УДК 621.327:621.391 DOI: 10.26565/2519-2310-2022-1-04

ВИКОРИСТАННЯ ПАРАМЕТРІВ ДОВЖИН СЕРІЙ, ЯК ЕЛЕМЕНТА МІЖБЛОЧНОГО МУЛЬТИПЛЕКСУ ДАНИХ СТЕГАНОАЛГОРИТМУ

Микита Гончаров, Юлія Лєсная

Харківський національний університет імені В.Н. Каразіна, майдан Свободи 4, Харків, 61022, Україна
worldxdark@gmail.com, xa12284109@student.karazin.ua

Надійшла: липень 2022. Прийнята: липень 2022.

Анотація: *Розглянуто особливості використання параметрів довжин серій та кількості сформованих опорних блоків, як елементів комплексного ключа екстрактора даних, гібридного стеганоалгоритму. Наведено результати атаки (зламу) тестових зображень, які отримані для стеків вибірки різної довжини (різної бази перестановок діючих параметрів серій). Зроблено висновок про провідну роль параметра «довжина серій» при реалізації процедур міжблокового мультиплексування стеганоконтента. Наголошено, що одночасне використання дворівневого мультиплексу даних, значно розширює можливості протистояння спробам атак контенту. Встановлено, що застосування блоків з більшою розмірністю, істотно зменшує роль параметра «довжин серій», як основного елемента для руйнування структури вихідних зображень. Констатується, що збільшення довжини стека вибірки серій, розширює потенційну комбінаторику мультиплексування для діючих пар параметрів серій, та в більшої мірі руйнує кореляційні зв'язки елементів вихідного масиву даних. За результатами моделювання зроблено висновок, що використання різних способів розгортки серій забезпечує ще одну позицію в структурі ключа екстрактора даних.*

Ключові слова: *зображення; стеганографія; контейнер; контент; згладжування зображень; візуальна помітність викривлень; кодування серій; мультиплексування.*

1. Вступ

Ця робота відображає деякі результати, отримані в ході проведення моделювання основних процедур міжблокового мультиплексування даних, у рамках відпрацювання загальної концепції малоресурсного гібридного стеганографічного алгоритму [1]. На даному етапі робіт основна увага приділена дослідженню одержуваних ефектів, при використанні в якості елементів мультиплексування, діючих параметрів масиву серій [2]: – кількості опорних блоків (ОБ) зображень та параметра довжин серій ОБ. Можливість формування базового масиву серій ОБ певною мірою, забезпечується за рахунок проведення відповідних процедур «згладжування» зображень (*або передобробки вихідних даних*), що реалізованих на першому етапі роботи дослідного алгоритму [1]. У рамках моделювання процедур першого етапу було досліджено три варіанта згладжування вихідних зображень [1], які дозволяють отримати необхідний результат [2] за кількістю блоків ідентичного змісту (*тобто серій ОБ*) для різних типів вихідних даних *(у даному випадку, тестових зображень)*.

В якості тестових зразків даних використані напівтонові зображення трьох різних типів: - зображення типу *«портрет»*; - зображення типу *«пейзаж»* та зображення типу «*мнемосхема»*. Основна відмінність між ними полягає в характерних значеннях ймовірності перепаду яскравості між сусідніми елементами (пікселями) зображень кожного типу [3-5].

2. Основна частина

Для демонстрації отриманих ефектів, використана *полегшена* версія дослідного алгоритму, що передбачає підтримку виключно міжблокового рівня мультиплексу діючих параметрів масиву довжин серій (*ОБ та їх довжин серій)*, який було сформовано за результатами етапу згладжування тестових зображень. Маска перестановок, котра використовується для цих параметрів, визначається відповідною позицією елемента в структурі композиційного ключа екстрактора даних, та забезпечує міжблоковий рівень мультиплексування діючих параметрів масиву довжин серій для обраної розмірності ОБ [1-2].

Для імітації шифрування контенту використано короткий стек вибірки серій ОБ, з довжиною в 4 серії. Реалізований механізм перестановок діючих параметрів серій ОБ (*на вузької базі*) наведено на рис.1.

Зразки використаних тестових зображень наведенні на рис. 2-5*(з)*.

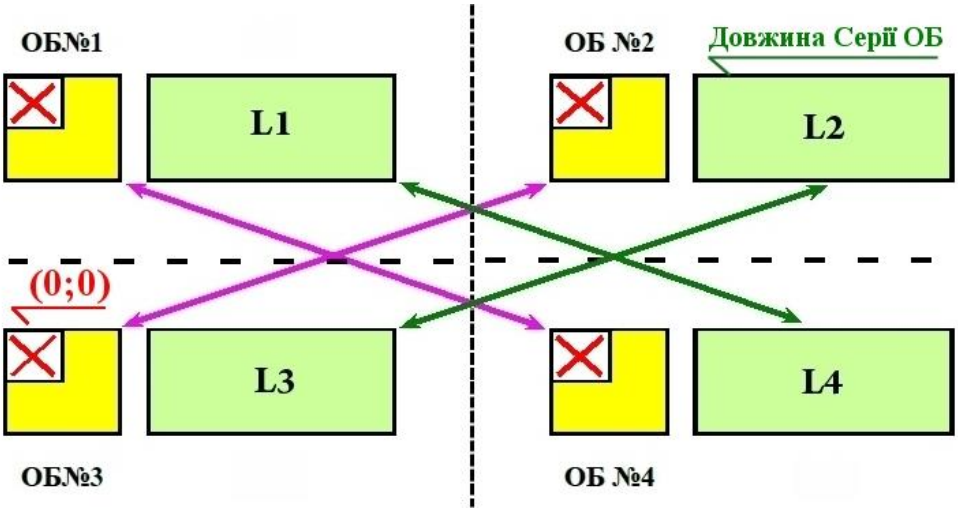


Рис. 1 – Тестова маска перестановок на вузькій базі вибірки

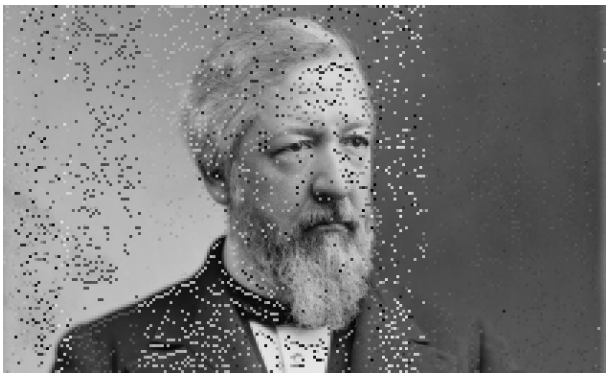
В межах проведеного моделювання передбачалося, що атакуючий правильно визначив діючий параметр довжини стека *(тобто, базу вибірки*) та принцип вибірки серій (*спосіб розгортки серій*), проте послідовно помилявся у визначенні діючих значень 2-х ключових параметрів, що залишилися: 1 – параметра зсуву ОБ *(жовті блоки на рис.1*); 2 – параметра довжини серій ОБ (*зелені блоки* - ***L)***. Відповідно до двох зазначених варіантів розвитку атаки на рис. 2-4*(а-е)* наведені результати несанкціонованого вилучення контенту при хибному доборі зазначених ключових параметрів:

- при неправильному доборі параметра зсуву ОБ – зразки *(а, в, д)*;
- при неправильному параметрі ***L*** – зразки (*б, г, е*).

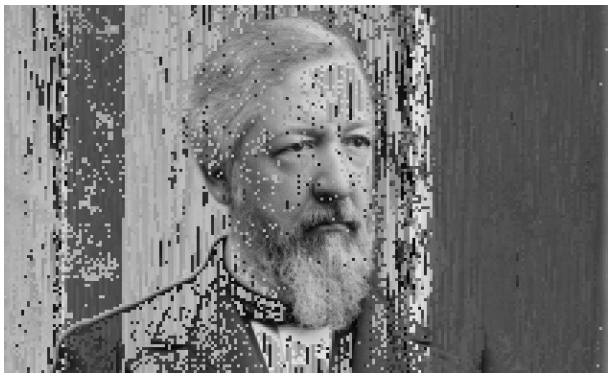
На рис. 5 представлена візуалізація отриманої різниці між вихідним та відновленим *(тобто. нелегітимно вилученим*) зображеннями для зазначених вище результатів атаки при різних розмірностях блоків *(зразки (а, в, д) проти зразків (б, г, е))*. При цьому, чим яскравіше (*біліше*) точка чи фрагмент зображень на рис. 5, то тем більше ступінь відмінності «зламаного» зображення від його оригіналу, і відповідно, чим темніше зазначений елемент/фрагмент, тим ближчі його параметри до вихідних значень оригіналу.

Ширину потенційної бази вибірки досліджуваних параметрів серій (*ОБ та* ***L)*** для різних типів зображень, різних розмірностей блоків та значень **Pz** (*значення порога загрубления яскравості елементів* зображення [1]) можна оцінити за діаграмами на рис 2-4 (*ж*). Значення параметра **Pz** = 3 вибиралося, як компроміс між урахуванням особливостей обраних тестових зображень та особливостями зорової системи людини [4-5], наприклад, властивістю зорового апарату людини виявляти на зображенні різні регулярні структури [3].

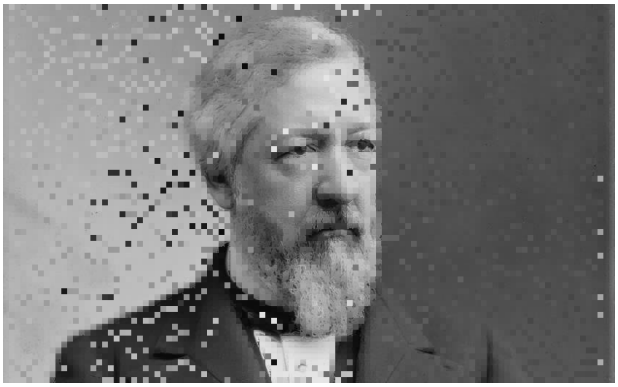
Вибір стека малої довжини обумовлений припущенням того, що *неправильний* підбір параметрів вилучення контенту *(спроба атаки*) на стеку з більш широкою базою буде призводити до більш руйнівних результатів при відновленні вихідних даних. Іншими словами, використання широкої бази перестановок поточних параметрів серій ОБ, призводить до більшого порушення просторової кореляції між елементами зображень [3-5].



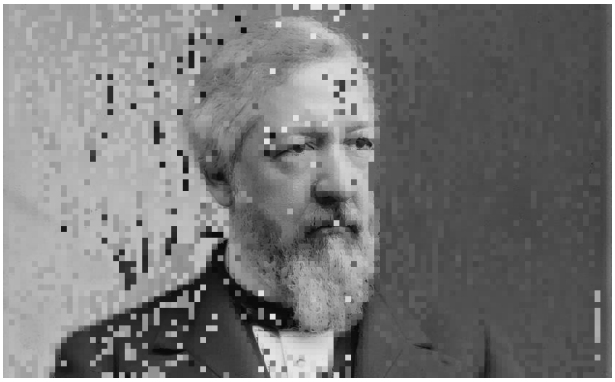
а) тіх ОБ розмірністю 4×4 ел.;



б) тіх Довжин серій, ОБ 4×4 ел.;

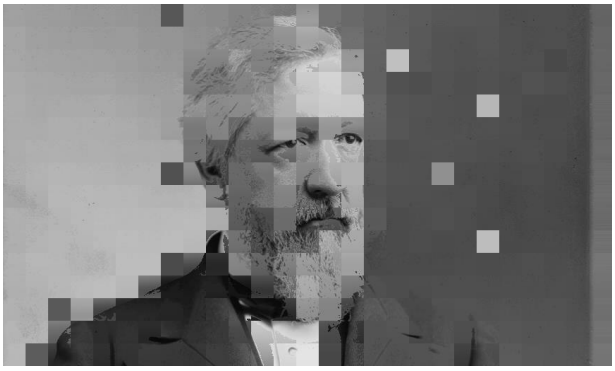


в) тіх ОБ розмірністю 8×8 ел.;

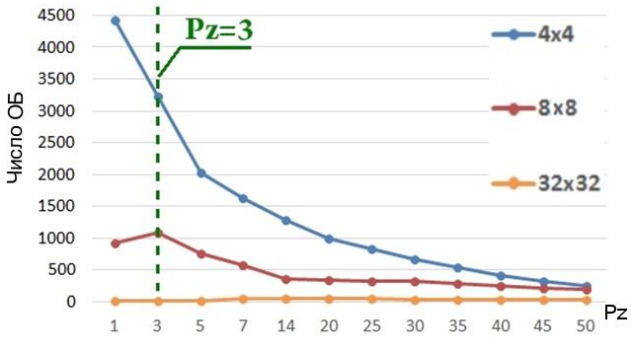


г) тіх Довжин серій, ОБ 8×8 ел.;

д) тіх ОБ розмірністю 32×32 ел.;



е) тіх Довжин серій, ОБ 32×32 ел.;



ж) кількість ОБ різної розмірності;

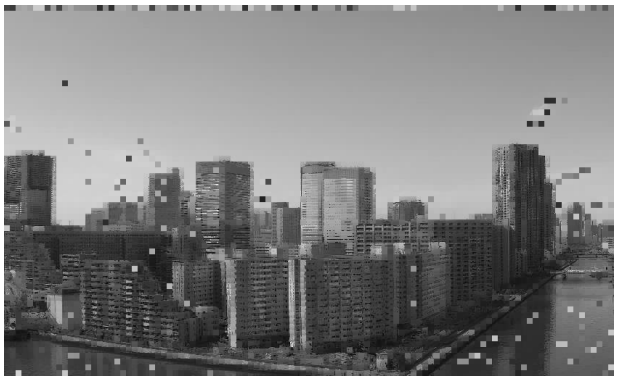
з) вихідне тестове зображення;

Рис. 2 – Результати атаки тестового зображення типу *«портрет»*
для різних розмірностей ОБ *(2-й Вар. згладжування)*

*Прим.: - «**тіх**», це скорочення терміну мультиплексування.*

а) тіх ОБ розмірністю 4×4 ел.;

б) тіх Довжин серій, ОБ 4×4 ел.;

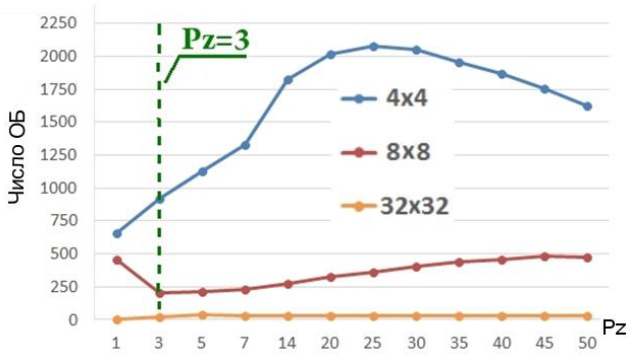


в) тіх ОБ розмірністю 8×8 ел.;

г) тіх Довжин серій, ОБ 8×8 ел.;

д) тіх ОБ розмірністю 32×32 ел.;

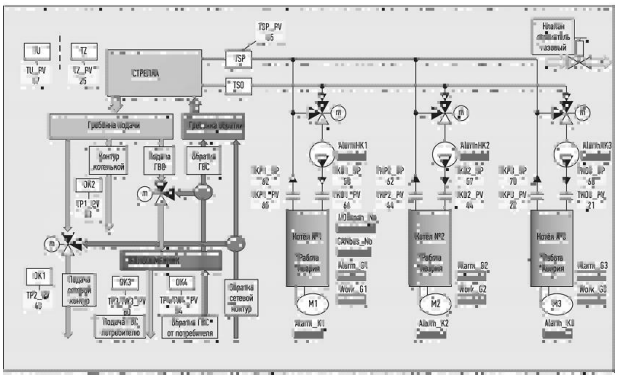
е) тіх Довжин серій, ОБ 32×32 ел.;



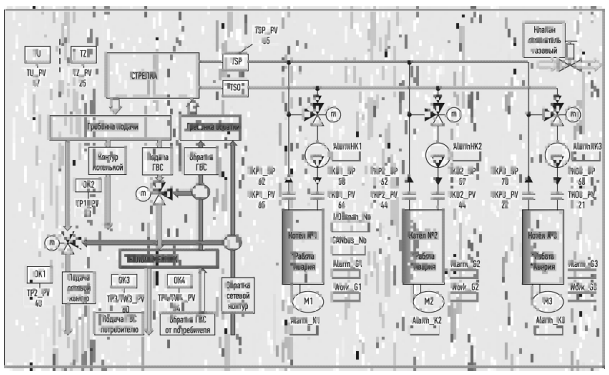
ж) кількість ОБ різної розмірності;

з) вихідне тестове зображення;

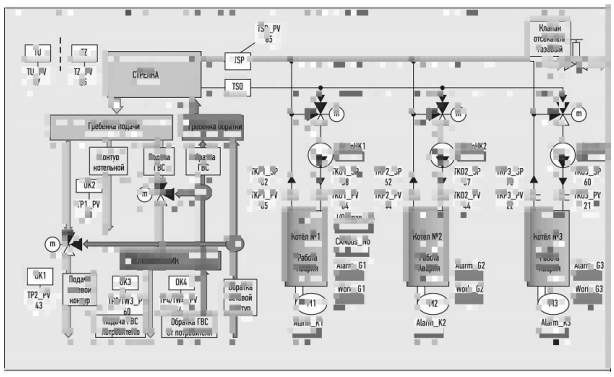
Рис. 3 – Результати обробки зображення типу *«пейзаж»*
для різних розмірностей ОБ *(2-й Вар. згладжування)*



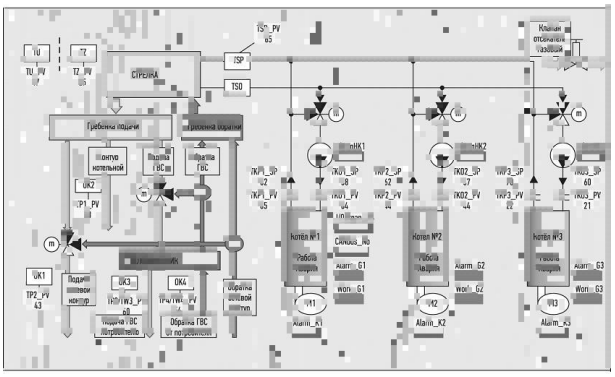
а) тіх ОБ розмірністю 4×4 ел.;



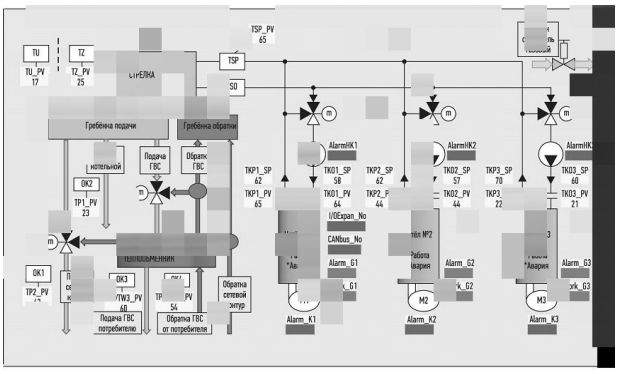
б) тіх Довжин серій, ОБ 4×4 ел.;



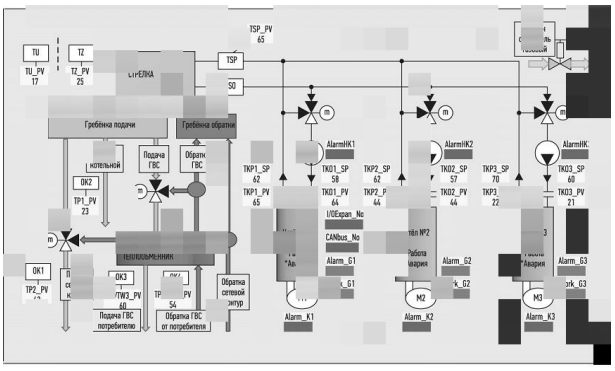
в) тіх ОБ розмірністю 8×8 ел.;



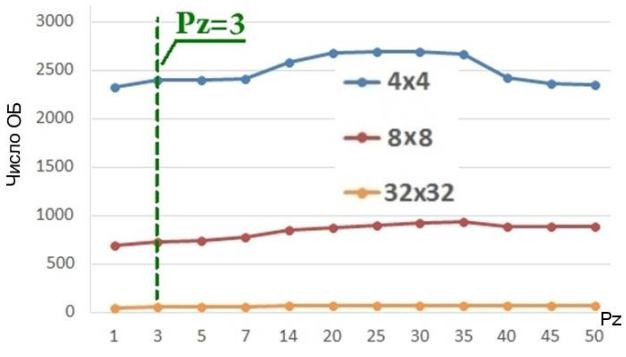
г) тіх Довжин серій, ОБ 8×8 ел.;



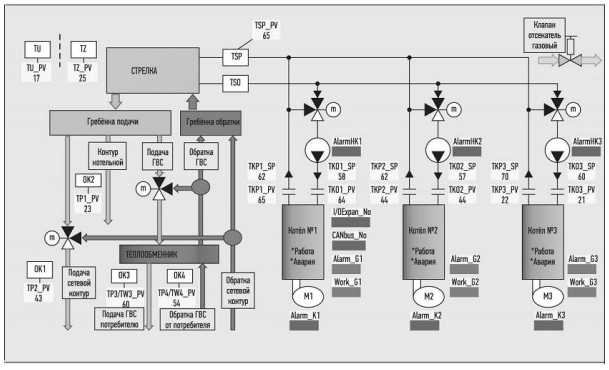
д) тіх ОБ розмірністю 32×32 ел.;



е) тіх Довжин серій, ОБ 32×32 ел.;



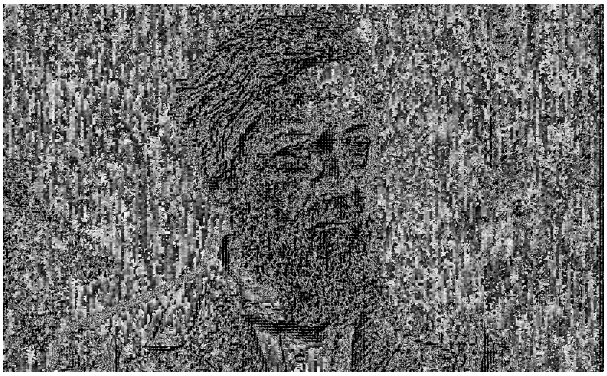
ж) кількість ОБ різної розмірності;



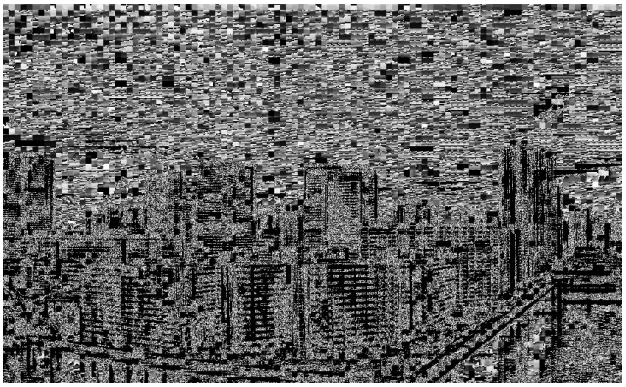
з) вихідне тестове зображення;

Рис. 4 – Результати атаки зображення типу «*мнемосхема*»
для різних розмірностей ОБ *(3-й Вар. згладжування)*

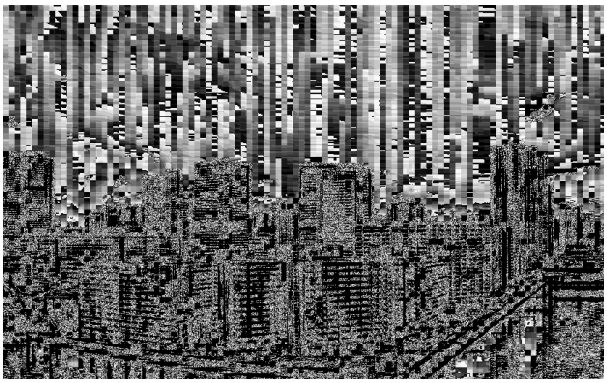
а) тіх ОБ, «портрет», 4×4 ел.;



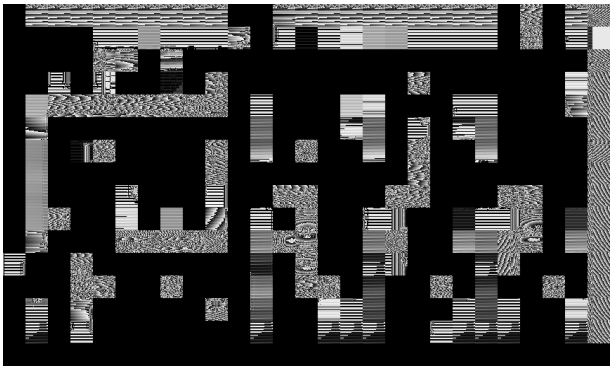
б) тіх Серій, «портрет», 4×4 ел.;



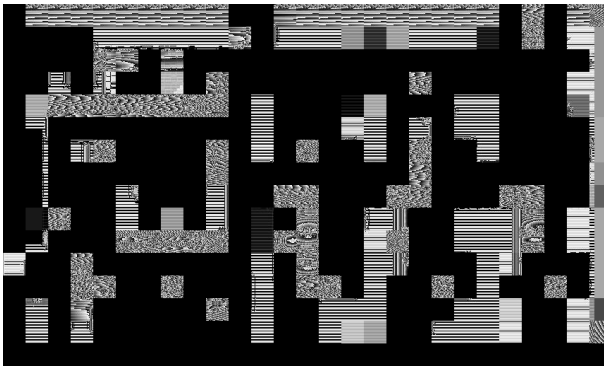
в) тіх ОБ, «пейзаж», 8×8 ел.;



г) тіх Серій, «пейзаж», 8×8 ел.;



д) тіх ОБ, «мнемосхема», 32×32 ел.;



е) тіх Серій, «мнемосхема», 32×32 ел.;

Рис. 5 – Візуалізація різниці вихідних та «атакованих» зображень
для різних розмірностей та комбінацій підбору діючих параметрів мультиплексу
((а-г) - 2-й Вар. згладжування; (д-е) - 3-й Вар. згладжування)

Таким чином, якщо на малій довжині стека одержуваний ефект (*тобто руйнування вихідних даних)* буде помітним, то при використанні стеку з більшою комбінаторикою перестановок цей процес стане ще більш очевидним. Відповідна тестова маска перестановок параметрів серій ОБ, що діють, на довгому стеку вибірки, представлена на рис. 6. В даному випадку базовий масив серій ОБ, був розділений на дві частини (напівстекеки) рівної довжини, між елементами яких проводяться відповідні маніпуляції.

Характерні результати атаки (*тобто, спроб нелегітимного вилучення стеганоконтенту*) для 2-х вибраних типів тестових зображень при використанні стеків вибірки різної довжини *(див. рис.1 та 6)*, представлені нижче, на рис.7.

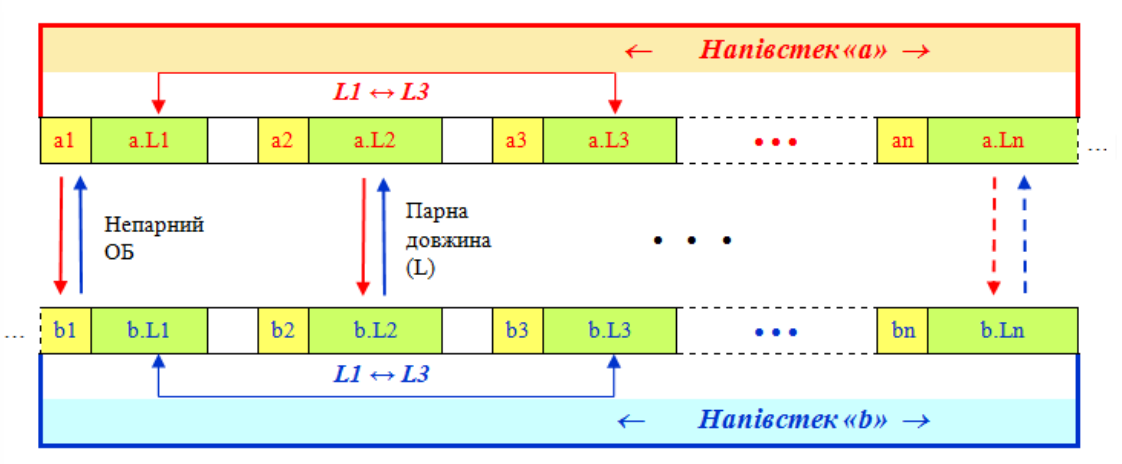


Рис. 6 – Тестова маска перестановок на широкій базі вибірки параметрів серій

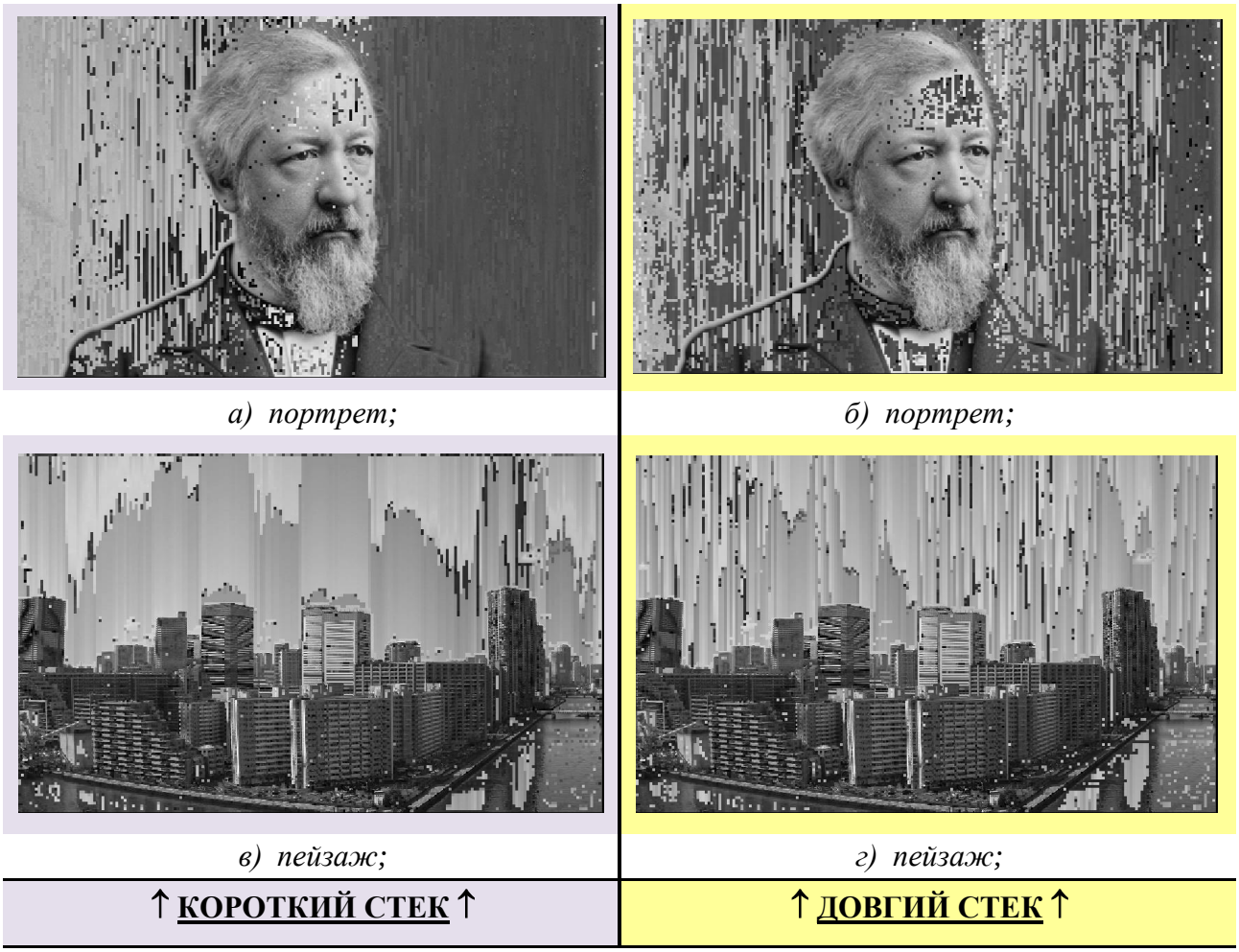


Рис. 7 – Результати «атаки» тестових зображень для стеків різної розмірності при $P_Z = 7$ (для *ОБ 4×4 ел.*)

Як було зазначено вище, основною метою представленого матеріалу є демонстрація можливості використання метода кодування довжин серій, як основи механізму міжблокового мультиплексування діючих параметрів довжин серій ОБ для забезпечення протидії спробам нелегітимної екстракції стеганоконтенту. Наочним підтвердженням цієї можливості є попарне порівняння зразків зображень, які представлені на рис. 2-4 *(а-е)*, де добре проявля-

ються вертикальні регулярні структури блоків (*т.зв. доріжки*), що є наслідком хибного підбору використаних ключових параметрів на довжині стека *всього* в 4 серії. А порівняння тестових зразків, що представлені на рис.7, переконливо демонструє різницю в наслідках атаки контенту, для захисту якого використовуються стеки різної довжини, що формують різне комбінаторне поле для діючих параметрів серій ОБ.

Важливо підкреслити, що на даному циклі моделювання було вимкнуто функцію внутрішньоблокового мультиплексування даних (*зсув значень (0;0) на рис.1*), що добре видно по практично неспотвореним високодетальним областям тестових зображень на рис.7 (*фрагменти обличчя з зачіскою, очима і бородою, та частини зображень, з деталями будівель*). Крім того, варто звернути увагу на те, що спосіб формування базового масиву серій може бути дуже різним, а його конкретна реалізація, також є одним із елементів використовуваної ключової послідовності. Як видно із рис. 2-4 *(б, г),* у даному випадку була використана розгортка по стовпцях (*про що свідчать вертикальні доріжки з помилково відновлених блоків на рис.7).* При цьому поєднання різних способів вибірки діючих пар довжин серій (*ОБ і параметру його довжини*) та різних принципів організації самої розгортки серій (*по стовпцях, зигзаг тощо*), додатково розширює спектр можливих станів відповідного елемента в структурі складеного ключа екстрактора даних (*див. рис.2 в роботі* [1]).

3. Висновки

1. Використання параметра «довжин серій» ОБ, дозволяє отримати набагато більш суттєвіший ефект, ніж при реалізації міжблокового мультиплексу тільки за рахунок зсуву ОБ (*див. рис. 2-5, зразки (а, в, д) проти зразків (б, г, е)).*

2. Результати експериментів підтвердили припущення, що поєднання одразу 2-х зазначених параметрів серій [4], помітно ускладнює процес підбору діючих компонентів складеного ключа екстрактора даних. Причому основну роль відіграє саме параметр довжин серій (*інтенсивність та кількість блоків білого кольору на рис. 5 у зразках (а, в, д) проти (б, г, е)).*

3. Одночасне використання дворівневого мультиплексу даних *(ОБ і параметра довжин серій на 1-му рівні, та значень середньої яскравості ОБ (елемент (0; 0) на рис. 1) на 2-му рівні мультиплексу),* значно розширює можливості протистояння спробам атак контенту.

4. Збільшення довжини стека вибірки серій розширює комбінаторику мультиплексу діючих параметрів серій ОБ, та в більшої мірі руйнує кореляційні зв'язки елементів вихідного масиву даних. Цей ефект виразно підтверджується значним збільшенням щільності розміщення серій блоків різного відтінку у фонових областях тестових зображень, що використовують широку базу перестановок *(див. порівняння зображень в різних стовбцях на рис 7).*

5. Збільшення розмірності ОБ для всіх типів зображень призводить до зменшення загальної кількості серій, що звужує базу можливих перестановок (рис. 2-4(*ж*)). За сукупністю показників, найбільш збалансованою розмірністю блоків є діапазон від 4 до 8 елементів.

6. Застосування блоків більшої розмірності (див. Рис. 2-4 *(д-е)*) значною мірою зменшує роль параметра довжин серій, як основного елемента для «руйнування» вихідної структури зображень. У даному випадку кількість серій, що формуються, практично для всіх типів зображень, знаходиться в одному діапазоні (*нижній помаранчевий графік на рис. 2-4 (ж)).*

7. Використовуваний варіант згладжування вхідних даних значно меншою мірою визначає кількість отриманих серій, ніж це робить параметр загрублення **Pz**. При заданих розмірностях блоків збільшення порогового значення яскравості сусідніх елементів (*зміщення пунктирної зеленої лінії «***Pz***» вправо на рис. 2-4(ж))* помітно змінює можливу комбінаторику перестановок, особливо для блоків малої та середньої розмірності *(4×4 та 8 ×8 ел.).*

8. Використання для контенту та контейнера блоків різної розмірності (*тобто режиму «несиметричної обробки»*) створює хороші вихідні співвідношення для подальшої інкапсуляції контенту.

9. При взаємному порівнянні результатів впливу параметра **Pz** і розмірності блоків на кількість та довжину формованих серій ОБ, безумовним лідером є параметр розмірності блоків (рис. 2-4(*ж*)).

10. Збільшення значення **Pz** в області малої візуальної помітності (*тобто до 7 градацій яскравості*) для різних типів даних, за вектором процесу, що спостерігається (*збільшення або зменшення кількості ОБ*), дає різні результати [2], що пояснюється відмінністю статистичних властивостей обраних тестових зображень [4]. При цьому у всіх випадках спостерігається збільшення помітності спотворень відновлюваних зображень.

11. Впровадження різних способів розгортки серій забезпечує ще одну позицію в структурі ключа екстрактора даних.

Список літератури

[1] Лесная, Ю., Гончаров, Н., & Малахов, С. (2021). Отработка концепта многоуровневого мультиплекса данных гибридного стеганоалгоритма. Збірник наукових праць SCIENTIA. (Vol.2), 48-55. Вилучено з https://ojs.ukrlogos.in.ua/index.php/scientia/article/view/17666

[2] Гончаров О., Лєсная Ю., Погоріла К., Богданова Є., Малахов С. Дослідження параметру «серій опорних блоків», як елементу композитного ключа екстрактора даних стеганоалгоритму // Problems of science and practice, tasks and ways to solve them. Proceedings of the XX International Scientific and Practical Conference. Warsaw, Poland. 2022. Pp. 779-785. Вилучено з https://isg-konf.com/problems-of-science-and-practice-tasks-and-ways-to-solve-them-two/

[3] Ярославский Л. П. (1979). Введение в цифровую обработку изображений. Москва: Сов. Радио.

[4] Прэтт У. (1985). Цифровая обработка изображений (Д. С. Лебедева, пер. с англ.). т. 1,2. Москва: Мир.

[5] Зубарев Ю.Б., Дворкович В.П. Цифровая обработка телевизионных и компьютерных изображений. – М.: МЦНТИ, 1997. – 212 с.

Received: on July 2022. Accepted: on July 2022.

Authors:
Mykyta Honcharov, student of the Department of Security of Information Systems and Technologies, V. N. Karazin Kharkiv National University, Ukraine.
ORCID ID https://orcid.org/0000-0002-9790-7260
E-mail: worldxdark@gmail.com
Yuliia Liesnaia, 4rd year student, Faculty of Computer Science, V.N. Karazin Kharkiv National University, Ukraine.
E-mail: xa12284109@student.karazin.ua

Using the parameters of the series lengths as an element of the interblock data multiplex of the steganoalgorithm.

Annotation. The peculiarities of using the parameters of runs lengths and the number of generated anchor blocks as elements of the composite key of the data extractor for a hybrid steganoalgorithm, are considered. The results of attacking (hacking) test images obtained for sample stacks of different lengths (forming a different base of permutations of valid runs parameters) are presented. It has been concluded that the "runs length" parameter plays the leading role in implementing inter-block multiplexing procedures for steganocontent. It has been emphasized that the simultaneous use of 2-level data multiplexing significantly extends the capabilities to withstand content attack attempts. It has been found that the use of blocks of higher dimensionality, significantly reduces the role of the "runs length" parameter in breaking the structure of the original images. It is stated that increasing the length of a sample stack of runs expands the potential combinatorial multiplexing of the valid pairs of runs parameters and destroys the correlation between the elements of the original data set to a greater extent. Based on the simulation results, it has been concluded that the introduction of different methods of runs scanning provides additional position in the structure of the key of the data extractor.

Keywords: image; steganography; container; content; image smoothing; visual visibility of distortions; series coding; multiplexing.