В. И. Коржик, К. А. Небаева

ОСНОВЫ СТЕГАНОГРАФИИ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

СПб ГУТ)))

САНКТ-ПЕТЕРБУРГ
2014

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
ОБРАЗОВАТЕЛЬНОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ
им. проф. М.А. БОНЧ-БРУЕВИЧА»

В. И. Коржик
К.А. Небаева

ОСНОВЫ СТЕГАНОГРАФИИ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ПРАКТИЧЕСКИМ ЗАНЯТИЯМ

СПб ГУТ)))

САНКТ-ПЕТЕРБУРГ
2014

УДК 004.9
ББК 3973я73????
К 66??????????????????

Рецензент
Р.Р. Биккенин доктор технических наук, профессор

Рекомендовано к печати
редакционно-издательским советом университета

Коржик, В.И.
К66 Основы стеганографии: методические указания к практическим занятиям. / *В.И. Коржик, К.А. Небаева.* – СПб. : Издательство «Теледом» ГОУВПО СПбГУТ, 2014. – XX с.

Предназначены для подготовки и проведения практических занятий для специальностей 210403 «Защищенные системы связи», 210700 «Инфокоммуникационные технологии и системы связи», 090900 «Информационная безопасность» при изучении дисциплин «Основы стеганографии», «Технологии стеганографии», «Технологии стеганографии в системах инфокоммуникаций».

В процессе выполнения практических занятий студенты закрепляют полученные на лекциях знания по курсу «Основы стеганографии» и подготавливаются к проведению цикла лабораторных работ.

В данном пособии охвачены все основные разделы лекционного материала. Текст пособия полностью согласован с электронной версией курса «Основы стеганографии» http://ibts.sut.ru/materialy/

Оглавление

Введение

Методическое пособие предназначено для помощи преподавателям, проводящим практические занятия по курсу «Основы стеганографии», а также студентам, их выполняющих. Предполагается, что студенты, перед тем как приступить к практическим занятиям, прослушали лекции по темам, соответствующим каждому занятию.

Основная цель практических занятий состоит в закреплении лекционного материала и подготовке к выполнению лабораторных работ. Каждое практическое занятие рассчитано на два академических часа. В процессе проведения занятий, преподаватель фиксирует индивидуальную активность студентов на этом занятии, которое будет учитываться при выставлении итоговой оценки после экзамена (или зачета). Задачи, отмеченные «*» являются заданиями повышенной сложности.

При подготовки к занятию студенты могут пользоваться материалами электронного курса лекций «Основы стеганографии».

Практическое занятие 1
Стегосистема с вложением информации в наименьшие значащие биты (СГ-НЗБ)

Вопросы для проверки знаний

1. Какую задачу решают стегосистемы (СГ)?
2. В какие покрывающие объекты (ПО) может вкладываться дополнительная информация?
3. Какими параметрами можно оценить эффективность заданий СГ?
4. Чем стеганография отличается от криптографии?
5. Как формулируется предположение Кирхгоффа для стеганографии?
6. Как реализуется вложение и извлечение информации для СГ-НЗБ (НЗБ-замена, ± 1 НЗБ)?
7. Преимущества и недостатки СГ-НЗБ?
8. Визуальный метод обнаружения СГ-НЗБ.
9. Обнаружение СГ-НЗБ по критерию χ^2.
10. Обнаружение СГ-НЗБ по методу парно-выборочного стегоанализа.

Задачи

1. Предположим, что для 8-битового в каждом пикселе цифрового изображения значение яркости некоторых пикселей будут равны: 1, 7, 112, 253, 255.
   Какие значения яркостей этих пикселей получатся после погружения в эти пиксели двоичной информации 10110 для методов НЗБ-замены и ± 1 НЗБ?
2. Сколько (в среднем) бит информации можно погрузить по методу СГ-НЗБ в цифровое изображение размером 200*300 пикселей при вероятностях погружения P в каждый пиксель: 1; 0.5; 0.1; 0.01 ?
3. Предположим, что на части цифрового изображения имеется прямой вертикальный контур, для которого значение яркости слева равна 16, а справа 153. Каковы будут значения яркости на этом контуре после его преобразования к двоичному виду с вложением, соответствующем НЗБ случайной (зашифрованной) двоичной последовательностью без вложения? В каком случае сохраняется контур?
4. Предположим, что часть гистограммы цифрового изображения имеет следующий вид:

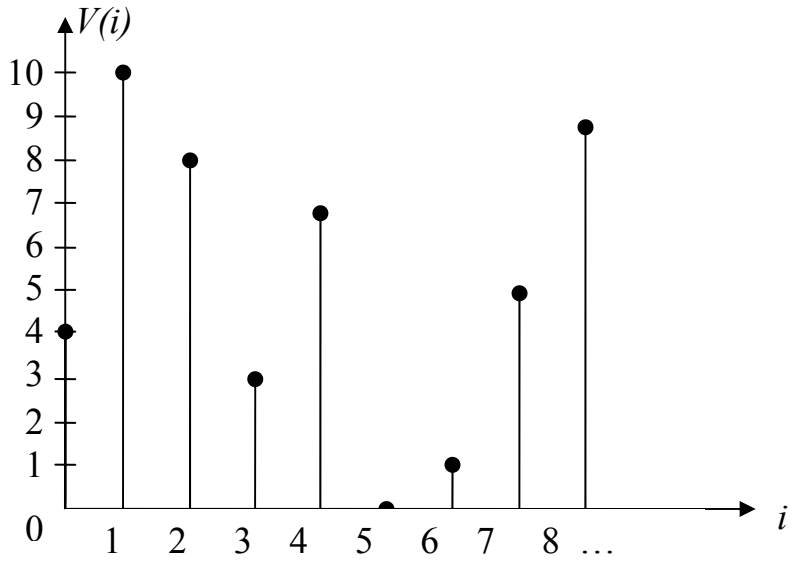

Рисунок 1 – Гистограмма цифрового изображения

Какой вид будет иметь гистограмма после вложения по методу СГ-НЗБ (замещения) равновероятной и взаимонезависимой последовательности информационных бит?

Практическое занятие 2
Матричные методы погружения

Вопросы для проверки знаний

1. Как влияет количество измененных НЗБ на обнаруживаемость СГ-НЗБ?
2. Как формально записать погружение в НЗБ в терминах n – общее количество НЗБ, k – количество погружаемых бит, ρ – количество изменений НЗБ после погружения информации.
3. Что такое покрывающая функция и как она может быть использована при вложении в НЗБ.
4. Какой величиной оценивается эффективность погружения для СГ-НЗБ.
5. * Как формулируется основная теорема матричного погружения с использованием двоичных систематических (n,k) – кодов?
6. Процедуры матричного погружения и извлечения информации для СГ-НЗБ.
7. * Граница для средней эффективности погружения.
8. Верхняя граница для максимального количества погружаемых бит для кодов с известным радиусом покрытия.
9. Как определяется класс двоичных кодов Хэмминга.
10. Какой вид имеет проверочная матрица для двоичного кода Хэмминга (7,4)?

11. Описать процедуры матричного вложения по извлечению информации при использовании для этой цели двоичного кода Хэмминга (7,4)?
12. Как выполняется погружение информации в стегосистеме F5?

Задачи

1. Построить таблицу относительной нагрузки $\alpha_p = p/(2^p - 1)$, $p = 1,2,\dots,10$, и эффективности погружения $l_p = p/(1 - 2^{-p})$ для класса двоичных кодов Хэмминга с параметрами $(2^p - 1, 2^p - 1 - p)$
   Пояснить наглядный смысл полученной зависимости.
2. Предположим, что для матричного погружения двоичный код Хэмминга (7,4). Пусть отсчеты яркости семи последовательных пикселей ???????? изображения имеют вид: $g = (11,10,15,17,13,21)$. Найти значения яркостей пикселей после вложения в эти 7 пикселей трех бит информации m=(0,1,1). Пояснить смысл полученного результата.

Практическое занятие 3
Стегосистема, использующая широкополосные сигналы, формируемые по секретному стегоключу (СГ-ШПС)

Вопросы для проверки знаний

1. Почему СГ-НЗБ не устойчива к атаке удаления вложений информации даже при невозможности обнаружения вложения?
2. Как выполняется вложение информации в СГ-ШПС?
3. Что такое информированный и слепой декодер?
4. Как выполняется извлечение информации информированным и слепым декодером?
5. Как зависит вероятности ошибки при извлечения информации в случае информированного и слепого декодера от параметров СГ-ШПС и атаки?
6. Каким образом осуществляется обнаружение СГ-ШПС?

Задачи

1. Рассчитать вероятность ошибки извлечения биты информации информированным и слепым декодером, если СГ-ШПС имеет следующие параметры: дисперсия (ПО-изображение) $\sigma_c^2 = 350$, глубина погружения $\alpha = 5$, дисперсия шума при атаке $\sigma_\varepsilon^2 = 25$,

количество пикселей, в которые погружается один бит информации N=5. Во сколько раз нужно увеличить количество пикселей N, в которые погружается 1 бит информации, чтобы для слепого декодера получить при извлечении такую же вероятность ошибки, как и для информированного декодера?

Указание. При расчетах можно использовать следующую верхнюю границу для функции $Q(x)$:

$$Q(x) \le e^{-\frac{x^2}{2}}$$

2. Предположим, что для обнаружение СГ-ШПС используется статистика [1]:

$$\Gamma = \frac{1}{2N\sigma_c^2}\sum_{n=1}^{N}(C(n+1) - C(n))^2,$$

где N – общее количество пикселей изображения,
$\sigma_c^2 = Var\{C(n)\}$,
Причем при вложении информации используется модифицированный метод СГ-ШПС:
$C_w(n) = \beta C(n) + \alpha(-1)^b \pi(n), n = 1,2,\dots,N$
где $\beta = \sqrt{1 - \frac{\alpha^2}{\sigma_c^2}}$.

Требуется рассчитать среднее значение этой статистики при отсутствии вложения информации [1]:

$$E\{\Gamma\} = 1 - R_c(n, n+1)$$

и при наличии вложения

$$E\{\Gamma\} = 1 - \beta^2 R(n, n+1)$$

где $R(n, n+1)$ – коэффициент корреляции между смежными пикселями покрывающего изображения.

Расчет производить при выборе следующих параметров:
$R(n, n+1) = 0{,}999; 0{,}99; 0{,}9; 0{,}5$, $\alpha = 5, \sigma_c^2 = 2500$.

Сделать вывод о возможности (или нет) обнаружения СГ-ШПС по данной статистике.

Практическое занятие 4.
Лингвистические стегосистемы (СГ-Л)

Вопросы для проверки знаний

1. Что такое лингвистические стегосистемы?
2. Каков принцип вложения и извлечения информации в СГ-Л на основе использования базы синонимов?
3. Каков принцип вложения и извлечения информации в СГ-Л на основе редактирования исходного текста?
4. Являются ли СГ-Л необнаруживаемыми?
5. Какова скорость вложения информации в СГ-Л?

6. Могут ли СГ-Л противостоять атакам удаления вложенной информации?

Задачи

Используя метод абсолютных синонимов произвести вложение цепочки бит «0110» в следующий текст: «На территории США опасным стихийным бедствием являются торнадо». Какова скорость вложения информации в данном примере?

Практическое занятие 5
Идеальные и почти идеальные стегосистемы

Вопросы для проверки знаний

1. Что такое идеальные и почти идеальные СГ?
2. Какой метод погружения обеспечивает получение идеальной СГ, если отсчеты ПО являются одинаково распределенным и взаимонезависимым гауссовскими величинами.
3. В чем состоит принцип погружения информации для модельно обусловленных СГ, использующих идеальное сжатие ПО?
4. В чем состоит принцип погружения информации для СГ с адаптивным квантованием (СГ-АК)?
5. В чем состоит принцип погружения информации для СГ с сохранением статистики (СГ-Р) и почему он не является практически реализуемым в полном объеме?
6. * Чем отличается метод погружения информации с использованием решетчатых кодов (проект HUGO) с весовым коэффициентом, полученным по Марковской модели (SPAM) от матричного погружения на основе линейных блоковых кодов?

Задачи

Для модели SG-R рассчитать двоичную последовательность после погружения информационной цепочки «1001» в двоичную последовательность (ПО) abaabaabbbabbaaaabbbb.

Найти другу двоичную последовательность ПО, которая при вложении той же информационной цепочки «1001» приведет к прежней двоичной стегоцепочке.

Практическое занятие 6
Стегосистема для каналов с шумом (СГ-Ш)

Вопросы для проверки знаний

1. Что означает понятие СГ в каналах с шумом?
2. В чем состоит задача обнаружения в каналах с шумом?
3. Может ли быть известен стегоаналитику покрывающий объект в случае сценария канала с шумом?
4. * Почему относительную энтропию и расстояние Бхаттачариа удобно использовать для оценки необнаруживаемости СГ в каналах с шумом?
5. Описать две основные модели каналов с шумом, использующих для оценки эффективности СГ-Ш.
6. Почему целесообразно использовать СГ с распределенным вложением (с вложением, рассредоточенным во времени)
7. Скорость вложения информации для СГ-Ш.

Задачи

1. Пусть максимальное количество допустимых для вложения бит m для обеспечения необнаруживаемости СГ, соответствующей относительной энтропии D=0,1 и вероятности ошибки извлечения бита $P_e = 10^{-3}$ при количестве отсчетов $N = 10^7$ и вероятности ошибки юита в ДСК $P_0 = 0{,}01$

Указание. Можно воспользоваться следующими соотношениями [7]:

$$P_e \le \left[\left(2\sqrt{P_0(1-P_0)} - 1\right)P_w + 1\right]^{N/m}$$

где P_w – вероятность единицы в псевдослучайной вкладываемой последовательности.

2. Найти максимальное количество допустимых для вложения бит m для обеспечения необнаруживаемости СГ, соответствующей относительной энтропии D=0,1 при количестве отсчетов $N = 10^7$ и требованием к вероятности ошибки при извлечении бита $P_e \le 10^{-4}$.

Указание. Можно воспользоваться следующим соотношением [1]:

$$P_e \le Q\left(1{,}29\sqrt{\frac{(ND)^{1/2}}{m}}\right)$$

Какое отношение мощностей ПО и вложение требуется в этом случае? Допустимо ли оно при цифровой реализации процедуры вложения?

Практическое занятие 7.
Слепой стегоанализ для обнаружения СГ.

Вопросы для проверки знаний

1. Что такое слепой стегоанализ (СГА)и чем он отличается от целевого СГА?
2. Надо ли при слепом СГА знать алгоритм вложения информации в анализируемую СГ?
3. Что такое «калибровка» СГ?
4. Каков критерий при построении гиперкривой для разделения ПО и СГ по методу «оперных векторов»(SVM)?
5. Можно ли с помощью SVM различить не только ПО и СГ, но и алгоритм погружения?
6. Как следует выбирать векторные функционалы для SVM?
7. Определить типы SVM: линейные сепарабельные, линейные несепарабельные, нелинейные.
8. Что такое функционалы SPAM?
9. Что такое ROC-кривая и как она оценивает эффективность СГА?

Задачи

1. Пусть точки, показанные на рис. 2 соответствуют ПО (белые точки) и СГ (черные точки). Провести линии, соответствующие оптимальному разделению этих двух классов для линейной сепарабельной, несепарабельной и нелинейной SVM.

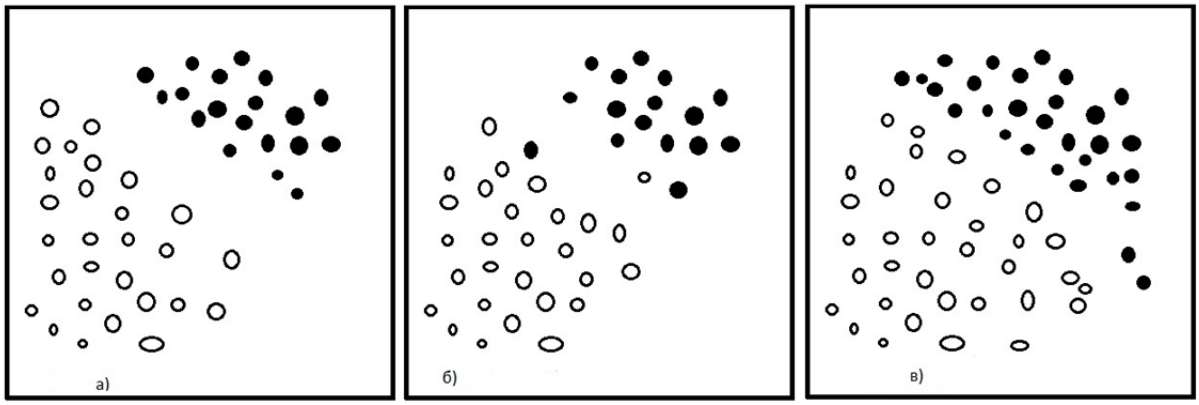

Рисунок 2 – Иллюстрация типов SVM для двумерного пространства признаков. а) сепарабельная, б) несепарабельная, в) нелинейная SVM

Указание. Начертить прямые (кривые) линии, которые максимизируют «зазор» (margin) между разделяемыми _____ или минимизируют вероятность ошибки.

2*. Доказать, что если существует гиперплоскость разделяющая выборки ПО и СГ, то будут справедливы следующие неравенства:

$\mathbf{w}^T\mathbf{x}_i + b \geq 1$, если $y_i = 1$ ($\mathbf{x}_i -$ соответствует СГ),

$\mathbf{w}^T\mathbf{x}_i + b \leq -1$, если $y_i = -1$ ($\mathbf{x}_i -$ соответствует ПО),

причем «зазор» (margin) между областями равен $2/\|\mathbf{w}\|$, причем его максимизация эквивалентна минимизации следующего Лагранжиана относительно вектора **w** и параметра *b*.

Практическое занятие 8
Погружение ЦВЗ с помощью информированного кодера

Вопросы для проверки знаний

1. Что такое информированный кодер и в каких случаях целесообразно его использовать.
2. Чем отличается модель обычной телекоммуникационной системы для каналов с помехами от модели системы ЦВЗ при наличии атаки по удалению ЦВЗ?
3. В чём состоит идея использования кодовой книги в качестве реализации информированного кодера?
4. Описать метод вложения и извлечения ЦВЗ на основе концепции улучшения сигнала (УШПС).
5. В каком случае использование УШПС и слепого декодера дает почти такой же результат, как и использование обычного ШПС и информированного декодера?
6. Описать метод вложения и извлечения ЦВЗ на основе концепции квантованной проективной модуляции (КПМ).
7. * В каком случае метод КПМ оказывается более предпочтительным, чем метод УШПС?

Задачи

1. Рассчитать вероятность ошибки P_e при извлечении бита информации из ЦВЗ, если используется метод УШПС при выборе следующих параметров: $\sigma_c = 60, \alpha = 3, \sigma_\varepsilon = 5, N = 5000$. Сравнить с вероятностью ошибки P_e при использовании обычного ШПС и слепого декодера.
   Указание. Можно воспользоваться соотношением [1]:
   $$P_e = Q\left(\sqrt{\frac{N - \eta_w}{\eta - 1}}\right)$$
   где $\eta = \frac{\eta_w}{\eta_a}$; $\eta_w = \frac{\sigma_c^2}{\alpha^2}$; $\eta_a = \frac{\sigma_c^2}{\alpha^2 + \sigma_\varepsilon^2}$.
2. Рассчитать вероятность ошибки P_e при извлечении бита информации из ЦВЗ, если используется метод КПМ при выборе следующих параметров: $\sigma_c = 50, \alpha = 5, \sigma_\varepsilon = 4, N = 3000$.
   Сравнить с вероятностью ошибки P_e при использовании обычного ШПС и слепого декодера.
3. *Если N задано, а также заданы σ_c и σ_ε, то при каком выборе параметра α метод КПМ оказывается лучше метода УШПС (по вероятности ошибки P_e, если задано ограничение на $\eta_w = \frac{\sigma_c^2}{\alpha^2} \geq \eta_0$.

Практическое занятие 9.
Оценка эффективности «изощренных» атак

Вопросы для проверки знаний

1. Что такое «изощренная» атака на ЦВЗ и возможны ли подобные атаки при создании преднамеренных помех в обычных системах связи?
2. Является ли эффективной атака фильтрации для удаления ЦВЗ при сохранении высокого качества ПО?
3. Что такое «_____ЦВЗ» и зачем они могут применяться?
4. при каких условиях атака по удалению ЦВЗ с использованием оценки элементов ПСП, образующих ЦВЗ может оказаться эффективнее атаки аддитивным шумом?
5. Как можно оценить элементы ПСП, из которых составляется ЦВЗ?

Задачи

1. Рассчитать вероятность ошибки извлечения биты информации P для декодера, при выборе следующих параметров: $\alpha = 5$, $\sigma_c = 5$, $N_0 = 1000, P_{es} = 0{,}1$ $\alpha' = \alpha$.
2. * Построить зависимость величины P от параметра α' в условиях задачи 1 и найти оптимальные значения α' (возможно неравное α). Объяснить смысл полученного результата.

Практическое занятие 10.
Устойчивость систем ЦВЗ-ШПС к коалиционным атакам.

Вопросы для проверки знаний

1. Какие два основных вида вложения ЦВЗ возможны в условиях коалиционных атак и какие цели при этом преследует собственник продукта?
2. Как выполняются атаки при двух основных видах вложения ЦВЗ?
3. Оценка эффективности обеспечения прав собственности и отслеживание участников коалиции при вложении ЦВЗ.
4. Как зависит требуемая длина ЦВЗ, обеспечивающего надежное обнаружение участников коалиции от ее размера?
5. * Что такое сферический алгоритм декодирования и как он может быть использован для обнаружения участников коалиции при вложении ЦВЗ?

Задачи

1. Пусть собственник продукта легально рассылает его нескольким пользователям с различными вложениями ЦВЗ, запрещая, однако, дальнейшее нелегальное распространения этого продукта. Предположим, что k пользователей (по существу – пираты) объединились в коалицию с целью выработки преобразования от своих копий некоторого продукта, подлежащего нелегальному распространению. Какова будет вероятность P_e при размере коалиции k=30?

Указание. Для расчета P_e можно использовать следующую формулу:

$$P_e = Q\left(\frac{1}{2}\sqrt{\frac{N}{k^2\eta}}\right)$$

где $\eta = \frac{\eta_w}{\eta_a}$; $Q(x) \le e^{-\frac{x^2}{2}}$.

Практическое занятие 11.
Использование ЦВЗ для аутентификации ПО.

Вопросы для проверки знаний

1. Чем аутентификация при помощи ЦВЗ отличается от обычной аутентификации цифровых сообщений?
2. Какие требования предъявляются к аутентификации при помощи ЦВЗ?
3. Чем отличается точная и селективная аутентификация?
4. В чем заключается концепция аутентификации при помощи модульных операций? Каковы недостатки этого метода?
5. * Является ли вероятность ошибки извлечения биты информации при модульном погружении монотонно убывающей функции коэффициента глубины погружения? Если нет, то почему?
6. Описать основную идею аутентификации с использованием ЦВЗ при выполнении процедуры сжатия без потерь.
7. Описать основную идею аутентификации с использованием ЦВЗ при выполнении преобразования гистограмм.

Задачи

Пусть значения яркостей пикселей на одной из групп G равна (0,1,2,3,4). Найти значения дискриминантной функции на этой группе и на группе преобразованной по правилу $0 \leftrightarrow 1, 2 \leftrightarrow 3, 3 \leftrightarrow 4, 4 \leftrightarrow 5$. Будет ли

эта группа регулярной, сингулярной или неиспользуемой? Задать значений яркостей на группе G, состоящей из четырех пикселей, когда эта группа окажется сингулярной.

Практическое занятие 12.
Пропускная способность систем СГ и ЦВЗ.

Вопросы для проверки знаний

1. каково определение пропускной способности для стегосистем и чем оно отличается от пропускной способности для обычных телекоммуникационных систем по Шеннону?
2. Как рассчитывается пропускная способность стегосистем для каналов с шумом?
3. Что означает тот факт, что пропускная способность СГ для каналов с шумом равна нулю? Означает ли это, что в этом случае невозможно вложить секретную и надежную информацию?
4. Как определить пропускную способность для систем с ЦВЗ?
5. Перечислите основные свойства пропускной способности для систем ЦВЗ.
6. Поясните формулу для пропускной способности систем ЦВЗ при модели ПО в ДСК.
7. Поясните формулу для пропускной способности систем ЦВЗ при модели гауссовского ПО.
8. Почему значение пропускной способности СГ и ЦВЗ выведенные для моделей ДСК и гауссовского ПО являются важными и для практически реальных моделей ПО?

Задачи

Рассчитать максимально допустимую скорость вложения информации для модели СГ в канале с шумом при следующих исходных данных: $D = 0{,}1; n = 100$. В каких единицах измеряется скорость вложения?

Указание. Можно воспользоваться соотношением [1]:

$$R \leq 1{,}2\sqrt{\frac{D}{n}}$$

Литература

Основная:

1. Коржик, В.И. Курс лекций: «Основы стеганографии». [Электронный ресурс] / В.И. Коржик. – Электрон. дан. – Режим доступа: http://zss.sut.ru/materialy/#, свободный. – Загл. с экрана.

2. Fridrich, J. Steganography in Digital Media: Principles, Algoritms and Applications / J. Fridrich. – 2010. – 441 p.

Дополнительная:

3.Грибунин В.Г. Цифровая стеганография/В.Г. Грибунин, И.Н. Оков, И.В. Туринцев //— М.: Солон-Пресс, 2002.

3.Barni M.Watermarking system Engineering /M. Barni, F. Bartolini // Maral Dekker, 2004.

4. Cox I., et all, Digital Watermarking / MK, 2002.

Валерий Иванович Коржик
Ксения Андреевна Небаева

ОСНОВЫ СТЕГАНОГРАФИИ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ЛАБОРАТОРНЫМ РАБОТАМ

Ответственный редактор ***В.И. Коржик***

Редактор ***Л.А. Медведева***

План 2014 г., п. 7

Подписано к печати 24.10.2011
Объем 32усл. печ. л. Тираж?? экз. Зак. ??
Издательство СПбГУТ. 191186 СПб., наб. р. Мойки, 61
Отпечатано в СПбГУТ

В. И. Коржик
К. А. Небаева

ОСНОВЫ СТЕГАНОГРАФИИ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

САНКТ-ПЕТЕРБУРГ
2014