Лабораторная работа № 12

Исследование стеганографического метода на основе преобразования наименее значащих бит

Цель: изучение стеганографического метода осаждения/ извлечения тайной информации с использованием электронного файла-контейнера на основе преобразования наименее значащих бит (НЗБ), приобретение практических навыков программной реализации данного метода

Задачи:

1.Закрепить теоретические знания из области стеганографического преобразования информации, моделирования стеганосистем, классификации и сущности методов цифровой стеганографии.

2. Изучить алгоритм осаждения/извлечения тайной информации на основе метода НЗБ (LSB – Least Significant Bit), получить опыт практической реализации метода.

3. Разработать приложение для реализации алгоритма осаждения/ извлечения тайной информации с использованием электронного файла-контейнера на основе метода НЗБ.

4. Познакомиться с методиками оценки стеганографической стойкости метода НЗБ.

5. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

12.1 ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

12.1.1. Основные определения, классификация и сущность стеганографических методов

Сведения, необходимые для понимания сущности вопросов, относящихся к предметной области, в достаточном объеме изложены в [2] (гл. 7); полезно также ознакомиться с содержанием книги [52].

Здесь мы остановимся на основных понятиях, которыми будем далее оперировать.

Определение 1. *Стеганографическая система* (*stegosystem, стегосистема* или *стеганосистема* – в русскоязычной тематической литературе используются оба сокращения) – совокупность средств и методов, которые используются для формирования *скрытого канала* передачи (или хранения) информации.

При этом скрытый канал организуется на базе и внутри открытого канала с использованием особенностей восприятия информации. «Скрытость» канала передачи тайной информации отличает стеганографии от криптографии: в первом случае тайной является сам факт наличия канала (передачи информации).

Определение 2. Абстрактно *стеганографическая система* обычно определяется, как некоторое множество отображений одного пространства (множества возможных сообщений, *M*) в другое пространство (множество возможных стеганосообщений, *S*, и наоборот.

Основные компоненты стеганосистемы:

контейнер, *C* (файл-контейнер или электронный документ произвольного формата), в котором размещается (осаждается, скрывается) тайное сообщение, *M*; именно контейнер является упомянутым скрытым каналом;

тайное сообщение, *M*, осаждаемое в контейнер для передачи или хранения (например, с целью доказательства или защиты авторских прав на документ-контейнер [2, 53-56]; здесь речь может идти о невидимых цифровых водяных знаках, ЦВЗ);

ключи или *ключевая информация*, *K* системы, выполняющие ту же функцию, что и криптографические ключи; ключей может быть несколько, в соответствии с этим современные стеганосистемы характеризуют как *многоключевые*: один ключ отождествляется с методом осаждения/извлечения тайной информации, другой – с выбором элементов (например, битов) контейнера для его модификации при осаждении тайной информации, третий (или третьи) – для предварительного (перед осаждением) преобразования тайной информации (например, на основе помехоустойчивого кодирования, сжатия или зашифрования) и т. д. [2, 57, 58];

контейнер с осажденным сообщением или *стеганоконтейнер*, *S*, который передается по *открытому каналу*, также являющемуся важным компонентом анализируемой системы; стеганоконтейнер будем именовать также *стеганосообщением*;

для полноты упомянем также субъектов системы: *отправителя* и *получателя*.

В зависимости от формата документа-контейнера *цифровую* (или *компьютерную*) стеганографию подразделяют на классы [2, 52, 59-64]:

- аудиостеганография,
- видеостеганография,
- графическая стеганография,
- текстовая стеганография

и др.

Определение 3. Стеганографической системой $\sum$ будем называть совокупность сообщений $\mathbf{M}$, контейнеров $\mathbf{C}$, ключей $\mathbf{K}$, стеганосообщений (заполненных контейнеров) $\mathbf{S}$ и преобразований (прямого $\mathbf{F}$ и обратного $\mathbf{F}^{-1}$), которые их связывают:

$$\sum = (\mathbf{M}, \mathbf{C}, \mathbf{K}, \mathbf{S}, \mathbf{F}, \mathbf{F}^{-1}). \qquad (12.1)$$

Как видим, сущностью рассматриваемой системы является тайное хранение или передача одной информации в другой информации, которая является открытой.

При построении стеганосистемы должны, таким образом, учитываться следующие основные положения:

- свойства контейнера должны быть модифицированы так, чтобы изменение невозможно было выявить при визуальном контроле; это требование определяет качество сокрытия внедряемого сообщения: для обеспечения беспрепятственного прохождения стеганосообщения по каналу связи оно никоим образом не должно привлечь внимание атакующего;
- противник (интруз) имеет полное представление о стеганографической системе и деталях ее реализации; единственной информацией, которая остается ему неизвестной, является ключ, с помощью которого только его держатель может установить факт присутствия и содержание скрытого сообщения;
- если противник каким-то образом узнает о факте существования скрытого сообщения, это не должно позволить ему извлечь подобные сообщения до тех пор, пока ключ хранится в тайне;
- потенциальный противник должен быть лишен каких-либо технических и иных преимуществ в распознавании или раскрытии содержания тайных сообщений.

Информацию об основных видах атак на стеганосистемы можно найти, например, в [2].

12.1.2 Метод НЗБ и особенности его реализации

Большинство исследований в предметной области посвящено использованию в качестве стеганоконтейнеров изображений (текст также можно рассматривать как изображение). Это обусловлено следующими причинами:

- относительно большим объемом цифрового представления изображений, что позволяет внедрять большой объем данных;
- заранее известным размером контейнера, отсутствием ограничений, накладываемых требованиями реального времени;
- наличием в большинстве реальных изображений текстурных областей, имеющих шумовую структуру и хорошо подходящих для встраивания информации;
- слабой чувствительностью человеческого глаза к незначительным изменениям цветов изображения, его яркости, контрастности, содержанию в нем шума, искажениям вблизи контуров;
- хорошо разработанными в последнее время методами цифровой обработки изображений.

Метод НЗБ основывается на ограниченных способностях зрения или слуха человека, вследствие чего людям тяжело различать незначительные вариации цвета или звука. Рассмотрим это на примере 24-битного растрового RGB-изображения. Как известно, каждая точка кодируется 3-мя байтами. Каждый байт определяет интенсивность красного (Red), зеленого (Green) и синего (Blue) цветов. Совокупность интенсивностей цвета в каждом из 3-х каналов определяет оттенок пикселя.

Представим пиксель тремя байтами в битовом виде, как это показано на рис. 12.1.

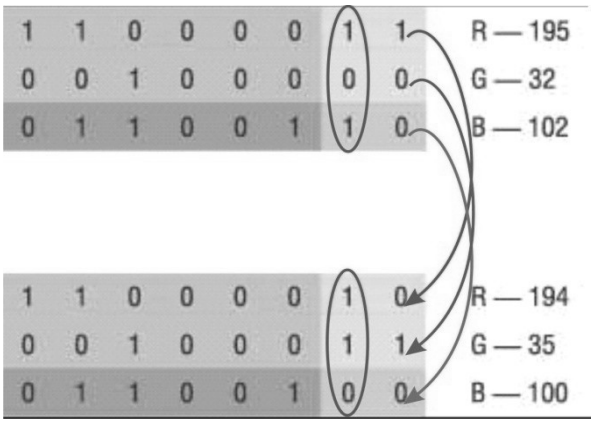


Рисунок 12.1 Пример, показывающий принцип реализации метода LSB

Младшие биты (выделены бледным, справа) дают незначительный «вклад» в изображение по сравнению со старшими.

Замена одного или даже нескольких младших бит для человеческого глаза будет почти незаметна, поскольку реально человек может различать около полторы сотни цветовых оттенков.

Рассмотрим простейший пример.

Пример 1. Контейнером *C* выступает обычная буква «А». Цвет текста-контейнера «А» – черный: данный цвет представлен в MS Office Word как (00000000, 00000000, 00000000), т.е. (0, 0, 0), см. рис. 12.2, а.

а) б)

Рисунок 12.2 Диалоговое окно MS Wordc указанием цветовых координат символов текста: (0, 0, 0) – а), (1, 1, 1) – б)

Необходимо внедрить секретное сообщение: $M = 111$ в текст-контейнер *C*, используя текстовый процессор MS Word. Мы решаем задачу чисто механически: изменяем младший из символов цветового кода в каждом канале (рис. 12.2), т. е. в десятичном виде это можно

представить как (1, 1, 1), а в двоичном – (00000001, 00000001, 00000001). Результат осаждения секретного сообщения (111) в текст-контейнер «А» показан на рисунке 12.3: «пустой» контейнер (*C*) никак визуально не отличается от стеганоконтейнера (*S*).

На рис. 12.4 приведены четыре строки символов, цветовые координаты которых соответствуют различным числам, с изменением вплоть до пятого (справа-налево) бита цветового кода (см. табл. 12.1). В таблице номера столбцов с кодами соответствуют позиции символа в строке на рис. 12.4. Цветовые оттенки символов (в строках) на рис. 12.4 едва различимы, притом только в четвертой строке.

Подчеркнем, что именно визуальный анализ графического объекта является основой наиболее часто используемой (прежде всего, в силу трудозатрат) методики стеганографического анализа.

а) не модифицированный символ-контейнер; б) модифицированный символ-контейнер (со встроенным секретным сообщением «111»)

Рисунок 12.3 Пример практической реализации метода НЗБ

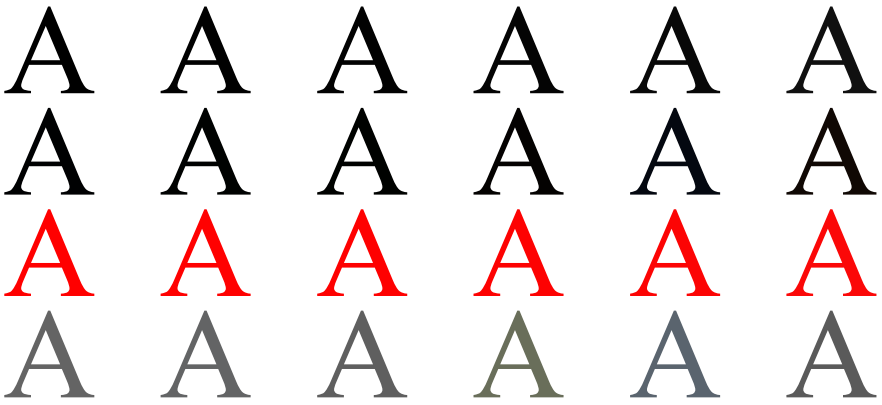


Рисунок 12.4 Одинаковый символ (А) с различной кодировкой цвета

Таблица 12.1 Цветовая кодировка символов на рис. 12.4

Столбец Строка	1	2	3	4	5	6
1	0,0,0	1,1,1	2,2,2	4,4,4	8,8,8	16,16,16
2	0,1,2	0,2,1	0,2,0	4,2,1	4,8,16	16,8,4
3	255,0,0	255,1,1	254,1,1	253,2,2	252,4,4	251,8,8
4	100,100,100	99,100,101	95,95,95	105,110,90	90,100,110	90,90,90

При этом проявляется еще одно важное обстоятельство: примерно в 50% случаев бит, который мы хотим записать, и бит в изображении-контейнере будут совпадать и изменять ничего не нужно.

Понятно, что графические контейнеры в реальной стеганографии много сложнее рассмотренных примеров.

Одним из простейших и понятных для решения наших задач является формат BMP (BitMaP) – одна из форм представления растровой графики. Изображение представляется в виде матрицы пикселов, где каждая точка характеризуется тремя параметрами: *x*-координатой, *y*-координатой и цветом кодом на основе RGB-модели. Все операции графического ввода-вывода на экран монитора (принтера и на некоторые другие устройства) в конечном итоге осуществляются в этом формате. Для работы с этим форматом в ОС Windows предусмотрено много специальных функций и структур API, которые помогают производить все необходимые операции на достаточно высоком логическом уровне.

Контейнеры на основе BMP-формата разделяют на два класса: «чистые» и зашумленные. В первых прослеживается связь между младшими и остальными битами элементов цвета, а также видна зависимость самих младших битов между собой. Осаждение сообщения в такой контейнер нарушает такие зависимости, что легко выявляется аналитиком. Если же картинка зашумлена (например, получена со сканера или фотокамеры), то определить осажденное сообщение сложнее. Таким образом, в качестве файлов-контейнеров для метода LSB рекомендуется использовать файлы, которые не были созданы на компьютере изначально.

Другим из растровых форматов используемых в стеганографии контейнеров является формат PNG (Portable Network Graphics). По качеству цветового отображения данный формат превосходит JPEG (Joint Photographic Experts Group) и GIF (Graphics Interchange Format), но размер файла будет на 30-40% больше.

Вышеприведенные табличные и иллюстративные данные, а также опыт специалистов показывают, что при модификации даже 3-4 младших разрядов состояние графического стеганоконтейнера у экспертов подозрений не вызывает при визуальном его контроле.

Исходя из такой оценки, следует соотносить объем осаждаемого сообщения, V_M с объемом V_C используемого контейнера. Например, если размер изображения 500х500 = 250 000 пикселов, а с учетом используемой 3-хцветовой модели имеем 750 000 единиц цветовых координат. Если мы планируем модифицировать только самые младшие биты всех цветовых каналов матрицы, то максимальный объем осаждаемого сообщения ($V_{M\,max}$) не должен превышать 750 тыс. бит.

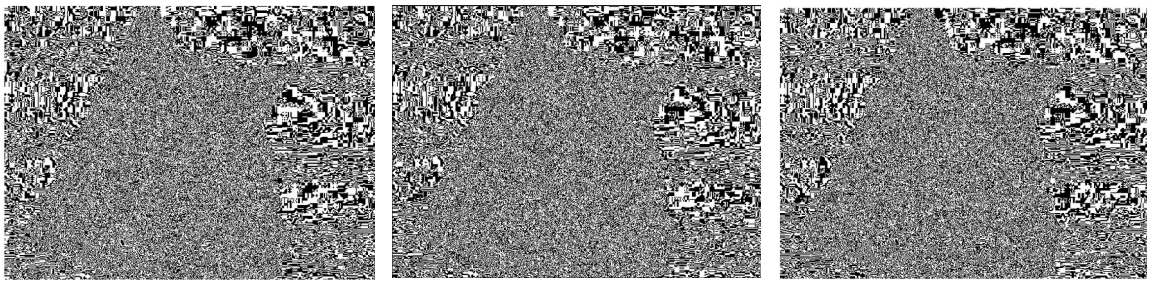
Посмотрим далее на некоторые технические детали и особенности реализации метода НЗБ при использовании в качестве контейнера изображения в формате PNG.

Незаполненный контейнер имеет вид, показанный на рис. 12. 5.

Далее возьмем самый младший бит в каждом цветовом канале и отобразим на этом основании «самый нижний слой» исходного изображения в черно-белых (иначе нельзя) оттенках: нулевое значение младшего бита соответствует белому цвету, единичное – черному. Полученные мозаики показаны на рис. 12.6: а) соответствует красному цветовому каналу, б) – зеленому, в) – синему. Даже внимательный сравнительный анализ трех картинок показывает практически полную их идентичность.

Источник: https://habr.com/ru/post/422593/

Рисунок 12.5 Вид «пустого» контейнера



а) б) в)

Рисунок 12.6 Черно-белое отображение младших разрядов «пустого» контейнера в красном канале (а), **в зеленом канале** (б), **в синем канале** (в)

Теперь наложим изображения на рис. 12.6 со следующими кодовыми параметрами пикселов в каждом канале: 1 – 255 (11111111 – в бинарном коде; т.е. классический красный, либо классический зеленый, либо классический синий), 0 – 0 (00000000 – в бинарном коде). Результат показан на рис. 12.7.

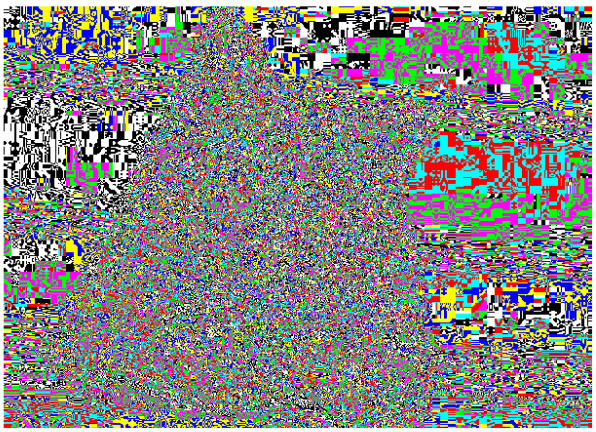


Рисунок 12.7 Цветовое отображение младших разрядов «пустого» контейнера в красном

Именно последняя картинка может отождествляться с самым младшим или наименее значащим битом исходного, т. е. пустого контейнера. Такое изображение для стеганоконтейнера может служить основе для выполнения операций стеганоанализа. Для того, чтобы этот процесс затруднить, сообщение осаждается в цветовые каналы пикселов в нерегулярной, а в псевдослучайной последовательности. Такая последовательность должна рассматриваться как один из элементов ключа стеганосистемы.

Алгоритм реализации, как мы видим из примеров, достаточно прост. Прежде всего, нужно определиться с содержанием сообщения *M*, а далее – выбрать контейнер с учетом наших вышеприведенных

оценок. В задачах по защите прав интеллектуальной собственности нужно идти от обратного, т.е., прежде всего, иметь в виду готовый контейнер. В обоих случаях нужно сопоставлять объемы контейнера и осаждаемого сообщения.

Для затруднения стеганоанализа порядок и количество осаждаемых бит в различные цветовые каналы можно подчинить различным правилам. В качестве аналога можно использовать, например, подход, реализованный в [65]. Для операций «размазывания» сообщения по контейнеру могут применяться ключевой файл и пароль в виде, например, текстов, которые символы которых заменяются числами и в совокупности определяют местоположение пикселя для записи в него части секретного сообщения. Простейшая реализация этого подхода показана на рис. 12.8.

Обычно для выполнения операций стеганографического анализа применяется метод «χ-квадрат». Особенности практического применения такого и других методов (атак) можно найти в [66].

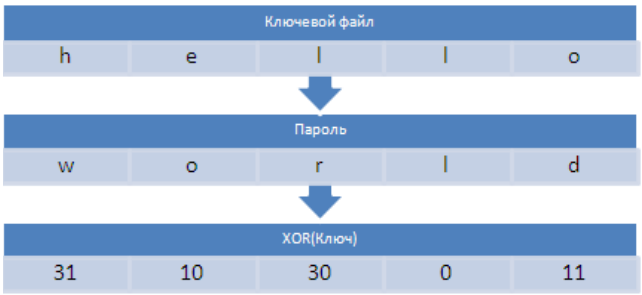


Рисунок 12.8 Пояснение к расчету местоположения пикселов для осаждения/извлечения сообщения

12.2 ПРАКТИЧЕСКОЕ ЗАДАНИЕ

1. Перед выполнением основного задания целесообразно познакомится со структурой, интерфейсом и функциональными особенностями доступных и заслуживающих внимания приложений, в которых реализован метод НЗБ. К ним можно отнести следующие:

- *openstego* (https://github.com/syvaidya/openstego) – может применяться не только для осаждения данных, но и для ЦВЗ; использует *RandomLSB* – псевдослучайный принцип осаждения; поддерживает шифрование (дополнительный ключ), имеет также GUI; осаждение реализуется командой *openstego embed-mf secret.txt -cf cover.png -p password -sfstego.png*, извлечение – *openstego extract -sf*

openstego.png -p abcd -xf output.txt; как видим, работает с PNG-контейнерами;

- *Stegano* (https://github.com/cedricbonhomme/Stegano) – работает не только с классическим LSB; имеет гибкую настройку, может также использоваться как модуль Python; осаждение реализуется командой *stegano-lsb hide --input cover.jpg-f secret.txt -e UTF-8 --output stego.png* , извлечение – *stegano-lsb reveal -i stego.png -e UTF-8 -o output.txt* ; работает также с PNG-контейнерами;
- *LSB-Steganography* (https://github.com/RobinDavid/LSB-Steganography) – приложение написано на Python; работает с PNG- и BMP-контейнерами.

2. Разработать собственное приложение, в котором должен быть реализован метод НЗБ. При этом:

- выбор файла-контейнера – по согласованию с преподавателем;
- реализовать два варианта осаждаемого/извлекаемого сообщения:
  - ✓ собственные фамилия, имя и отчество,
  - ✓ текстовая часть отчета по одной из выполненных лабораторных работ;
- реализовать два метода (на собственный выбор) размещения битового потока осаждаемого сообщения по содержимому контейнера;
- сформировать цветовые матрицы (по аналогии с рис. 12.7), отображающие каждый задействованный для осаждения уровень младших значащих бит контейнера;
- выполнить визуальный анализ (с привлечением коллег в качестве экспертов) стеганоконтейнеров с различным внутренним содержанием; сделать выводы на основе выполненного анализа.

3. Результаты выполнения работы оформить в виде отчета по установленным правилам.

ВОПРОСЫ ДЛЯ КОНТРОЛЯ И САМОКОНТРОЛЯ

1. Охарактеризовать цели, задачи и области применения стеганографии.
2. В чем состоят сходства и различия между стеганографией и криптографией?
3. Дать определение стеганографической системы. Охарактеризовать составные части стеганосистемы и их взаимосвязь.

4. Основные классификационные критерии методов стеганографии.

5. Пояснить сущность основных атак на стеганосистемы.

6. Изобразить структурную схему стеганографической системы.

7. Сущность метода НЗБ. Области его применении.

8. Изобразить алгоритмы осаждения и извлечения сообщений на основе метода НЗБ при передаче этих сообщений.

9. Изобразить алгоритмы осаждения и извлечения сообщений на основе метода НЗБ при решении задачи защиты прав интеллектуальной собственности на электронный контент.

52. Грибунин, В.Г. Цифровая стеганография. Аспекты защиты/ В.Г. Грибунин, И.Н. Оков, И.В. Туринцев. – М.: Солон-Пресс, 2002. – 272 с.

53. Urbanovich, P. Text steganography application for protection and transfer of the information/ P. Urbanovich, K. Chourikau, A. Rimorev, N. Urbanovich. – Przeglad elektrotechniczny. – 2012. – № 8. – P. 342–344. https://elib.belstu.by/handle/123456789/24783

54. Шутько, Н. П. Защита авторских прав на электронные текстовые документы методами стеганографии / Н. П. Шутько // Труды БГТУ. - Минск : БГТУ, 2013. - № 6 (162). - С. 131-134. https://elib.belstu.by/handle/123456789/9708

55. Шутько, Н. П. Защита авторских прав на текстовые документы на основе стеганографической модификации цвета символов текста / Н. П. Шутько, П. П. Урбанович // Информационные технологии: материалы 83-й научно-технической конференции профессорско-преподавательского состава, научных сотрудников и аспирантов (с международным участием), Минск, 4-15 февраля 2019 г. / отв. за изд. И. В. Войтов; УО БГТУ. – Минск : БГТУ, 2019. – С. 41-43. https://elib.belstu.by/handle/123456789/28369

56. Text steganography application for protection and transfer of the information / Pavel Urbanovich, Konstantin Chourikov, Andrey Rimorev, Nadzeya Urbanovich // Przegląd elektrotechniczny. – 2010. – R. 86. – № 7.– P. P. 342–344. https://elib.belstu.by/handle/123456789/24783

57. Urbanovich, P. Theoretical Model of a Multi-Key Steganography System / P. Urbanovich, N. Shutko // Recent Developments in Mathematics and Informatics. Contemporary Mathematics and Computer Science. Vol. 2, Chapter 11. – Lublin: KUL, 2016. – P. 181-202

https://elib.belstu.by/handle/123456789/24543

58. Шутько, Н. П. Моделирование стеганографической системы в задачах по охране авторских прав / Н.П. Шутько, Н.И. Листопад, П.П. Урбанович // Восьмая Междунар. научно-техн. конф. «Информационные технологии в промышленности» (ITI`2015): тезисы докладов. – Минск: ОИПИ НАН Беларуси, 2015. – С. 30-31.

https://elib.belstu.by/handle/123456789/25880

59. Блинова, Е. А. Сравнительные особенности использования стеганографических методов в электронных картах / Е. А. Блинова, П. П. Урбанович // X Международная научно-техническая конференция «Информационные технологии в промышленности, логистике и соци-

альной сфере» (ITI*2019) : тезисы докладов, Минск, 23-24 мая 2019 г. - Минск : ОИПИ НАН Беларуси, 2019. - С. 22-25.
https://elib.belstu.by/handle/123456789/29368

60. Блинова, Е. А. Стеганографический метод на основе встраивания дополнительных значений координат в изображения формата SVG / Е. А. Блинова, П. П. Урбанович // Труды БГТУ. Сер. 3, Физико-математические науки и информатика. – Минск: БГТУ, 2018. – № 1 (206). – С. 104-109.
https://elib.belstu.by/handle/123456789/25323

61. Блинова, Е. А. Стеганографический метод на основе встраивания дополнительных значений координат в пространственные данные, хранящиеся в базе данных / Е. А. Блинова, П. П. Урбанович // Информационные технологии: тезисы докладов 82-й научно-технической конференции профессорско-преподавательского состава, научных сотрудников и аспирантов (с международным участием), Минск, 1-14 февраля 2018 г. / Белорусский государственный технологический университет. – Минск: БГТУ, 2018. – С. 8-9.
https://elib.belstu.by/handle/123456789/24678

62. Сущеня, А. А. Применение форматов электронных книг при передаче конфиденциальной информации методами компьютерной стеганографии / А. А. Сущеня, П. П. Урбанович // Информационные технологии: материалы 83-й научно-технической конференции профессорско-преподавательского состава, научных сотрудников и аспирантов (с международным участием), Минск, 4-15 февраля 2019 г. / отв. за изд. И. В. Войтов; УО БГТУ. – Минск: БГТУ, 2019. – С. 39-40.
https://elib.belstu.by/handle/123456789/28378

63. Колмаков, М. В. Особенности применения стеганографических методов в альтернативных потоках файловой системы NTFS / М. В. Колмаков, Е. А. Блинова // Информационные технологии: тезисы докладов 82-й научно-технической конференции профессорско-преподавательского состава, научных сотрудников и аспирантов (с международным участием), Минск, 1-14 февраля 2018 г. / Белорусский государственный технологический университет. - Минск : БГТУ, 2018. - С. 23-24. https://elib.belstu.by/handle/123456789/24667

64. Сущеня, А. А. Программное средство стеганографического преобразования текстов-контейнеров на основе языка разметки XML / А. А. Сущеня // 69-я научно-техническая конференция учащихся, студентов и магистрантов, 2-13 апреля 2018 г., Минск: сборник научных работ : в 4 ч. Ч. 4 / Белорусский государственный технологический университет. - Минск : БГТУ, 2018. – С. 81-84.

https://elib.belstu.by/handle/123456789/27087

65. Пласковицкий, В. А. Шифрование кодов программ на основе ключа, задаваемого рекуррентными математическими соотношениями / В. А. Пласковицкий, П. П. Урбанович // Труды БГТУ. – Минск: БГТУ, 2012. – № 6 (153).– С. 146-148.
https://elib.belstu.by/handle/123456789/3234

66. Andreas Westfeld and Andreas Pfitzmann. Attacks on Steganographic Systems
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.5975&rep=rep1&type=pdf