

Лабораторная работа 4: Реализация и исследование методов НЗБ-стегоанализа изображений

Задания

В лабораторной работе необходимо выполнить стегоанализ методов НЗБ-встраивания и ± 1 -встраивания различными методами, в зависимости от показателя заполненности контейнера q (3.7). Параметры встраивания, а также способы расчёта признаков определяются вариантом задания.

Входными данными, необходимыми для выполнения лабораторной работы, являются K полутоновых изображений одного размера. В качестве входных данных предлагается использовать коллекцию BOWS-2. Краткий вариант (1000 изображений формата TIFF) можно скачать по [ссылке](#), полный вариант (10000 изображений формата PGM) – по [ссылке](#).

Порядок выполнения лабораторной работы:

1. Реализовать процедуру расчёта вектора признаков, используемых для стегоанализа (вектор признаков должен быть выбран студентами самостоятельно исходя из заданного метода стегоанализа).
2. Выполнить имитацию работы стеганографической системы для первых $K/2$ изображений: заполнить долю q битовой плоскости p каждого изображения разными реализациями равномерного белого шума. Вторую половину изображений не менять.
3. Произвести обучение классификатора по выборке, содержащей первые 70 % изображений каждого из двух типов (со встраиванием и без встраивания). То есть общий объём обучающей выборки составляет $K \cdot 0,7$.
4. Применить обученный классификатор на оставшихся 30 % изображений и оценить качество классификации.
5. Повторить пп. 2-4 для других долей q . Построить график.

В п. 2 для нечётных вариантов позиции для встраивания ЦВЗ следует выбирать последовательно, для чётных – псевдослучайно.

Для всех вариантов необходимо собрать статистику для ряда значений q (например, от 0.1 до 1 с шагом 0.1 или от 0.2 до 1 с шагом 0.2).

Для оценки качества классификации в лабораторной работе предлагается использовать один из двух показателей (Accuracy или F1) по своему усмотрению. Ниже приводятся формулы их расчёта, основанные на данных таблицы 1.

Мера Accuracy рассчитывается как

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

а F-мера (или F_1 мера) – по формуле

$$F_1 = \frac{2}{\frac{1}{Pr} + \frac{1}{R}}, \quad (2)$$

где

$$Pr = \frac{TP}{TP + FP}, \quad (3)$$

$$R = \frac{TP}{TP + FN}. \quad (4)$$

Величины (3), (4) называются соответственно Precision и Recall (на русском – точность и полнота).

Табл. 1 – Виды ошибок классификации

	Правильная классификация	Неправильная классификация
Информация встроена	TP	FP
Информация не встроена	TN	FN

В работе предлагается использовать один из трёх методов стегоанализа: метод частоты переходов, метод длин серий и метод пар значений. Важно учесть следующие нюансы:

- Все методы, включая метод пар значений, должны быть реализованы по схеме «расчёт признаков» + «управляемая классификация». Даже метод пар значений.
- Для методов «частота переходов» и «длины серий» желательно использовать серпантинную развёртку плоскости изображения.
- Конкретный вектор признаков нужно сформировать самостоятельно. Нет одного «правильного» варианта: нужно руководствоваться здравым смыслом.

- Если встраивание осуществляется методом ± 1 -встраивания, то метод длин серий стоит адаптировать (считать серии по двум битовым плоскостям сразу).
- Вместо перечисленных трёх методов стегоанализа можно выбрать альтернативный метод (что будет оцениваться повышенным баллом) – один из 4-х методов так называемых структурных методов стегоанализа (*aump*, *sp*, *triples*, *ws*). Реализации этих методов можно найти по [ссылке](#) и переписать на Python. Разумеется, в сути реализуемого метода тоже надо разобраться и обстоятельно рассказать преподавателю о выбранном методе и принципах его работы.

Для вариантов 5-24 предлагается также дополнительное исследование. Оно не является обязательным, но за него начисляется дополнительный балл. Предусмотрены 5 различных вариантов исследований:

1. Выполнить сравнительное исследование не менее 5 разных моделей классификаторов для решения поставленной задачи стегоанализа. Гиперпараметры выбирать логичным образом.
2. Тем или иным способом выполнить сравнительное исследование значимости различных признаков (компонент выбранного вектора признаков) применительно к произвольному фиксированному классификатору.
3. Выбрать одну базовую модель классификатора на свой вкус и провести ROC-анализ в зависимости от её гиперпараметров.
4. Сравнить не менее двух различных вариантов формирования вектора признаков, оставаясь в рамках заданного вариантом базового метода стегоанализа.
5. Выполнить сравнительное исследование влияния развёртки двумерной области на качество решения задачи стегоанализа. Для этого сравнить 4 вида развёрток: построчную, серпантинную, Гильберта-Пеано, а также зигзагообразную. Первые три рассмотрены в параграфе 9.2, а зигзагообразная – в параграфе 6.3 при описании СВИ-14 (Cox et al.).

Варианты заданий

Основные параметры исследований отражены в таблице ниже. В столбце «р» указан номер битовой плоскости.

№ (var)	p	Метод встраивания	Метод стегоанализа	Дополнительное исследование
1	1	НЗБ	Частота переходов	
2	1	НЗБ	Частота переходов	
3	1	± 1	Частота переходов	
4	1	± 1	Частота переходов	
5	1	НЗБ	Длины серий	1
6	1	НЗБ	Длины серий	1
7	1	НЗБ	Пары значений	1
8	1	НЗБ	Пары значений	1
9	3	НЗБ	Пары значений	2
10	2	НЗБ	Пары значений	2
11	2	НЗБ	Пары значений	3
12	2	НЗБ	Длины серий	2
13	2	НЗБ	Длины серий	2
14	3	НЗБ	Длины серий	3
15	1	± 1	Длины серий	3
16	1	± 1	Длины серий	5
17	1	± 1	Пары значений	4
18	1	± 1	Пары значений	4
19	3	± 1	Пары значений	5
20	2	± 1	Пары значений	5
21	2	± 1	Пары значений	3
22	2	± 1	Длины серий	4
23	2	± 1	Длины серий	4
24	3	± 1	Длины серий	5

Контрольные вопросы

Для данной лабораторной работы ответы на контрольные вопросы не являются отдельной обязательной для сдачи процедурой. Список вопросов приводится для того, чтобы студенты ориентировались в теоретическом материале. В случае сомнений в понимании сути выполненной работы, преподаватель может задать какой-либо вопрос из представленного списка при сдаче практической части лабораторной работы.

1. Задача стегоанализа и различные подходы к её решению.
2. В чём состоит СВИ-2 (Стеганографическое НЗБ-встраивание)?
3. В чём состоит СВИ-3 (± 1 -встраивание)?
4. Опишите возможные способы заполнения контейнера в СВИ-2 (Стеганографическое НЗБ-встраивание). Что такое заполненность контейнера?
5. Опишите общий принцип стегоанализа НЗБ-систем и перечислите рассмотренные методы атак на подобные системы.
6. Как осуществляется расчёт числа переходов, почему оно позволяет выявить наличие встроенной информации?
7. Как осуществляется расчёт числа серий, почему оно позволяет выявить наличие встроенной информации? Проиллюстрируйте ответ графиками.
8. В чём состоит метод пар значений в общих чертах?
9. Приведите несколько примеров векторов признаков на основе числа серий.
10. Опишите показатели качества решения задачи стегоанализа, используемые в данной лабораторной работе.
11. Перечислите развёртки двумерных областей, используемые в данной лабораторной работе. Как вы думаете, какие из них лучше других и почему?