УДК 519.6
DOI: 10.26102/2310-6018/2023.41.2.013

Модификация гистограммного метода для стеганоанализа изображений со значительной глубиной искажения

Р.А. Солодуха✉

Воронежский государственный университет инженерных технологий, Воронеж, Российская Федерация
standartal@list.ru✉

Резюме. Актуальность темы исследования обусловлена необходимостью противодействия скрытым каналам передачи данных в форме файловой стеганографии в ведомственных и корпоративных компьютерных сетях. Статья посвящена формированию вектора признаков на основе гистограммы яркости для выявления стеганографии, искажающей несколько битовых плоскостей пространственной области изображения. Предполагается, что данный вид стеганографии наиболее вероятен для использования внутренним нарушителем, так как не требует глубоких базовых познаний в сфере информационных технологий, реализован в программных продуктах сегмента freeware, позволяет осуществить вложение до 50 % от размера контейнера. Для верификации результатов выполнен численный эксперимент. Приведено описание исходных данных и методики эксперимента. Датасеты получены в среде MatLab. Для обеспечения воспроизводимости эксперимента датасеты представлены в Kaggle. Применяется процедура машинного обучения на основе машины опорных векторов (SVM-регрессия). На основе экспериментальных данных рассчитаны базовые метрики результативности машинного обучения по предложенному вектору признаков для BPCS- и LSB-стеганоанализа. Показана зависимость ошибки регрессии для вектора признаков, учитывающего разные битовые срезы. С помощью полученных оценок аналитик может принять решение о включении признаков в комплексный вектор выявления стегановложения.

Ключевые слова: стеганоанализ, вектор признаков, BPCS-стеганография, LSB-стеганография, стеганографический канал, машинное обучение, машина опорных векторов, регрессия.

Для цитирования: Солодуха Р.А. Модификация гистограммного метода для стеганоанализа изображений со значительной глубиной искажения. *Моделирование, оптимизация и информационные технологии.* 2023;11(2). URL: https://moitvivt.ru/ru/journal/pdf?id=1315 DOI: 10.26102/2310-6018/2023.41.2.013

Modification of steganalytic histogram method for images with deep distortion

R.A. Solodukha✉

Voronezh State University of Engineering Technologies, Voronezh, the Russian Federation
standartal@list.ru✉

Abstract. The relevance of the research is due to the need to counteract hidden data transmission channels in the form of file steganography in institutional and corporate computer networks. The article is devoted to the formation of a feature vector based on the brightness histogram to identify the steganography that distorts several bit planes of the spatial domain in the image. It is assumed that this type of steganography is most likely to be used by inner violator because it does not require deep knowledge in the field of information technology. Additionally, it is implemented in software products of the freeware segment and helps to payload up to 50 % of the container size. A numerical experiment was performed to verify the results. The description of the initial data and the experimental methodology is given. Datasets are obtained by MatLab. To ensure reproducibility of the experiments, the datasets

and MatLab scripts are presented in Kaggle. The machine learning procedure based on SVM regression is applied. Based on experimental data, the basic metrics of machine learning effectiveness of feature vectors for BPCS- and LSB-steganalysis are calculated. The dependence of the regression error for feature vectors based on combinations of different bit planes is shown. With the help of the obtained estimates, the analyst can include one features or another in the complex vector.

Keywords: steganalysis, feature vector, reliability, BPCS-steganography, LSB-steganography, steganography channel, machine learning, support vector machine, regression.

For citation: Solodukha R.A. Modification of steganalytic histogram method for images with deep distortion. *Modeling, Optimization and Information Technology*. 2023;11(2). URL: https://moitvivt.ru/ru/journal/pdf?id=1315 DOI: 10.26102/2310-6018/2023.41.2.013 (In Russ.).

Введение

По данным на 2022 год [1, 2] насчитывается около пяти десятков стеганографических программных средств сегмента freeware. Такое многообразие позволяет потенциальному инсайдеру выбрать наиболее подходящий инструмент для того типа контейнеров, которым он может обменяться с внешней средой, не вызывая подозрения со стороны службы безопасности и системы предотвращения утечек данных (DLP - Data Leak Prevention).

Исторически сложилось так, что защита ведомственных компьютерных сетей, в первую очередь, была ориентирована на противодействие угрозам извне. На внутренние угрозы стали обращать внимание начиная с 2000-х годов. Законодательно это отразилось в международных и российских стандартах (например, проверка каналов связи для выявления передачи скрытой информации приведена в разделе «12.5.4 Утечка информации» в стандарте ГОСТ Р ИСО/МЭК 27002-2012). Современные DLP-системы в состоянии перехватывать весь трафик, выходящий за пределы сети, сканируя его на наличие скрытых данных. Теоретически, подобные системы способны отследить структурные стеганографические каналы и сигнатурный стеганоконтент, однако задача обнаружения цифровой стеганографии несравнимо сложнее [3].

Можно предположить, что для инсайдера, пользующегося контейнерами-изображениями, наиболее важной характеристикой стеганографического канала является пропускная способность, так как большой поток исходящих изображений более подозрителен и труднообъясним, нежели разовая передача изображения низкого качества. Таким образом, из средств цифровой стеганографии наиболее вероятно использование BPCS- или LSB (Least Significant Bits)-стеганографии (последней с глубиной искажения в несколько бит). Целью данной работы является формирование и оценка результативности вектора признаков для обнаружения такой стеганографии на основе статистического гистограммного метода Pair of Values.

BPCS-стеганография

Алгоритм BPCS (Bit Plane Complexity Segmentation) – стеганографии предложен Eiji Kawaguchi и Richard Eason [4] в 1998 г. Если традиционные стеганографические методы позволяют внедрить 10-15 % от размеров контейнера, то BPCS поднимает эту планку до 50 %. Эта техника использует свойство человеческого зрения, при которой не различаются разные графические объекты в сложных бинарных шаблонах.

Контейнер разбивается на информативные и шумоподобные блоки в каждой битовой плоскости, затем шумодобные блоки заменяются стегановложением.

Для BPSC-стеганографии битовые плоскости представляются кодом Грея. Авторы называют это переходом от Pure Binary Coded (PBC) bit planes к Canonical Gray Coded (CGC) bit planes. Это связано с тем, что хотя PBC-плоскости обеспечивают больше

места для стегановложения, CGC-плоскости менее подвержены эффекту «Hamming cliff», когда небольшое изменение в значении цвета пикселя приводит к значительным изменениям в его битовом представлении. Например, $127_{10} \rightarrow 0111111_2$, а $128_{10} \rightarrow 10000000_2$.

Для оценки шумоподобности блока используется несколько метрик. Базовая метрика основывается на «black-and-white border (BWB) complexity» – сложности черно-белых границ и представляет собой суммарное количество переходов между 0 и 1 по строкам и столбцам блока. Например, один ноль в окружении единиц дает BWB–сложность 4. Сложность блока оценивается приведенной величиной $\alpha = \frac{BWB}{2 \cdot n \cdot (n-1)}$, где n – размер блока (данное значение получается из BWB для шаблона «шахматка»).

Если ноли и единицы в блоке расположены с определенной периодичностью, в строках или столбцах, то такие блоки для внедрения не подходят, но BWB их принимает за сложные.

Визуальный анализ гистограмм яркости изображения до и после BPCS-преобразования с помощью программы Qtech-HV02 [5] показывает достаточно сильные искажения, что иллюстрирует Рисунок 1 (файл № 8 из коллекции BOSSBase 1.01 [6]), что предполагает осуществление гистограммной атаки.

На Рисунках 1-2 по оси абсцисс отложено значение яркости, по оси ординат количество пикселей.

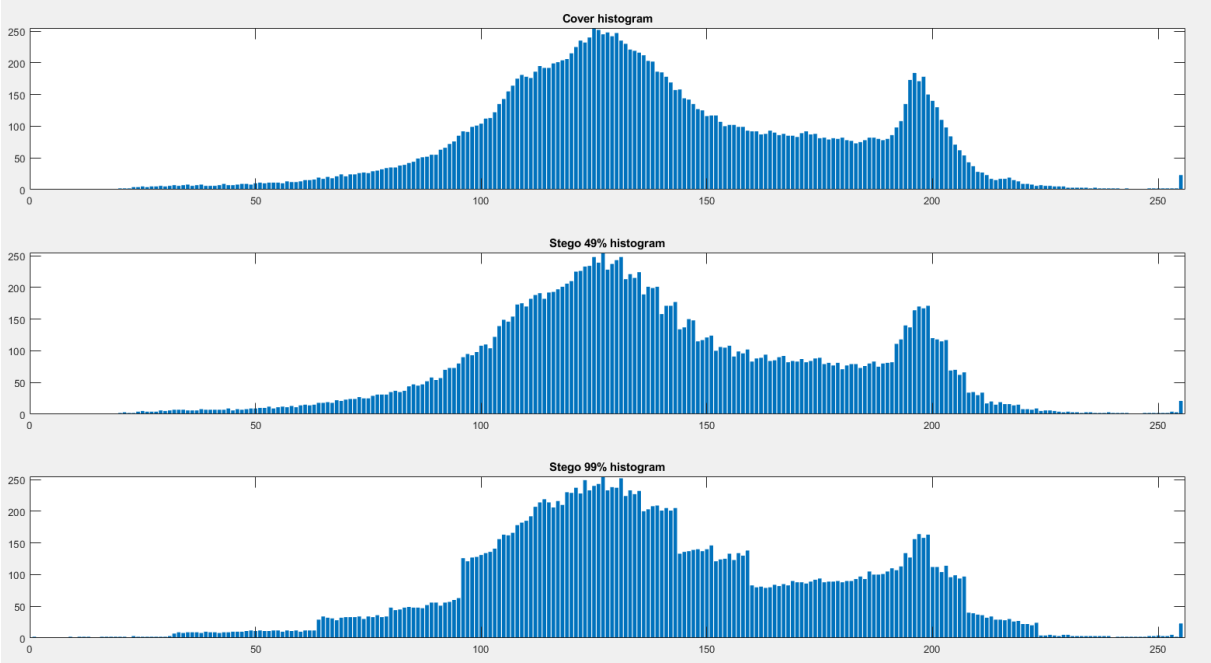

Рисунок 1 – Гистограммы исходного изображения, модифицированного Qtech-HV02, с вложением 49 %, с вложением 99 %
Figure 1 – Histograms of empty image modified by Qtech-HV02 with 49 % payload, with 99 % payload

LSB-стеганография

Метод LSB (Least Significant Bits) – наиболее распространен среди стеганографических методов в пространственной области изображений.

Метод LSB Replacement (LSBR) предложен Е. Адельсоном в 1990 г. [7]. Суть метода заключается в замене наименее значимых (младших) битов в пространственной области исходного изображения битами стегановложения. В основе лежит избыточность

цветового представления цифрового изображения как, изначально, объекта аналоговой природы и погрешность оцифровки. Также существует модификация метода LSB, которую называют ±1 встраиванием, или LSB Matching (LSBM). Данный алгоритм осуществляет псевдослучайные операции инкремента и декремента с целью приведения младших бит контейнера в соответствие с битами сообщения. LSBM разработан в качестве противодействия гистограммной атаке.

Алгоритм LSBR, использующий для стегановложения 4 младших битов, реализован в программе CryptArkan [8], искажения гистограммы иллюстрирует Рисунок 2.

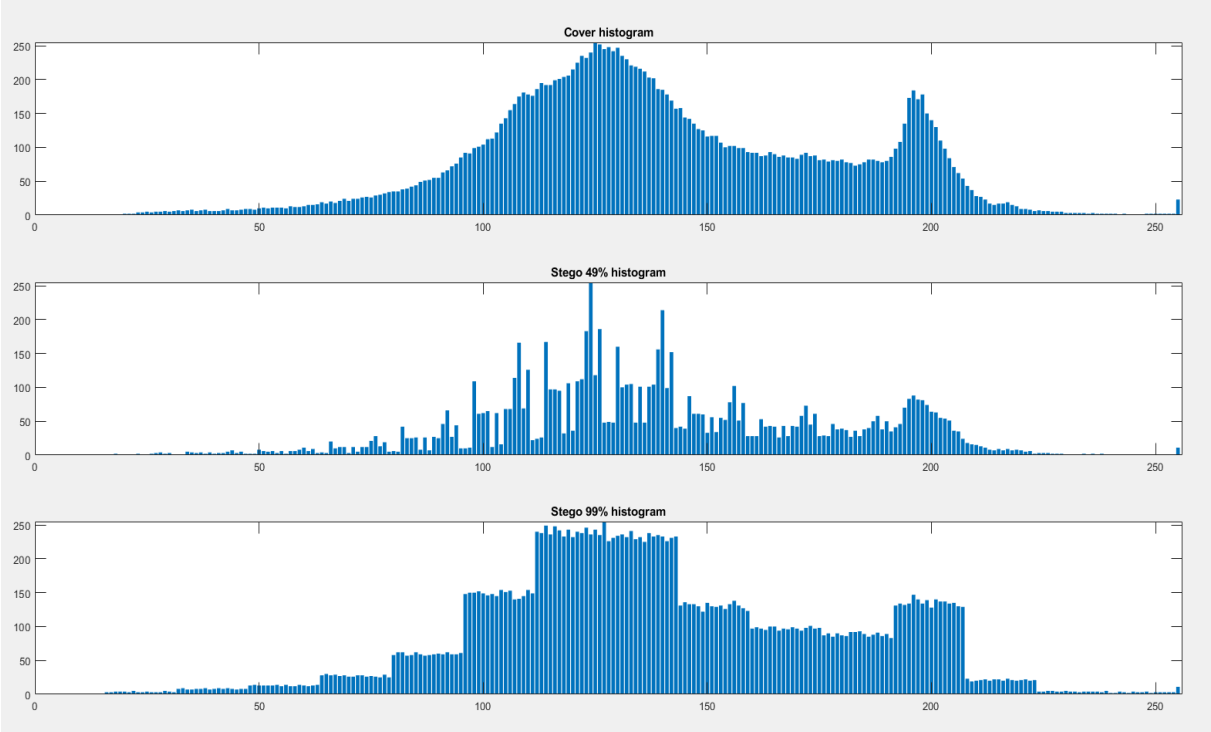

Рисунок 2 – Гистограммы исходного изображения, модифицированного CryptoArkan с вложением 49 %, с вложением 99 %
Figure 2 – Histograms of empty image modified by CryptoArkan with 49 % payload, with 99 % payload

Модификация метода гистограммного анализа

Pair of Values (PoVs) является одним из первых стеганоаналитических методов, выявляющих LSB-replacement на основе гистограммной атаки [9].

Метод основан на сравнении частот появления определенных значений яркости в естественных и заполненных изображениях. При реализации LSB-replacement значение яркости пиксела заполненного контейнера с вероятностью 0.5, в случае равновероятного распределения LSB, что характерно для предварительно заархивированного или зашифрованного вложения.

Пусть $h_c = \{C_i\}$, $h_s = \{S_i\}$ – гистограммы яркости одного из цветовых каналов в цветовой модели RGB, пустого и заполненного контейнера, соответственно:

$$h_c(n) = |\{(i,j)|c(i,j) = n\}|, h_s(n) = |\{(i,j)|s(i,j) = n\}|,$$

где $c(i,j), s(i,j)$ – значения пикселей с координатами (i,j). Пусть глубина цвета на канал 8 бит/пиксел, тогда цветовая гамма изображения содержит 256 оттенков, $0 \le i \le 127$. Тогда верно:

$$S_{2i} = C_{2i} + p^0 \cdot C_{2i+1} - p^1 \cdot C_{2i} = p^0(C_{2i} + C_{2i+1}), \quad (1)$$

$$S_{2i+1} = C_{2i+1} + p^1 \cdot C_{2i} - p^0 \cdot C_{2i+1} = p^1(C_{2i} + C_{2i+1}), \quad (2)$$

где p^0, p^1 – частоты появления ноля и единицы во встраиваемой битовой строке. Заметим, что при $p^0 = p^1$ верно $S_{2i} = S_{2i+1}$.

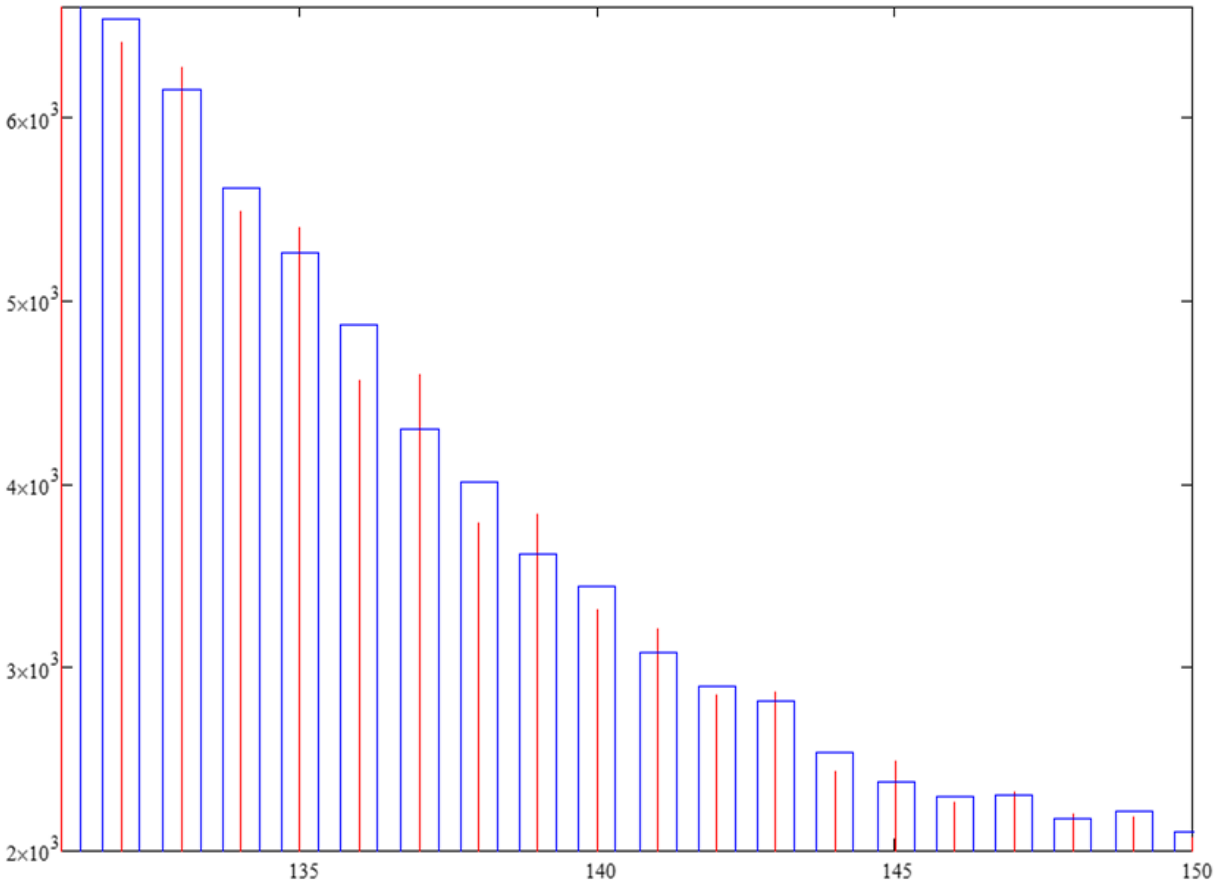

Рисунок 3 – Пример участка гистограммы, S – линии, C – прямоугольники
Figure 3 – Example of a histogram part, S – lines, C – rectangulars

Пары S_{2i}, S_{2i+1}. – «Pair of Values», индексы которых различаются только значением наименее значащего бита, представляют собой частоты появления младшего четного и старшего нечетного значений яркости.

Согласно [10], если $\tilde{n}_i = \frac{(h_{2i}+h_{2i+1})}{2}$, $n_i = \frac{h_{2i}}{2}$ (или $n_i = \frac{h_{2i+1}}{2}$), $0 \le i \le 127$, то справедлива гипотеза о том, что для изображения с максимальным вложением и равновероятным нахождением ноля и единицы в плоскости LSB и стегановложении должно выполняться $\tilde{n}_i \approx n_i$. Маркером наличия вложения является значение статистической значимости различий между фактическими и теоретическими значениями выборки по критерию Пирсона (χ^2). В качестве теоретических частот берутся значения $\tilde{n}_i$ в качестве наблюдаемых n_i. Величина χ^2_{k-1}оценивает разницу между двумя распределениями по *k-1* числу степеней свободы:

$$\chi^2_{k-1} = \sum_{i=1}^{k} \frac{(n_i - \tilde{n}_i)^2}{\tilde{n}_i}. \quad (3)$$

Сравнение двух распределений осуществляют согласно следующей формуле:

$$p = 1 - \frac{1}{2^{\frac{k-1}{2}} G\left(\frac{k-1}{2}\right)} \int_0^{\chi^2_{k-1}} e^{\frac{-x}{2}} x^{\frac{k-1}{2}-1} dx, \quad (4)$$

где G – гамма-функция Эйлера. Вопрос корректности использования теоремы Пирсона рассмотрен в [8].

Суть модификации в следующем: когда стеганоалгоритм затрагивает более старшие битовые плоскости, сравнению подлежат уже не соседние элементы гистограммы, а находящиеся на расстоянии 2^{m-1}. Под m понимается номер плоскости LSB: $LSB_1 – 2^0$, $LSB_2 – 2^1$, $LSB_3 – 2^2$, $LSB_4 – 2^3$. Применение PoVs к каждому битовому срезу дает один элемент вектора признаков. Механизм вычисления элементов гистограммы, подлежащих попарному сравнению для LSB_1, LSB_2 приведен на Рисунке 4, для LSB_3, LSB_4 – на Рисунке 5. Значение PoVs для LSB_5 формируется аналогично с $i=0..7, j=0..15$.

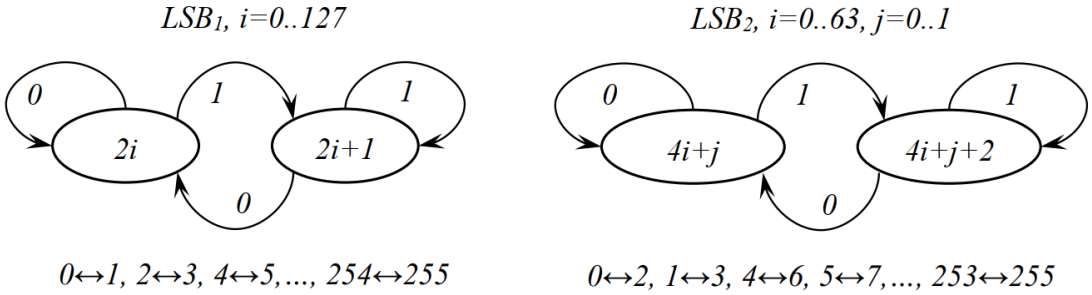

Рисунок 4 – Диаграммы переходов в первой-второй младших битовых плоскостях
Figure 4 – Transition diagrams in the first-second least bit planes

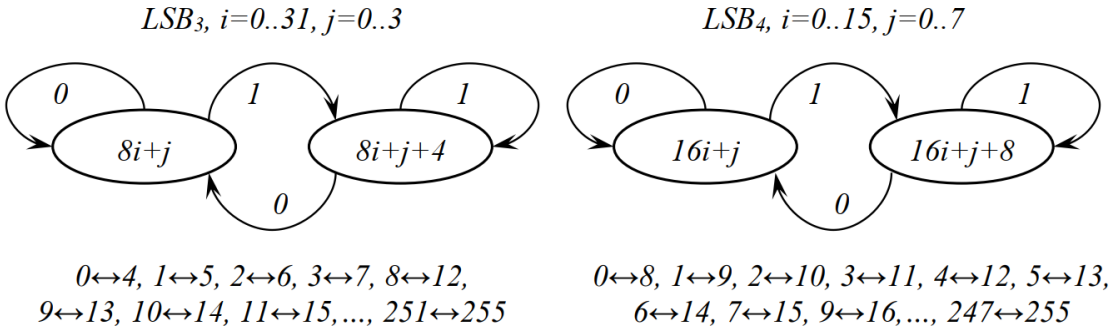

Рисунок 5 – Диаграммы переходов в третьей-четвертой младших битовых плоскостях
Figure 5 – Transition diagrams in the third-fourth least bit planes

Каждый элемент вектора признаков представляет собой значение, вычисленное по (3) для соответствующей LSB_m. Например, для $m=3$: $\tilde{n}_i = \frac{(h_{8i+j}+h_{8i+j+4})}{2}$, $n_i = \frac{h_{8i+j}}{2}$ (или $n_i = \frac{h_{8i+j+4}}{2}$), $0 \le i \le 31$, $0 \le j \le 3$. Преобразование (4) не применяется, так как бинаризирует результаты, что приводит к нецелесообразности применения регрессии.

В итоге, с помощью модифицированного PoVs формируется 5D вектор признаков для BPCS в реализации Qtech-HV02 и 4D для реализации CryptArkan.

Программный эксперимент и результаты

Для эксперимента использованы первые 1000 8-битных полутоновых изображений размером 512x512 пикселей в формате PGM (Portable Grey Map) из коллекции BOSSbase 1.01. Для дальнейшей работы файлы были преобразованы в формат BMP. Формирование контейнеров осуществлено с помощью скриптов AutoIt [11].

Контейнеры заполнялись с шагом 10 % от максимального размера вложения от 9 до 99 %. Таким образом, выборка составила 11000 контейнеров. Схема подготовки эксперимента соответствовала изложенной в [12]. Учитывая значительное количество классов – 11, в качестве прогнозной модели выбрана регрессионная, а именно, машина опорных векторов с гауссовским ядром. В качестве среды машинного обучения использовался MATLAB Regression Learner с регрессором Fine Gaussian из группы SVM с настройками по умолчанию. Выборка делилась на обучающую / тестирующую в соотношении 75 / 25. В качестве метрик результативности машинного обучения использованы: коэффициент детерминации (R-Squared) и среднеквадратическая ошибка (RMSE).

Результаты экспериментов приведены в Таблице 1. Видно, как с ростом количества признаков распознавание улучшается, что подтверждает эффективность предложенной модификации. Следует отметить, что применение иных моделей, доступных в MATLAB Regression Learner, в том числе ансамблевых, не приводило к улучшению метрик результативности. Датасеты доступны в Kaggle: https://www.kaggle.com/datasets/romansolodukha/povs10005lsbsqtech4lsbscryptarkan.

Таблица 1 – Результаты работы модифицированного PoVs
Table 1 – Result of modified PoVs

Стегано-программа	Метрика	LSB_1 (1D)	LSB_{1-2} (2D)	LSB_{1-3} (3D)	LSB_{1-4} (4D)	LSB_{1-5} (5D)
Qtech-HV02	RMSE	*31*	*29*	*28,5*	*27,4*	*26,5*
	R-Squared	*0,03*	*0,11*	*0,19*	*0,24*	*0,3*
CryptoArkan	RMSE	*24,6*	*22,6*	*20,3*	*18*	-
	R-Squared	*0,4*	*0,49*	*0,59*	*0,67*	-

Заключение

Предложена модификация гистограммного метода для анализа изображений с глубиной искажения в несколько бит. Основным результатом является экспериментальное подтверждение того, что выявленный механизм отображения искажений в битовых слоях изображения на гистограмму яркости верен. Также следует отметить, что, хотя модификация улучшает распознавание, полученные на ее основе векторы признаков недостаточно эффективны при автономном использовании. Один из вариантов применения разработанных векторов признаков – в составе комплексного вектора для анализа глубоких искажений изображения при атаке на основании известной стеганопрограммы / алгоритма.

В перспективе планируется провести эксперименты по формированию векторов признаков на базе других стеганоаналитических алгоритмов [13] с целью стеганоанализа изображений с искажением нескольких битовых слоев.

СПИСОК ИСТОЧНИКОВ

1. Питолин А.В., Преображенский Ю.П., Чопоров О.Н. Исследование возможностей использования стеганографических способов защиты информации. *Моделирование, оптимизация и информационные технологии.* 2018;6(2). Доступно по: https://moit.vivt.ru/wp-content/uploads/2018/04/PitolinSoavtors_2_18_1.pdf (дата обращения: 10.12.2022).
2. Свидетельство о регистрации базы данных № 2022620647 Российская Федерация. Трасологическая и таксономическая информация о стеганографических приложениях, использующих в качестве контейнеров файлы растровых графических форматов / Р.А. Солодуха, А.А. Волков, А.Г. Кромских, А.О. Ефимов.
3. Солодуха Р.А. Концепция формирования системы противодействия стеганографическим каналам в компьютерных сетях органов внутренних дел. *Вестник ВИ МВД России*. 2021;1:132–142.
4. Kawaguchi E., Eason R. Principle and applications of BPCS-steganography. *Multimedia Systems and Applications*.1998;3528:464–473.
5. Qtech Hide & View Download Page [Электронный ресурс]. Режим доступа: http://datahide.org/BPCSe/QtechHV-download-e.html (дата обращения 10.12.2022).
6. Image Database BOSSbase 1.01 [Электронный ресурс]. Режим доступа: http://dde.binghamton.edu/download/ImageDB/BOSSbase_1.01.zip (дата обращения: 10.01.2022).
7. Adelson E. Digital Signal Encoding and Decoding Apparatus. – U.S. Patent. – No. 4939515, 1990.
8. CryptArkan encrypts and hides data files and directories [Электронный ресурс]. Режим доступа: https://cryptarkan.software.informer.com/download/ (дата обращения 10.12.2022).
9. Westfeld A., Pfitzmann A. Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and S-Tools-and Some Lessons Learned. *3rd International Workshop on Information Hiding*. 2000.
10. Солодуха Р.А., Атласов И.В. Модификация метода pair of values для атаки на основании известного стегановложения. *Вестник ВИ МВД России.* 2012;3:194–200.
11. Свидетельство о государственной регистрации программы для ЭВМ № 2022682838 Российская Федерация. Автоматизация стеганографических приложений / Р.А. Солодуха.
12. Минайчев А.А., Мезенцев А.О., Яндашевская Э.А. Разработка системы стегоанализа цифровых изображений на основе нейросетевого классификатора. *Моделирование, оптимизация и информационные технологии*. 2022:10(2). Доступно по: https://moitvivt.ru/ru/journal/pdf?id=1196 DOI: 10.26102/2310-6018/2022.37.2.020.
13. Солодуха Р.А., Атласов И.В., Кубасов И.А. *Стеганоанализ цифровых изображений: технологии, алгоритмы, программная реализация*: монография. Воронеж: Воронежский институт МВД России; 2022. 172 с.

REFERENCES

1. Pitolin A.V. Preobrazhenskii Yu.P. Choporov O.N. A study of the possibilities of using steganographic methods of information protection. *Modelirovanie, optimizaciya i informacionnie tehnologii = The scientific journal modeling, optimization and information technology*. 2018;6(2). URL: https://moit.vivt.ru/wp-content/uploads/2018/04/PitolinSoavtors_2_18_1.pdf (In Russ.). (accessed on 10.12.2022).

2. Russia Patent № 2022620647. Trasologicheskaya i taksonomicheskaya informaciya o steganograficheskih prilojeniyah_ ispolzuyuschih v kachestve konteinerov faili rastrovih graficheskih formatov / R.A. Soloduha, A.A. Volkov, A.G. Kromskih, A.O. Efimov. (In Russ.).
3. Soloduha R.A. Koncepciya formirovaniya sistemy protivodeystviya steganograficheskim kanalam v komputernyh setyah organov vnutrennih del. *Vestnik VI MVD Rossii*. 2021;1:132–142. (In Russ.).
4. Kawaguchi E., Eason R. Principle and applications of BPCS-steganography. *Multimedia Systems and Applications*, 1998;3528:464–473.
5. Qtech Hide & View Download Page. URL: http://datahide.org/BPCSe/QtechHV-download-e.html (accessed on 10.12.2022).
6. Image Database BOSSbase 1.01 URL: http://dde.binghamton.edu/download/ImageDB/BOSSbase_1.01.zip (accessed on 10.01.2022).
7. Adelson E. Digital Signal Encoding and Decoding Apparatus. – U.S. Patent. – No. 4939515, 1990.
8. CryptArkan encrypts and hides data files and directories. URL: https://cryptarkan.software.informer.com/download/ (accessed on 10.12.2022).
9. Westfeld A., Pfitzmann A. Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and S-Tools-and Some Lessons Learned. *3rd International Workshop on Information Hiding*. 2000.
10. Soloduha R.A., Atlasov I.V. Modificaciya metoda pair of values dlya ataki na osnovanii izvestnogo steganovlozeniya. *Vestnik VI MVD Rossii*. 2012:3:194-200. (In Russ.).
11. Russia Patent № 2022682838. Avtomatizaciya steganograficheskih prilojenii / R.A. Soloduha. (In Russ.).
12. Minaychev A.A., Mezentcev A.O., Yandashevskaya E.A. Development of a steganalysis system for digital images based on a neural network classifier. *Modeling, Optimization and Information Technology*. 2022:10(2). URL: https://moitvivt.ru/ru/journal/pdf?id=1196 DOI: 10.26102/2310-6018/2022.37.2.020 (In Russ.). (accessed on 10.12.2022).
13. Soloduha R.A., Atlasov I.V., Kubasov I.A. *Steganoanaliz cifrovyh izobrazhenij: tekhnologii, algoritmy, programmnaya realizaciya*: monografiya. Voronezh: Voronezhskij institut MVD Rossii; 2022. 172 p. (In Russ.).

ИНФОРМАЦИЯ ОБ АВТОРЕ / INFORMATION ABOUT THE AUTHOR

Солодуха Роман Александрович, кандидат технических наук, доцент, доцент Воронежского государственного университета инженерных технологий, Воронеж, Российская Федерация.
e-mail: standartal@list.ru
ORCID: 0000-0002-3878-4221

Roman Aleksandrovich Solodukha, Candidate of Technical Sciences, Assistant Professor, Assistant Professor at Voronezh State University of Engineering Technologies, Voronezh, the Russian Federation.

Статья поступила в редакцию 25.01.2023; одобрена после рецензирования 20.04.2023; принята к публикации 26.05.2023.

The article was submitted 25.01.2023; approved after reviewing 20.04.2023; accepted for publication 26.05.2023.