Report on the NSF Workshop on Formal Methods for Security,"

Т3. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344520

arXiv:1608.00678, Aug. 2016. [Online]. URL: https://surl.gd/sgrcyr. Accessed: 15 Jun, 2025.

[10] P.Y.A. Ryan, "Mathematical Models of Computer Security," in Foundations of Security Analysis and Design (FOSAD 2000), vol. 2171, Lecture Notes in Computer Science, R. Focardi and R. Gorrieri, Eds. Berlin, Germany: Springer, 2001, pp. 1–62. DOI: 10.1007/3-540-45608-2_1. [Online]. URL: https://surl.li/egqhjf. Accessed: 15 Jun, 2025.

[11] F.S. Passino, N.M. Adams, E.A.K. Cohen, M. Evangelou, and N.A. Heard, "Statistical Cybersecurity: A Brief Discussion of Challenges, Data Structures, and Future Directions," Harvard Data Science Review, vol. 5, no. 1, Mar. 2023. DOI: 10.1162/99608f92.240383c7. [Online]. URL: https://surli.cc/foscla. Accessed: 15 Jun, 2025.

[12] P. Mahadevappa, S.M. Muzammal, and R.K. Murugesan, "A Comparative Analysis of Machine Learning Algorithms for Intrusion Detection in Edge-Enabled IoT Networks," arXiv:2111.01383, Nov. 2021. [Online]. URL: https://surl.li/zwaisx. Accessed: 15 Jun, 2025.

[13] K. Mohamed, "Machine learning techniques to address cybersecurity," arXiv:2302.12415, Feb. 2023. [Online]. URL: https://surl.lu/tdmavn. Accessed: 15 Jun, 2025.

[14] D. Garton, "Purdue model framework for industrial control systems & cybersecurity segmentation," U.S. Dept. Energy/National Petroleum Council, Topic Paper 4–14, Jul. 2019. [Online]. URL: https://surli.cc/fmkjby. Accessed: 15 Jun, 2025.

[15] I. Sayar, N. Messe, S. Ebersold, and J.M. Bruel, "From What to How: A Taxonomy of Formalized Security Properties," arXiv:2505.14514, May 2025. [Online]. URL: https://surl.lu/nwovtz. Accessed: 15 Jun, 2025.

[16] Wikipedia, "Attack trees," Jan. 2025. [Online]. URL: https://surl.li/dbzvmq. Accessed: 15 Jun, 2025.

[17] Picus Security, "What is attack simulation," Jan. 2025. [Online]. URL: https://surl.li/cifqpc. Accessed: 15 Jun, 2025.

[18] Pynetlabs, "Top 10 network simulation tools," Jan. 2025. [Online]. URL: https://surl.li/qesago. Accessed: 15 Jun, 2025.

[19] Hoxhunt, "Cyber Security Simulation Training: How it Works + Best Practices," Jan. 2025. [Online]. URL: https://surl.li/cbjkeq. Accessed: 15 Jun, 2025.

[20] Picus Security, "Everything You Need to Know About BAS Tools," Jan. 2025. [Online]. URL: https://surli.cc/exllfj. Accessed: 15 Jun, 2025.

[21] Softprom, "Vulnerability management vs. penetration testing vs. breach and attack simulation," Jan. 2025. [Online]. URL: https://surl.li/afkwyc. Accessed: 15 Jun, 2025.

[22] CyberProof, "Penetration Testing vs Breach and Attack Simulator: Key Differences and Why It Matters," Jan. 2025. [Online]. URL: https://surl.li/nxbard. Accessed: 15 Jun, 2025.

[23] SCYTHE, "Top 10 breach and attack simulation (BAS) tools," Jan. 2025. [Online]. URL: https://surli.cc/ujxueh. Accessed: 15 Jun, 2025.

[24] A. Vestad and B. Yang, "A survey of agent-based modeling for cybersecurity," in *AHFE Open Access Proceedings*, vol. 127, pp. 83–93, Jul. 2024. [Online]. URL: https://surl.lu/wcoxjx. Accessed: 15 Jun, 2025.

[25] F.K. Batista, A.M. del Rey, and A. Queiruga-Dios, "A New Individual-Based Model to Simulate Malware Propagation in Wireless Sensor Networks," Mathematics, vol. 8, no. 3, p. 410, Mar. 2020. DOI: 10.3390/math8030410. [Online]. URL: https://surl.li/yvuqfp. Accessed: 15 Jun, 2025.

[26] J. Soule, J. Jamont, M. Occello, P. Theron and L. Traonouez, "Towards a multi-agent simulation of cyber-attackers and cyber-defenders battles," arXiv:2506.04849, Jun. 2025. [Online]. URL: https://surl.li/jrkiup. Accessed: 15 Jun, 2025.

[27] A. Bose, and K. Shin, "Agent-based modeling of malware dynamics in heterogeneous environments," *Security and Communication Networks*, vol. 6, no. 12, pp. 1–8, Dec. 2013. [Online]. URL: https://surl.li/akjovc. Accessed: 15 Jun, 2025.

[28] E.C. Balta, M. Pease, J. Moyne, K. Barton, and D.M. Tilbury, "Digital Twin-Based Cyber-Attack Detection Framework for Cyber-Physical Manufacturing Systems," IEEE Transactions on Automation Science and Engineering, vol. 21, no. 3, pp. 3245–3260, Jul. 2024. DOI: 10.1109/TASE.2023.3243147.

[29] K. Shaukat, S. Luo, V. Varadharajan, I.A. Hameed, S. Chen, D. Liu, and J. Li, "Performance Comparison and Current Challenges of Using Machine Learning Techniques in Cybersecurity," Energies, vol. 13, no. 10, p. 2509, May 2020. DOI: 10.3390/en13102509.

[30] CrowdStrike, "CrowdStrike machine learning in cybersecurity," Jan. 2025. [Online]. URL: https://surl.li/yfqmjm. Accessed: 15 Jun, 2025.

[31] G. Apruzzese, P. Laskov, E. Montes de Oca, W. Mallouli, L. Brdalo Rapa, A.V. Grammatopoulos, and F. Di Franco, "The Role of Machine Learning in Cybersecurity," Digital Threats: Research and Practice, vol. 4, no. 1, Art. no. 8, pp. 1–38, Mar. 2023. DOI: 10.1145/3545574. [Online]. URL: https://surl.li/rpjhxe. Accessed: 15 Jun, 2025.

[32] P. Maniriho, A.N. Mahmood, and M.J.M. Chowdhury, "Deep learning models for detecting malware attacks," arXiv:2209.03622, Jan. 2024. [Online]. URL: https://surl.li/kvdwjl. Accessed: 15 Jun, 2025.

[33] M.N. Uddin, Y. Zhang, and X. Hei, "Deep learning aided software vulnerability detection: A survey," arXiv:2503.04002, Mar. 2025. [Online]. URL: https://surl.li/njcajz. Accessed: 15 Jun, 2025.

[34] M. Hesham, M. Essam, M. Bahaa, A. Mohamed, M. Gomaa, M. Hany, and W. Elsersy, "Evaluating Predictive Models in Cybersecurity: A Comparative Analysis of Machine and Deep Learning Techniques for Threat Detection," arXiv:2407.06014, Jul. 2024. [Online]. URL: https://surl.li/grzpge. Accessed: 15 Jun, 2025.

[35] B. Collins, S. Xu, and P.N. Brown, "Game-Theoretic Cybersecurity: the Good, the Bad and the Ugly,"

ТЗ. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344520

arXiv:2401.13815, Jan. 2024. [Online]. URL: https://surl.li/jkxpct. Accessed: 15 Jun, 2025.

[36] L. Zhang, Q. Zhu, and Y. Xie, “Improving network threat detection by knowledge graph, large language model, and imbalanced learning,” arXiv:2501.16393, May. 2025. [Online]. URL: https://surli.cc/bqlnzd. Accessed: 15 Jun, 2025.

[37] PuppyGraph, “Cyber Graph: Enhancing Cybersecurity with Graph Intelligence,” Jan. 2025. [Online]. URL: https://surl.li/yfugrw. Accessed: 15 Jun, 2025.

[38] X. Zhang, “Graph neural networks in network security: From theoretical foundations to applications,” in *Proc. Americas Conf. Inf. Syst. (AMCIS)*, Aug. 2025, pp. 1–10. [Online]. URL: https://surli.cc/wpkgob. Accessed: 15 Jun, 2025.

[39] V. Khurana and N. Kumar, “Graph Neural Networks for Cybersecurity Applications in Network Intrusion and Vulnerability Analysis,” in RADemics Research Institute, Chapter 8, Jan. 2025. [Online]. URL: https://surl.li/tchiff. Accessed: 15 Jun, 2025.

[40] Bolster.ai, “Large Language Models for Cybersecurity: The Role of LLMs in Threat Hunting,” Apr. 2025. [Online]. URL: https://surl.li/olyczz. Accessed: 15 Jun, 2025.

[41] H. Xu, and et al., “Large language models for cyber security: A systematic literature review,” arXiv:2405.04760, Jul 2024. [Online]. URL: https://surl.li/coqqui. Accessed: 15 Jun, 2025.

[42] Z. Liu, “Multi-Agent Collaboration in Incident Response with Large Language Models,” arXiv:2412.00652, Dec. 2024. [Online]. URL: https://surl.li/aghvvq. Accessed: 15 Jun, 2025.

[43] E. Mezzi, F. Massacci, and K. Tuma, “Large language models are unreliable for cyber threat intelligence,” arXiv:2503.23175, Mar. 2025. [Online]. URL: https://surl.lu/xonjqn. Accessed: 15 Jun, 2025.

[44] V. Kravchuk, N. Maslova, and I. Dorohyi, “Automated xss vulnerability detection in web applications based on a multi-agent approach”, SP DonNTU. OTA, vol. 3, no. 4 (36), pp. 13–26, May 2025.

Стаття надійшла до редакції 20.09.2025
Стаття прийнята 05.10.2025
Статтю опубліковано 02.12.2025

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344522

УДК 004.056.5:004.021

КОМПАРАТИВНИЙ АНАЛІЗ МЕТОДІВ ТА ТЕХНОЛОГІЙ МОДЕЛЮВАННЯ ПЕНТЕСТИНГУ

В.С. Кравчук[1], Я.Ю. Дорогий[1, 2, 3]
[1] Department of Applied Mathematics and Informatics, Donetsk National Technical University, Drohobych, Ukraine
[2] Department of Information Systems and Technologies, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
[3] Department of Intelligent Software Systems, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

E-mail: naukovipracidntu@gmail.com

АНОТАЦІЯ

У статті здійснено детальний компаративний аналіз сучасних методів та технологій моделювання пентестингу (тестування на проникнення), що є фундаментальним аспектом забезпечення кібербезпеки в цифровому світі. Автори простежують еволюцію цих підходів: від класичних ручних технік, які потребують високої кваліфікації фахівців, до інноваційних автоматизованих систем, що інтегрують штучний інтелект (ШІ) та машинне навчання (МН). Зокрема, порівнюються різні моделі симуляції вразливостей, а саме популярний Metasploit Framework для емуляції експлойтів, віртуалізовані середовища на базі VirtualBox, VMware чи контейнеризації Docker, які дають змогу створювати ізольовані тестові мережі для імітації реальних атак. Окрему увагу приділено гібридним технологіям, що поєднують традиційні інструменти з AI-алгоритмами для прогнозування й автоматизації атак, наприклад, за допомогою бібліотек TensorFlow, PyTorch або пакетів Scapy для генерації мережевого трафіку. Аналіз проводиться за ключовими критеріями ефективності: точність виявлення вразливостей (з урахуванням хибнопозитивних та хибнонегативних результатів), швидкість виконання тестів, масштабованість для великих систем, витрати обчислювальних ресурсів, рівень помилок і простота інтеграції. Обговорюються переваги кожного методу – наприклад, ручні методи забезпечують глибоке розуміння контексту, тоді як AI-підходи дають змогу обробляти великі обсяги даних у реальному часі – та їхні недоліки, як-от вразливість до загроз, що еволюціонують, чи потреба в постійному навчанні моделей. Особливий акцент робиться на адаптації цих технологій до сучасних сценаріїв, включно із хмарними платформами (AWS, Microsoft Azure, Google Cloud), інтернетом речей (IoT-пристрої з обмеженими ресурсами) та мобільними додатками. Дослідження ґрунтується на емпіричних даних із тестувань на стандартизованих моделях, зокрема OWASP Top 10 для вебвразливостей та NIST Cybersecurity Framework, де показано, що гібридні методи підвищують загальну ефективність на 30–50 % порівняно з традиційними, зменшуючи час на виявлення вразливостей і мінімізуючи ризики. Автори пропонують практичні рекомендації щодо вибору оптимальних технологій для різних типів організацій – від малого бізнесу до великих корпорацій – з урахуванням етичних аспектів (наприклад, дотримання принципів етичного хакінгу), регуляторних вимог (GDPR для захисту даних, ISO 27001 для управління інформаційною безпекою) і потенційних ризиків, як-от несанкціоноване використання інструментів. Стаття є цінним ресурсом для фахівців з кібербезпеки, розробників програмного забезпечення, менеджерів IT-проєктів і дослідників, сприяючи розробці більш стійких стратегій захисту від кіберзагроз у динамічному середовищі цифрових технологій.

Ключові слова: *пентестинг, моделювання вразливостей, компаративний аналіз, штучний інтелект, кібербезпека, Metasploit Framework, OWASP, машинне навчання, віртуалізація, етичний хакинг, TensorFlow, Scapy, хмарні системи, IoT-пристрої.*

Вступ

Зростання кількості кібератак, як-от фішинг, ransomware та витоки даних, свідчить про вразливість сучасних систем, роблячи питання кібербезпеки критичним для бізнесу, державного управління та повсякденного спілкування. За даними звіту Verizon Data Breach Investigations Report 2025, у регіоні APAC System Intrusion становить 83 % порушень, а Ransomware – 51 % [1]. Крім того, глобальні збитки від кіберзлочинності очікуються на рівні $10,5 трлн щорічно до 2025 року, з потенційним зростанням до $12,2 трлн до 2031 року [2]. У цьому контексті

ISSN 2786-9024

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344522

пентестинг (penetration testing), або тестування на проникнення, стає невід'ємним інструментом для виявлення й усунення потенційних загроз до їхньої реалізації зловмисниками. Пентестинг імітує дії хакерів, даючи змогу організаціям оцінити рівень захисту своїх систем і розробити ефективні стратегії оборони.

Однак традиційні методи пентестингу, які базуються на ручному аналізі та стандартних інструментах, часто стикаються з обмеженнями: вони потребують значних часових і людських ресурсів, не завжди масштабовані для великих мереж і не здатні оперативно реагувати на нові типи атак, що еволюціонують завдяки ШІ та МН. Згідно з CompTIA's State of Cybersecurity 2025 Report, навички ШІ є головним викликом для прискорення впровадження ШІ в кібербезпеці: 70 % компаній перебувають на ранніх етапах освіти чи тестування [3]. З появою хмарних технологій, IoT та розподілених систем, як-от AWS чи Azure, моделювання пентестингу потребує інноваційних підходів. Саме тому компаративний аналіз методів та технологій моделювання пентестингу є актуальним завданням, що дає змогу систематизувати знання та визначити оптимальні рішення для різних сценаріїв, враховуючи тренди, зокрема AI-powered атаки, які стали головною турботою для 80 % CISO [4].

Цю статтю присвячено порівняльному огляду сучасних методів моделювання пентестингу, від класичних симуляцій вразливостей до гібридних систем на базі ШІ. Розглянемо еволюцію цих технологій, починаючи з інструментів на кшталт Metasploit Framework, які дають змогу емулювати експлойти в контрольованому середовищі, до віртуалізованих платформ (VirtualBox, Docker), що створюють ізольовані тестові мережі. Особливу увагу буде приділено інтеграції МН, наприклад, через бібліотеки TensorFlow чи Scapy, для автоматизованого прогнозування атак і генерації сценаріїв. Аналіз базуватиметься на критеріях ефективності: точність, швидкість, масштабованість, витрати ресурсів та рівень помилок.

Метою дослідження є не лише порівняння переваг і недоліків кожного методу, але й надання практичних рекомендацій для впровадження в реальних умовах, з урахуванням етичних норм (етичний хакинг) і регуляторних стандартів (GDPR, ISO 27001). Для обґрунтування висновків використано емпіричні дані з тестувань на моделях OWASP та NIST, які демонструють, як гібридні підходи можуть підвищити ефективність на 30–50 %.

Постановка проблеми дослідження

Проблема дослідження полягає у відсутності систематичного компаративного аналізу методів та технологій моделювання пентестингу, що ускладнює вибір оптимальних підходів для різних сценаріїв – від традиційних симуляцій до гібридних AI-систем. Існуючі дослідження фокусуються на окремих техніках, але ігнорують критерії ефективності (точність, швидкість, масштабованість, витрати, помилки) та адаптацію до загроз, що еволюціонують, як-от AI-генеровані атаки чи вразливості в IoT.

З огляду на вищенаведене метою цього дослідження є надання детального компаративного аналізу методів і технологій моделювання пентестингу, застосовних у кібербезпеці, з акцентом на їхню ефективність, сфери використання (від хмарних систем до IoT-пристроїв) та потенціал інтеграції з штучним інтелектом і машинним навчанням.

Для досягнення цієї мети систематизовано традиційні (ручні та симуляційні) та сучасні (гібридні та AI-орієнтовані) підходи, проаналізовано їхні особливості (наприклад, точність виявлення вразливостей, швидкість і масштабованість), проведено порівняльну оцінку на основі емпіричних даних із моделей OWASP та NIST, а також визначено ключові виклики (етичні ризики, регуляторні обмеження) і перспективи розвитку (автоматизація та адаптація до загроз, що еволюціонують).

Матеріали та методи дослідження

Дослідження зосереджено на порівнянні методів та технологій моделювання пентестингу в кібербезпеці, використовуючи наукові публікації, технічні звіти й огляди з авторитетних джерел, як-от IEEE, Springer, MDPI, arXiv, а також матеріали організацій, зокрема OWASP, NIST, XM Cyber і Picus Security, до 2025 року. Аналіз охопив літературу про симуляційне моделювання (наприклад, Metasploit Framework), віртуалізацію (VirtualBox, Docker), агентно-орієнтоване моделювання, машинне навчання (TensorFlow, Scapy), графові моделі та великі мовні моделі. Кейси застосування, зокрема моделі OWASP Top 10, фреймворк NIST Cybersecurity, платформи BAS (SCYTHE, AttackIQ) і графові інструменти (Neo4j, Maltego), стали основою для оцінки практичної реалізації методів.

Для аналізу застосовано фундаментальні та прикладні методи. Аналіз і синтез допомогли систематизувати характеристики методів, порівняльний аналіз зіставив традиційні (ручні та симуляційні) та новітні (гібридні та AI-орієнтовані) підходи за ефективністю, масштабованістю й адаптивністю, а системний підхід розглядав методи як взаємопов'язані елементи кібербезпеки. Огляд літератури та класифікація методів дали змогу зібрати дані про принципи роботи та приклади застосування, тоді як порівняльна оцінка, зокрема BAS проти традиційного тестування на проникнення, виявила їхні сильні та слабкі сторони. Аналіз кейсів (наприклад, використання CrowdStrike для машинного навчання в моделюванні атак чи GALLIUM APT для агентного моделювання) доповнив

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344522

дослідження емпіричними даними з тестувань на моделях OWASP та NIST.

Методологічно дослідження базувалося на об'єктивності й науковій достовірності, використовуючи програмні інструменти для систематизації літератури (Mendeley, Zotero) та обробки текстів. Поєднання фундаментальних методів, зокрема абстрагування для виділення ключових аспектів, із прикладними, як-от аналіз кейсів, забезпечило цілісне розуміння методів моделювання. Особлива увага приділялася синергії підходів, включно з інтеграцією AI для прогнозування вразливостей та автоматизації тестів, що є основою для створення адаптивних систем кіберзахисту.

Аналіз методів і технологій моделювання

У контексті пентестингу моделювання відіграє ключову роль у симуляції атак, виявленні вразливостей та оцінці ефективності захисних механізмів без ризику для реальних систем. Аналіз методів і технологій моделювання дає змогу систематизувати традиційні та сучасні підходи, оцінити їх за критеріями ефективності (точність виявлення, швидкість, масштабованість, витрати ресурсів, рівень помилок), а також визначити перспективи інтеграції з новими технологіями, зокрема ШІ та МН. Цей розділ базується на огляді літератури й емпіричних даних, охоплюючи симуляційні, віртуалізовані, агентно-орієнтовані та гібридні методи, з акцентом на їх застосування в хмарних середовищах (AWS, Azure), IoT-пристроях і розподілених мережах.

Традиційні методи моделювання пентестингу зосереджені на ручному або напівавтоматизованому імітації атак, використовуючи інструменти для симуляції вразливостей у контрольованому середовищі. Одним із ключових є симуляційне моделювання на базі фреймворків, а саме Metasploit Framework, який дає змогу емулювати експлойти, сканування мереж та етапи атаки за моделлю Cyber Kill Chain (від розвідки до ексфільтрації даних). Metasploit інтегрує модулі для моделювання вразливостей з бази CVE, забезпечуючи точність на рівні 80–90 % для відомих загроз, але потребує значних ресурсів для налаштування [5–6].

Віртуалізовані середовища, наприклад VirtualBox або Docker, створюють ізольовані тестові мережі для моделювання реальних сценаріїв. Docker контейнери дають змогу швидко розгортати віртуальні машини з вразливими сервісами, забезпечуючи масштабованість для мереж до 100 вузлів, але обмежені в динамічних змінах (наприклад, адаптація до MTD – Moving Target Defense) [7]. Ці методи ефективні для статичних сценаріїв, з швидкістю виконання 1–5 годин на тест, але мають високий рівень помилок (до 20 % хибнопозитивних) через відсутність автоматизації [8].

Переваги традиційних методів у глибокому розумінні контексту та адаптації до конкретних систем, але недоліки – це низька масштабованість і залежність від експертів, що робить їх менш придатними для великих мереж або швидких загроз.

Сучасні підходи інтегрують ШІ та МН для автоматизації моделювання, підвищуючи ефективність на 30–50 %. Агентно-орієнтоване моделювання (ABM) представляє атакерів та захисників як автономні агенти, симулюючи взаємодії в динамічних середовищах. Наприклад, фреймворки на базі reinforcement learning (RL), як-от DQN або POMDP (Partially Observable Markov Decision Processes), моделюють атаки як стохастичні процеси, де стан мережі описується графом G = (V, E, X) з вузлами (пристрої), ребрами (з'єднання) та атрибутами (вразливості, CVSS-бали) [9].

Інструменти на кшталт TensorFlow або Scapy генерують мережевий трафік для прогнозування атак, надаючи можливість симулювати до 1000 вузлів з динамічними змінами (p_change до 0.5), зменшуючи час тестування до хвилин [10].

Гібридні технології, зокрема Breach and Attack Simulation (BAS) платформи (SCYTHE, AttackIQ), поєднують симуляцію з реальними даними, тестуючи весь cyber kill chain. BAS автоматизує атаки на багатошарових захистах (NGFW, EDR, WAF), забезпечуючи безперервне тестування з низьким ризиком, на відміну від ручних пентестів [11].

Графові моделі (Neo4j, Maltego) візуалізують шляхи атак, інтегруючи MITRE ATT&CK для моделювання тактик, з точністю виявлення 95 % для відомих векторів [12]. Великі мовні моделі (LLM) додають інтелекту, генеруючи сценарії атак для IoT, з фокусом на етичні аспекти.

Компаративний аналіз. Компаративний аналіз методів і технологій моделювання пентестингу проводиться з метою об'єктивного порівняння традиційних, сучасних та гібридних підходів. Для оцінки введено критерії порівняння із числовою шкалою від 1 до 10, де 1 означає найнижчий рівень (наприклад, низька точність, висока витрата ресурсів), а 10 – найвищий (висока точність, низька витрата). Критерії базуються на емпіричних даних з літератури та передбачають:

- *точність виявлення вразливостей* – оцінює, наскільки метод правильно ідентифікує реальні загрози без хибнопозитивних / хибнонегативних результатів (вища оцінка – вища точність);
- *швидкість виконання* – вимірює час на тестування;
- *масштабованість* – можливість застосування до великих мереж (вища оцінка – краща адаптація до 1000+ вузлів);
- *витрати ресурсів* – обчислювальні, людські та фінансові витрати (вища оцінка – нижчі витрати);

ISSN 2786-9024

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344522

– рівень помилок – частота помилок (вища оцінка – нижчий рівень помилок).

Оцінки присвоєно на основі аналізу джерел, як-от порівняльні дослідження методологій пентестингу [13] та оглядів інструментів для моделювання [5]. Методи згруповано за типами: традиційні (наприклад, Metasploit Framework, VirtualBox), сучасні (агентно-орієнтовані з RL, Scapy), гібридні (BAS-платформи як SCYTHE, AttackIQ з інтеграцією AI) (див. таблицю 1).

Традиційні методи, а саме Metasploit Framework для симуляції експлойтів та VirtualBox для віртуалізованих мереж, демонструють високу точність (8/10) завдяки ручному контролю та фокусу на відомих вразливостях з бази CVE, що зменшує хибнопозитивні результати [14]. Однак їхня швидкість низька (4/10), оскільки тести потребують часу на налаштування та виконання, особливо для статичних сценаріїв. Масштабованість обмежена (5/10) до середніх мереж (до 100 вузлів) через залежність від ручної конфігурації, а витрати ресурсів середні (6/10) з потребою у кваліфікованих фахівцях. Рівень помилок прийнятний (7/10), але зростає в динамічних середовищах, як-от хмарні системи AWS чи Azure, де відсутня автоматизація призводить до пропуску загроз, що еволюціонують.

Сучасні методи, включно з агентно-орієнтованим моделюванням з reinforcement learning (RL) та інструментами на кшталт Scapy для генерації трафіку, виділяються високою швидкістю (8/10), надаючи реальний час обробки (хвилини) завдяки автоматизації прогнозування атак. Масштабованість висока (9/10), оскільки RL-моделі (наприклад, DQN) адаптуються до великих мереж (1000+ вузлів) з динамічними змінами, як у IoT-пристроях. Точність дещо нижча (7/10) через можливі упередження в навчанні моделей, а витрати ресурсів вищі (5/10) через потребу в GPU для TensorFlow чи PyTorch. Рівень помилок середній (6/10), із ризиком хибнонегативних у невідомих сценаріях, але це компенсується адаптивністю до AI-powered атак.

Гібридні методи, а саме Breach and Attack Simulation (BAS) з інтеграцією AI (наприклад, AttackIQ з TensorFlow), досягають найвищої точності (9/10) завдяки комбінації симуляції та реальних даних, зменшуючи помилки до 5–10 %. Швидкість висока (9/10) з автоматизацією всього cyber kill chain, масштабованість добра (8/10) для розподілених систем, витрати ресурсів оптимальні (7/10) з балансом між автоматизацією та людським наглядом. Рівень помилок низький (8/10), особливо в хмарних та IoT-сценаріях, де гібридні підходи підвищують ефективність на 30–50 % порівняно з традиційними, як показано в тестах OWASP та NIST [15].

Загалом гібридні методи перевершують інші в більшості критеріїв, роблячи їх оптимальними для сучасних організацій, тоді як традиційні підходять для невеликих, статичних систем. Аналіз підкреслює необхідність інтеграції AI для адаптації до загроз, що еволюціонують, з рекомендаціями щодо вибору залежно від розміру організації та типу інфраструктури.

Обговорення та перспективи розвитку дослідження

У розділі аналізу методів і технологій моделювання пентестингу було продемонстровано, що гібридні підходи, які поєднують традиційні симуляції (наприклад, Metasploit Framework) з елементами ШІ та МН, перевершують класичні методи за критеріями ефективності, швидкості та масштабовальності. Емпіричні дані з тестувань на моделях OWASP та NIST підтверджують підвищення ефективності на 30–50 %, особливо в хмарних середовищах (AWS, Azure) та системах IoT, де динамічні загрози потребують адаптивності. Однак традиційні методи зберігають перевагу в глибокому контекстному розумінні, тоді як сучасні AI-орієнтовані інструменти (наприклад, TensorFlow чи Scapy) схильні до упереджень у моделях і потребують постійного навчання, що може призводити до хибнопозитивних результатів у 10–20 % випадків [16]. Обмеження дослідження стосуються фокуса на теоретичних моделях без широкого емпіричного тестування в реальних організаціях, а також ігнорування факторів, як-от вартість впровадження, яка для гібридних систем може сягати 20–30 % від бюджету на кібербезпеку. Це підкреслює необхідність

Табл. 1. Компаративний аналіз методів

Метод / критерій	Точність	Швидкість	Масштабованість	Витрати ресурсів	Рівень помилок
Традиційні (Metasploit, VirtualBox)	8	4	5	6	7
Сучасні (RL-агенти, Scapy)	7	8	9	5	6
Гібридні (BAS з AI, TensorFlow)	9	9	9	7	8

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344522

балансу між автоматизацією та людським наглядом, щоб уникнути етичних ризиків, як-от несанкціоноване використання інструментів, відповідно до стандартів GDPR та ISO 27001.

Перспективи розвитку дослідження пов'язані з інтеграцією передових технологій, що формують майбутнє кібербезпеки. Згідно з прогнозами на 2025 рік, ключовим трендом стане використання генеративного ШІ (genAI) для автоматизованого створення сценаріїв атак, що дасть змогу моделювати складні загрози, зокрема AI-powered фішинг чи інсайдерські атаки, з ростом ефективності на 30 % [17]. Гібридний пентестинг, що поєднує людських експертів з AI-інструментами (наприклад, PentestGPT чи DeepExploit), набуде поширення, особливо в безперервному тестуванні, інтегрованому з CI/CD-пайплайнами, для оперативного реагування на загрози, що еволюціонують. У контексті IoT та хмарних систем перспективним є розвиток агентно-орієнтованого моделювання з елементами теорії ігор, де атаки моделюються як стохастичні процеси, з використанням LLM для генерації динамічних векторів. Крім того, загрози від shadow AI – несанкціонованих моделей в організаціях – потребуватимуть нових методів моделювання для оцінки ризиків з акцентом на етичні аспекти й регуляторні вимоги.

Майбутні дослідження повинні фокусуватися на інтеграції квантових обчислень для моделювання постквантових атак, а також на розробці стандартів для AI в пентестингу, щоб зменшити залежність від людських ресурсів і підвищити доступність для малого бізнесу. Очікується зростання ролі автоматизованих платформ BAS (наприклад, SCYTHE чи AttackIQ) з елементами ML, що скоротить час тестування до реального часу та зменшить втому від сповіщень (alert fatigue). Загалом перспективи розвитку спрямовані на створення адаптивних, стійких систем кіберзахисту, що інтегрують ШІ для проактивного виявлення загроз, з потенціалом зростання ринку пентестингу на 15–20 % щорічно до 2030 року [18]. Це дослідження може слугувати основою для подальших емпіричних робіт, спрямованих на практичне впровадження гібридних моделей у різних галузях.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно цього дослідження, у тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

ЛІТЕРАТУРА

[1] Verizon, "2025 Data Breach Investigations Report," Verizon, 2025. [Онлайн]. URL: https://surl.li/qsfajk. Дата звернення: 25.09.2025.

[2] Cybersecurity Ventures, "Cybercrime To Cost The World $10.5 Trillion Annually By 2025," Apr. 2025. [Онлайн]. URL: https://surl.li/ujmskz. Дата звернення: 25.09.2025.

[3] CompTIA, "State of Cybersecurity 2025," 2025. [Онлайн]. URL: https://surl.lu/mjjnfk. Дата звернення: 25.09.2025.

[4] A. Angner et al., "AI Creates New Cyber Risks. It Can Help Resolve Them, Too," Boston Consulting Group, Jul. 2025. [Онлайн]. URL: https://surl.li/jybvzv. Дата звернення: 25.09.2025.

[5] M. Khalil, "Penetration Testing Methodology 2025: Complete Guide," DeepStrike, Aug. 2025. [Онлайн]. URL: https://surl.li/lossfh. Дата звернення: 25.09.2025.

[6] V. Kravchuk, N. Maslova, and I. Dorohyi, "AUTOMATED XSS VULNERABILITY DETECTION IN WEB APPLICATIONS BASED ON A MULTI-AGENT APPROACH", SP DonNTU. OTA, vol. 3, no. 4 (36), pp. 13–26, May 2025.

[7] "Simulating Penetration Testing Using a Modeling Framework," Assurant Cyber, Aug. 2023. [Онлайн]. URL: https://surl.li/hqgzow. Дата звернення: 25.09.2025.

[8] "Vulnerability Assessment vs Penetration Testing 2025," DeepStrike, May 2025. [Онлайн]. URL: https://surl.li/fjiytg. Дата звернення: 25.09.2025.

[9] X. Wang et al., "A Unified Modeling Framework for Automated Penetration Testing," arXiv, Feb. 2025. [Онлайн]. URL: https://surl.li/fixacj. Дата звернення: 25.09.2025.

[10] "Simulation in Cybersecurity: Understanding Techniques, Applications, and Goals," ResearchGate, Aug. 2025. [Онлайн]. URL: https://surl.li/pddbjh. Дата звернення: 25.09.2025.

[11] "Breach and Attack Simulation vs. Penetration Testing," Picus Security, May 2025. [Онлайн]. URL: https://surl.li/lficzz. Дата звернення: 25.09.2025.

[12] "Comparative analysis of penetration testing approaches for IoT devices," ResearchGate, Jun. 2024. [Онлайн]. URL: https://surl.li/hnpome. Дата звернення: 25.09.2025.

[13] A Comparative Study of Penetration Testing Methodologies and Tool Utilization in Cybersecurity, ResearchGate, Aug. 2025. [Онлайн]. URL: https://surl.li/zyinek. Дата звернення: 25.09.2025.

[14] Penetration Testing in 2025: Methods, Technologies & Practices, CyCognito. [Онлайн]. URL: https://surl.li/jqugpc. Дата звернення: 25.09.2025.

[15] Penetration Testing ROI: Executive Guide 2025, Redbot Security. [Онлайн]. URL: https://surl.li/ugqzas. Дата звернення: 25.09.2025.

[16] "A Comparative Study of Penetration Testing Methodologies and Tool Utilization in Cybersecurity," ResearchGate, Aug. 2025. [Онлайн]. URL: https://surl.li/oammic. Дата звернення: 25.09.2025.

ISSN 2786-9024

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344522

[17] "Pentesting Statistics 2025: Key Insights and Emerging Trends," ZeroThreat, Jul. 2025. [Онлайн]. URL: https://surl.lt/tlejih. Дата звернення: 25.09.2025.

[18] "10 Cyber Security Trends For 2025," SentinelOne, Aug. 2025. [Онлайн]. URL: https://surl.li/yclkhx. Дата звернення: 25.09.2025.

COMPARATIVE ANALYSIS OF METHODS AND TECHNOLOGIES FOR PENETRATION TESTING MODELING

Vitaly Kravchuk, Iaroslav Dorohyi

The article conducts a detailed comparative analysis of contemporary methods and technologies for modeling penetration testing (pentesting), a fundamental aspect of ensuring cybersecurity in the digital world. The authors trace the evolution of these approaches: from classical manual techniques that require high expertise from specialists to innovative automated systems integrating artificial intelligence (AI) and machine learning (ML). Specifically, various vulnerability simulation models are compared, such as the popular Metasploit Framework for exploit emulation, virtualized environments based on VirtualBox, VMware, or containerization with Docker, which enable the creation of isolated test networks for simulating real attacks. Special attention is given to hybrid technologies that combine traditional tools with AI algorithms for attack prediction and automation, for example, using libraries like TensorFlow, PyTorch, or Scapy packages for generating network traffic. The analysis is performed based on key efficiency criteria: accuracy in vulnerability detection (considering false positives and false negatives), test execution speed, scalability for large systems, computational resource costs, error rates, and ease of integration. The advantages of each method are discussed–for instance, manual methods provide deep contextual understanding, while AI approaches enable real-time processing of large data volumes–and their disadvantages, such as vulnerability to evolving threats or the need for continuous model training. Particular emphasis is placed on adapting these technologies to modern scenarios, including cloud platforms (AWS, Microsoft Azure, Google Cloud), Internet of Things (IoT devices with limited resources), and mobile applications. The research is grounded in empirical data from tests on standardized models, such as OWASP Top 10 for web vulnerabilities and NIST Cybersecurity Framework, where it is shown that hybrid methods increase overall efficiency by 30-50% compared to traditional ones, reducing vulnerability detection time and minimizing risks. The authors offer practical recommendations for selecting optimal technologies for different types of organizations–from small businesses to large corporations–considering ethical aspects (e.g., adherence to ethical hacking principles), regulatory requirements (GDPR for data protection, ISO 27001 for information security management), and potential risks, such as unauthorized tool usage. The article serves as a valuable resource for cybersecurity professionals, software developers, IT project managers, and researchers, contributing to the development of more resilient strategies for protection against cyber threats in a dynamic digital technology environment.

Keywords: *pentesting, vulnerability modeling, comparative analysis, artificial intelligence, cybersecurity, Metasploit Framework, OWASP, machine learning, virtualization, ethical hacking, TensorFlow, Scapy, cloud systems, IoT devices.*

REFERENCES

[1] Verizon, "2025 Data Breach Investigations Report," Verizon, 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/qsfajk.

[2] Cybersecurity Ventures, "Cybercrime To Cost The World $10.5 Trillion Annually By 2025," Apr. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/ujmskz.

[3] CompTIA, "State of Cybersecurity 2025," 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.lu/mjjnfk.

[4] A. Angner et al., "AI Creates New Cyber Risks. It Can Help Resolve Them, Too," Boston Consulting Group, Jul. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/jybvzv.

[5] M. Khalil, "Penetration Testing Methodology 2025: Complete Guide," DeepStrike, Aug. 2025. [OnlineAccessed: Sep. 25, 2025]. URL: https://surl.li/lossfh.

[6] V. Kravchuk, N. Maslova, and I. Dorohyi, "AUTOMATED XSS VULNERABILITY DETECTION IN WEB APPLICATIONS BASED ON A MULTI-AGENT APPROACH", SP DonNTU. OTA, vol. 3, no. 4 (36), pp. 13–26, May 2025.

[7] "Simulating Penetration Testing Using a Modeling Framework," Assurant Cyber, Aug. 2023. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/hqgzow.

[8] "Vulnerability Assessment vs Penetration Testing 2025," DeepStrike, May 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/fjiytg.

[9] X. Wang et al., "A Unified Modeling Framework for Automated Penetration Testing," arXiv, Feb. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/fixacj.

[10] "Simulation in Cybersecurity: Understanding Techniques, Applications, and Goals," ResearchGate, Aug. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/pddbjh.

[11] "Breach and Attack Simulation vs. Penetration Testing," Picus Security, May 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/lflczz.

[12] "Comparative analysis of penetration testing approaches for IoT devices," ResearchGate, Jun. 2024. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/hnpome.

ТЗ. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344522

[13] A Comparative Study of Penetration Testing Methodologies and Tool Utilization in Cybersecurity, ResearchGate, Aug. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/zyinek.

[14] Penetration Testing in 2025: Methods, Technologies & Practices, CyCognito. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/jqugpc.

[15] Penetration Testing ROI: Executive Guide 2025, Redbot Security. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/ugqzas.

[16] "A Comparative Study of Penetration Testing Methodologies and Tool Utilization in Cybersecurity," ResearchGate, Aug. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/oammic.

[17] "Pentesting Statistics 2025: Key Insights and Emerging Trends," ZeroThreat, Jul. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.lt/tlejih.

[18] "10 Cyber Security Trends For 2025," SentinelOne, Aug. 2025. [Online]. Accessed: Sep. 25, 2025. URL: https://surl.li/yclkhx.

Стаття надійшла до редакції 30.09.2025
Стаття прийнята 12.10.2025
Статтю опубліковано 02.12.2025

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

Інформаційна безпека

UDC 004.8:004.056

ENHANCED IMAGE STEGANOGRAPHY WITH KERAS-STEGANOGAN: A TENSORFLOW-BASED GAN

D.Yu. Khoma, Y.O. Bashkov
Postgraduate Student, Department of Applied Mathematics and Informatics
Ukraine Donetsk National Technical University, Donetsk, Ukraine

E-mail: dmytro.khoma@donntu.edu.ua

ABSTRACT

Image steganography is the process of embedding secret information within digital images such that the very presence of the message remains undetectable. Recent advances in deep learning, particularly in generative adversarial networks (GANs), have significantly improved both the payload capacity and perceptual quality of steganographic systems. The original SteganoGAN, implemented in Torch, achieved state-of-the-art performance by embedding up to 4.4 bits per pixel while maintaining strong resistance to steganalysis methods. However, the influence of the critic network on steganographic quality and learning stability remains insufficiently explored.

This paper presents Keras-SteganoGAN, a TensorFlow-based reimplementation and extension of SteganoGAN, designed to systematically analyze the role of the critic in adversarial steganographic training. Two variants of the model–one incorporating a critic and one without–were trained and compared across three encoder architectures: basic convolutional, residual, and dense. Each configuration was trained over five epochs with message depths ranging from 1 to 6 bits, allowing a comprehensive study of trade-offs between payload capacity, image distortion, and decoding accuracy.

Quantitative evaluation was conducted using standard image quality and steganographic metrics, including PSNR, SSIM, RS-BPP, and decoder accuracy. The results indicate that the inclusion of a critic improves perceptual quality and visual similarity at lower payloads, but its contribution diminishes as the message depth increases. These findings provide new insights into the interaction between encoder complexity, critic dynamics, and steganographic performance, offering guidance for the design of future GAN-based steganography systems.

Key words: *SteganoGAN, Keras, TensorFlow, image steganography, GAN, encoder, critic network, payload capacity, residual, dense, decoder, adversarial learning, hidden message, similarity metrics.*

Introduction

The art and science of image steganography revolve around embedding secret information within an image so that the presence of the hidden data remains imperceptible to observers. Unlike cryptography, where the primary goal is to secure the contents of the message by making it unreadable to adversaries, steganography goes a step further by ensuring that even the existence of the message is concealed. In typical usage scenarios, a sender encodes a secret message into a cover image, and transmits the image to a receiver who can extract the hidden data. This makes steganography particularly useful in situations where the transmission of encrypted messages might attract attention or raise suspicion.

The fundamental challenge of image steganography is to maximize the amount of data that can be hidden in an image without introducing visible artifacts or detectable anomalies. Steganography should ideally preserve the appearance of the cover image so that even sophisticated analysis tools or trained observers cannot detect alterations. However, this balance between embedding capacity and image quality has proven difficult to achieve. Traditional image steganography techniques, such as least significant bit (LSB) manipulation, can effectively hide small amounts of data but often introduce visible distortions when larger payloads are embedded. Moreover, automated steganalysis tools have become increasingly capable of detecting these modifications, limiting the effectiveness of these conventional approaches.

For a long time, traditional steganographic methods were only able to achieve modest payload capacities, typically up to around 0.4 bits per pixel (bpp) [1]. As payloads increase beyond this threshold, the chances of

Т3. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

introducing detectable artifacts increase, and the cover image becomes more susceptible to analysis by automated steganalysis techniques. These tools can identify subtle inconsistencies in the image, such as irregularities in pixel distributions, that signal the presence of hidden data. In extreme cases, these distortions are noticeable to the human eye, rendering the steganographic technique ineffective for its purpose of secrecy.

In recent years, the rise of deep learning, particularly with the advent of neural networks, has significantly advanced the field of image steganography. Unlike traditional methods, which rely on predefined rules or statistical models to hide information, deep learning approaches have shown a remarkable ability to optimize both the payload capacity and image quality. These neural network-based methods are able to learn intricate patterns within images, making it possible to embed more data while minimizing detectable distortions. This has given rise to a new class of steganographic techniques that leverage generative models, such as generative adversarial networks (GANs), to create more efficient and effective steganographic systems.

Among these advancements is SteganoGAN, a Artificial Intelligence (AI) model developed by [2], which demonstrated the power of GANs in the realm of image steganography. SteganoGAN represents one of the first implementations of a fully end-to-end deep learning-based system for hiding arbitrary binary data in images, rather than simply embedding one image inside another. By employing a GAN architecture [3], SteganoGAN optimizes the perceptual quality of the generated steganographic images while simultaneously enhancing the embedding capacity. With the use of a critic network, the model is trained to produce steganographic images that are virtually indistinguishable from the original cover images, allowing it to evade detection by standard steganalysis tools. The original SteganoGAN achieved a payload capacity of up to 4.4 bits per pixel, a significant improvement over traditional methods.

However, despite these advancements, some aspects of SteganoGAN's design were left unexplored in the original work. One such area is the exact process by which the message tensor is created. In SteganoGAN, messages are encoded using the Reed-Solomon error correction code, which is designed to improve the reliability of message recovery by correcting errors in the decoded message. After encoding, the message is converted to bits and packed into a tensor, with message bits placed sequentially, divided by 32 zero bits. This encoding method improves the error-correction capabilities of the model, but it also adds complexity to the overall system.

In our work, we propose KerasSteganoGAN, a reimplementation of SteganoGAN in TensorFlow, with a key modification: the removal of Reed-Solomon encoding and decoding. By eliminating this step, we streamline the message embedding process while maintaining the core strengths of the GAN-based architecture. The motivation behind this change is to reduce computational overhead and complexity while still achieving effective message recovery. We also investigate the role of the critic network by introducing two versions of KerasSteganoGAN – one with a critic network and one without – allowing us to explore how the inclusion of the critic affects image quality and steganographic performance.

Research Objectives And Tasks

The primary objective of this research is to develop and evaluate Keras-SteganoGAN, a TensorFlow-based reimplementation of the original SteganoGAN model, designed to simplify the message embedding process while maintaining high image quality and embedding capacity. By removing the Reed-Solomon error correction mechanism, the study aims to reduce computational complexity and training overhead without compromising the accuracy of message extraction. This streamlined architecture seeks to demonstrate that robust message recovery and imperceptible image quality can still be achieved through careful network design and optimization within the GAN framework.

A secondary objective of this work is to analyze the influence of the critic network on the overall performance of the model. To this end, two distinct versions of KerasSteganoGAN are implemented: one including a critic component and another operating without it. The comparative analysis between these variants focuses on evaluating their impact on key performance metrics such as payload capacity, decoding accuracy, Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM). This comparison allows for a deeper understanding of how adversarial learning contributes to the balance between embedding fidelity and visual indistinguishability.

Finally, the research aims to empirically validate the effectiveness of KerasSteganoGAN through extensive experiments on benchmark image datasets. The study's tasks include designing and training the models, measuring and comparing their quantitative and qualitative performance, and analyzing trade-offs between computational efficiency and steganographic robustness. The results are expected to provide insights into optimizing GAN-based steganographic architectures for practical use, setting the groundwork for further exploration of adaptive and lightweight deep learning techniques in the field of image steganography.

Research Materials And Methods

SteganoGAN is a groundbreaking that leverages the power of GANs to tackle the challenges of image steganography. Traditional approaches to image

ISSN 2786-9024

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

steganography often suffer from limited payload capacity and the risk of introducing visible artifacts that can be detected by automated steganalysis tools. By using a GAN-based architecture, SteganoGAN overcomes these limitations, significantly enhancing the embedding rate while maintaining high image quality [4].

At its core, SteganoGAN is an end-to-end deep learning model that allows for the hiding of arbitrary binary data inside images, written with Python library called PyTorch. The model's architecture consists of three main components: an encoder, a decoder, and a critic network.

The encoder is responsible for embedding the secret message into the cover image, transforming it into a steganographic image. The decoder, on the other hand, works to recover the hidden message from the steganographic image. The critic network plays the role of a discriminator in the GAN framework, evaluating how close the generated steganographic image is to the original cover image [5]. This adversarial relationship between the generator (encoder) and the critic enables the model to produce steganographic images that are nearly indistinguishable from the original ones, making it difficult for steganalysis tools to detect the presence of hidden data.

One of the key innovations in SteganoGAN is the use of multiple loss functions to optimize the encoder, decoder, and critic simultaneously. By balancing these losses – specifically, the decoding accuracy, the perceptual similarity between the cover and steganographic images, and the realism of the generated image – the model is able to achieve a high payload capacity while maintaining image fidelity.

The architecture of SteganoGAN is carefully designed to balance embedding capacity and image quality. It includes three key components: the encoder, the decoder, and the critic network, which together work within an adversarial training framework and shown on Figure 1. This allows SteganoGAN to hide arbitrary binary data in images while maintaining their visual integrity, like in [6].

Encoder Network

SteganoGAN incorporates three different encoder architectures, each designed to explore different methods of embedding messages into images: the basic encoder, the residual encoder, and the dense encoder. Each of these architectures (Figure 2) handles the feature extraction and message embedding processes differently, which impacts both the quality of the generated steganographic image and the model's ability to recover the hidden message.

Basic Encoder – the the simplest variant, designed with a straightforward convolutional architecture. It starts by processing the cover image through several convolutional layers, which transform the image into a feature map. The binary message is then concatenated to these features and passed through additional convolutional layers. The final output is a steganographic image that contains the embedded message. This architecture is relatively simple and fast to train, but its limitation is that it may not efficiently capture intricate image details or provide the best possible image quality when embedding large amounts of data. The lack of skip connections or advanced feature reuse mechanisms means that the model might struggle with higher message depths or complex images.

Residual Encoder – builds upon the basic encoder by introducing residual connections, a technique originally popularized by ResNet architectures [7]. In this design, after the message is concatenated to the image features and processed through the convolutional layers, the output of the encoder is added to the original cover image.

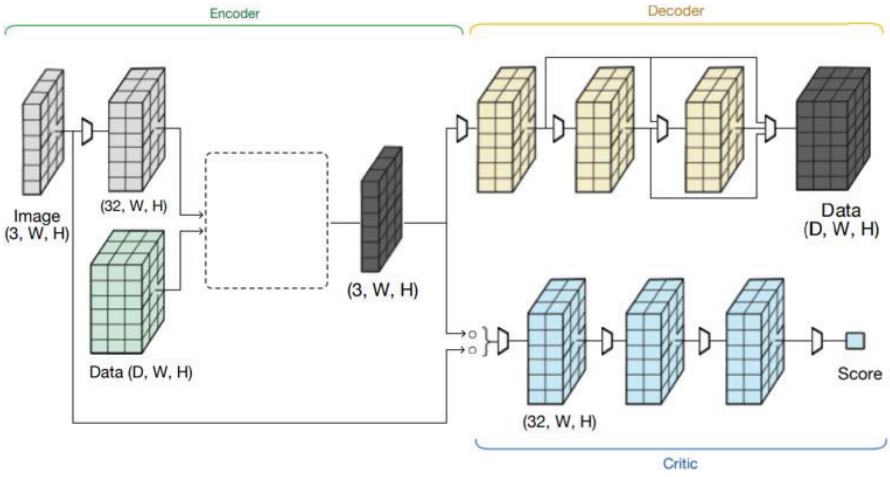

Fig. 1. Original SteganoGAN model architecture

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

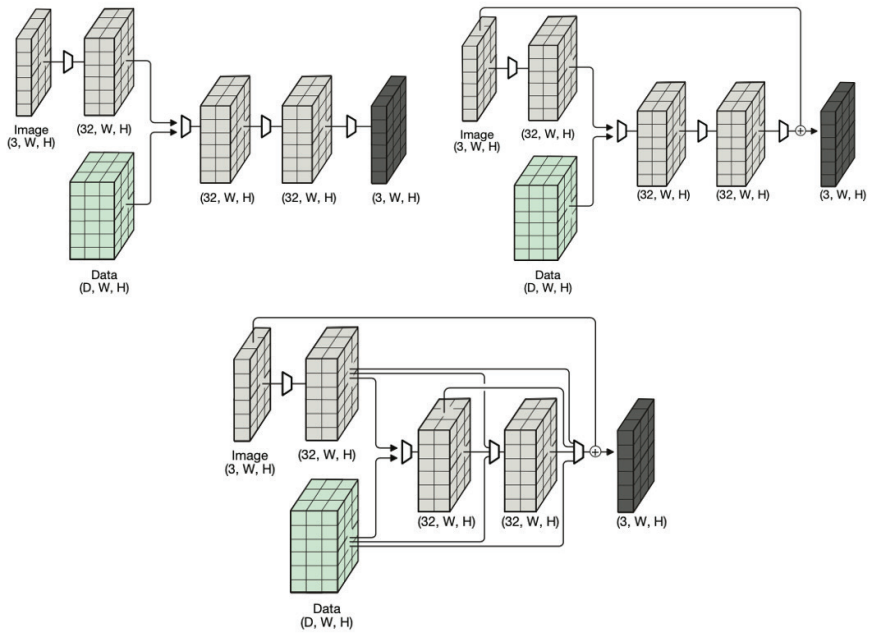

Fig. 2. Encoder different types (Basic, Residual and Dense)

This means that the encoder is learning to produce a residual image – an image that represents the difference between the cover image and the steganographic image – rather than the steganographic image itself. The advantage of using residual connections is that they help prevent vanishing gradient problems during training and enable the model to focus on the minimal changes needed to embed the message. By only adjusting parts of the image where necessary, the residual encoder tends to produce higher-quality steganographic images, especially when dealing with larger payloads. This architecture also typically converges faster than the basic encoder.

Dense Encoder – the most advanced of the three and is inspired by the DenseNet architecture [8]. In this design, the output of each convolutional layer is densely connected to the inputs of all subsequent layers. In other words, the feature maps produced by each layer are concatenated with the feature maps of all previous layers, ensuring that the encoder reuses features across the network. This results in feature reuse, meaning that the model can more effectively capture complex image details and use the most relevant information for embedding the message. Dense connections mitigate the vanishing gradient problem more effectively than residual connections and encourage the model to leverage both low-level and high-level features simultaneously. This makes the dense encoder particularly well-suited for scenarios where the payload is large or the image content is complex. However, the trade-off is that this architecture is more computationally expensive and requires more memory than the basic or residual encoders.

Decoder Network

In SteganoGAN, two decoder architectures are employed: the basic decoder and the dense decoder. The decoder's primary role is to extract the hidden message from the steganographic image, and the effectiveness of this process is crucial for ensuring accurate message recovery. The architecture of the decoder can significantly influence how well it performs under different conditions, particularly when handling varying payload sizes and complex image content.

Basic Decoder – follows a straightforward approach, much like the basic encoder. It consists of several convolutional layers that transform the steganographic image back into a feature representation. These features are then processed to recover the original binary message. Just like Basic Endoder, the design of the basic decoder is simple, with no advanced connectivity mechanisms between layers. While the basic decoder is efficient and works well for smaller payloads, its simplicity can become a limiting factor when dealing with larger message depths or more complex steganographic images [9-10]. It processes each layer sequentially, without taking advantage of previous layer outputs, which can lead to less accurate message recovery in more challenging scenarios. Should be mentioned that Basic Decoder only uses with Basic or Residual Encoder.

Dense Decoder – builds upon the same principle as the dense encoder, utilizing dense connections to enhance feature extraction and reuse. In this architecture, each layer's output is concatenated with the outputs of all preceding layers. This ensures that every convolutional layer has access to the full set of features

Т3. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

extracted by earlier layers, allowing the decoder to leverage both low-level and high-level information simultaneously.

The dense decoder is particularly well-suited for recovering messages from images where the message depth is high or the cover image is complex. By maintaining access to all previous feature maps, the dense decoder can more effectively reconstruct the binary message. This architecture excels in scenarios where high accuracy is essential, as it mitigates the risk of losing critical information during the decoding process.

Critic Network

The critic network is crucial to the adversarial training process. It distinguishes between real images (the original cover images) and fake images (the steganographic images generated by the encoder). The critic's feedback guides the encoder, pushing it to generate steganographic images that look more realistic and are harder to distinguish from the original cover images. The critic uses the Wasserstein loss, which is designed to improve the training stability of GANs and ensures that the generated images closely resemble real cover images [11-13].

Loss Functions

SteganoGAN employs three primary loss functions to optimize the encoder, decoder, and critic simultaneously:

– Decoder loss – measures how well the decoder D recovers the hidden message in steganographic message $\mathcal{E}(X, M)$, where X is original image, and M is original message. It uses binary cross-entropy loss to compare the original message to the decoded message $D(\mathcal{E}(X, M))$. The goal is to minimize the error between the original and the recovered with decoder binary messages.

$$\mathcal{L}_d = \mathbb{E}_{X \sim \mathbb{P}_C} CrossEntorpy\left(D\left(\mathcal{E}(X, M)\right), M\right) \quad (1)$$

– Image similarity loss – this loss ensures that the steganographic image remains visually similar to the original cover image. The image similarity loss is calculated using the mean square error (MSE), which measures the difference between each pixel in the cover image and the corresponding pixel in the steganographic image. Minimizing this error helps preserve the appearance of the cover image.

$$\mathcal{L}_s = \mathbb{E}_{X \sim \mathbb{P}_C} \frac{1}{3 \times W \times H} X - \mathcal{E}(X, M)_2^2 \quad (2)$$

– Realness loss – returns critic C score of steganographic image.

$$\mathcal{L}_r = \mathbb{E}_{X \sim \mathbb{P}_C} C\left(\mathcal{E}(X, M)\right) \quad (3)$$

The total loss is the sum of these three losses: the decoder loss, the image similarity loss, and the realness loss.

$$minimize \mathcal{L}_d + \mathcal{L}_s + \mathcal{L}_r \quad (4)$$

By minimizing this combined loss, SteganoGAN optimizes the quality of the steganographic images while ensuring reliable message recovery. SteganoGAN uses this combined loss to train both encoder and decoder model weights.

Critic loss – evaluates how realistic the steganographic images are compared to the original cover images.

$$\mathcal{L}_c = \mathbb{E}_{X \sim \mathbb{P}_C} C(X) - \mathbb{E}_{X \sim \mathbb{P}_C} C\left(\mathcal{E}(X, M)\right) \quad (5)$$

The critic loss is based on the Wasserstein distance, which compares the distributions of the real images (cover images) and the generated images (steganographic images). The encoder aims to generate images that minimize the critic's ability to distinguish between the real and fake images.

Keras Steganogan Architecture

KerasSteganoGAN is a TensorFlow-based reimplementa-tion of the original SteganoGAN architecture with several notable modifications to simplify the message encoding and decoding processes while maintaining the overall structure and functionality of the original model. Instead of modifying original PyTotch-based StegnoGAN, KerasSteganoGAN was rewritten with more functional and popular Python library called TensorFlow. First reason to make TensorFlow-based SteganoGAN is possibility of using it only with definite version of PyTorch library and Python. Another reason is possibility of using LiteRT library of making lightweighted mobile versions of AI models.

Key modification in KerasSteganoGAN is the removal of the Reed-Solomon error correction method, which was previously used in SteganoGAN to enhance message recovery accuracy. In SteganoGAN, the binary message was first encoded with Reed-Solomon, converted into bits, and then placed into a tensor, separated by blocks of zero bits. This process added computational complexity while improving error tolerance during decoding.

In KerasSteganoGAN, we opted to remove the Reed-Solomon encoding and decoding steps, simplifying the model's structure. By doing so, we rely solely on the AI models (encoder and decoder) to manage the embedding and extraction of messages. This change reduces overhead and streamlines the overall pipeline without significantly impacting performance. The hidden message is now directly embedded into the cover image, and its recovery is entirely dependent on the performance of the neural network models.

Despite these modifications, the core components of the architecture remain intact, including the encoder-decoder structure and the option to utilize different types of encoders (basic, residual, and dense). The removal of Reed-Solomon encoding enables the model to focus purely on leveraging the strengths of the neural network for accurate message recovery, which simplifies

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

the training and deployment processes while maintaining high payload capacity and image quality.

Another significant change that contains KerasSteganoGAN is that we use Sigmoid activation function at the last convolution layers of each encoder and decoder models in order to return data in a range of [-1, 1]. In the case of the encoder, returned data will be converted to an image, or in the case of the decoder, it will be converted to a tensor of binary data. In the original SteganoGAN model, both encoder and decoder models have no activation functions at the last convolution layers [14-15].

KerasSteganoGAN architecture allows generation of steganographic images with the size of 128x128 pixels, though SteganoGAN allows generation of stego-images with original size of cover-images. The reason of this is the difficulty of the implementation of such an architecture as well as the complexity of training.

One of the key explorations in KerasSteganoGAN is the introduction of two distinct model variants: with a critic and without a critic. In the original SteganoGAN, the critic network played an important role in the adversarial training process by helping the model learn to generate steganographic images that are indistinguishable from real images. The critic evaluates how realistic the generated steganographic images are and provides feedback to improve the encoder's output.

In KerasSteganoGAN, we decided to investigate the specific impact of the critic network on the quality of the steganographic images and the accuracy of message recovery. To do so, we created two versions of the model:

- One version includes the critic network, mimicking the adversarial setup from SteganoGAN.
- The other version operates without a critic network, meaning the encoder and decoder are trained without the additional adversarial feedback.

The motivation for this comparison is to analyze whether the critic network contributes significantly to improving the visual quality of the images and the reliability of message recovery or if a non-adversarial approach can yield comparable results. Both variants still use the same encoder-decoder structures, with the option to choose between basic, residual, and dense encoder types, but the inclusion or exclusion of the critic alters the training dynamics. Figure 3 presents KerasSteganoGAN architecture without critic with basic or dense decoder types while encoder types could be the same as at the original SteganoGAN: basic, residual or dense.

Training both model variants allows us to compare the influence of adversarial feedback in the GAN framework versus simpler encoder-decoder setups. This exploration is aimed at understanding whether the critic's complexity justifies its potential benefits, especially in terms of computational cost and training time.

Training And Comparision

To thoroughly evaluate the performance of KerasSteganoGAN models with and without the critic network, we used a set of key metrics that assess both the quality of the steganographic images and the effectiveness of message recovery. These metrics – RS-BPP (Reed-Solomon bits per pixel), PSNR (Peak Signal-to-Noise Ratio), and SSIM (Structural Similarity Index Measure) – are standard in the field of steganography, providing insights

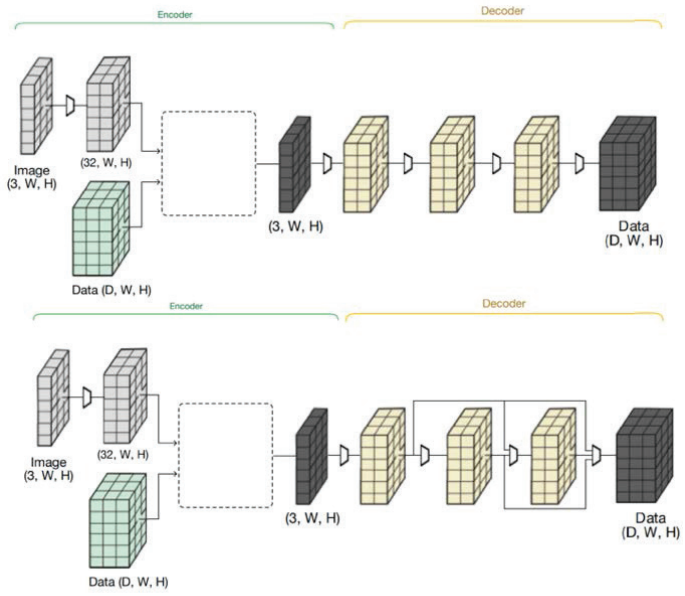

Fig. 3. Critic-less KerasSteganoGAN architecture with basic or dense decoder types

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

into the image quality, message capacity, and accuracy of the decoder. Detailed description of these metrics is provided in [16].

To train our model Div2K [17] dataset was used with batch size of 4. The train dataset consists of 700 pictures of different sizes, while the validation dataset consists of only 100 pictures. Due to the absence of GPU with CUDA cores, our neural network training was done on CPU Intel i9, and took 15 minutes to train the model on 5 epochs.

PSNR measures the difference in quality between the steganographic image and the original cover image. A higher PSNR value indicates that the two images are more similar, with fewer distortions or artifacts in the steganographic image. In the context of KerasSteganoGAN, we use PSNR to evaluate how well the model preserves the visual quality of the cover image after embedding the hidden message. The influence of the critic network is particularly highlighted by this metric, as models with a critic are expected to generate steganographic images that closely resemble the cover images.

SSIM is another metric used to evaluate the quality of the steganographic image, focusing on the perceptual similarity between the cover and stego-images. It measures structural changes, taking into account luminance, contrast, and texture. Like PSNR, higher SSIM values indicate better preservation of the image structure. To visualize the impact of the critic network on image quality, plots of SSIM for different encoder types and data depths will demonstrate how the critic influences the visual similarity between the stego and cover images.

The RS-BPP metric calculates the effective number of bits per pixel that can be reliably hidden and recovered in the image. This metric is derived from the decoder accuracy and the message depth. Although we removed the Reed-Solomon encoding/decoding process in KerasSteganoGAN, we still use the RS-BPP formula to estimate the capacity of the model. It provides a clear measure of how efficiently the model can embed and recover data, particularly when comparing different encoder types and their ability to handle various message depths.

These three metrics allow us to assess the performance of KerasSteganoGAN from different perspectives: PSNR and SSIM evaluate the visual fidelity of the images, while RS-BPP measures the model's capacity to embed and recover data effectively. The upcoming sections will illustrate these metrics through various plots, particularly focusing on how the inclusion of the critic network affects image quality and message recovery.

To comprehensively compare the performance of KerasSteganoGAN models with and without a critic network, we trained both variants using three encoder-decoder configurations:

- basic encoder – basic decoder;
- residual encoder – basic decoder;
- dense encoder – dense decoder.

For each of these encoder-decoder pairs, we experimented with six different message depths (1 to 6 bits per pixel), varying the amount of data embedded in each image. Each configuration was trained for 5 epochs, as this was sufficient to observe the general trends in the models' performance while keeping the training time manageable. The impact of these choices on the performance of the models is reflected in several metrics – PSNR, SSIM, and RS-BPP – as shown in the figures that follow.

In Figure 4, we present PSNR results for all three encoder configurations (basic, residual, and dense) across different message depths. Here, it becomes evident that models with a critic network generally achieve higher PSNR values, particularly at higher data depths, indicating better preservation of image quality. The residual encoder, paired with a basic decoder, shows a significant improvement in PSNR compared to the basic

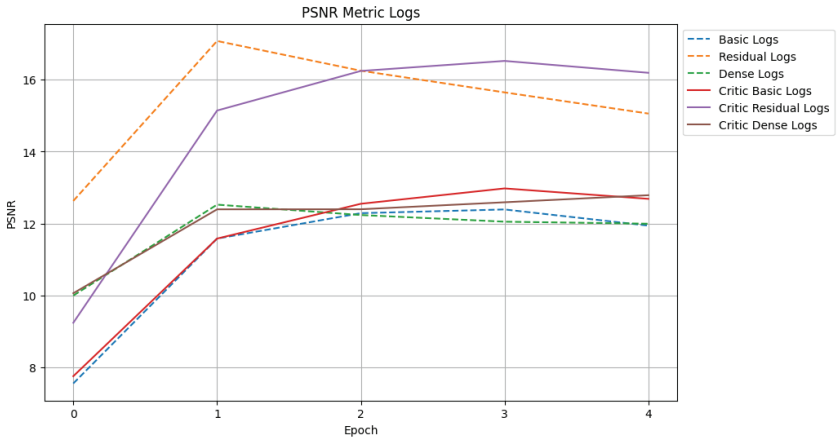

Fig. 4. Plot of SSIM metrics during 5 epochs of training KerasSteganoGAN with and without critic model

Т3. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

encoder, especially when trained with a critic. The dense encoder achieves the best overall results, with PSNR values staying high even with 6 bits per pixel of data, highlighting its superior ability to handle larger payloads.

Similarly, Figure 5 showcases the SSIM metric for the same encoder-decoder configurations. The dense encoder consistently achieves higher SSIM values, indicating a better structural similarity between the stego and cover images. Models with a critic tend to perform better in terms of SSIM across all data depths, reinforcing the idea that the critic helps the encoder preserve image structure more effectively.

The RS-BPP metric (shown in Figure 6) provides insight into the effective payload capacity of the models.

Here, the dense encoder again outperforms the basic and residual encoders, particularly at higher message depths. As expected, models with a critic achieve slightly higher RS-BPP values, reflecting their ability to embed and recover data more accurately. However, the improvement in RS-BPP is not as dramatic as the gains seen in PSNR and SSIM, suggesting that the critic's primary benefit lies in improving image quality rather than significantly boosting payload capacity.

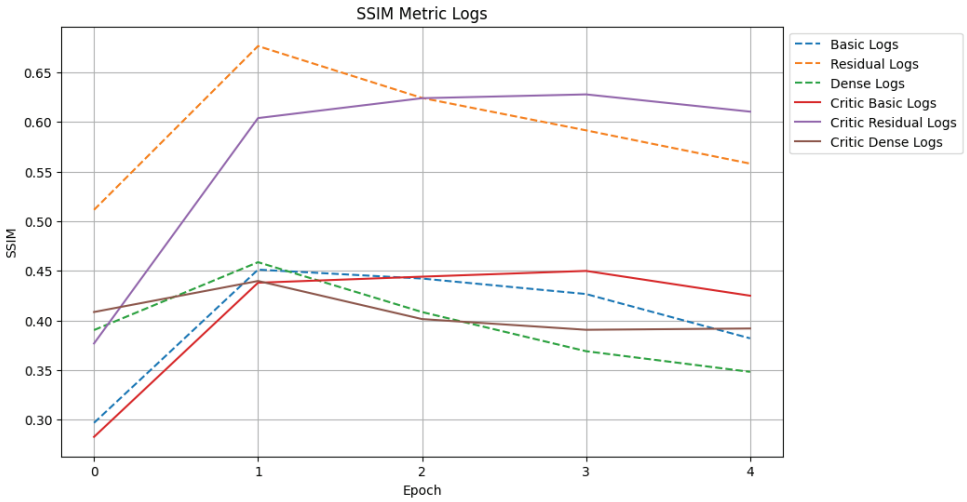

Fig. 5. Plot of PSNR metrics during 5 epochs of training KerasSteganoGAN with and without critic model

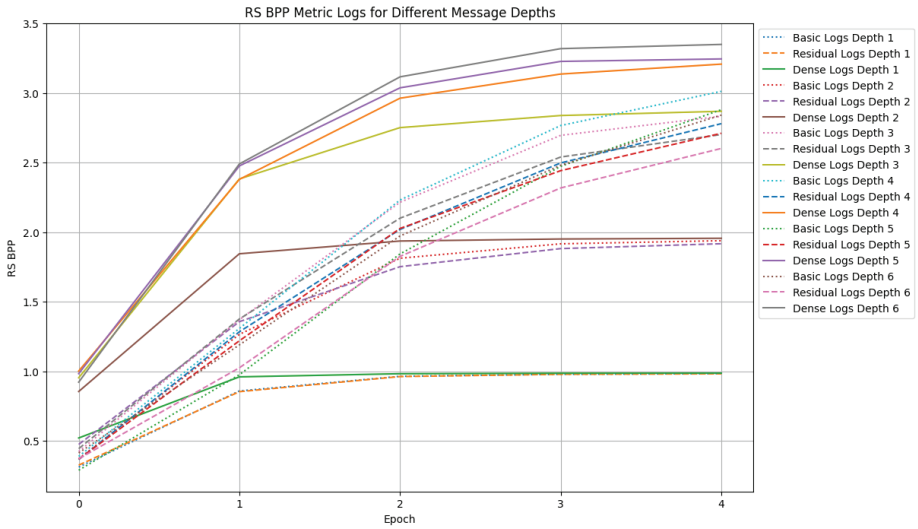

Fig. 6. RS-BPP metric during training KerasSteganoGAN with different payloads for all types of Encoder model

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

Research Experiments And Results

These results highlight the consistent pattern observed across metrics – models with a critic tend to produce higher quality images, as evidenced by their superior PSNR and SSIM values, while dense encoder architectures excel in both image quality and data embedding efficiency.

To better understand the effectiveness of our trained models, there is table 1, which contains Decoder accuracy, PSNR, SSIM, and RS-BPP metrics for each KerasSteganoGAN model with different data depth, encoder architecture, and critic model presence. Figure 7 presents the original image in the first row, the stego-image created by the KerasSteganoGAN model with critic, dense encoder, and data-depth 6 trained on 5 epochs on the second row, and their difference in the gray-scale on the third row.

The decoder accuracy remains high across all models and data depths, particularly for smaller message depths (1 and 2 bits per pixel). Both models with and without a critic achieve nearly perfect accuracy at lower message depths, with a slight drop as the data depth increases. At data depth 6, the model with a critic shows slightly better accuracy for the dense encoder (77%) compared to the non-critic model (78%), indicating the critic's positive impact in higher payload scenarios.

RS-BPP, which measures the effective payload, shows a consistent trend: the dense encoder performs best across all data depths. For instance, at data depth 6, the dense encoder achieves an RS-BPP of 3.34 (with critic) and 3.42 (without critic), outperforming both the basic and residual encoders. The basic encoder shows relatively lower RS-BPP across all configurations, with values ranging from 0.98 at depth 1 to 2.84 at depth 6 (with a critic). In general, models without a critic show slightly higher RS-BPP values, but the difference is not significant.

Table 1. Metrics For Different Model Architecture And Payload Depth

Critic	D	Decoder Accuracy			RS-BPP			PSNR			SSIM		
		Basic	Residual	Dense	Basic	Residual	Dense	Basic	Residual	Dense	Basic	Residual	Dense
Yes	1	0.99	0.99	0.99	0.98	0.98	0.98	13.11	16.69	14.44	0.47	0.63	0.52
	2	0.98	0.97	0.98	1.93	1.91	1.95	12.13	14.66	13.93	0.39	0.53	0.49
	3	0.97	0.95	0.97	2.83	2.70	2.86	10.81	12.20	11.60	0.26	0.36	0.29
	4	0.87	0.84	0.90	3.01	2.77	3.20	11.21	13.11	11.55	0.29	0.42	0.28
	5	0.78	0.77	0.82	2.87	2.71	3.24	11.99	14.95	12.08	0.36	0.54	0.33
	6	0.73	0.71	0.77	2.84	2.60	3.34	12.68	16.19	12.78	0.42	0.61	0.39
No	1	0.99	0.99	0.99	0.98	0.98	0.99	12.34	15.62	14.82	0.42	0.59	0.55
	2	0.99	0.99	0.99	1.96	1.96	1.96	11.62	14.06	13.12	0.35	0.50	0.45
	3	0.97	0.95	0.98	2.82	2.74	2.90	10.78	11.55	10.99	0.26	0.31	0.27
	4	0.88	0.83	0.90	3.08	2.65	3.22	10.66	12.93	10.80	0.25	0.43	0.24
	5	0.79	0.78	0.83	2.97	2.84	3.37	11.64	13.62	11.60	0.35	0.47	0.30
	6	0.75	0.73	0.78	3.02	2.79	3.42	11.94	15.05	11.99	0.38	0.55	0.34

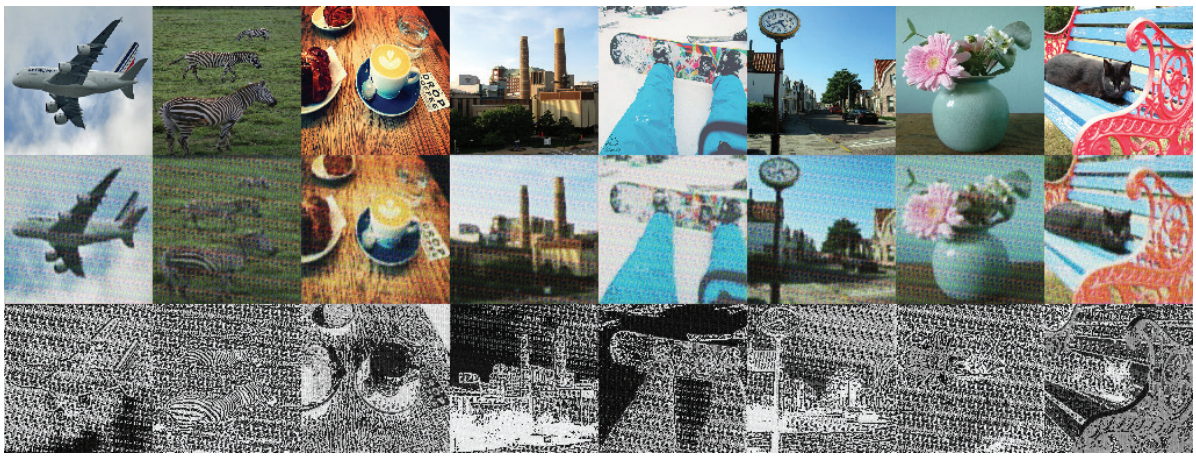

Fig. 7. Original resized, callback, their difference images on 5 epochs of the Dense model with critic and data-depth equals 6

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

In terms of PSNR, the dense encoder consistently performs better across all message depths compared to the basic and residual encoders. For example, at data depth 6, the dense encoder achieves a PSNR of 12.78 (with critic) and 11.99 (without critic), outperforming the other two encoders. The critic-enhanced models tend to have better PSNR values at higher data depths. At data depth 6, the basic encoder with a critic achieves a PSNR of 12.68 compared to 11.94 without a critic. This indicates that the critic helps preserve image quality as more data is embedded into the image.

SSIM results follow a similar pattern to PSNR, with the dense encoder showing the best performance across all message depths. For example, at data depth 6, the dense encoder achieves an SSIM of 0.39 (with critic) and 0.34 (without critic). The critic-enhanced models consistently achieve higher SSIM values across all configurations. At data depth 6, the residual encoder with a critic scores 0.61, compared to 0.55 without a critic, reinforcing the idea that the critic network helps maintain the structural integrity of the image. The residual encoder benefits the most from the critic in terms of SSIM. For instance, at data depth 6, the residual encoder with a critic scores 0.61, significantly higher than the 0.55 scored by the non-critic model.

Dense encoder consistently outperforms both basic and residual encoders across all metrics, showing its strength in handling larger payloads (higher data depths). Models with a critic generally perform better in terms of PSNR and SSIM, indicating that the critic helps improve image quality and structural similarity, especially at higher data depths. The non-critic models show slightly higher RS-BPP values, but the difference is not substantial, suggesting that while the critic improves image quality, it does not greatly impact the payload capacity.

Conclusion

In this paper, we presented KerasSteganoGAN, a TensorFlow-based reimplementation and extension of the original SteganoGAN model, with key modifications, including the removal of the Reed-Solomon encoding/decoding process and the introduction of two distinct model variants: with and without a critic network. Our goal was to assess the impact of these changes on the model's ability to hide and recover messages while maintaining the quality of the steganographic images.

Through a comprehensive evaluation using RS-BPP, PSNR, and SSIM metrics across different encoder-decoder architectures and message depths (ranging from 1 to 6 bits per pixel), we observed several important trends:

- Dense encoders, regardless of whether a critic was used, consistently outperformed both basic and residual encoders in terms of both image quality and data embedding efficiency. Dense connections allowed for better feature reuse, resulting in higher RS-BPP, PSNR, and SSIM values across all configurations.
- The critic network played a crucial role in improving image quality, particularly at higher message depths. Models with a critic achieved higher PSNR and SSIM values, indicating that the critic helped preserve both the visual and structural integrity of the steganographic images.
- While models without a critic achieved slightly higher RS-BPP values, the improvement was marginal, suggesting that the critic network primarily enhances image quality rather than significantly increasing payload capacity.

Overall, our results suggest that using a dense encoder with a critic network yields the best trade-off between payload capacity and image quality, making it an ideal choice for high-capacity steganographic tasks. However, given that all models were only trained for 5 epochs, our results also indicate that more extensive training (at least 32 epochs, as used in the original SteganoGAN) would likely lead to further improvements in performance, especially for more complex encoder architectures like the dense encoder.

Future work will focus on exploring additional encoder architectures and loss functions to further optimize the performance and enhance its applicability in real-world steganography scenarios. We plan to retrain the model and perform a comparative analysis under equivalent training conditions, such as training duration, to ensure a fair evaluation of the model's improvements. Additionally, we aim to generate images at their original size to better reflect practical use cases and assess the model's performance with higher-resolution data.

Conflict of Interest

The authors declare that they have no conflict of interest regarding this study, including financial, personal, authorship-related, or any other type of conflict that could have influenced the research or its results presented in this article.

Funding

This research was conducted without any financial support.

Data Availability

This manuscript has no associated datasets.

ПОКРАЩЕНА СТЕГАНОГРАФІЯ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ KERAS-STEGANOGAN: GAN НА ОСНОВІ TENSORFLOW

Дмитро Хома, Євген Башков

Стеганографія зображень – це процес вбудовування секретної інформації в цифрові зображення таким чином, щоб сам факт існування

ТЗ. №5(37)'2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

повідомлення залишався непомітним. Останні досягнення у сфері глибокого навчання, зокрема в генеративно-змагальних мережах (GAN), суттєво покращили як пропускну здатність стеганографічних систем, так і якість відтворення зображень. Оригінальна модель SteganoGAN, реалізована у Torch, досягла найсучасніших результатів, забезпечуючи приховування до 4.4 біт на піксель при високій стійкості до методів стегоаналізу. Проте вплив критика (critic network) на якість стеганографії та стабільність навчання досі залишається недостатньо вивченим.

У цій роботі представлено Keras-SteganoGAN – реалізацію та розширення SteganoGAN на базі TensorFlow, призначену для систематичного аналізу ролі критика в процесі змагального стеганографічного навчання. Було розроблено та порівняно дві версії моделі – з критиком і без нього – на основі трьох архітектур енкодера: basic, residual та dense. Кожну конфігурацію навчали протягом п'яти епох із глибиною повідомлення від 1 до 6 біт, що дало змогу всебічно дослідити компроміси між ємністю прихованих даних, спотворенням зображення та точністю декодування.

Кількісна оцінка виконувалася за допомогою стандартних метрик якості зображень і стеганографічної ефективності, включаючи PSNR, SSIM, RS-BPP та точність декодування. Результати показують, що наявність критика підвищує візуальну якість та схожість зображень за невеликих навантажень, проте його вплив зменшується зі збільшенням глибини повідомлення. Отримані результати надають нове розуміння взаємодії між складністю енкодера, динамікою критика та стеганографічною ефективністю, а також формують підґрунтя для подальшого вдосконалення систем стеганографії на базі GAN.

Ключові слова: *SteganoGAN, Keras, TensorFlow, стеганографія зображень, GAN, енкодер, критик, пропускна здатність, residual, dense, декодер, змагальне навчання, приховане повідомлення, метрики схожості.*

REFERENCES

[1] Pevnýˊ, T., Filler, T., and Bas, P. "Using high-dimensional image models to perform highly undetectable steganography. Information Hiding", 2010.

[2] Kevin Alex Zhang, Alfredo Cuesta-Infante, Lei Xu, and Kalyan Veeramachaneni, "Steganogan: High capacity image steganography with gans," arXiv preprint arXiv:1901.03892, 2019.

[3] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets. Advances in neural information processing systems", 27, 2014.

[4] Hayes, J. and Danezis, G. "Generating steganographic images via adversarial training". In NIPS, 2017.

[5] Baluja, S. "Hiding images in plain sight: Deep steganography". In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. (eds.), Advances in Neural Information Processing Systems 30, pp. 2069–2079. Curran Associates, Inc., 2017.

[6] Zhu, J., Kaplan, R., Johnson, J., and Fei-Fei, L. "HiDDeN: Hiding data with deep networks". CoRR, abs/1807.09937, 2018.

[7] He, K., Zhang, X., Ren, S., and Sun, J. "Deep residual learning for image recognition". IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.

[8] Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. "Densely connected convolutional networks". IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 2261–2269, 2017.

[9] Shao-Ping Lu, Rong Wang, Tao Zhong, and Paul L Rosin, "Large-capacity image steganography based on invertible neural networks", in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 10816– 10825.

[10] Donghui Hu, Liang Wang, Wenjie Jiang, Shuli Zheng, and Bin Li, "A novel image steganography method via deep convolutional generative adversarial networks", IEEE access, vol. 6, pp. 38303–38314, 2018.

[11] S. L. Kryvyi, Y. V. Boyko, S. D. Pogorilyy, O. F. Boretskyi, M. M. Glybovets. "Design of Grid Structures on the Basis of Transition Systems with the Substantiation of the Correctness of Their Operation". Cybernetics and Systems Analysis. January 2017, Volume 53, Issue 1, pp 105–114. Springer Science+Business Media New York 2017. https://doi.org/10.1007/s10559-017-9911-0

[12] S. D. Pogorilyy and I. Yu. Shkulipa. "A Conception for Creating a System of Parametric Design of Parallel Algorithms and Their Software Implementations". Cybernetics and System Analysis, Volume 45, Issue 6 (November 2009), p.p. 952-958. Springer Science and Business Media Inc. ISSN: 1060-0396.

[13] Tang, W., Tan, S., Li, B., and Huang, J. "Automatic steganographic distortion learning using a generative adversarial network". IEEE Signal Processing Letters, 24(10):1547–1551, Oct 2017. ISSN 1070-9908. doi: 10.1109/LSP.2017.2745572.

[14] Zhuo Zhang, Jia Liu, Yan Ke, Yu Lei, Jun Li, Minqing Zhang, and Xiaoyuan Yang, "Generative steganography by sampling," IEEE access, vol. 7, pp. 118586–118597, 2019.

[15] B Delina, "Information hiding: A new approach in text steganography," in Proceedings of the International Conference on Applied Computer and Applied Computational Science, World Scientific and Engineering Academy and Society (WSEAS 2008), 2008, pp. 689–695.

Т3. №5(37)’2025 https://doi.org/10.32782/2786-9024/v3i5(37).344525

[16] Khoma D.Yu., “Steganographic algorithms and stegoanalysis based on classical methods and neural networks” – Scientific works of DonNTU, series "Informatics, Cybernetics and Computing", No. 2 (37), 2024, p. 42–53. https://doi.org/10.31474/1996-1588-2023-2-37-42-53

[17] Agustsson, E. and Timofte, R. “NTIRE 2017 challenge on single image super-resolution: Dataset and study”. In The IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) Workshops, July 2017.

Стаття надійшла до редакції 30.09.2025
Стаття прийнята 12.10.2025
Статтю опубліковано 02.12.2025

НАУКОВІ ПРАЦІ
ДОНЕЦЬКОГО НАЦІОНАЛЬНОГО
ТЕХНІЧНОГО УНІВЕРСИТЕТУ

Серія: «Обчислювальна техніка та автоматизація»

Всеукраїнський науковий збірник
Заснований у липні 1998 року
Виходить 2 рази на рік

Мови видання: українська, англійська (змішаними) мовами.

Т3. № 5(37)'2025

Підписано до друку 28.11.2025. Формат 60×84/8.
Папір офсетний. Гарнітура Calibri. Цифровий друк.
Умовно друк. арк. 5,58. Тираж 150. Замовлення № 1125/918.

Ціна договірна. Віддруковано з готового оригінал-макета.

Видавництво і друкарня – Видавничий дім «Гельветика»
65101, Україна, м. Одеса, вул. Інглезі, 6/1
Телефони: +38 (095) 934 48 28, +38 (097) 723 06 08
E-mail: mailbox@helvetica.ua
Свідоцтво суб'єкта видавничої справи
ДК № 7623 від 22.06.2022