UDC 004.67(075.8) DOI: 10.26565/2519-2310-2022-2-02

MODELING STEGANOCONTENT EXTRACTION ATTEMPTS WITH DIFFERENT LENGTHS STACK SAMPLING SERIES OF IMAGES BLOCKS

Honcharov Mykyta, Pavlova Larysa, Lesnaya Yulia

V.N. Karazin National University, Kharkiv, Ukraine
worldxdark@gmail.com, l.v.pavlova@karazin.ua, xa12284109@student.karazin.ua

Received: August 2022. Accepted: September 2022

Abstract: *The results obtained by using different lengths of sample stacks of runs in simulating the attempts of unauthorized extraction (attack) of steganocontent "protected" by implementing the mechanism of inter-block multiplexing of the parameters of the run lengths of image blocks have been considered in the article. The relationship between the parameters of processing the content (namely, halftone images) and the number of series, as well as the combinatorics of the component elements of the obtained pairs of series parameters, which are the objects of inter-block multiplexing has been demonstrated. It is concluded 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 current parameters of the series reliance (base) blocks, in breaking the structure of the original images (those. original content). It is noted that use of two levels of multiplexing of output data at once significantly increases the resistance of the content to attempts at its unauthorized extraction, leading to large distortions in the attacked image, in case of incorrect selection of the active processing parameters.*

Keywords*: run-lengths encoding; steganography; content; hacking; stack.*

1. Introduction

This paper presents the results of modeling the procedures of adapting the method of run-length encoding to implement inter-block multiplexing of steganocontent data, as the main method of preventing the illegitimate extraction of data (*in this case image-content*) from a steganocontainer. These experimental results have been obtained as the part of the research aimed at developing the general concept of a low-resource hybrid steganographic algorithm [1-2]. It is important to emphasize that at this stage of modelling preliminary smoothing of the original images has not performed, which slightly increases the total amount of series in the original (base) image-content array. However, in the current prototype of the algorithm various methods of smoothing the original images are used at the stage of data preprocessing, which allows us to obtain the required result from the number of blocks of identical content when the certain criteria for visual detecting distortions is given. Halftone images of three different types, where the main difference is the characteristic values of the probability of brightness gradient between adjacent image elements, have been used as test data samples [3].

2. Main part

In order to analyze the obtained effects, a simplified version of the inter-block multiplexing of data which is limited by the combinatorics of two elements of the composite key of the data extractor (*the number of blocks and the run lengths*) has been used. Furthermore, to simulate the countering of hacking attempts a simplified version of the masks of the inter-block multiplexing of data for two stacks of different size has been implemented. In other words, during the simulation the attacker is assumed to determine the method of scanning the series and the effective parameter of stack length (*runs sampling base*) correctly, but to be mistaken in determining the current parameters of the displacement of base block (***BB***) and the lengths series BB. The consequences of such an attack are well identified by the vertical "*tracks*" of blocks of different brightness in the background areas of the test image in Fig. 1-2. Thus, if that effect is noticeable when a demo stack of a small length is used, it will be even more noticeable for a wider base.

In the first case, four series were used as a stack, therefore a base of mutual permutations of runs parameters was small [4]. In the second case, the length of the sampling stack was equal to the total number of formed series BB, and the mutual multiplexing of series parameters was carried out between its two halves (*half-stacks*), as well as within each of them. In both cases, the *"destruction"* of the initial pairs of the parameters was carried out: BB – the BB series length [5].

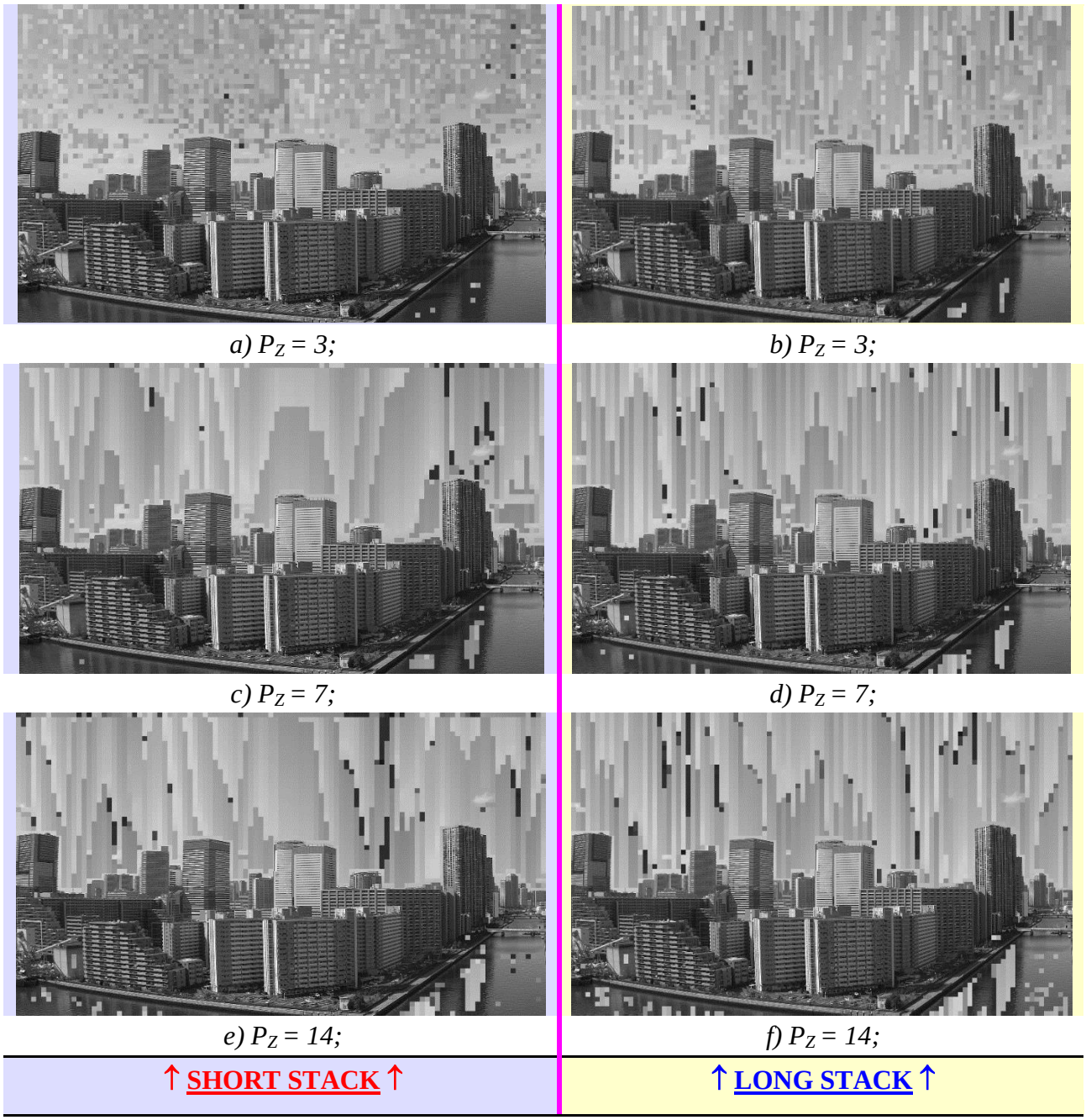


Fig. 1 - The results of attack of the *"landscape"* type test image for stacks of different size and different P_Z *(for BB 8×8 el. scan by "columns")* [7]

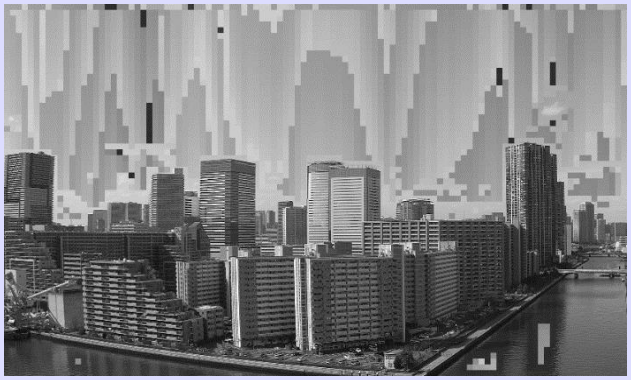
It should be noted that the results of a failed attack (*i.e., an unauthorized extraction of a test image*) presented below were obtained when the source massif had been scanned column by column [6]. The term "*scanning*" in this context should be understood as a method of traversing and consequent extracting the current parameters of the BB runs from the base massif of image – content series. During the simulation, the function of intra-block multiplexing of data [1] was turned off, which is clearly visible in the practically undistorted highly-detailed areas of the test image in Fig. 1-3 (a *part of the image with urban buildings*).

a) BB 4× 4 el., "columns";

b) BB 4 × 4 el., "columns";

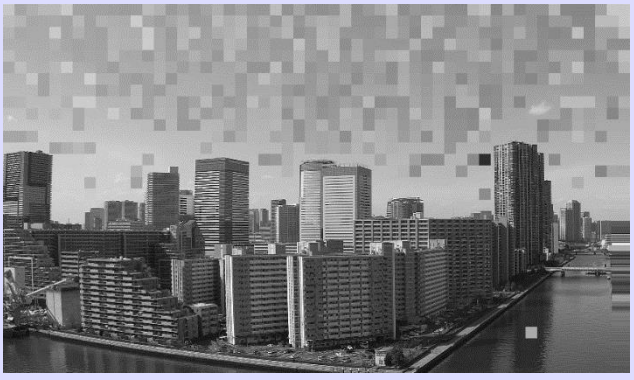
a') BB 4× 4 el., "line by line";

b') BB 4 × 4 el., "line by line";



c) BB 8×8 el.;

d) BB 8×8 el.;



e) BB 16 × 16 el.;

f) BB 16 × 16 el.;

↑ **SHORT STACK** ↑ ↑ **LONG STACK** ↑

Fig. 2 - Attack results of the test image for stacks of different dimensions
(samples a-f, scan by "columns" at P_Z = 5; samples a'-b', scan by "line by line" for P_Z = 7)

In other words, mutual obfuscation of significant elements for different BBs was not carried out. In addition, to present the total number of series subjected to the inter-block multiplexing visually, all series BB were marked with white *(see Fig. 3 (b, d, f))*. It is important to emphasize that in the case of Fig. 3, a sample stack of a small size was used (*4 series*), which did not affect the sizes of the areas of the test images for which the procedure of inter-block multiplexing of the effective parameters of the runs under the specified limits of the value of P_Z (*where* $P_Z \leq 14$) was implemented.

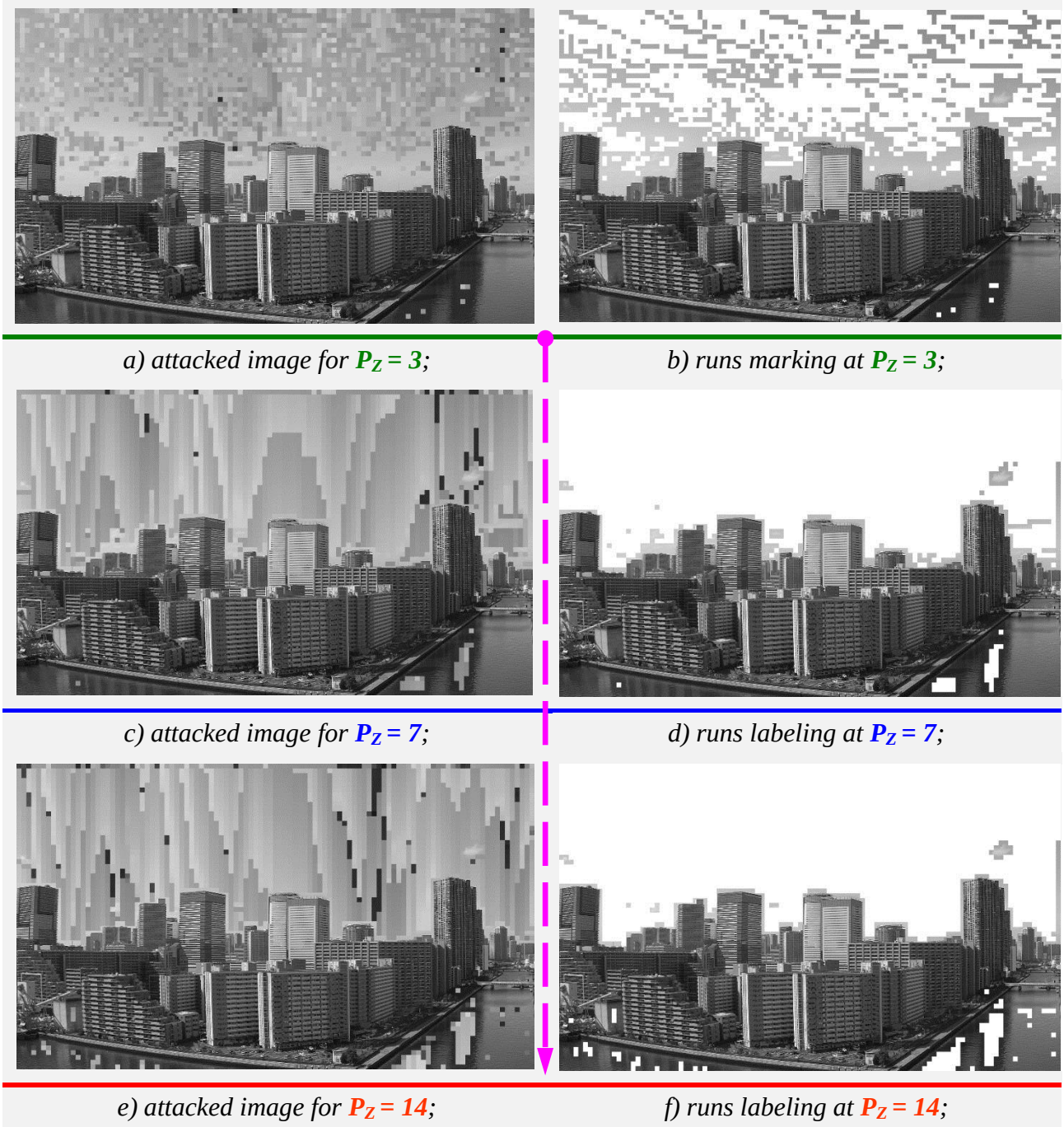


Fig. 3 - The result of image attack *(a, c, e)* and labeling of all series BB *(b, d, f)* for different values P_Z *(for BB 8×8 el., scan by "columns")* [7]

3. Conclusions

1. Increasing the length of the sample stack of runs expands the combinatorics of inter-block multiplexing for both parameters of the formed runs, which significantly destroys the correlations

between the elements of the original array of series of the content. This effect is clearly confirmed by a significant increase in the density of placement of series of different shades in the background areas of the attacked test image, which uses a wide base of permutations *(see comparison of Fig. 1(c) and Fig. 1(d))*.

2. Regardless the size of a sample stack, an increase in the value of the coarsening threshold P_Z results in an increase in the length of the formed BB series *(see the comparison of image columns (a, c, e) and (b, d, f) in Fig. 1)*. In addition, $P_Z \leq 7$ should be considered acceptable for the vast majority of images. In doing so $P_Z = 14$ (*for 256 brightness levels*) should be considered critical (limiting) for most realistic images (*such as portrait and landscape*) [3]. The use of large values leads P_Z to serious degradation of the original data (*red zon in Fig. 1(e)* [8]).

3. Regardless the size of a sample stack, an increase in the size of the BB results in a simultaneous decrease in the number of the series, and in their average length *(clearly visible in Fig. 2, samples (b) and (f))*. This leads to a reduction in the combinatorics of permutations for the current parameters of the series within the limits of the adopted multiplexing masks.

4. With equal values of P_Z, using a longer stack undeniably increases the possibility of resisting attempts of illegitimate extraction (*selection of current multiplexing parameters*) of content (*see of samples comparison (a-b), (c-d) and (e-f), in Fig. 2*).

5. With an increase the threshold value of the acceptable difference in the brightness of the elements of the adjacent blocks of the image (i.e., the P_Z) increasing, the total number of series decreases, and their length increases *(see samples a, c, e, in Fig. 3)*.

6. As the value of P_Z increases by more than 7 gradations of brightness *(when quantizing of the elements is 8 bit/el.)* in the background areas of the test images, there is an erroneous "dropping" of blocks with uncharacteristic brightness of constituent elements *(see chains of black runs in Fig. 3(c) and Fig. 3(e))*. Such an effect does not conform to acceptable level of content distortions, especially for the less informative background areas.

7. It's obvious that the use of two levels of multiplexing [1] of output data at once significantly increases the resistance of the content to attempts at its unauthorized extraction, leading to large distortions in the attacked image, in case of incorrect (erroneous) selection of the current processing algorithm parameters.

8. The dimensionality of the BB and the method of organizing the sweep of the blocks series *(see samples a-b and a′-b′, in Fig. 2)*, are elements of the composite key of the data extractor [1], which determine the current procedure for the implementation of interblock data processing procedures *(1st level of protection)*, as a tools for legitimizing access to content data.

References

[1] Лесная, Ю., Гончаров, Н., & Малахов, С. (2021). ОТРАБОТКА КОНЦЕПТА МНОГОУРОВНЕВОГО МУЛЬТИПЛЕКСА ДАННЫХ ГИБРИДНОГО СТЕГАНОАЛГОРИТМА. Збірник наукових праць SCIENTIA. Вилучено із https://ojs.ukrlogos.in.ua/index.php/scientia/article/view/17666

[2] Гончаров, М., Лєсная, Ю., & Малахов, С. (2021). Дослідження властивостей прототипу гібридного стеганоалгоритму. Комп'ютерні науки та кібербезпека, (2), 45-56. https://doi.org/10.26565/2519-2310-2021-2-05

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

[4] Гончаров, Н., & Малахов, С. (2022). Использование параметра длин серий, как элемента межблочного мультиплекса данных стеганоалгоритма. Збірник наукових праць ΛΌГОΣ, 180-187. https://doi.org/10.36074/logos-08.07.2022.050

[5] Бутаков Е. А., Островский В. И., & Фадеев И. Л. (1987). Обработка изображений на ЭВМ. Москва: Радио и связь.

[6] Гончаров, Н., Лесная, Ю., & Малахов, С. (2022). Адаптация принципа кодирования длин серий для противодействия попыткам неавторизованной экстракции стеганоконтента. Grail of Science, (17), 241-247. https://doi.org/10.36074/grail-of-science.22.07.2022.042

[7] Гончаров, М., Лєсная, Ю., & Малахов, С. (2022). МОДЕЛЮВАННЯ СПРОБ ЕКСТРАКЦІЇ СТЕГАНОКОНТЕНТА ПРИ РІЗНІЙ ДОВЖИНІ СТЕКУ ВИБІРКИ ПАРАМЕТРІВ СЕРІЙ. *Grail of Science*, (18-19), 173–177. https://doi.org/10.36074/grail-of-science.26.08.2022.31

[8] Гончаров Н., Лєсная Ю., Семёнов А., Малахов С. Моделирование атаки стеганоконтента на коротком стеке выборки параметров серий при грубых оценках подобия исходных данных. // The main prospects for the development of science in modern life. Proceedings of the XXXVI International Scientific and Practical Conference. Warsaw, Poland. 2022. Pp. 344-347 URL: https://isg-konf.com/the-main-prospects-for-the-development-of-science-in-modern-life/

Надійшла: серпень 2022. Прийнята: вересень 2022.

Автори:

Микита Гончаров, студент факультету комп'ютерних наук (магістрат), Харківський національний університет імені В.Н. Каразіна, Україна.

ORCID ID https://orcid.org/0000-0002-9790-7260

E-mail: worldxdark@gmail.com

Лариса Павлова, ст. викладач кафедри іноземних мов професійного спрямування факультету іноземних мов, Харківський національний університет імені В.Н. Каразіна, Україна.

ORCID ID https://orcid.org/0000-0002-5854-4209

E-mail: l.v.pavlova@karazin.ua

Юлія Лєсная, студентка факультету комп'ютерних наук (магістрат), Харківський національний університет імені В.Н. Каразіна, Україна.

E-**mail**: xa12284109@student.karazin.ua

Моделювання спроб вилучення стеганоконтенту з різною довжиною стеку вибірки серій, блоків зображень.

Анотація. Розглянуто результати, які отримані при використанні стеків вибірки серій різної довжини, при моделюванні спроб несанкціонованого вилучення (атаки) стеганографічного контенту, що «захищається» за допомогою реалізації механізму міжблокового мультиплексування діючих параметрів серій опорних блоків зображення. Підтверджується взаємозв'язок між параметрами обробки контенту (напівтонових зображень) та кількістю серій, а також комбінаторикою складових елементів діючих пар параметрів серій, що є об'єктами міжблокового мультиплексування. Зроблено висновок, що одночасне використання 2-рівневого мультиплексування даних, значно розширює можливості протистояти спробам атаки контенту. Встановлено, що використання блоків більшої розмірності істотно знижує роль поточних параметрів серій опорних (базових) блоків, стосовно порушення структури вихідних зображень (вихідного контенту). Відзначено, що використання одразу двох рівнів мультиплексу вихідних даних, в значної мірі посилює стійкість контенту до спроб його неавторизованого вилучення, обумовлюючи великі спотворення в атакованому зображенні, в разі хибного підбору діючих параметрів обробки.

Ключові слова: кодування довжин серій; стеганографія; контент; атака; стек.