УДК 519.6
doi:10.31799/1684-8853-2024-3-2-10
EDN: FOOKRY

Научные статьи
Articles

Повышение точности стеганоанализа пространственной области изображений за счет дополнительных стегановложений

***Р. А. Солодуха*[a]**, *канд. техн. наук, доцент, orcid.org/0000-0002-3878-4221, standartal@list.ru*
[a]*Воронежский государственный университет инженерных технологий, Революции пр., 19, Воронеж, 394036, РФ*

Введение: *большинство стеганоаналитических алгоритмов используют стеганографический контейнер в исходном виде, пытаясь найти следы произведенного ранее воздействия. В то же время в случае атаки на основании известного стеганоалгоритма/стеганопрограммы, располагая даже модифицированным контейнером, аналитик может наблюдать закономерности в характере изменений контейнера при стегановложениях различного размера.* ***Цель:*** *сформировать векторы признаков на базе известного стеганоалгоритма и дополнительных вложений для выявления стеганографии в пространственной области изображений.* ***Результаты:*** *с помощью эмулятора показаны расхождения в корреляции значений стеганоалгоритма Triples analysis и глубины искажения контейнера. Разработан вектор признаков для выявления стеганографии пространственной области изображения, его эффективность подтверждена численным экспериментом с использованием регрессионной модели машинного обучения в среде MatLab. Для обеспечения воспроизводимости эксперимента датасеты и программный код представлены в Kaggle. На основе экспериментальных данных рассчитаны базовые метрики результативности машинного обучения. Подтверждено наличие статистических закономерностей отклика контейнера на дополнительные вложения, получены зависимости точности стеганоанализа от размера вектора признаков.* ***Практическая значимость:*** *на примере алгоритмов Bit Plane Complexity Segmentation и Least Significant Bits показана зависимость ошибки регрессии для векторов признаков различного размера. С помощью полученных оценок аналитик может варьировать точность/размер векторов признаков в зависимости от доступных вычислительных мощностей и размера обучающего множества.*

Ключевые слова *— стеганоанализ, вектор признаков, Bit Plane Complexity Segmentation, Least Significant Bits, стеганография, машинное обучение, метод опорных векторов, регрессия, дополнительные вложения, пространственная область.*

Для цитирования: Солодуха Р. А. Повышение точности стеганоанализа пространственной области изображений за счет дополнительных стегановложений. *Информационно-управляющие системы*, 2024, № 3, с. 2–10. doi:10.31799/1684-8853-2024-3-2-10, EDN: FOOKRY
For citation: Solodukha R. A. Increasing the accuracy of spatial domain steganalysis through additional embeddings. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 3, pp. 2–10 (In Russian). doi:10.31799/1684-8853-2024-3-2-10, EDN: FOOKRY

Введение

Доступность программ, реализующих файловую стеганографию (обзор приведен в работе [1]), позволяет любому пользователю компьютера осуществить несанкционированную передачу информации ограниченного доступа из ведомственной/корпоративной компьютерной сети. Наиболее популярными и простыми в использовании контейнерами для цифровой стеганографии [2] являются изображения [3–6] (в том числе векторная графика [7]), аудио- [8, 9] и видеофайлы [10]. При этом графические файлы можно легко замаскировать под составляющие деловой переписки и передать посредством сервиса электронной почты [11, 12].

По данным, приведенным в исследовании компании «СёрчИнформ» (Исследование уровня информационной безопасности в компаниях России и СНГ за 2020 год. https://static.searchinform.ru/uploads/sites/1/2022/05/issledovaniya-2021.pdf), именно электронная почта является «самым популярным каналом для слива данных в компаниях — на них приходится 45 % утечек в России и 41 % в СНГ». Для госсектора отмечено, что 50 % утечек персональных данных происходит посредством электронной почты (Итоги 2023 года. Исследование осведомленности и отношения сотрудников организаций государственного сектора к проблемам защиты персональных данных. https://static.searchinform.ru/uploads/sites/1/2023/10/issledovaniya-gr-2023-itog.pdf).

Системы противодействия утечкам (Data Leakage Prevention, DLP) способны выявить структурную файловую стеганографию и вложения, совершенные программным обеспечением, оставляющим сигнатуру [13]. Цифровой стеганоанализ несравнимо сложнее, и, несмотря на значительное количество методов стеганоанализа [14], такая функциональность в DLP-системах не заявлена. Это может быть связано как с отсут-

ствием спроса из-за непонимания заказчиками серьезности угрозы, так и со сложностью технической реализации проверки и принятия решения в онлайн-режиме, априорной невозможностью получить вывод в категорической форме.

Работы в направлении обнаружения стеганоконтента проводила компания McAfee. Веб-приложение Steganography Analysis Tool (Steganography defense initiative. https://web.archive.org/web/20210420075148/https://www.mcafee.com/enterprise/ru-ru/downloads/free-tools/steganography.html) позволяло проанализировать графический файл на наличие стеганографии. На данный момент страница приложения недоступна, что свидетельствует либо о потере компанией интереса к данному направлению, либо к его засекречиванию.

Примерами популярных стеганоалгоритмов, реализующих вложения в пространственную область изображения, являются BPCS (BitPlane Complexity Segmentation) [15] с глубиной искажения 5 бит или LSB (Least Significant Bits) [16] с глубиной искажения 1 бит, имеющие программные реализации в сегменте freeware как в скомпилированном виде, так и в виде исходного кода на github.com. Доступность потенциальному нарушителю и распространенность определяют выбор данных алгоритмов для экспериментальной части статьи.

Настоящая статья является развитием работы [17], где применялся стеганоаналитический алгоритм RS [18] для групп различного размера [19] и дополнительные вложения. А также продолжает направление [20, 21] по формированию и проверке эффективности векторов признаков с возможностью управления соотношением точность/ресурсоемкость, что важно для потокового режима работы DLP-систем.

Целью данной работы является формирование и анализ эффективности векторов признаков на основе дополнительных вложений для обнаружения BPCS- и LSB-стеганографии, выявление зависимости точности стеганоанализа от размера вектора признаков.

Обоснование идеи исследования

Искажения стеганографических контейнеров имеют закономерности, определяемые характеристиками изображения и стеганоалгоритмом. Стандартным способом увеличения точности стеганоанализа является обработка контейнера различными алгоритмами, т. е. увеличение размерности вектора признаков. При этом контейнер остается неизменным.

Пусть $S(\mathbf{I}, \mathbf{p})$ — стеганографическая функция, где $\mathbf{I}$ — изображение, а $\mathbf{p}$ — стегановложение (битовая строка), $\mathbf{I}' = S(\mathbf{I}, \mathbf{p})$ — модифицированное изображение. Задача статистического стеганоанализа состоит в выявлении взаимосвязи между специфическими характеристиками (признаками наличия стегановложения) $\mathbf{G}_{\mathbf{I}'}$ и размером $\mathbf{p}$.

Дополнительное вложение, осуществленное в $\mathbf{I}'$, также влияет на признаки наличия стегановложения, и этим влиянием можно управлять, варьируя размер дополнительного вложения. На этапе формирования обучающей выборки технология дополнительных вложений предполагает последовательное осуществление первичного $\mathbf{p}_1$ и дополнительного $\mathbf{p}_2$ вложений с размерами $|\mathbf{p}_1|$ и $|\mathbf{p}_2|$, получение контейнеров $\mathbf{I}' = S(\mathbf{I}, \mathbf{p}_1)$, $\mathbf{I}'' = S(\mathbf{I}', \mathbf{p}_2)$. При этом для $\mathbf{p}_3 = \mathbf{p}_1||\mathbf{p}_2$, где $||$ — конкатенация, имеет место соотношение $S(\mathbf{I}, \mathbf{p}_3) \neq S(\mathbf{I}', \mathbf{p}_2)$. Предполагается, что учет различных комбинаций размеров $\mathbf{p}_1$ и $\mathbf{p}_2$ способствует построению более точной регрессионной зависимости между $\mathbf{G}_{\mathbf{I}''}$ и $|\mathbf{p}_1|$, нежели между $\mathbf{G}_{\mathbf{I}'}$ и $|\mathbf{p}_1|$. Предикторами являются $\mathbf{G}_{\mathbf{I}''}$ и $|\mathbf{p}_2|$, зависимая переменная — $|\mathbf{p}_1|$.

Рассмотрим ситуацию с позиций практики стеганоанализа. Аналитику на исследование поступает файл, для которого требуется определить размер вложения, выполненный известной стеганопрограммой. В терминах настоящей статьи аналитик должен сформировать из исследуемого файла $\mathbf{I}'$ несколько файлов $\mathbf{I}''$ с известным размером дополнительного вложения $|\mathbf{p}_2|$, получить вектор признаков $\mathbf{G}_{\mathbf{I}''}$ и определить размер первичного вложения $|\mathbf{p}_1|$ с помощью ранее обученной регрессионной модели.

Идея искажения исходного контейнера, в том числе путем дополнительных вложений, реализована в ряде работ [22–24].

Набор признаков, полученный путем вычисления 23 функционалов от коэффициентов дискретного косинусного преобразования, описан в [22]. Каждый функционал применяется к изображению $\mathbf{J}_1$ и его калиброванной версии $\mathbf{J}_2$. Калиброванный признак рассчитывается как разность $F(\mathbf{J}_1) - F(\mathbf{J}_2)$, если F — скаляр, как L_1 — норма $||\mathbf{F}(\mathbf{J}_1) - \mathbf{F}(\mathbf{J}_2)||$, если $\mathbf{F}$ — вектор или матрица. Калиброванное JPEG-изображение получается следующим образом. Изображение разворачивается из частотного в пространственное представление, обрезается на несколько пикселей по обоим направлениям, опять сжимается в JPEG с прежними параметрами. Калиброванное изображение сохраняет свойства исходного на макроуровне. При этом коэффициенты дискретного косинусного преобразования изменяются за счет переформирования блоков 8×8, но сохраняют влияние процедуры компрессии. Таким образом, калиброванный набор признаков не чувствителен к визуальному контенту изображения, но чувствителен к изменениям при стегановложении.

Дополнительное вложение используется в работе [23] для обхода проблемы “cover source mismatch” при известных алгоритме и размере вложения. Предлагаемый метод заключается в создании «искусственного» обучающего набора, который формируется путем двукратного применения стеганографического алгоритма к исходным контейнерам (как пустым, так и заполненным). В работе показано, как наличие трех множеств: исходного, с дополнительным заполнением, с повторным дополнительным заполнением — позволяет осуществить классификацию «без учителя».

Эффективность распознавания сверточных нейронных сетей ухудшается, если в качестве контейнера использовано уменьшенное за счет интерполяции значений соседних пикселей (downsampling) изображение [24]. В качестве меры противодействия предлагается обучать сверточные нейронные сети на уменьшенных изображениях (полученных как из пустых, так и заполненных контейнеров) с дополнительным одно- и двукратным вложением. При атаке на основании известного стеганоалгоритма точность классификации увеличивается на 34,8 %.

Таким образом, работы [17, 23, 24] свидетельствуют о том, что в случае атаки на основании известного стеганоалгоритма атакующему становится доступна статистика поведения контейнера при различных размерах вложений, пусть и с некоторым «смещением».

Сопоставление характеристик, отображающих степень искажения контейнера, с элементами вектора признаков

Произведем теоретический расчет количества модифицированных пикселей монохромного изображения с учетом дополнительного вложения для алгоритма LSB Replacement, использующего псевдослучайную, без повторений генерацию координат пикселей для модификации.

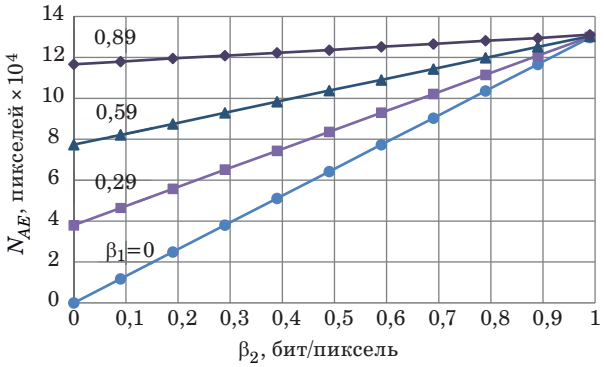
Предположим, что соотношение единиц и нулей в последних битах изображений и встраиваемой битовой строке одинаково, что будет приводить к изменению значений половины пикселей, содержащих скрываемые данные. Обозначим N количество пикселей контейнера, β_1 — размер первичного стегановложения, β_2 — размер дополнительного стегановложения [бит/пиксель], $\mu = 1/2$ пиксель/бит — коэффициент модификации (сколько пикселей изменяется при сокрытии 1 бита). Тогда количество измененных пикселей контейнера после первичного стегановложения (Payload) $N_P = \mu\beta_1 N$. При дополнительном стегановложении изменяются значения пикселей, как модифицированных первичным стегановложением, так и находящихся в исходном состоянии. При этом модифицированные пиксели частично возвращаются в исходное состояние. Количество изменений после дополнительного вложения (Additional Embedding) $N_{AE} = \mu\beta_2(N - N_P) + (N_P - \mu\beta_2 N_P)$.

Подставляя N_P и μ, получим $N_{AE} = (\beta_1 + \beta_2 - \beta_1\beta_2)N/2$.

В графическом представлении получаем семейство прямых, что подтверждено с помощью эмулятора случайного LSB Replacement. На рис. 1 представлено усредненное количество изменений первых 10 файлов из коллекции BOSSbase (Image Database BOSSbase 1.01. http://dde.binghamton.edu/download/ImageDB/BOSSbase_1.01.zip) в допущении, что все пиксели (512 × 512) доступны для модификации стеганоалгоритмом (в реальности часть пикселей отводится под метаданные, параметры алгоритма, хеш пароля и т. п.).

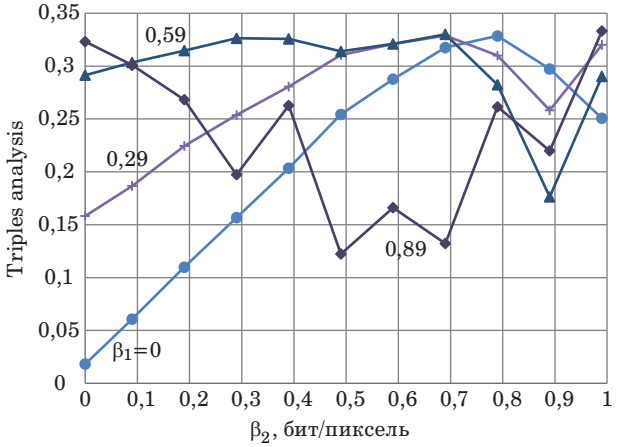
Стеганоалитические предикторы должны коррелировать с количеством изменений. Чем сильнее корреляция, тем точнее результаты. На рис. 2 представлены значения стеганоаналитического алгоритма Triples analysis (TA) [25]. Можно заметить, что кривые TA перестают коррелировать с изменениями, если совокупный размер вложения превышает 50 % [16]. Данный алгоритм показал средние результаты в исследованиях [20, 21] и хорошо подходит для тестирования в рамках настоящей статьи.

Triples analysis является обобщением Sample pairs analysis [26], идея которого состоит в анализе мощности множеств пар соседних пикселей, разности которых принимают одинаковые значения в естественном и модифицированном изображении. Все сводится к решению квадратного уравнения относительно размера вложения, где коэффициенты формируются из значений мощ-



■ ***Рис. 1.*** Усредненное количество модифицированных пикселей

■ ***Fig. 1.*** Average count of modified pixels



■ ***Рис. 2.*** Усредненные значения Triples analysis
■ ***Fig. 2.*** Average values of Triples analysis

ности множеств. ТА оперирует не парами, но тройками смежных пикселей, с решением кубического уравнения.

Формирование вектора признаков

В традиционных стеганоаналитических моделях [14] обучение проводится на выборке, состоящей из пустых и заполненных путем эмуляции стеганографических алгоритмов контейнеров. Контейнеры заполняются с некоторыми фиксированными размерами вложения, как правило, определяемыми в бит на пиксель [бит/пиксель]. Шаг заполнения выбирается исходя из требований к точности прогноза и разделимости получаемых классов.

Если выборка формируется посредством стеганографических приложений, то размер вложения целесообразно учитывать в процентах от максимально возможного для конкретного контейнера (данная информация отображается в стеганографических приложениях, задействованных в экспериментальной части работы). Например, в [20, 21] применяются размеры вложений {9, 19, 29, …, 99} процентов от максимально возможного. Далее понятие «размер вложения» имеет аналогичное содержание.

Пусть имеется набор из N контейнеров $\mathbf{F} = \{\mathbf{f}_n\}$, $n \in [1, N]$. В каждый контейнер реализованы первичные вложения $\mathbf{P} = \{\mathbf{p}_i\}$, $i \in [1, |\mathbf{P}|]$ ($|\mathbf{P}|$ – мощность множества $\mathbf{P}$), получен набор контейнеров $\mathbf{F}_P$, $\mathbf{F}_P = \{\mathbf{f}_n^i, \mathbf{f}_n\}$. Таким образом, для дополнительного вложения исходными являются $|\mathbf{P}|$ файлов с первичным вложением и исходный контейнер.

Затем в каждый контейнер $\mathbf{F}_P$ реализуются дополнительные вложения $\mathbf{A} = \{\mathbf{a}_j\}$, $j \in [1, |\mathbf{A}|]$, получен набор контейнеров $\mathbf{F}_{AE}$. При этом для $\mathbf{p} = \mathbf{a}$, $\mathbf{p} \in \mathbf{P}$, $\mathbf{a} \in \mathbf{A}$ комбинации размеров первичного и дополнительного вложений (0, $\mathbf{a}$) и ($\mathbf{p}$, 0) со статистической точки зрения можно считать идентичными, но для сохранения общности в наименовании файлов обучающего множества они формируются отдельно. Таким образом, $\mathbf{F}_{AE} = \{\mathbf{f}_n^{i,j}, \mathbf{f}_n^i, \mathbf{f}_n^j, \mathbf{f}_n\}$, $|\mathbf{F}_{AE}| = |\mathbf{P}| \cdot |\mathbf{A}| + |\mathbf{P}| + |\mathbf{A}| + 1$. В частном случае $|\mathbf{P}| = |\mathbf{A}|$, рассматриваемом в данной статье, $|\mathbf{F}_{AE}| = (|\mathbf{P}| + 1)^2$.

Предположим, что после применения стеганоаналитической функции S' к множеству $\mathbf{F}_{AE}$ каждому контейнеру сопоставлен набор признаков размером D, $S'(\mathbf{F}_{AE}) \to \mathbf{G}_{AE}$, $\mathbf{G}_{AE} = \{g_n^{i,j,d}\}$, $i \in [0, |\mathbf{P}|]$, $j \in [0, |\mathbf{A}|]$, $d \in [1, D]$, где $i, j = 0$ – индексы, обозначающие отсутствие первичного ($\mathbf{p}_0 \in \varnothing$) или дополнительного вложений ($\mathbf{a}_0 \in \varnothing$) соответственно. Таким образом, задача состоит в нахождении функционала $\Phi(\mathbf{G}_{AE}) \to \{|\mathbf{p}_i|\}$, где $|\mathbf{p}_i|$ – размер $\mathbf{p}_i$, $|\mathbf{p}_0| = 0$. В практическом контексте $\{g^{j,d}\}$ представляет собой матрицу размером $|A| \times D$, вытянутую в вектор:

$$\left(g^{0,1}, \ldots, g^{0,D}, g^{1,1}, \ldots, g^{1,D}, \ldots, g^{|A|,1}, \ldots, g^{|A|,D}\right),$$

которая содержит данные для машинного обучения.

Экспериментальная часть

В качестве программной реализации (программы доступны в Kaggle: https://www.kaggle.com/datasets/romansolodukha/steganoprograms) алгоритма Bit Plane Complexity Segmentation использована Qtech-HV02. Для LSB выбрана модификация LSB Replacement с псевдослучайным выбором пикселей для модификации в реализации The Third Eye. Применена реализация стеганоалгоритма ТА (Structural LSB Detectors. http://dde.binghamton.edu/download/structural_lsb_detectors) на языке MatLab, размещенная на сайте Digital Data Embedding Laboratory (Binghampton University) и модифицированная для применения на разных битовых плоскостях, остальные алгоритмы запрограммированы самостоятельно.

В качестве источника контейнеров использованы первые 1000 файлов коллекции BOSSbase 1.01 с переконвертацией PGM → → BMP24. В полученных файлах все три цветовых канала идентичны, поэтому для анализа использован канал красного цвета.

Для автоматизированного заполнения контейнеров использован скрипт AutoIt с шагом 10 % от максимального размера вложения от 9 до 99 % как для первичного, так и для дополнительного вложения ({|**p**|} = {|**a**|} = {9, 19, 29, …, 99}), выборка составила 121 000 контейнеров.

После преобразования стеганоаналитических данных в вектор признаков получен датасет (доступен в Kaggle: https://www.kaggle.com/datasets/romansolodukha/triples-for-bpsc-lsb-with-ae) из 11 000 элементов, что адекватно решаемой задаче [27].

В качестве прогнозной модели выбран стандартный регрессор Medium Gaussian с ядром $\sqrt{D}$ (D — количество предикторов) на базе метода опорных векторов (Support Vector Machine, SVM) из среды машинного обучения MatLab Regression Learner с настройками по умолчанию. Выбор регрессионной модели обучения обусловлен значительным количеством классов (одиннадцатью). Стандартные настройки не оптимизировались, так как цель исследования не в поиске максимально результативной модели обучения для полученного вектора признаков (как, например, в [28, 29]), а в оценке прироста точности распознавания при учете дополнительных вложений.

В качестве методики машинного обучения использована 5-fold кросс-валидация. В качестве метрик результативности [30] использованы коэффициент детерминации (R-Squared, R^2) и среднеквадратичная ошибка (Root Mean Square Error, RMSE).

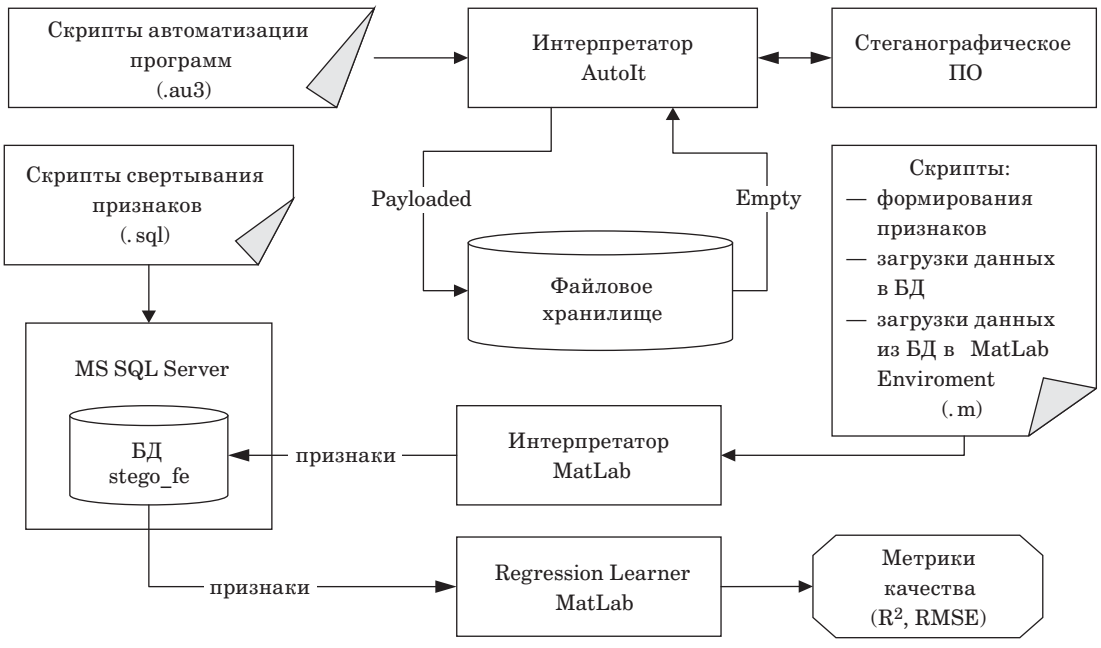
Выбор модели и метрик машинного обучения обусловлен использованием их в ряде подобных работ [23, 30–32], в частности в предшествующих работах [20, 21]. Поскольку речь идет об эффективности вектора признаков, модель машинного обучения, метрики и исходные изображения зафиксированы для адекватного сравнения.

Стенд для проведения эксперимента (рис. 3) собран на базе офисного компьютера Intel i5-12400 2,5 GHz, SSD 500 GB, RAM 32 GB под управлением Windows 10 Pro с установленным программным обеспечением MatLab R2017b, MS SQL Server 2019, AutoIt v3. Для данной конфигурации время преобразования данных в вектор признаков составило 17 с на 1000 строк, время одного вычисления ТА и сохранения результата в базу данных (БД) — 0,3 с.

Одной из задач исследования является оценка влияния на точность распознавания количества и размера учитываемых в векторе признаков дополнительных вложений. Для ранжирования признаков использованы алгоритмы Minimum Redundancy Maximum Relevance и Regressional ReliefF, встроенные в MatLab Regression Learner, которые в целом подтвердили убывание значимости признаков с возрастанием размера первичного вложения, наблюдаемое на рис. 2.

В этой связи план эксперимента содержит последовательное включение элементов в результирующий вектор признаков по мере увеличения размера дополнительного вложения, что отражено в таблице (например, столбец 1D содержит результаты без учета дополнительных вложений, столбец 11D — все дополнительные вложения, столбец 4D — дополнительные вложения 9, 29, 39 %).

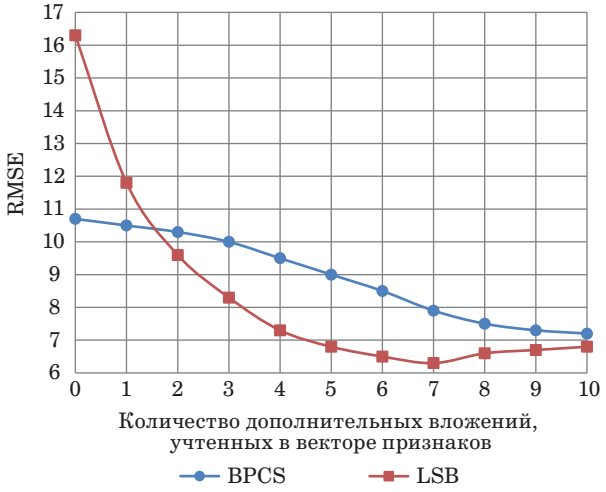
Видно, что с ростом количества признаков распознавание для BPCS плавно улучшается



■ ***Рис. 3.*** Схема численного эксперимента
■ ***Fig. 3.*** Program experiment scheme

■ Результаты применения векторов признаков разного размера
■ Results of applying different dimension feature vectors

Стеганография	Метрика	Вектор										
		1D	2D	3D	4D	5D	6D	7D	8D	9D	10D	11D
BPCS (D = 5)	RMSE	10,7	10,5	10,3	10,0	9,5	9,0	8,5	7,9	7,5	7,3	7,2
	R^2	0,88	0,89	0,89	0,9	0,91	0,92	0,93	0,94	0,94	0,95	0,95
LSB (D = 1)	RMSE	16,3	11,8	9,6	8,3	7,3	6,8	6,5	6,3	6,6	6,7	6,8
	R^2	0,73	0,86	0,91	0,93	0,95	0,95	0,96	0,96	0,96	0,95	0,95



■ ***Рис. 4.*** Зависимость RMSE от размера вектора признаков
■ ***Fig. 4.*** Dependence of RMSE on feature vector dimension

с максимальным приростом R^2 на 0,07 (кривая *BPCS* на рис. 4). Для LSB прирост R^2 составил 0,23, и минимум RMSE достигнут при учете дополнительных вложений размером 9–69 % (кривая *LSB* на рис. 4). RMSE уменьшилось для BPCS в 1,5 раза, для LSB в 2,6 раза.

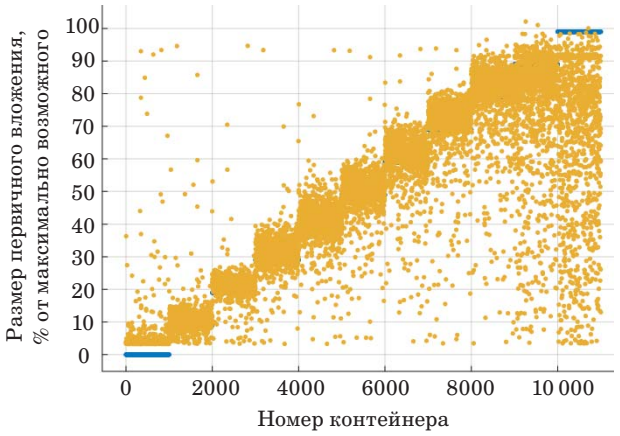
Таким образом, векторы признаков с максимальной достижимой точностью:

BPCS – $\left(g^{0,1}, ..., g^{0,5}, g^{1,1}, ..., g^{1,5}, ..., g^{10,1}, ..., g^{10,5}\right)$, размер вектора – 55;

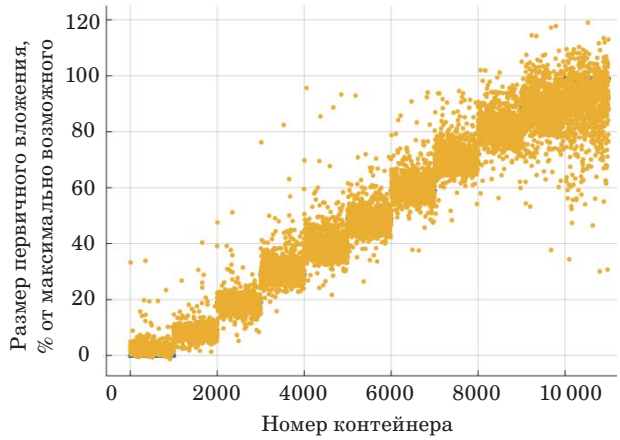
LSB – $\left(g^0, g^1, ..., g^7\right)$, размер вектора – 8.

Визуализированные данные по предсказанию размера вложения, выполненного для контейнеров, модифицированных LSB, представлены на рис. 5, 6. Контейнеры упорядочены по возрастанию размера вложения, синие отрезки – истинные значения.

Значительный прирост точности распознавания находится в области вложений, размер которых более 80 % от максимально возможного. Также за счет выхода прогнозных значений за область допустимых значений [0, 100] результаты можно уточнить, приравнивая к нулю отрицательные прогнозы и к 100 – прогнозы, данное значение превышающие.



■ ***Рис. 5.*** Предсказание без учета дополнительных вложений (один предиктор Triples analysis)
■ ***Fig. 5.*** Prediction without additional embeddings (one Triples analysis predictor)



■ ***Рис. 6.*** Предсказание с учетом дополнительных вложений (восемь предикторов Triples analysis)
■ ***Fig. 6.*** Prediction with additional embeddings (eight Triples analysis predictors)

Заключение

На основе учета результатов применения стеганоаналитического алгоритма ТА к контейнерам с дополнительным заполнением сформированы векторы признаков для цифрового стеганоанализа изображений, модифицированных алгоритмами BPCS и LSB.

Корректировка полученных векторов признаков выполнена по результатам численного эксперимента по определению размера стегановложения. Использована технология машинного обучения, реализованная в среде MatLab, — SVM-регрессия, с оценками коэффициента детерминации и среднеквадратичной ошибки в качестве метрик, что позволяет сравнить полученный результат с аналогичными работами.

Наблюдаемое улучшение распознавания (по метрике RMSE: BPCS — 1,5 раза, LSB — 2,6) подтверждает наличие статистических закономерностей отклика контейнера на дополнительные вложения.

Также получены зависимости оценки точности распознавания от размера вектора признаков, что позволяет аналитику управлять балансом между достоверностью и ресурсоемкостью обнаружения.

В дальнейших исследованиях по данной тематике предполагается провести аналогичный численный эксперимент для частотных областей изображений.

Литература

1. **Герлинг Е. Ю., Ахрамеева К. А.** Обзор современного программного обеспечения, использующего методы стеганографии. *Экономика и качество систем связи*, 2019, № 3 (13), с. 51–58. EDN: KEFWXI
2. **Верещагина Е. А., Золкин А. Л., Капецкий И. О.** *Совершенствование методов аудио-, видео- и сетевой стеганографии*: монография. М., РУСАЙНС, 2023. 140 с.
3. **Савельева М. Г., Урбанович П. П.** Метод стеганографического преобразования web-документов на основе растровой графики модели RGB. *Труды БГТУ. Серия 3: Физико-математические науки и информатика*, 2022, № 2 (260), с. 99–107. doi:10.52065/2520-6141-2022-260-2-99-107, EDN: OMAOWS
4. **Бречко А. А., Булгакова М. И.** Способ скрытия информационного взаимодействия. *Известия ТулГУ. Технические науки*, 2022, № 5, с. 152–158. doi:10.24412/2071-6168-2022-5-152-159
5. **Пономарев И. В., Строкин Д. И.** Стеганографические методы встраивания и обнаружения сокрытых сообщений, использующие gif-изображения в качестве файлов-контейнеров. *Известия Алтайского государственного университета*, 2022, № 1 (123), с. 112–115. doi:10.14258/izvasu(2022)1-18
6. **Мельман А. С., Петров П. О., Шелупанов А. А., Аристов А. В., Похолков Ю. П.** Встраивание информации в JPEG-изображения с маскировкой искажений в частотной области. *Доклады Томского государственного университета систем управления и радиоэлектроники*, 2020, т. 23, № 4, с. 45–50. doi:10.21293/1818-0442-2020-23-4-45-50
7. **Николайчук А. Н., Урбанович П. П.** Стеганографический метод на основе использования особенностей отображения элементов в формате SVG. *Труды БГТУ. Серия 3: Физико-математические науки и информатика*, 2023, № 1 (266), с. 64–70. doi:10.52065/2520-6141-2023-266-1-11
8. **Воронцова Н. В., Миляева И. В.** Стеганографическая защита информации. *Известия Тульского государственного университета. Технические науки*, 2020, № 12, с. 86–95. EDN: YWQQLM
9. **Рублёв Д. П., Макаревич О. Б., Федоров В. М.** Метод стеганографического встраивания сообщений в аудиоданные на основе вейвлет-преобразования. *Известия ЮФУ. Технические науки*, 2009, № 11, с. 199–205. EDN: LAUDHN
10. **Радаев С. В., Басов О. О., Мясин К. И., Мотиенко А. И.** Встраивание стеганографических сообщений в видеофайлы формата MPEG-4. *Экономика. Информатика*, 2018, т. 45, № 4, с. 769–781. doi:10.18413/2411-3808-2018-45-4-769-781
11. **Солодуха Р. А.** Концепция формирования системы противодействия стеганографическим каналам в компьютерных сетях органов внутренних дел. *Вестник Воронежского института МВД России*, 2021, № 1, с. 131–142.
12. **Мисюков Г. И.** Извлечение текстовой информации из изображений модифицированного текста. *Инженерный вестник Дона*, 2023, № 8 (104). http://ivdon.ru/ru/magazine/archive/n8y2023/8625 (дата обращения: 24.03.2024).
13. **Солодуха Р. А.** О возможностях сигнатурного анализа в цифровой стеганографии. *Вестник Воронежского института ФСИН России*, 2016, № 1, с. 52–57.
14. **Вильховский Д. Э.** Обзор методов стеганографического анализа изображений в работах зарубежных авторов. *Математические структуры и моделирование*, 2020, № 4 (56), с. 75–102. doi:10.24147/2222-8772.2020.4.75-102
15. **Kawaguchi E., Eason R.** Principle and applications of BPCS-steganography. *Multimedia Systems and Applications*, 1998, vol. 3528, pp. 464–473.
16. **Powel B. A.** Securing LSB embedding against structural steganalysis. *Journal of Computer Security*, 2021, vol. 30, iss. 42022, pp. 517–539. doi:https://doi.org/10.3233/JCS-200123
17. **Солодуха Р. А.** Использование дополнительного заполнения графических контейнеров для уточнения результатов RS-VGS-стеганоанализа. *Вестник Воронежского института МВД России*, 2014, № 1, с. 87–94.

18. **Fridrich J., Goljan M., Du R.** Reliable detection of LSB steganography in color and grayscale images. *Proc. of the Workshop on Multimedia and Security: Association for Computing Machinery*, New York, 2001. doi:https://doi.org/10.1145/1232454.1232466
19. **Solodukha R. A., Atlasov I. V.** Modification of RS-steganalysis to attacks based on known stego-program. *2017 Second Russia and Pacific Conf. on Computer Technology and Applications (RPC)*, Vladivostok, Russia, 2017, pp. 176–179. doi:10.1109/RPC.2017.8168093
20. **Солодуха Р. А.** Статистический стеганоанализ фотореалистичных изображений с использованием градиентных путей. *Вопросы кибербезопасности*, 2022, № 1(47), с. 26–36. doi:10.21681/2311-3456-2022-1-26-36.
21. **Солодуха Р. А.** Стеганоанализ изображений, модифицированных алгоритмом Bit Plane Complexity Segmentation. *Информационно-управляющие системы*, 2023, № 2, с. 27–38. doi:10.31799/1684-8853-2023-2-27-38, EDN: DXURBZ
22. **Fridrich J.** Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. *Information Hiding, 6th Intern. Workshop*, Toronto, Canada, May 23–25, 2004, Lecture Notes in Computer Science, 2005, vol. 3200, pp. 67–81. doi:10.1007/978-3-540-30114-1_6
23. **Lerch-Hostalot D.**, **Megias D.** Unsupervised steganalysis based on artificial training sets. *Engineering Applications of Artificial Intelligence*, 2016, vol. 50, pp. 45–59. http://dx.doi.org/10.1016/j.engappai.2015.12.013
24. **Kato H., Osuge K., Haruta S., Sasase I.** A preprocessing by using multiple steganography for intentional image downsampling on CNN-based steganalysis. *IEEE Access*, 2020, vol. 8, pp. 195578–195593. doi:10.1109/ACCESS.2020.3033814
25. **Ker A.** A general framework for structural steganalysis of LSB Replacement. *Proc. of the Information Hiding*, 2005, pp. 296–311.
26. **Dumitrescu S., Wu X., Memon D.** On steganalysis of random LSB embedding in continuous-tone images. *IEEE Intern. Conf. on Image Processing*, 2002, vol. 3, pp. 641–644.
27. **Парасич А. В., Парасич В. А., Парасич И. В.** Формирование обучающей выборки в задачах машинного обучения. Обзор. *Информационно-управляющие системы*, 2021, № 4, с. 61–70. doi:10.31799/1684-8853-2021-4-61-70
28. **Сирота А. А., Дрюченко М. А., Иванков А. Ю.** Стегоанализ цифровых изображений с использованием методов поверхностного и глубокого машинного обучения: известные подходы и новые решения. *Вестник ВГУ. Серия: Системный анализ и информационные технологии*, 2021, № 1, с. 33–52. doi:10.17308/sait.2021.1/3369
29. **Полунин А. А., Яндашевская Э.А.** Использование аппарата сверточных нейронных сетей для стегоанализа цифровых изображений. *Труды ИСП РАН*, 2020, № 4, с. 155–163. doi:10.15514/ISPRAS-2020-32(4)-11
30. **Лебедев И. С.** Адаптивное применение моделей машинного обучения на отдельных сегментах выборки в задачах регрессии и классификации. *Информационно-управляющие системы*, 2022, № 3, с. 20–30. doi:10.31799/1684-8853-2022-3-20-30
31. **Shankar D. D., Azhakath A. S.** Minor blind feature based Steganalysis for calibrated JPEG images with cross validation and classification using SVM and SVM-PSO. *Multimed Tools Appl*, 2021, vol. 80, pp. 4073–4092. https://doi.org/10.1007/s11042-020-09820-7
32. **Kheddar H., Hemis M., Himeur Y., Megias D., Amirae A.** Deep learning for steganalysis of diverse data types: A review of methods, taxonomy, challenges and future directions. *Neurocomputing*, 2024, vol. 581, p. 127528. https://doi.org/10.1016/j.neucom.2024.127528

UDC 519.6
doi:10.31799/1684-8853-2024-3-2-10
EDN: FOOKRY

Increasing the accuracy of spatial domain steganalysis through additional embeddings

R. A. Solodukha[a], PhD, Tech., Associate Professor, orcid.org/0000-0002-3878-4221, standartal@list.ru
[a]Voronezh State University of Engineering Technologies, 19, Revolucii Ave., 394036, Voronezh, Russian Federation

Introduction: Most steganalytical algorithms use the steganographic container in its original form, trying to reveal traces of payload. In the case of an attack based on a known steganographic algorithm or program the analyst can observe patterns in the changes of the container caused by various payload values even in a modified container. **Purpose:** To develop feature vectors based on known steganalytical algorithm and additional embeddings to reveal steganography in image spatial domain. **Results:** We show discrepancies in the correlation of Triples analysis results with the depth of container distortion. We develop a feature vector to detect spatial domain steganography. We verify its effectiveness by the numerical experiment using a machine learning regression model in MatLab. To ensure reproducibility of the experiments the datasets and scripts are presented in Kaggle. With the reference to the experimental data we confirm the presence of statistical patterns in the container's response to additional embeddings. We also obtain dependences of the steganalysis accuracy and the feature vector dimension. **Practical relevance:** For Bit Plane Complexity Segmentation and Least Significant Bits algorithms, the dependence of the regression error on different dimensions feature vectors is shown. Using the obtained estimates, the analyst can vary the accuracy/dimension of feature vectors accorfing to the available computing power and the size of the training set.

Keywords – steganalysis, feature vector, Bit Plane Complexity Segmentation, Least Significant Bits, steganography, machine learning, SVM-regression, additional embeddings, spatial domain.

For citation: Solodukha R. A. Increasing the accuracy of spatial domain steganalysis through additional embeddings. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 3, pp. 2–10 (In Russian). doi:10.31799/1684-8853-2024-3-2-10, EDN: FOOKRY

References

1. Gerling E., Ahrameeva K. The review of the modern software using steganography methods. *Ekonomika i kachestvo sistem svyazi*, 2019, no. 3 (13), pp. 51–58 (In Russian). EDN: KEFWXI
2. Vereshchagina E. A., Zolkin A. L., Kapeckij I. O. *Sovershenstvovanie metodov audio-, video- i setevoj steganografii* [Improvement of audio-, video- and network steganography]. Moskow, RUSAJNS Publ., 2023. 140 p. (In Russian).
3. Saveleva M. G., Urbanovich P. P. Method of steganographic transformation of web-documents based on raster graphics and RGB model. *Proceedings of BSTU. Issue 3, Physics and Mathematics. Informatics*, 2022, no. 2 (260), pp. 99–107 (In Russian). doi:10.52065/2520-6141-2022-260-2-99-107, EDN: OMAOWS
4. Brechko A. A., Bulgakova M. I. A method of hiding information communications. *Izvestiya TulGU. Tekhnicheskie nauki*, 2022, no. 5, pp. 152–158 (In Russian). doi:10.24412/2071-6168-2022-5-152-159
5. Ponomarev I. V., Strokin D. I. Steganographic methods for embedding and detecting hidden messages using GIF images as container files. *Izvestiya Altajskogo gosudarstvennogo universiteta*, 2022, no. 1 (123), pp. 112–115 (In Russian). doi:10.14258/izvasu(2022)1-18
6. Melman A. S., Petrov P. O., Shelupanov A. A., Aristov A. V., Pokholkov Y. P. Embedding information into JPEG images with distortion masking in frequency domain. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2020, vol. 23, no. 4, pp. 45–50 (In Russian). doi:10.21293/1818-0442-2020-23-4-45-50
7. Nikolaichuk A. N., Urbanovich P. P. A steganographic method based on the use of the features of elements displaying in SVG format. *Proceedings of BSTU. Issue 3, Physics and Mathematics. Informatics*, 2023, no. 1 (266), pp. 64–70 (In Russian). doi:10.52065/2520-6141-2023-266-1-11
8. Vorontsova N. V., Milyaeva I. V. Steganographic information protection. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki*, 2020, no. 12, pp. 86–95 (In Russian). EDN: YWQQLM
9. Rublyov D. P., Makarevich O. B., Fedorov V. M. Steganographical method for messages embedding to audiodata based on the wavelet-transform. *Izvestiya SFedU. Engineering Sciences*, 2009, no. 11, pp. 199–205 (In Russian). EDN: LAUDHN
10. Radaev S. V., Basov O. O., Myasin K. I., Motienko A. I. Embedding steganographic messages into MPEG-4 video files. *Economics. Information Technologies*, 2018, vol. 45, no. 4, pp. 769–781 (In Russian). doi:10.18413/2411-3808-2018-45-4-769-781
11. Solodukha R. A. Conception of forming the steganographic channels counteraction system in the internal affairs computer networks. *The Bulletin of Voronezh Institute of the Federal Penitentiary Service of Russia*, 2016, no. 1, pp. 52–57 (In Russian).
12. Misyukov G. I. Extraction text information from modified text image. *Inzhenernyj vestnik Dona*, 2023, no. 8 (104) (In Russian). Available at: http://ivdon.ru/ru/magazine/archive/n8y2023/8625 (accessed 29 March 2024).
13. Solodukha R. A. The possibilities of the signature analysis for the digital steganography. *The Bulletin of Voronezh Institute of the Ministry of Internal Affairs of Russia*, 2021, no. 1, pp. 131–142 (In Russian).
14. Vilkhovskiy D. E. A survey of steganalysis methods in the papers of foreign authors. *Mathematical Structures and Modeling*, 2020, no. 4 (56), pp. 75–102 (In Russian). doi:10.24147/2222-8772.2020.4.75-102
15. Kawaguchi E., Eason R. Principle and applications of BPCS-steganography. *Multimedia Systems and Applications*, 1998, vol. 3528, pp. 464–473.
16. Powel B. A. Securing LSB embedding against structural steganalysis. *Journal of Computer Security*, 2021, vol. 30, iss. 42022, pp. 517–539. doi:https://doi.org/10.3233/JCS-200123
17. Solodukha R. A. Additional embedding in graphic stego-container for RS-VGS-steganalysis results refinement. *The Bulletin of Voronezh Institute of the Ministry of Internal Affairs of Russia*, 2014, no. 1, pp. 87–94 (In Russian).
18. Fridrich J., Goljan M., Du R. Reliable detection of LSB steganography in color and grayscale images. *Proc. of the Workshop on Multimedia and Security: Association for Computing Machinery*, New York, 2001. doi:https://doi.org/10.1145/1232454.1232466
19. Solodukha R. A., Atlasov I. V. Modification of RS-steganalysis to attacks based on known stego-program. *2017 Second Russia and Pacific Conf. on Computer Technology and Applications (RPC)*, Vladivostok, Russia, 2017, pp. 176–179. doi:10.1109/RPC.2017.8168093
20. Solodukha R. A. Statistical steganalysis of photorealistic Images using gradient paths. *Voprosy kiberbezopasnosti*, 2022, no. 1(47), pp. 26–36 (In Russian). doi:10.21681/2311-3456-2022-1-26-36
21. Solodukha R. A. Steganalysis of Bit Plane Complexity Segmentation algorithm. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 2, pp. 27–38 (In Russian). doi:10.31799/1684-8853-2023-2-27-38, EDN: DXURBZ
22. Fridrich J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. *Information Hiding, 6th Intern. Workshop*, Lecture Notes in Computer Science, 2005, vol. 3200, pp. 67–81. doi:10.1007/978-3-540-30114-1_6
23. Lerch-Hostalot D., Megias D. Unsupervised steganalysis based on artificial training sets. *Engineering Applications of Artificial Intelligence*, 2016, vol. 50, pp. 45–59. http://dx.doi.org/10.1016/j.engappai.2015.12.013
24. Kato H., Osuge K., Haruta S., Sasase I. A Preprocessing by using multiple steganography for intentional image downsampling on CNN-based steganalysis. *IEEE Access*, 2020, vol. 8, pp. 195578–195593. doi:10.1109/ACCESS.2020.3033814
25. Ker A. A general framework for structural steganalysis of LSB Replacement. *Proc. of the Information Hiding*, 2005, pp. 296–311.
26. Dumitrescu S., Wu X., Memon D. On steganalysis of random LSB embedding in continuous-tone images. *IEEE Intern. Conf. on Image Processing*, 2002, vol. 3, pp. 641–644.
27. Parasich A. V., Parasich V. A., Parasich I. V. Training set formation in machine learning tasks. Survey. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 4, pp. 61–70 (In Russian). doi:10.31799/1684-8853-2021-4-61-70
28. Sirota A. A., Dryuchenko M. A., Ivankov A. Yu. Steganalysis of digital images by means of shallow and deep machine learning: existing approaches and new solutions. *Proceedings of Voronezh State University. Series: Systems Analysis and Information Technologies*, 2021, no. 1, pp. 33–52 (In Russian). doi:10.17308/sait.2021.1/3369
29. Polunin A. A., Yandashevskaya E. A. Using of convolutional neural networks for steganalysis of digital images. *Proc. of the Institute for System Programming of the RAS*, 2020, vol. 32, iss. 4, pp. 155–164 (In Russian). doi:10.15514/ISPRAS–2020–32(4)–11
30. Lebedev I. S. Adaptive application of machine learning models on separate segments of a data sample in regression and classification problems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 20–30 (In Russian). doi:10.31799/1684-8853-2022-3-20-30
31. Shankar D. D., Azhakath A. S. Minor blind feature based Steganalysis for calibrated JPEG images with cross validation and classification using SVM and SVM-PSO. *Multimed Tools Appl*, 2021, vol. 80, pp. 4073–4092. https://doi.org/10.1007/s11042-020-09820-7
32. Kheddar H., Hemis M., Himeur Y., Megías D., Amirae A. Deep learning for steganalysis of diverse data types: A review of methods, taxonomy, challenges and future directions. *Neurocomputing*, 2024, vol. 581, p. 127528. https://doi.org/10.1016/j.neucom.2024.127528