УДК 004.056.5:518

Борисенко И.И., Кобозева А.А.
Borisenko I.I., Kobozeva A.A.

ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ СТЕГАНОГРАФИЧЕСКОГО МЕТОДА, ОСНОВАННОГО НА РЕШЕНИИ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

PRACTICAL ACHIEVEMENT OF STEGANOGRAPHIC METHOD BASED ON LINEAR ALGEBRAIC EQUATIONS SYSTEMS SOLUTION

Аннотация

Областью исследования настоящей работы является компьютерная стеганография. Данная статья посвящена реализации нового стеганографического метода, основанного на решении систем линейных алгебраических уравнений, предложенного Кобозевой А.А. ранее. Полученные в работе практические результаты позволяют подтвердить теоретически обоснованный рост устойчивости метода за счет двухэтапного декодирования, а также возможность его применения для любого основного сообщения.

Abstract

The research area of this work is computer steganography. The main aim of this article is achievement of new steganographic method based on linear algebraic equations systems solution, which offered by Kobozeva A.A. earlier. The received practical results allow to confirm growth of stability of this method due to two-stage decoding, and also its application for any basic message.

1. Введение

Одной из актуальных, но нерешенных на сегодняшний день проблем является задача защиты авторских прав, прав интеллектуальной собственности, а также конфиденциальных данных, имеющих цифровой формат, от несанкционированного доступа. Важным направлением в решении этого вопроса является разработка методов сокрытия информации, в частности, цифровой стеганографии [1,2].

Все стеганографические методы характеризуются тем, что скрываемое сообщение, или дополнительная информация (ДИ), встраивается в некоторый объект, или основное сообщение (ОС), не привлекающий внимания, который затем открыто пересылается адресату по каналу связи. Эффективность любого стеганографического метода оценивается, исходя из совокупности требований, среди которых важное место занимают надежность восприятия после погружения скрываемого сообщения, эффективность декодирования ДИ при заданных помехах, устойчивость декодирования к возмущающим воздействиям в канале связи, пропускная способность.

В [3] был предложен и теоретически обоснован новый стеганографический метод организации пересылки и декодирования секретного сообщения x, основанный на решении систем линейных алгебраических уравнений (СЛАУ), главной задачей которого являлось обеспечение дополнительной «защиты» скрываемой информации, или повышение устойчивости процесса декодирования к возмущающим воздействиям, применимый для любого ОС.

Сообщение b, подлежащее непосредственному погружению, формировалось с использованием матрицы F ОС и секретного сообщения x. Первым шагом декодирования являлось извлечение b каким-либо известным устойчивым алгоритмом из полученного стегосообщения. Выделение информации, представляющей непосредственный интерес для адресата, осуществлялось на втором этапе при решении СЛАУ, матрица которой отвечала возмущенной матрице ОС. Устойчивость этого этапа зависела от числа обусловленности F и обеспечивалась для произвольного ОС за счет формирования виртуального диагонального преобладания в матрице системы. К детальному рассмотрению идеи метода мы вернемся в п.2.

Целью настоящей статьи является разработка и реализация предложенного метода, дающая возможность практически подтвердить

а) более высокую устойчивость процесса двухэтапного декодирования к возмущающим воздействиям в канале связи по сравнению с непосредственной пересылкой и декодированием секретной информации;

б) независимость результатов работы нового метода (объема правильно восстановленной информации) от числа обусловленности матрицы ОС.

Предлагаемая реализация призвана решить вопросы, оставшиеся в [3] открытыми:

1)вопрос организации непосредственного решения СЛАУ, требующего $\underline{O}(n^2)$ арифметических операций, где n - размерность системы, независимо от вида матрицы ОС;

2) увеличение пропускной способности.

Данная работа является очередным шагом авторов по пути использования вычислительной линейной алгебры в области цифровой стеганографии.

2. Формирование СЛАУ для декодирования информационного сообщения

Не ограничивая общности рассуждений, в качестве ОС, или контейнера, будем рассматривать произвольное монохромное изображение, моделью которого является квадратная $n \times n$-матрица F. ДИ (вектор x) имеет вид числовой последовательности, содержащей n или менее (в этом случае последовательность дополняется до нужной длины незначащими элементами, а длина информационной части является частью секретного ключа) элементов, каждый из которых принадлежит множеству $-1,1$. По матрице ОС и вектору ДИ формируется новый вектор b, размерности n:

$$b = Fx, \tag{1}$$

который погружается в F вместо несущего нужную информацию x, являющегося решением СЛАУ

$$Fx = b\ . \tag{2}$$

Важным здесь является тот факт, что встраиваемое сообщение адаптируется к виду контейнера (формула (1)). Сразу заметим, что проделать это погружение непосредственно затруднительно, т.к. элементы вектора b могут иметь очень большие значения.

Декодирование секретного сообщения x происходит на втором этапе при решении СЛАУ $F_B x_{пр} = b_B$, где F_B, b_B - возмущенные матрица F и вектор b соответственно, $x_{пр}$ - полученное после декодирования приближенное значение x.

В [3] показано, что упомянутый выше метод будет устойчивым к возмущающим воздействиям в канале связи, т.е. позволит адресату декодировать полученную информацию с малой результирующей ошибкой, если матрица изображения, используемого в качестве ОС, будет иметь малое число обусловленности Скила, а также

предложен практический метод обеспечения этого условия для произвольного изображения, никак это изображение не изменяющий реально.

Идеальной с точки зрения устойчивости алгоритма декодирования [3] является диагональная матрица F ОС. Очевидно, качественная картина не будет сильно нарушена, если для элементов F будет выполняться соотношение

$$f_{ii} >> \sum_{j=1,\, j\neq i}^{i-1} f_{ij},\ i=\overline{1,n}. \tag{3}$$

Но для реальных изображений (3) выполняется редко. Моделирование выполнения условия (3) для матрицы F проведем виртуально, никак не затрагивая матрицу исходного изображения. Для этого матрице F поставим в соответствие нижнюю треугольную матрицу $\overline{F}$ той же размерности, элементы которой определим следующим образом:

$$\overline{f}_{ij} = \begin{cases} 0, \textit{если } f_{ij} \le 127, \\ 1, \textit{если } f_{ij} > 127 \end{cases},\ j=\overline{1,n-1},\ i=\overline{j+1,n}, \tag{4}$$

$$\overline{f}_{11} = m,\ \overline{f}_{ii} = m\sum_{j=1}^{i-1} \overline{f}_{ij},\ i=\overline{2,n},\ m\in N \tag{5}$$

где m – натуральное число, выбор которого призван обеспечить для матрицы $\overline{F}$ наличие свойств, близких к (3), а логика требования треугольного вида $\overline{F}$ станет ясна ниже.

Формирование вектора правой части системы уравнений для определения x теперь произведем не с помощью матрицы исходного контейнера в соответствии с (1), а при помощи $\overline{F}$ следующим образом.

1. Вычисляем вектор $\bar{b} = \overline{F}x$. Тогда x есть решение СЛАУ

$$\overline{F}x = \bar{b}, \tag{6}$$

матрица которой $\overline{F}$ по построению имеет малое число обусловленности Скила. Заметим, что за счет (5) элементы $\bar{b}$ все еще могут оказаться достаточно большими.

2. По $\bar{b}$ генерируем вектор b^* по следующему правилу:

$$b_i^* = \begin{cases} \bar{b}_i, \textit{если } -15 \le \bar{b}_i \le 15, \\ -15, \textit{если } \bar{b}_i < -15 \\ 15, \textit{если } \bar{b}_i > 15 \end{cases},\ i=\overline{1,n}, \tag{7}$$

а СЛАУ (6) для восстановления x заменяем на

$$\overline{F}x = b^*. \tag{8}$$

Заметим, что матрица системы (8) является нижней треугольной, а, значит, количество арифметических операций при решении СЛАУ определяется как $\underline{O}(n^2)$. Эта

цель преследовалась при кодировании матрицы исходного изображения матрицей треугольного вида.

Остановимся более подробно на целесообразности формирования b^* в соответствии с (7). Кодирование матрицы системы (4), (5), в основном, уменьшает значения элементов матрицы, что дает возможность соответствующего уменьшения и элементов вектора $\bar{b} = \bar{F}x$ по сравнению с $b = Fx$ (конечно, если m не очень велико), однако все же не гарантирует их малости. Если в качестве правой части системы, используемой для восстановления x рассматривать b (система (2)) или $\bar{b}$ (система (6)), то перед погружением такого вектора для обеспечения надежности восприятия после встраивания, очевидно, потребуется какой-либо способ его кодирования, например, перевод в бинарную последовательность, что может значительно увеличить длину погружаемой числовой последовательности. Это приведет к некорректности сравнения результатов декодирования x при его непосредственной и опосредованной, при помощи правой части системы, пересылке. Следовательно, элементы погружаемого вектора должны быть настолько малыми, чтобы избежать необходимости какого-либо кодирования для уменьшения мощности множества, содержащего их возможные значения. Однако при выборе конкретных пороговых значений в (7), 15 и -15, учитывается не только это требование. Обсуждение этого вопроса продолжим в п.5.

При практической реализации метода встает вопрос выбора значения *m*. Очевидно, *m* для обеспечения устойчивости декодирования должно быть как можно больше, хотя реально не может увеличиваться до бесконечности; с другой стороны, *m* непосредственно участвует в формировании элементов вектора $\bar{b}$, по которому генерируется b^*. Чем больше будет *m*, тем больше будут значения элементов $\bar{b}$, тем большее возмущение получит вектор b^* на этапе формирования и, если бы речь шла о классическом решении СЛАУ, то такие возмущения могли бы привести к неприемлемой погрешности декодирования даже с учетом малого числа обусловленности Скила матрицы системы. Однако, как будет показано ниже, используемый нами подход к решению системы позволяет уйти от отрицательных последствий возмущения вектора правой части на этапе формирования. Таким образом, приоритетным остается обеспечение устойчивости декодирования за счет выбора *m*, причем достижение этой цели, как будет показано и обосновано в п.5, происходит при сравнительно небольших значениях *m*.

3. Новый подход к решению СЛАУ для декодирования информационного вектора в предлагаемой реализации метода.

После формирования вектора b^*, он погружается в матрицу F ОС. При пересылке стеганографическое сообщение подвергается возмущающим воздействиям в канале связи. Выделение нужного информационного вектора *x* происходит при решении неоднородной СЛАУ

$$\overline{F}_B \, x_{np} = b_B^* \,, \qquad (9)$$

где $b_B^* = b^* + \delta b^*$ - возмущенный вектор b^*, матрица $\overline{F}_B$ получается по возмущенной матрице $F_B = F + \delta F$ ОС аналогично (4), (5), а, значит, является нижней треугольной матрицей, независимо от вида F_B, что позволяет затратить на второй этап декодирования, решение СЛАУ (9), лишь $\underline{O}(n^2)$ арифметических операций, что чрезвычайно важно при решении систем большой размерности.

Для большинства изображений при малых возмущениях в канале связи $\overline{F}_B \approx \overline{F}$, а $\left\| \overline{F}_B - \overline{F} \right\| \approx 0$. Действительно, отличие этих матриц может быть в тех элементах, которые

являются результатом кодирования элементов исходной матрицы, значения которых находятся в окрестности 127. Таким образом, с незначительным допущением в предположении малых возмущений можно считать, что система (9) отличается от системы (6) лишь вектором правой части.

В любом случае, очевидно, $x_{np} \neq x$. Учитывая вид множества, которому принадлежат элементы *x*, заметим, что для осуществления декодирования нас не столько интересуют непосредственные значения элементов x_{np}, сколько их знак. Окончательный шаг декодирования отвечает формуле:

$$\bar{x}_i = sign(x_{np})_i, \; i = \overline{1,n}. \qquad (10)$$

Вектор $\bar{x}$ назовем *sign-решением* системы (9), а непосредственную реализацию алгоритма, предложенную выше, будем называть CM-SIGN (реализация стеганографического метода sign-решения системы). Заметим, что использование формулы (10) при декодировании, вообще говоря, допускает неограниченно большие погрешности при решении (9), которые никак не повлияют на результат декодирования (10), т.е. $\|x_{np} - x\|$ может быть сколь угодно велика, если при этом выполняются условия: $sign(x_i) = sign(x_{np})_i, \; i = \overline{1,n}$. Таким образом, даже очень большие возмущения правой части системы при формировании b^*, о которых говорилось выше, сохраняющие знаки элементов вектора, не отражаются на результате декодирования, проводимого в соответствии с (9), (10). Такой подход к решению системы является новым и никогда ранее не рассматриваемым, имеющий геометрическую интерпретацию (изложение которой выходит за рамки настоящей работы), дающий возможность, как показывают результаты вычислительного эксперимента, получить большой объем правильно восстановленной информации даже при больших возмущениях входных данных.

4. Модификация метода, увеличивающая пропускную способность

Недостатком предлагаемого метода остается ограниченность пропускной способности, поскольку длина информационного вектора *x* не превосходит n. Увеличение объема передаваемой информации может быть достигнуто следующим образом.

Пусть матрица A размерности $n \times n$ отвечает ОС. Разобьем ее на квадратные блоки фиксированного небольшого размера, например, 8×8 (такая размерность используется в проведенном вычислительном эксперименте), как это делается в одном из наиболее полных и популярных стандартов сжатия неподвижных изображений – стандарте JPEG [4]. Пусть F - один из таких блоков. К каждому из блоков применим предложенный выше CM-SIGN. Заметим, что количество арифметических операций при работе с F - константа, не зависящая от размерности матрицы исходного изображения. Обозначим ее k. Количество блоков определяется как $kol = \left[\frac{n}{8}\right]^2$, тогда общее количество арифметических операций при работе со всем изображением равно

$$kol\ k = \underline{O}(n^2),$$

а суммарный объем погружаемой информации определится как

$$8\,kol = \underline{O}(n^2),$$

т.е. оказывается на порядок больше, чем при рассмотрении матрицы изображения целиком.

Такую модификацию предложенного метода далее будем называть блоковой реализацией стеганографического метода sign-решения системы (БСМ-SIGN).

5. Результаты вычислительного эксперимента.

Основной целью вычисительного эксперимента является практическое подтверждение теоретически обоснованной большей устойчивости к возмущающим воздействиям в канале связи предложенного стеганографического метода с двухэтапным декодированием по сравнению с методом, использующим непосредственную пересылку информационного вектора. Критерием для такого сравнения является объем правильно восстановленной информации в том и другом случае при одинаковых условиях проведения эксперимента.

Вычислительный эксперимент проводился в среде MATLAB. Возмущения в канале связи моделировались при помощи аддитивного гауссовского шума, наложение которого осуществлялось стандартной процедурой *imnoise*, со следующими параметрами: математическое ожидание полагалось везде равным нулю, а среднеквадратическое отклонение σ принимало значения 0.0001, 0.0005, 0.0008, 0.001, 0.01. Дальнейшее увеличение уровня шума приводило к потере информативности возмущенного изображения и поэтому не рассматривалось.

Для демонстрации результатов эксперимента в данной работе были выбраны два изображения, использованных в качестве ОС: POUT (рис.1 (а)) и SATURN (рис.2). Выбор сделан не случайно. Непосредственно устанавливается, что матрица первого изображения является хорошо обусловленной, а число обусловленности матрицы второго изображения бесконечно большое.

а б в

а - исходное изображение POUT; б – стегосообщение, сформированное на основе POUT с использованием CM-SIGN; в – стегосообщение, претерпевшее возмущение в виде аддитивного гаусовского шума, используемое для декодирования

Рисунок 1 – Изображение POUT и его преобразования

Погружение ДИ в ОС проводилось в пространственную область аддитивно в выделенный предварительно контур, причем непосредственное погружение x и погружение b^* происходило в одни и те же пиксели контура. В CM-SIGN пороговые значения -15 и 15 в формуле (7) определялись экспериментально, исходя, во-первых, из требования обеспечения надежности восприятия после погружения скрываемого сообщения (см. рис.1 (б)), а также с учетом особенности, о которой будет сказано ниже. На полученные стегосообщения накладывался шум с одинаковыми параметрами

(например, рис.1 (в)), после чего происходило одноэтапное или двухэтапное (CM-SIGN) декодирование вектора x (декодирование b^* проводилось тем же способом, что и x).

В результате вычислительного эксперимента была установлена зависимость между уровнем возмущающих воздействий в канале связи и процентом правильно восстановленной информации, а также практически подтверждена независимость результатов работы предложенного стегометода от свойств матрицы ОС. Увеличение процента декодирования достигалось путем варьирования m. Результаты проведенного исследования представлены на графиках зависимости процента правильно восстановленной информации от уровня шума, характеристикой которого является среднеквадратичное отклонение [4], при различных значениях m (рис.3). Для построения графиков каждый из экспериментов состоял из 100 независимых опытов, определяемых одинаковыми параметрами. В качестве результата эксперимента (процент правильно декодированной информации при заданном шуме и конкретном значении m) бралось среднее арифметическое значение по всем 100 опытам.

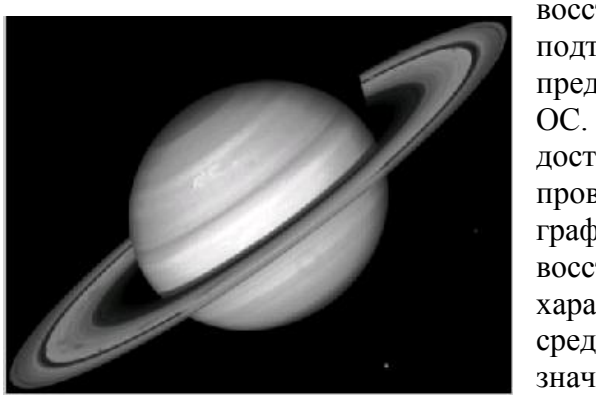
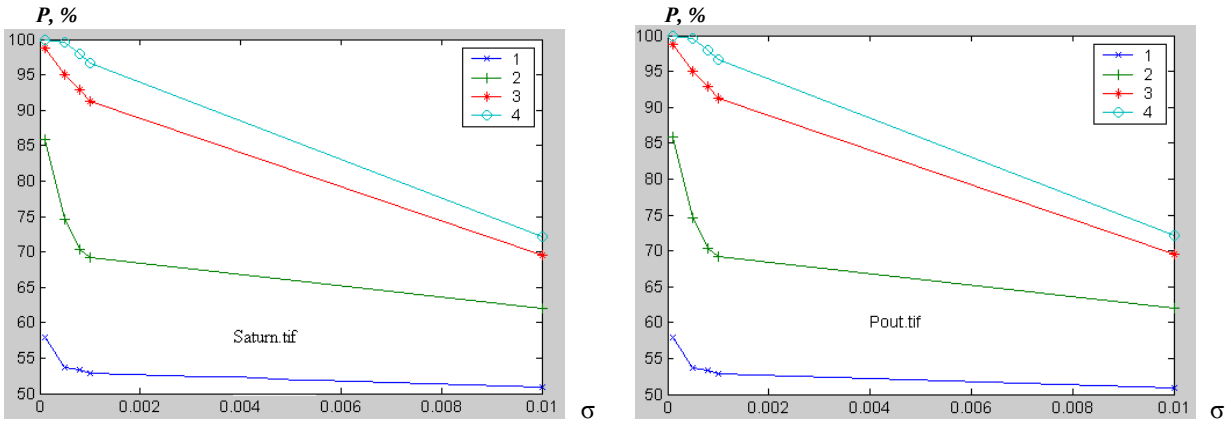


Рисунок 2 – изображение SATURN



а б

а – основное сообщение – SATURN; б – основное сообщение – POUT

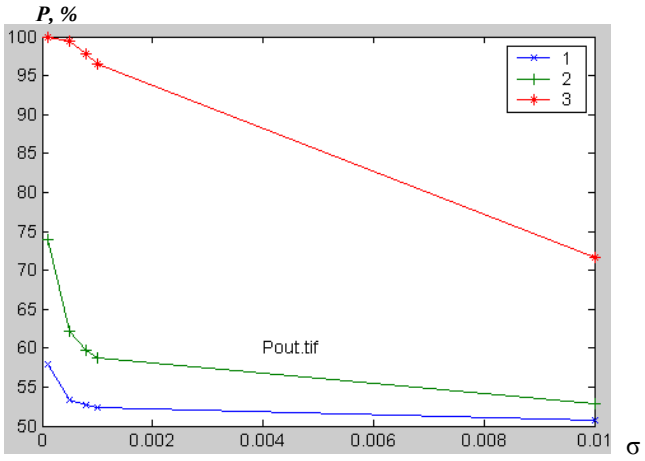
кривая 1 – отвечает непосредственной пересылке и декодированию информационного вектора x; кривая 2 – CM-SIGN при m=2; кривая 3 – CM-SIGN при m=10; кривая 4 – CM-SIGN при $m \geq 15$

Рисунок 3 – Зависимость объема правильно декодированной информации P, выраженного в процентах, от уровня шума, характеризуемого среднеквадратичным отклонением σ (математическое ожидание равно 0), при различных значениях m

Результат работы CM-SIGN, как и ожидалось, в силу виртуального моделирования свойства (3) для матрицы системы, не зависит от свойст матрицы изображения, используемого в качестве ОС (рис.3). CM-SIGN может использоваться для ОС с произвольной матрицей и является более устойчивым к возмущениям в канале связи по сравнению с алгоритмом, осуществляющим непосредственную пересылку информационного вектора, даже при достаточно малых значениях m, с точки зрения результатов декодирования.

Интересным является тот факт, что, как показывает эксперимент, кривая 4 на рис.3(а),(б), построенная для m=15, не будет изменяться при дальнейшем увеличении значения m. Этот странный на первый взгляд факт легко объясняется, исходя из особенностей выбранного нами «метода» решения СЛАУ, описанного в п.3. Поскольку нас интересуют знаки элементов вектора, получаемого при решении (9), то определяющим фактором успешного декодирования является сохранение знаков в векторе b_B^* по сравнению с b^*. Однако, как только диапазон возможных значений b^* перекроется шумом, нужные для декодирования знаки будут безвозвратно утеряны. И тут уже абсолютно не важно, насколько большим будет взято m. На итоговые значения b^* это никак не повлияет (при достаточно больших m все элементы $b^*_i = \pm 15$). Дальнейшее увеличение процента правильно восстановленной информации для достаточно большого квадратичного отклонения, характеризующего шум (возмущающие воздействия в канале связи), можно осуществить только путем расширения границ множества значений элементов b^*, однако это недопустимо при выбранном способе погружения (модули пороговых значений, равные 15, как установлено экспериментально, невозможно более увеличить, не нарушая требования надежности восприятия), и пока остается вопросом открытым.

Реализация БСМ-SIGN, проведенная в вычислительном эксперименте, дающая возможность на порядок увеличить пропускную способность, не изменила качественной картины работы сравниваемых методов, описанной выше, как и можно было предположить заранее. Результаты нашли свое отражение на рис.4 для изображения POUT (для второго изображения SATURN картина идентична).



кривая 1 – отвечает непосредственной пересылке и декодированию информационного вектора x; кривая 2 – БСМ-SIGN при m=2; кривая 3 – БСМ-SIGN при $m \geq 15$

Рисунок 4 – Зависимость объема правильно декодированной информации P, выраженного в процентах, от уровня шума, характеризуемого среднеквадратичным отклонением σ (математическое ожидание равно 0), при различных значениях m
в блоковой реализации СМ-SIGN для изображения POUT

6. Заключение.

Разработка и практическая реализация стеганографического метода, основанного на решении систем линейных алгебраических уравнений, предложенная в настоящей работе, дала возможность практически подтвердить достижение большей устойчивости нового метода к возмущающим воздействиям в канале связи, теоретически обоснованного в [3], по сравнению с методом, основанным на непосредственной пересылке и декодировании информационного вектора. Ценность СМ-SIGN заключается в том, что он может быть

использован для произвольного ОС, а свойства матрицы этого ОС не влияют на результат декодирования. Кроме того, предложенная численная реализация метода не требует большого числа арифметических операций. Их количество определяется как $\underline{O}(n^2)$.

Особого внимания заслуживает предлагаемый новый подход к решению СЛАУ, используемый при декодировании, допускающий бесконечно большие возмущения исходных данных без ущерба для результата восстановления информации.

В работе предлагается модификация БСМ-SIGN, дающая возможность на порядок увеличить пропускную способность, являвшуюся слабым звеном в предложенном первоначально методе.

Открытым пока остается вопрос увеличения диапазона значений элементов вектора b^* без ущерба для надежности восприятия после погружения ДИ, что позволило бы улучшить процент декодирования при достаточно больших возмущениях в канале связи. Этот вопрос требует дополнительных исследований, которые и ведутся в настоящий момент авторами статьи.

Литература

1. Хорошко В.А., Чекатков А.А. Методы и средства защиты информации. – К.: Юниор, 2003. - 501 с.
2. Конахович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. – К.: МК – Пресс, 2006.- 288 с.
3. Кобозева А.А., Коломийчук А.В. Стеганографический метод, основанный на решении систем линейных алгебраических уравнений.-«Праці УНДІРТ», 2006, № 1(45)-2(46), с.104-108.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений.- М.: Техносфера, 2005.- 1072 с.

Журнал «Праці УНДІРТ», №3(47), 2006,С. 78-83