

УДК 004.415.24/004.56.5

*Н.В. Кошкина*

## ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ МАТРИЦЫ СМЕЖНОСТИ ДЛЯ ВЫЯВЛЕНИЯ СТЕГАНОАУДИОКОНТЕЙНЕРОВ

### Введение

С развитием и распространением методов компьютерной стеганографии возрастает актуальность и сложность задачи выявления скрытых сообщений. Решение этой задачи информационной безопасности — основная цель стеганоанализа. Наиболее широкая область применимости у тех методов стеганоанализа, которые не опираются на знание использованного стеганографического программного продукта. Такие методы, как правило, базируются на поиске устойчивых закономерностей в изменениях статистики контейнеров (изображений, аудио-, видеосигналов и др.), вызванных их стеганообразованием, т.е. являются статистическими.

В работах [1, 2] рассматривается статистический аудиостеганоанализ, базирующийся на чувствительности к стегановнедрению матрицы смежности аудиосигнала, задаваемой как

$$P_f(\Delta d, d) = \#\{k \mid (d-1) \cdot \Delta d < |f(k+d) - f(k)| \leq d \cdot \Delta d\}, \quad (1)$$

где  $f(k)$  — проверяемый аудиосигнал,  $\#$  — количество пар отсчетов аудиосигнала, для которых разность амплитуд лежит в заданных пределах,  $d$  — интервал временной задержки между отсчетами в паре,  $\Delta d$  — разрешающая способность анализа.

Каждому проверяемому сигналу ставится в соответствие его характеристический вектор, полученный из  $P_f(\Delta d, d)$  при определенных значениях  $\Delta d$  и  $d$ . Характеристический вектор подается на вход ранее обученного на некоторой выборке пустых и заполненных контейнеров бинарного классификатора SVM [2], на выходе которого он получает одну из меток класса — «С» или «S». Метка «С» указывает на то, что проверяемый аудиосигнал является пустым контейнером, «S» — заполненным.

Подобный подход более развит в применении к стеганоанализу изображений [3–5]. Вопрос использования матрицы смежности для выявления стегановкладок в аудиоконтейнерах малоизучен. После реализации вышеописанного метода с помощью пакета Matlab и исследования получаемых на разных этапах стеганоанализа численных оценок осуществлена его модификация, относительно реализаций, описанных в работах [1, 2]. Например, в отличие от [1] предлагаемая в этой работе модификация не использует формулу  $\Delta d = \{\max[f(k)] - \min[f(k)]\} / 2000$  и гауссовское ядро SVM, поскольку в процессе исследований было определено, что более высокую точность обеспечивают  $\Delta d = 2^{-m}$  и линейное ядро SVM. После сравнительного анализа различных вариантов формирования характеристического вектора в данной работе в отличие от [2] предлагается использовать вектор,

полученный при разной разрешающей способности анализа  $\Delta d$ , что также улучшает точность выявления стеганоконтейнеров. Кроме того, для улучшения точности метода предлагается использовать переход в пространство вейвлет-коэффициентов и расчет матриц смежности детализирующих субполос.

В настоящей работе предложен путь определения не только факта наличия скрытого сообщения, но и его вероятной длины. Для разработанных модификаций метода исследована точность выявления стегановкладок в 8- и 16-битных наборах фрагментов аудиокниг.

### Постановка задачи

Для стеганографического сокрытия данных в аудиосигналах существующее программное обеспечение, как правило, реализует метод наименьшего значащего бита (НЗБ). В качестве исходных аудиоконтейнеров могут использоваться 8- или 16-битные сигналы, содержащие речь, музыку, чистые тоны, тишину, шум. Часть программ, например Steganos Privacy Suite 2012 или Invisible Secrets 2002, в случае коротких сообщений реализует их последовательное внедрение в младшие биты отсчетов сигнала-контейнера. Другие, например S-Tools 4.0 или Hide4PGP, распределяют биты коротких сообщений по всему объему контейнера. Причем при сокрытии данных в 8-битных сигналах в любом случае используется не больше одного младшего бита отсчета сигнала, иначе шум, вносимый стегановнедрением, будет слышимым. В 16-битных сигналах неслышимо могут модифицироваться до четырех младших битов, что, в частности, реализовано в программе Hide4PGP.

Точность статистических методов зависит от деталей реализации и условий экспериментов. В данном методе на точность стеганоанализа будут влиять параметры характеристического вектора и SVM. Численные эксперименты, определяющие точность данного метода, в работе [1] выполнены только для 450 речевых сигналов из базы данных CASIA98-99, оцифрованных с разрядностью 16 бит и частотой дискретизации 16 кГц. Также в [1] не указана наполненность стеганоконтейнеров в выполненных тестах и тип распределения коротких сообщений по контейнеру, что не позволяет оценить эффективность метода в полной мере. В [2] эффективность данного метода проверялась только для такого стеганопреобразования, как модуляция индекса квантования (МИК).

Для корректного применения метода на практике необходимо более детальное его исследование.

Цель работы — исследование при разных условиях стеганоанализа чувствительности к НЗБ-стегановнедрению матрицы смежности аудиосигналов; анализ, определение областей дифференцированного поведения, нахождение оптимальных параметров и развитие вышеописанного метода.

### Стеганоанализ 8-битных аудиосигналов

Для работы с 8-битными аудиосигналами с помощью программы Adobe Audition сформирован набор из тысячи одноминутных фрагментов аудиокниг (разные дикторы обоих полов, возможны участки тишины, фоновая музыка). НЗБ-стеганопреобразованием сигналов этого набора были образованы стеганонаборы с последовательным и распределенным внедрением и наполненностью контейнеров 100, 75, 50, 25, 12 и 6 %.

Определение параметров характеристического вектора  $\Delta d$  и  $d$  выполнялось визуальным анализом разностей матриц смежности наборов пустых и заполненных контейнеров  $G_f = P_{f_{\text{orig}}}(\Delta d, d) - P_{f_{\text{stego}}}(\Delta d, d)$ . Пример такого анализа для 50 % заполненных контейнеров при  $\Delta d = 2^{-m}$ ,  $m = 8,7,6$  приведен на рис. 1. На каждом графике изображены разности  $G_{f_1} \dots G_{f_{1000}}$  для  $d = 1 \dots 30$  ( $a$  — последовательное НЗБ-внедрение;  $b$  — распределенное НЗБ-внедрение).

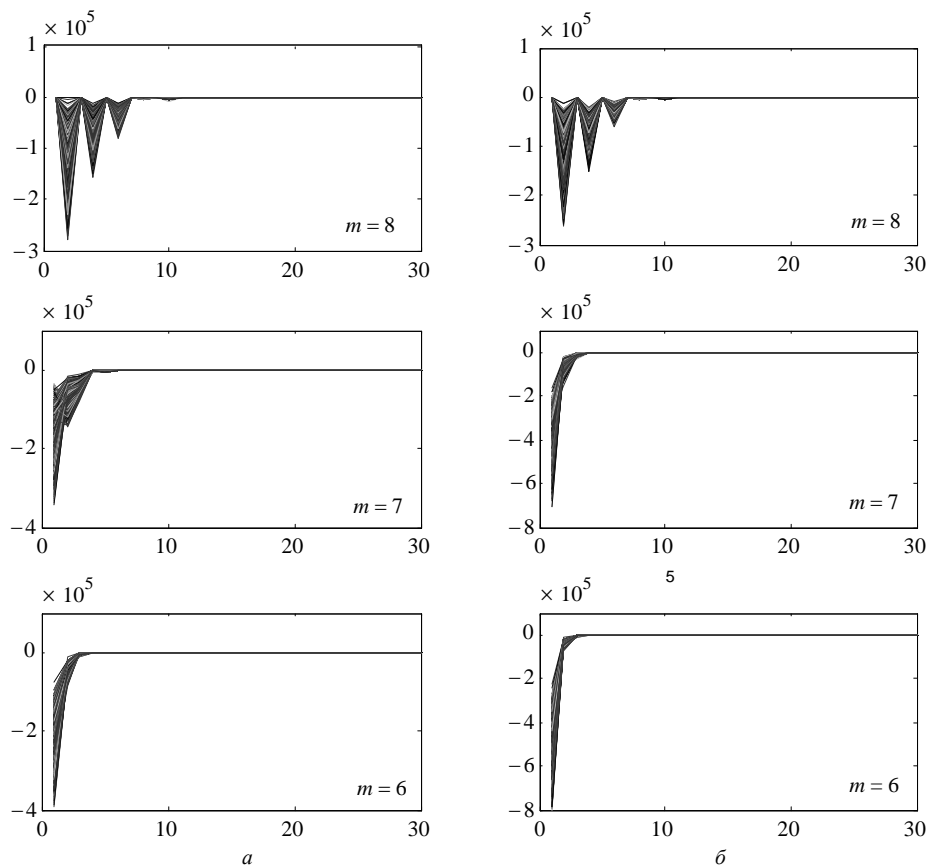


Рис. 1

Как видим, в матрицах смежности пустых и заполненных контейнеров наблюдаются устойчивые различия. Наибольшие различия наблюдаются при малых значениях  $d$ . При увеличении разрешающей способности анализа  $\Delta d$  различия концентрируются на меньшем множестве значений  $d$ .

Результатом визуального анализа разностей матриц смежности стал выбор в качестве характеристического вектора для 8-битных сигналов тридцатиэлементного вектора, полученного последовательным объединением строк матрицы  $P_f(\Delta d, d)$

размером  $3 \times 10$ , где  $\Delta d = 2^{-m}$ ,  $m = 8, 7, 6$ ,  $d = 1 \dots 10$ . Дальнейшие численные эксперименты показали, что добавление в характеристический вектор элементов матрицы смежности, полученных при  $d > 10$ , несущественно влияет на результирующую точность стеганоанализа, увеличивая при этом время расчетов.

Следующий этап — выбор ядра классификатора и формирование обучающей выборки для SVM. Обычно в стеганоанализе используют гауссовское ядро SVM. Однако в данном случае численные эксперименты на различных наборах тестовых сигналов и при разных параметрах тестирования показали, что более высокую точность обеспечивает линейное ядро классификатора с использованием квадратичного программирования для поиска разделяющей плоскости гиперплоскости.

В работе [6] показано, что обучающая выборка должна формироваться из пар «пустой»–«стеганоконтейнер», где стеганоконтейнер образован из стоящего с ним в паре пустого контейнера путем внедрения в него некоторого сообщения. Для выбора оптимальных параметров классификатора при стеганоанализе нежелательным является использование классической кросс-валидации, поскольку она разрывает эти пары. Соответственно этому в данной работе обучающая выборка формировалась из пар «пустой»–«стеганоконтейнер», а параметры классификатора подбирались по дискретной сетке значений.

Таблица 1

Наполненность стеганоконтейнеров контрольной выборки, %	Точность выявления последовательного внедрения методом НЗБ, %		Точность выявления распределенного внедрения методом НЗБ, %	
	Совпадение наполненности	Общая обучающая выборка	Совпадение наполненности	Общая обучающая выборка
100	100	100	100	100
75	99,9375	100	100	100
50	99,9375	100	100	99,9375
25	99,5625	99,8750	99,8125	99,9375
12	98,0625	95	99,8750	99,6875
6	93,3125	68,1875	99,2500	73,2500

Была проведена серия тестов, определяющих оптимальное количество сигналов в обучающей выборке. В них увеличивалось количество элементов обучающей выборки при неизменном количестве элементов контрольной, для которой обученный классификатор определяет метки класса согласно построенному решающему правилу. Причем на вход обученного классификатора по очереди подавались две контрольные выборки: 1) выборка пустых контейнеров, не использовавшихся при обучении SVM; 2) выборка стеганоконтейнеров фиксированной наполненности, также не использовавшихся при обучении SVM. Результирующая точность подсчитывалась как процент правильно расставленных меток на двух вышеупомянутых контрольных выборках. В целом такие тесты показали, что в ракурсе достигаемой точности стеганоанализа целесообразным является использование не менее 300 элементов в обучающей выборке. С увеличением количества элементов точность увеличивается (с какого-то момента незначительно), но увеличивается и время расчетов. Также исследовалась зависимость точности классификации от наполненности стеганоконтейнеров. Поскольку стеганоаналитик, как правило, не обладает информацией о длине скрытых сообщений, он может выбрать один из двух вариантов анализа:

1) использовать один классификатор, обученный на сигналах разной наполненности;

2) последовательно проверить некоторый контейнер на наборе бинарных классификаторов, обученных каждый на пустых контейнерах и стеганоконтейнерах одинаковой или близкой наполненности, так чтобы в совокупности набор охватывал все возможные варианты наполненности контейнеров.

Второй подход требует больше времени на анализ, но в целом позволяет более точно определять наличие скрытых сообщений. Так, точность стеганоанализа при совпадающей наполненности стеганоконтейнеров обучающей и контрольной выборок приведена в колонках «Совпадение наполненности» табл. 1 (обучающая выборка состоит из 200 пустых и 200 стеганоконтейнеров, контрольная — из 800 пустых и 800 стегано). В колонках «Общая обучающая выборка» этой таблицы приведена точность стеганоанализа при использовании обучающей выборки, состоящей из 200 пустых контейнеров и по 33 контейнера 100, 75, 50, 25, 12 и 6 % наполненности.

Исходя из результатов табл. 1, при проверке 8-битных аудиосигналов целесообразно последовательное использование обоих вышеописанных вариантов стеганоанализа: SVM-классификации с общей обучающей выборкой для выявления стеганоконтейнеров большой наполненности и SVM-классификации с обучением на стеганоконтейнерах малой наполненности для выявления относительно коротких скрытых сообщений.

Стеганографическое внедрение информации в сигнал можно рассматривать как добавление к нему шума. С помощью дискретного вейвлет-преобразования (ДВП) аудиосигнал делится на аппроксимирующую и детализирующую субполо-

сы. При этом шумовая составляющая больше отразится в детализирующих вейвлет-коэффициентах. Тогда, убрав аппроксимирующие коэффициенты из области анализа, в области детализирующих получим более выраженные различия между пустыми и заполненными контейнерами, что, в свою очередь, позволит более точно их различать.

Идея расчета коэффициентов характеристического вектора сигнала в области детализирующих вейвлет-коэффициентов также была реализована с помощью пакета Matlab. В табл. 2 представлена точность стеганоанализа при формировании характеристического вектора в области детализирующих коэффициентов одноуровневого и двухуровневого пакетного вейвлет-разложения. На вход SVM подавался двадцатиэлементный характеристический вектор, рассчитанный с разрешающей способностью  $\Delta d = 2^{-8}$ . Первые десять элементов этого вектора определялись для выделенной одноуровневым ДВП частотной субполосы ( $F_{\max/2}, F_{\max}$ ), т.е. деталей первого уровня. Вторые десять — для частотной полосы ( $F_{3\max/4}, F_{\max}$ ), выделенной соответственно в результате двухуровневого вейвлет-пакетного разложения. В скобках указана точность, полученная при использовании только первой половины элементов характеристического вектора и через запятую — только второй.

Как видим, анализ не всего аудиосигнала, а только детализирующих субполос улучшает точность классификации контейнеров. Наиболее сильно при этом улучшается точность выявления стеганоcontainers малой наполненности.

Самая высокая точность стеганоанализа достигается объединением характеристических векторов, рассчитанных во временной области и для детализирующих вейвлет-коэффициентов. Результат численных экспериментов для этого случая представлен в табл. 3.

Таблица 2

Наполненность стеганоcontainers контрольной выборки, %	Точность выявления последовательного внедрения методом НЗБ, %		Точность выявления распределенного внедрения методом НЗБ, %	
	Совпадение наполненности	Общая обучающая выборка	Совпадение наполненности	Общая обучающая выборка
100	100 (100, 100)	99,8750 (99,8125, 98,6250)	100 (100,100)	99,7500 (99,1250, 99,8125)
75	100 (100, 100)	99,8750 (99,8125, 98,6250)	100 (100,100)	99,7500 (99,1250, 99,7500)
50	100 (100, 100)	99,8750 (99,8125, 98,6250)	100 (100,100)	99,7500 (99,1250, 99,7500)
25	100 (100, 99,9375)	99,8750 (99,8125, 98,6250)	100 (100, 100)	99,7500 (99,1250, 99,8125)
12	100 (99,9375, 98)	99,8125 (98,4375, 97,2500)	100 (99,8750, 100)	99,7500 (99,1250, 99,8125)
6	99,1875 (97,3750, 92,8125)	90,6875 (79,1875, 71,6875)	95,8125 (94,3125, 91,3125)	89,7500 (82,8750, 84,9375)

Таблица 3

Наполненность стеганоcontainers контрольной выборки, %	Точность выявления последовательного внедрения методом НЗБ, %		Точность выявления распределенного внедрения методом НЗБ, %	
	Совпадение наполненности	Общая обучающая выборка	Совпадение наполненности	Общая обучающая выборка
100	100	99,6875	100	99,9375
75	100	99,6875	100	99,9375
50	100	99,6875	100	99,9375
25	100	99,6875	100	99,9375
12	99,9375	99,6250	100	99,9375
6	99	94,0625	99,9375	99,5000

## Стеганоанализ 16-битных аудиосигналов

НЗБ-стеганография может использоваться как для создания 8-битных стегано-аудио-контейнеров, так и для создания 16-битных. При этом шум, вносимый в 8-битный сигнал стегановнедрением, существенно выше шума, вносимого в 16-битный сигнал. Например, среднее значение SNR для одного из тестовых наборов, 100 % заполненных программой Hide4PGP 8-битных аудиоконтейнеров, равно 27 дБ, а для сигналов с таким же содержимым, но 16-битных — 56 дБ. Как следствие, выявить стегановложения в 16-битных аудио, труднее, чем в 8-битных.

Визуальный анализ строк матрицы смежности 16-битных сигналов и предварительная SVM-классификация их характеристических векторов, полученных при различных  $\Delta d$  и  $d$ , показали целесообразность использования 130-элементного характеристического вектора. Его первые 60 элементов рассчитаны при  $\Delta d = 2^{-16}$ , следующие 35 — при  $\Delta d = 2^{-15}$  и последние 35 — при  $\Delta d = 2^{-14}$ . Так же, как и в случае 8-битных контейнеров, эффективным является расчет характеристического вектора в области детализирующих вейвлет-коэффициентов.

В табл. 4 показана точность выявления Hide4PGP вложений в 16-битные 44 кГц речевые аудиосигналы. В данном случае использовалась общая обучающая выборка, состоящая из сигналов наполненности 0, 12, 25, 50, 75 и 100 %. Сигналы получены из разных источников. Количество сигналов обучающей выборки — 1600, контрольной — 6400.

Полученные численные оценки свидетельствуют о целесообразности использования характеристических векторов сигналов, рассчитанных как во временной области, так и для вейвлет-коэффициентов субполос  $(F_{\max/2}, F_{\max})$  и  $(F_{3\max/4}, F_{\max})$ .

Отметим, что, как и в случае с анализом 8-битных аудиосигналов, стеганоаналитик может использовать вариант последовательной проверки сигнала на наборе бинарных классификаторов, обученных каждый на стеганоконтейнерах одинаковой или близкой наполненности. Результат такого стеганоанализа для этого же тестового набора 16-битных сигналов представлен в табл. 5.

Таблица 4

Наполненность стегано-контейнеров контрольной выборки, %	Точность классификации, %			
	Анализ во временной области	Анализ в области детализирующих коэффициентов одноуровневого ДВП	Анализ в области детализирующих коэффициентов двухуровневого вейвлет-пакета	Анализ данных трех областей
100	94,5156	87,1094	85,2656	97,5313
75	93,0781	86,1719	82,2500	97,4375
50	90,0625	73,8281	77,3906	96,2969
25	86,2188	54,0313	75,2969	93,9844
12	52,3750	50,6563	73,8125	74,5469

Таблица 5

Наполненность стегано-контейнеров контрольной выборки, %	Точность классификации, %			
	Анализ во временной области	Анализ в области детализирующих коэффициентов одноуровневого ДВП	Анализ в области детализирующих коэффициентов двухуровневого вейвлет-пакета	Анализ данных трех областей
100	98,2031	98,9375	99,8281	99,8281
75	96,8125	96,2344	95,9531	99,7656
50	94,6250	77,8750	85,3125	98,8750
25	89,4531	60,0156	74,4531	96,2188
12	71,7969	57,0781	74,0625	83,0625

Чем больше длина внедренного сообщения, тем больше стабильных различий наблюдается в матрицах смежности пустого и заполненного контейнеров. Также будут наблюдаться различия в матрицах смежности двух стеганоконтейнеров. Например, матрица смежности стеганоконтейнера, заполненного на 5 %, будет более близка к матрице смежности пустого контейнера, чем стеганоконтейнера, заполненного на 100 %. Это наблюдение открывает путь к определению длины скрытого сообщения.

Для определения длины скрытого сообщения была построена совокупность бинарных SVM, обученных классифицировать два набора стеганоконтейнеров разной наполненности. Использовались те же 16-битные сигналы, что и в предыдущих экспериментах. Полученная точность классификации представлена в табл. 6. Анализ выполнялся во временной области и в частотных субполосах ( $F_{\max/2}$ ,  $F_{\max}$ ) и ( $F_{3\max/4}$ ,  $F_{\max}$ ). В случаях, когда наполненность контейнеров обучающей выборки с метками класса «S1» и «S2» не совпадала, обучающая выборка состояла из 800 пар контейнеров двух фиксированных наполненностей, полученных из одного и того же пустого. В остальных случаях для обучения брались 1600 разных стеганоконтейнеров одинаковой наполненности (как видим, в этих случаях точность SVM равна 50 %, т.е. сигналы нельзя поделить на два класса).

Результаты, приведенные в табл. 6, свидетельствуют о том, что два набора контейнеров фиксированной наполненности можно достаточно точно разделить. В этой серии экспериментов, а также в подобных для других наборов контейнеров, в том числе 8-битных, прослеживается такая закономерность: при сближении степеней наполненности контейнеров точность их классификации ухудшается и одновременно с этим точность всегда лучше для контейнеров большей наполненности.

Отметим, что кроме подхода, состоящего в построении совокупности бинарных классификаторов, обученных по дискретной сетке наполненностей контейнеров, и дальнейшем определении длины скрытого сообщения путем взвешенного голосования этих классификаторов, также был реализован подход, использующий многоклассовую SVM. Для тестирования многоклассовой SVM использовалась отдельная Matlab-функция. Но с ее помощью были получены результаты хуже, чем представленные в табл. 6.

Подобные эксперименты были выполнены не только для выявления стегановкладок Hide4PGP, но и других вариантов НЗБ-стегановнедрения. В частности, в табл. 7 приведена точность обнаружения последовательного и распределенного внедрения сообщений, в процессе которых в одном отсчете сигнала скрывалось не более одного бита (классический НЗБ). Использовался тот же, что и ранее, 16-битный исходный набор фрагментов аудиокниг, но максимальная наполненность стеганоконтейнеров в этих тестах в четыре раза меньше, чем в предыдущих. Выявить стеганоконтейнеры в этом случае труднее, так как модификаций, внесенных в оригиналы внедрением, меньше.

Таблица 6

Наполненность контейнеров обучающей выборки с меткой класса «S2», %	Наполненность контейнеров обучающей выборки с меткой класса «S1»					
	100 %	75 %	50 %	25 %	12 %	0 %
100	50	99,5625	99,5781	99,6094	99,5781	99,8281
75	99,5625	50	98,5781	98,8594	98,9375	99,7656
50	99,5781	98,5781	50	96,6875	96,7344	98,8750
25	99,6094	98,8594	96,6875	50	87,4688	96,2188
12	99,5781	98,9375	96,7344	87,4688	50	83,0625
0	99,8281	99,7656	98,8750	96,2188	83,0625	50

Таблица 7

Наполненность стеганоконтейнеров контрольной выборки, %	Точность классификации, %			
	Анализ во временной области	Анализ в области детализирующих коэффициентов одноуровневого ДВП	Анализ в области детализирующих коэффициентов двухуровневого вейвлет-пакета	Анализ данных трех областей
Выявление последовательного внедрения методом НЗБ				
100	88,5000	60	59,0625	88,7031
75	85,6250	59,2969	57,9688	86,2188
50	79,9844	56,6875	54,9844	79,9063
25	68,4375	53,4375	52,2656	69,7344
12	58,7031	51,3594	51,4063	60,4063
Выявление распределенного внедрения методом НЗБ				
50	77,9063	55,4219	55,9375	77,8906
25	59,7656	53,8125	54,0781	63,0156

Таблица 8

Наполненность стеганоконтейнеров контрольной выборки, %	Точность классификации, %			
	Анализ во временной области	Анализ в области детализирующих коэффициентов одноуровневого ДВП	Анализ в области детализирующих коэффициентов двухуровневого вейвлет-пакета	Анализ данных трех областей
100	95,8167 (94,3559)	99,8008 (69,7211)	98,8048 (65,4714)	99,8008 (95,5511)
50	93,2935 (82,0053)	99,3360 (60,6242)	94,1567 (57,0385)	99,3360 (84,7278)
25	71,4475 (61,9522)	96,9456 (53,7849)	86,8526 (53,5857)	97,6096 (63,7450)
12	57,9681 (55,4449)	84,9934 (51,3944)	71,3147 (51,5936)	90,9695 (55,7769)

Шум, вносимый стеганообразованием, будет более заметен на участках сигнала с малыми амплитудами. Поэтому в случае, когда НЗБ-внедрение происходило в сигналы со срезанными верхними частотами (например, вследствие mp3-сжатия), расчет элементов характеристического вектора в области вейвлет-коэффициентов становится более эффективным. Сжатие с потерями удаляет незначимые детали в спектре сигнала, что позволяет более точно классифицировать пустые и заполненные контейнеры в целом. Такой пример приведен в табл. 8. В ней представлены результаты экспериментов, где исходные широкополосные речевые аудиосигналы (тестовый набор 1250 фрагментов аудиокниг) подвергались mp3-сжатию с битрейтом 64 кбит/с. Распределенное НЗБ-внедрение осуществлялось во временной области этих сигналов после их обратного конвертирования в неупакованный формат. В скобках после каждого значения точности классификации контейнеров указана точность, полученная при тех же параметрах численных экспериментов, но для сигналов, которые не подвергались mp3-сжатию. Как видим, эта точность хуже, особенно для той части характеристического вектора, которая рассчитывается в области вейвлет-коэффициентов.

Аналогично исследовалось влияние шума на точность аудиостеганоанализа. Перед применением сокрытия согласно методу НЗБ к оригинальным тестовым сигналам добавлялся аддитивный гауссовский шум фиксированной мощности. Результаты тестов показали, что с увеличением мощности добавляемого шума точность выявления образованных после НЗБ-внедрения зашумленных стегано-контейнеров ухудшается.

**Заключение.** В целом полученные численные оценки свидетельствуют о том, что матрица смежности речевых аудиосигналов чувствительна к изменениям, вносимым различными вариантами метода НЗБ, и изменяется под их действием схожим образом, вследствие чего она может использоваться для аудиостеганоанализа. При этом следует учитывать, что обучение SVM-классификатора должно выполняться на сигналах с характеристиками, максимально сходными с характеристиками проверяемых.



Разработанный метод более точен при выявлении стеганоконтейнеров относительно большой наполненности. Дополнение характеристического вектора элементами, рассчитанными по формуле (1), но не во временной области, а в детализирующих вейвлет-субполосах, повышает точность стеганоанализа. Два набора стеганоконтейнеров фиксированной наполненности являются разделимыми с помощью SVM, что открывает путь к определению вероятной длины скрытых сообщений.

*H.V. Koshkina*

### ДОСЛІДЖЕННЯ ПРИДАТНОСТІ МАТРИЦІ СУМІЖНОСТІ ДЛЯ ВИЯВЛЕННЯ СТЕГАНОВАУДІОКОНТЕЙНЕРІВ

Досліджено чутливість рядків матриць суміжності аудіосигналів, обчислених в часовій області та області вейвлет-коефіцієнтів, до послідовного та розподіленого стеганографічних вкраплень за методом НЗБ. Виявлено залежності між наповненістю аудіоконтейнерів та точністю їх SVM-класифікації за особливостями рядків матриць суміжності. Запропоновано спосіб визначення довжини прихованого повідомлення шляхом зваженого голосування сукупності бінарних класифікаторів, що розрізняють стеганоконтейнери двох різних наповненостей.

*N.V. Koshkina*

### INVESTIGATION OF THE APPLICABILITY OF THE CO-OCCURRENCE MATRIX FOR DETECTING STEGANOAUDIO SIGNALS

The sensitivity of rows of the co-occurrence matrix of audio signals that were computed in the time and the wavelet coefficients domains to distributed and sequential steganographic embedding by LSB method has been investigated. The relations between the steganographic capacity and the accuracy of SVM classification based on rows features of co-occurrence matrices stego- and normal audios were found. The method for determination of the length of hidden messages by weighted voting of binary classifiers set, that distinguish between two stego audios with different steganographic capacity has been proposed.

1. *Qi Y., Wang Y., Yuan J.* Audio steganalysis based on co-occurrence matrix and PCA // IEEE Computer Society (Proc. of the Intern. Conf. on Measuring Technology and Mechatronics Automation 2009). — 2009. — **1**. — P. 433–436.
2. *Koshkina H.V.* Стеганоаналіз МІК-стеганографії на базі матриці суміжності та методу опорних векторів // Искусственный интеллект. — 2012. — № 4. — С. 567–577.
3. *Abolghasemi M., Aghainia H., Faez K., Mehrabi M.A.* LSB data hiding detection based on gray level co-occurrence matrix // International Symposium on Telecommunications. — 2008. — P. 656–659.
4. *Kekre H.B., Athawale A.A., Patki S.A.* Steganalysis of LSB embedded images using gray level co-occurrence matrix // International Journal of Image Processing. — 2011. — **5**, N 1. — P. 36–45.
5. *Fazli S., Zolfaghari-Nejad M.* A new steganalysis method for steganographic images on DWT domain // International Journal of Science and Engineering Investigations. — 2012. — **1**, N 2. — P. 1–4.
6. *Schwamberger V. Franz M.O.* Simple algorithmic modifications for improving blind steganalysis performance // Proceedings of the 12th ACM Multimedia & Security Workshop MMSec. — Rome. — 2010. — P. 225–230.

*Получено 01.08.2013*

Статья представлена к публикации чл.- корр. НАН Украины В.К. Задиракой.