

МЕТОД СТЕГАНОГРАФИЧЕСКОГО ПРЕОБРАЗОВАНИЯ ИНФОРМАЦИИ В ИЗОБРАЖЕНИЯХ JPEG НА ОСНОВЕ МАТРИЧНОГО ВСТРАИВАНИЯ ДАННЫХ

В статье предложен усовершенствованный метод стеганографического преобразования конфиденциальной информации в статических изображениях метода сжатия JPEG (стандарт JFIF) с использованием помехоустойчивых кодов на основе анализа стеганографической защиты данных в информационных сетях. Результаты имитационного моделирования подтверждают результативность данного подхода.

Стеганография, защита информации, помехоустойчивые коды.

On the basis of the analysis of steganographic protection of the data in information networks the advanced method of steganographic transformations of the confidential information to freeze frame images of a method of compression JPEG (standard JFIF) with use of noiseproof codes is offered in article. Results of imitating modeling confirm productivity of suggested approach.

Steganography, information protection, noiseproof codes.

Введение

Первостепенной задачей разработки стеганографической системы (СГС) является противодействие определению самого факта наличия скрытого сообщения, т. е. использования стеганографической системы. Основными средствами, которыми располагает противник, – визуальные и статистические атаки [1]. Для противодействия атакам на начальном этапе нужно оценивать степень пригодности контейнера для сокрытия. Надежная стеганографическая система должна приводить к ошибке второго рода для любого метода стегоанализа (СГА), т. е. в случаях, когда анализ модифицированного контейнера показывает, что контейнер пуст.

Модель стеганографической системы может быть представлена в виде

$$S = \langle M, C, K, K_C, Q, H_K, R_K, N_A, N_\Pi \rangle,$$

где M – множество закрытых сообщений; C – множество возможных контейнеров; K – множество криптографических ключей; K_C – множество стегоключей; Q – множество заполненных контейнеров; H_K, R_K – правила встраивания и извлечения сообщения; N_A, N_Π – преднамеренные информационно-технические воздействия и помехи в канале связи [2].

Одной из существенных характеристик, отражающих качество СГС, является результативность встраивания. В данной работе под результативностью встраивания e_p понимается показатель, определяющий среднее количество бит сообщения, приходящихся на изменение одного бита в исходном контейнере.

Матричное встраивание основывается на теории кодирования, при этом учитываются особые свойства помехоустойчивых кодов, которые позволяют передавать значительно больше информации по сравнению с обычной заменой бит контейнера битами сообщения [3]. Недостатком классического матрич-

ного встраивания является то, что при вложении данных не учитывается статистическая структура объекта-контейнера.

Основная часть

Усовершенствование матричного встраивания. Компьютерное моделирование показало, что использование помехоустойчивых (n, k, t) кодов с высокой корректирующей способностью t в целях матричного встраивания приводит к увеличению числа изменений в изображении. Это обусловлено тем, что за один этап внедрения (передача k бит информации в n битах контейнера) при вычислении синдрома ошибок может сложиться ситуация, требующая изменить не 0 или 1 бит для кодов с корректирующей способностью $t = 1$, а до $t > 1$ бит. Это приведет к большим отличиям заполненного контейнера от исходного. Однако матричное встраивание на основе помехоустойчивых кодов с малой корректирующей способностью уменьшает максимально допустимый объем контейнера [4].

Следовательно, выбирая различные коды и их параметры, удастся либо увеличить результативность встраивания, уменьшив максимальный объем контейнера, либо увеличить объем, теряя результативность. Выбор же для встраивания кодов с высокой корректирующей способностью позволяет увеличить максимально возможный объем закладки, практически не теряя (а в некоторых случаях значительно увеличивая) результативность, но при этом действительное количество внесенных изменений может оказаться велико [3]. Кроме того, встраивание на основании кодов с высокой корректирующей способностью требует либо решения систем полиномиальных уравнений, либо использования специальных таблиц подстановки, что делает процесс встраивания более ресурсоемким и длительным.

В связи с вышеизложенным, предлагается адаптивный метод матричного встраивания. Он заключается в том, что в зависимости от статистических характеристик контейнера и сообщения выбирается

оптимальная комбинация из нескольких кодов так, чтобы обеспечить наименьшие изменения контейнера. Такой подход позволяет использовать преимущества матричного встраивания, которые заключаются в высокой результативности встраивания и при этом максимально увеличивает возможный объем контейнера. Рассмотрим этот подход подробнее.

Адаптивное матричное встраивание разделяет исходный контейнер на фрагменты, которые, в свою очередь, состоят из одного или нескольких блоков 8x8 коэффициентов дискретного косинусного преобразования (ДКП). В пределах каждого такого фрагмента осуществляется оптимальный выбор используемого кода или нескольких кодов одновременно. После принятия решения о выборе параметров кода (кодов) происходит внедрение данных, а информация о параметрах сохраняется. Далее выбирается следующий фрагмент и встраивание продолжается. Сохранение информации осуществляется одним из двух путей: параметры кода преобразуются в стеганографический ключ, либо, если ключ не используется, в каждый блок коэффициентов ДКП вносится специальная метка, указывающая на параметры. При использовании метки учитывается параметр, называемый собственной четностью блока, при этом модификации подвергается один коэффициент ДКП из тех, которые не используются для передачи информации (коэффициенты со значением 0 или 1).

Особенностью данного метода является то, что он обладает высокой результативностью, рационально использует объем объекта-контейнера и, кроме того, обеспечивает минимальное действительное число изменений изображения. Параметрами данного метода является максимальное число блоков коэффициентов ДКП, объединяемых в фрагмент, а также вид и количество кодов, на основе которых функционирует адаптивный метод и которые образуют результирующий код, называемый адаптивным. При адаптивном матричном встраивании в пределах каждого фрагмента из одного или более блоков 8x8 коэффициентов ДКП несколько раз производится предварительное внедрение фрагмента сообщения с использованием различных комбинаций, применяемых для встраивания кодов, и вычисляется число изменений за фрагмент. После рассмотрения всех комбинаций кодов определяется такой их набор, который позволяет внедрить часть сообщения, внося наименьшее количество искажений в исходный объект. Далее фрагмент модифицируется выбранными кодами или кодом (если принято решение использовать только один код). В результате к контейнеру в целом для матричного встраивания применяется композиция из различных помехоустойчивых кодов, а не один код. Другими словами, при применении адаптивного матричного встраивания каждое решение на использование одного из используемых кодов приводит к тому, что в контейнер внедряется часть сообщения с результативностью, соответствующей данному коду, и для такого внедрения используется число бит контейнера, соответствующее длине кодовой комбинации данного кода. Результативность и покрывающая способность результирующего кода вычисляются как их среднее значение.

Чтобы оценить преимущества адаптивного метода матричного встраивания, введем дополнительную характеристику для каждого кода. Эта характеристика называется вероятностью сохранения кодовой комбинации и обозначается p_c . Величина p_c показывает, с какой вероятностью не будет изменена кодовая комбинация при матричном встраивании, когда используется данный код и обеспечивается результативность, с которой данный код производит внедрение информации. Вероятность сохранения кодовой комбинации рассчитывается как $p_c = \frac{k}{n \cdot e}$, где n , k – параметры кода, e – результативность помехоустойчивого кода.

Рассмотрим результативность e предложенного адаптивного метода. Пусть для встраивания используются i кодов $c_1, c_2, \dots, c_i$. Тогда результативность адаптивного метода, использующего эти коды, вычисляется как

$$e_a = \frac{\sum_{i=1}^n \eta_i \cdot e_i}{\sum_{i=1}^n \eta_i},$$

где η_i – количество решений использования данного кода, n – число кодов. Покрывающая способность адаптивного кода вычисляется аналогично:

$$\alpha_a = \frac{\sum_{i=1}^n \eta_i \cdot \alpha_i}{\sum_{i=1}^n \eta_i}.$$

В среднем, такой код имеет результативность внедрения $e_a = \frac{\sum_{i=1}^n e_i}{n}$ и покрывающую способность

$$\alpha_a = \frac{\sum_{i=1}^n \alpha_i}{n}.$$

Лучший и худший варианты определяются границами кода с лучшей (худшей) результативностью e_i и лучшей (худшей) покрывающей способностью α_i соответственно. Таким образом, для адаптивного метода матричного встраивания возможно зафиксировать значение α_a в определенных пределах, и, выбирая различные параметры и количество используемых кодов, получить необходимое улучшенное среднее значение результативности e_a . Под выбором комбинаций кодов подразумевается оптимальный их выбор с целью обеспечения требуемых параметров внедрения. Так, возможен выбор кодов для обеспечения зафиксированных значений результативности или эффективности, а так же для повышения одного из этих параметров. При моделировании адаптивного матричного встраивания и оценке результатов использовался способ перебора различных комбинаций кодов. Такой способ ориентирован на увеличение результативности матричного встраивания, но трудоемок при выборе большого числа кодов и числа блоков, объединяемых во фраг-

мент. Так же, предположительно, возможны варианты выбора кодов с учетом вероятностно-статистических характеристик контейнера и сообщения.

Крайне важным является то, что применение описанного метода матричного встраивания приводит к возникновению дополнительного преимущества. Его суть заключается в том, что неадаптивное матричное встраивание (n, k, t) -кодом осуществляется путем изменения от 0 до $t < k \ll n$ бит исходного контейнера в n битах, для передачи k бит сообщения. Код с корректирующей способностью $t = 1$, такой как код Хемминга, при одной операции встраивания изменяет 0 или 1 бит контейнера. Изменений не происходит, если фрагмент контейнера уже имеет синдром, соответствующий фрагменту сообщения. Аналогично для кодов с $t > 1$.

Пример 1. Для кода Хемминга (7, 3, 1) среднее число замен равно $0 \cdot \frac{1}{2^p} + 1 \cdot (1 - \frac{1}{2^p}) = \frac{7}{8}$, а эффективность внедрения $e_p = \frac{3}{1 - 2^{-p}} = \frac{3}{7/8} = 3,429 > 3$, где p – число бит в сообщении, $p = n - k$.

Для адаптивного метода результативность кода определяется действительной результативностью. Действительная результативность адаптивного кода находится как

$$e_d = e_a + e_a \cdot p_d, \quad (1)$$

p_d при этом определяется как сумма вероятностей нескольких независимых событий, которые заключаются в том, что каждая используемая кодовая комбинация не изменится. Для двух кодов $p_d = p_1 + p_2(1 - p_1) - (p_1 + p_2)/2$ или $p_d = 1 - (1 - p_1)(1 - p_2) - (p_1 + p_2)/2$. В общем случае

$$p_d = 1 - \prod_{i=1}^n (1 - p_i) - \sum_{i=1}^n p_i / n. \quad (2)$$

Адаптивное матричное встраивание значительно увеличивает вероятность того, что кодовая комбинация не изменится, а данные будут внедрены, что приводит к минимальному изменению в исходном контейнере. Из-за этого на основании (1) и (2) можно сделать вывод о том, что адаптивный метод матричного встраивания обладает повышенной действительной результативностью, величина прироста составляет $e_a \cdot p_d = e_a \cdot (1 - \prod_{i=1}^n (1 - p_i) - \sum_{i=1}^n p_i / n)$, а действительная результативность находится как

$$e_d = \frac{\sum_{i=1}^n e_i}{n} \cdot (1 + (1 - \prod_{i=1}^n (1 - p_i) - \sum_{i=1}^n p_i / n)), \quad (3)$$

где n – число кодов, p_i – вероятность сохранения комбинации i -го кода.

Пример 2. Рассмотрим код Хемминга (3,2) и код БЧХ (15,10). Для кода Хемминга (3,2) результативность $e_1 = 2,667$, $\alpha_1 = 0,667$, $p_c = \frac{k_1}{n_1 \cdot e_1} = 0,25$. Для

кода БЧХ (15,10) результативность $e_2 = 4,5$, $\alpha_2 = 0,667$, $p_c = \frac{k_2}{n_2 \cdot e_2} = 0,25$. Рассчитаем параметры

адаптивного кода, основанного на рассмотренных кодах. $e_a = \frac{\sum_{i=1}^2 e_i}{2} = 3,583$, $\alpha_a = \frac{\sum_{i=1}^2 \alpha_i}{2} = 0,667$. Далее получаем e_d согласно (3) как

$$e_d = e_a \cdot (1 + (1 - \prod_{i=1}^2 (1 - p_i) - \sum_{i=1}^2 p_i / n)) = 3,58 \cdot (1 + (1 - 0,638 - 0,199)) = 4,15.$$

Для лучшего случая e_d составит $e_d = 5,22$. Увеличение числа кодов приводит к дальнейшему увеличению действительной результативности внедрения информации. При этом выбор разных параметров кодов позволяет зафиксировать значение результативности или покрывающей способности в требуемых пределах. Так, в примере с кодами Хемминга (3,2) и БЧХ (15,10) покрывающая способность адаптивного кода не изменилась, а результативность увеличилась в среднем и лучшем случаях по сравнению с каждым из кодов. Для трех кодов код Хемминга (3,7) код БЧХ (15,8) и троичного кода Хемминга (4,2) с близкими показателями результативности средняя результативность адаптивного кода составит $e_d = 3,54$, а действительная $e_d = e_a \cdot 1,22 = 4,31$. В худшем и лучшем случае эта величина примет значения 4,18 и 4,45 соответственно. Покрывающая способность этого кода составит $\alpha_a = 0,584$. Сравнив значения e_d и α_a с исходными, заключим, что результативность кода выросла на величину от 18,0 до 25,7 %.

Метод адаптивного матричного встраивания. Внедрение данных методом адаптивного встраивания производится в несколько этапов:

1. Выбор изображения, формирование покрытия и секретного сообщения, выбор параметров кода.

2. Анализ контейнера и покрытия, выбор лучших параметров для встраивания и внедрение одним из кодов так, чтобы обеспечить наибольшую результативность.

3. Сохранение параметров в стеганографический ключ или в контейнер (сохранение параметров в контейнер осуществляется методом собственной четности) и передача стеганографического контейнера по открытому каналу получателю.

Данные для сравнения покрывающей способности кодов представлены в таблице. На рисунке показано сравнение результативности внедрения для неадаптивных и адаптивных кодов (средний случай). Из анализа данных, представленных на рисунке и в таблице, следует, что для рассмотренных помехоустойчивых кодов, при их совместном применении удается достичь повышения e_d до значения 5 и более

бит передаваемой информации на одно изменение контейнера. Полученное значение превосходит рассмотренные неадаптивные коды и является достаточно высоким. Это означает, что объем контейнера расходуется рационально. Утверждение, состоящее в

том, что адаптивные коды обладают высокой результативностью, обеспечивая при этом большую покрывающую способность, верно и для других комбинаций помехоустойчивых кодов.

Таблица

Покрывающая способность кодов

Код	<i>n</i>	<i>k</i>	<i>e</i>	<i>α</i>	<i>P_c</i>
адаптивный код 1					
Код БЧХ 15-8	15	8	3,65	0,5333	0,146
Код БЧХ 15-10	15	10	4,5	0,6667	0,148
<u><i>e_d</i> = 4,5859724</u>	1,125392		4,075	<u>0,6</u>	0,147
адаптивный код 2					
Код Хемминга 4-2(<i>q</i> =3)	4	2	3,55	0,79	0,14
Код БЧХ 15-8	15	8	3,65	0,5333	0,146
Код БЧХ 15-10	15	10	4,5	0,6667	0,148
<u><i>e_d</i> = 4,795402768</u>	1,22959		3,9	<u>0,663333</u>	0,144667
адаптивный код 3					
Код Хемминга 3-2	3	1	2,67	0,6667	0,25
Код БЧХ 15-8	15	8	3,65	0,5333	0,146
Код БЧХ 15-10	15	10	4,5	0,6667	0,148
<u><i>e_d</i> = 4,591144804</u>	1,272961		3,606667	<u>0,622222</u>	0,181333
адаптивный код 4					
Код Хемминга 7-3	7	3	3,43	0,4286	0,125
Код Хемминга 4-2(<i>q</i> =3)	4	2	3,55	0,79	0,14
Код БЧХ 15-8	15	8	3,65	0,5333	0,146
Код БЧХ 15-10	15	10	4,5	0,6667	0,148
<u><i>e_d</i> = 4,965382237</u>	1,312725		3,7825	<u>0,604643</u>	0,13975
адаптивный код 5					
Код Хемминга 13-3(<i>q</i> =3)	13	3	4,92	0,3623	0,05
Код Хемминга 7-3	7	3	3,43	0,4286	0,125
Код Хемминга 4-2(<i>q</i> =3)	4	2	3,55	0,79	0,14
Код БЧХ 15-8	15	8	3,65	0,5333	0,146
Код БЧХ 15-10	15	10	4,5	0,6667	0,148
<u><i>e_d</i> = 5,445785436</u>	1,358051		4,01	<u>0,556176</u>	0,1218
адаптивный код 6					
Код Хемминга 3-2	3	1	2,67	0,6667	0,25
Код Хемминга 13-3(<i>q</i> =3)	13	3	4,92	0,3623	0,05
Код Хемминга 7-3	7	3	3,43	0,4286	0,125
Код Хемминга 4-2(<i>q</i> =3)	4	2	3,55	0,79	0,14
Код БЧХ 15-8	15	8	3,65	0,5333	0,146
Код БЧХ 15-10	15	10	4,5	0,6667	0,148
<u><i>e_d</i> = 5,553986385</u>	1,466722		3,786667	<u>0,574591</u>	0,143167

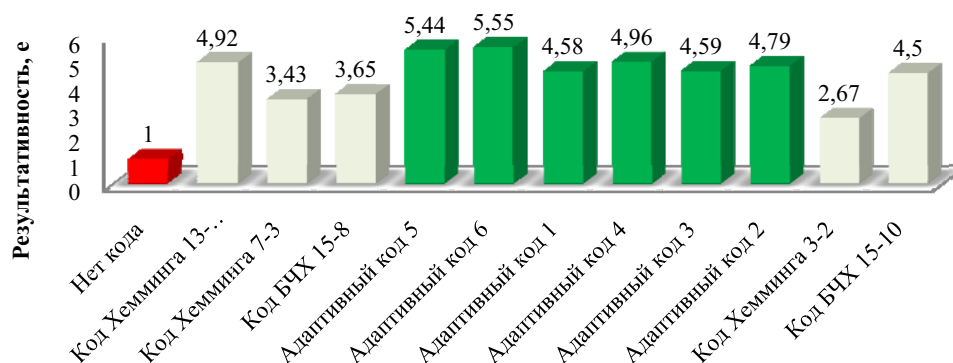


Рисунок. Результаты внедрения для неадаптивных и адаптивных кодов (средний случай)

Выводы

Итак, при оценке результативности предложенного метода и выбора параметров кодов были выявлены преимущества адаптивного матричного встраивания с точки зрения влияния на статистическую структуру контейнера. При этом было экспериментально подтверждено, что матричное встраивание не изменяет распределение вероятностей коэффициентов ДКП изображений. На основе простых помехоустойчивых кодов (Хемминга, БЧХ) были построены адаптивные коды для матричного встраивания стеганографических сообщений, проведено сравнение характеристик имеющихся и предложенных кодов, а также подтверждены предположения о том, что адаптивные коды имеют увеличенную результативность.

Разработано программное обеспечение, реализующее метод адаптивного матричного встраивания данных в изображения JPEG. Кроме того, осуществлен стегоанализ контейнеров по критериям: отношение сигнал-шум, критерий Хи-квадрат, относительная энтропия и др. По результатам исследований при отсутствии оригинального изображения вероятность ошибки I рода для системы стеганографического анализа составила 42 %, II рода – 22,3 % для внедряемого сообщения длиной 64 бита, что выше для

данных критериев по сравнению с неадаптивными методами.

Литература

1. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. М., 2016. 262 с.
2. Зеленкин С.Е. Модель робастной стеганографической системы // Наукоемкие технологии. М., 2010. №9. С. 84–86.
3. Fridrich J. *Steganography in Digital Media. Principles, Algorithms and Applications*. Cambridge Univ P, 2010. 462 с.
4. Winkler A. "Advances in Syndrome Coding Based on Stochastic and Deterministic Matrices for Steganography", pp. 25, 37–55, 67–77, Dissertation, Dresden 2012.

References

1. Gribunin V.G., Okov I.N., Turincev I.V. *Cifrovaia steganografiia* [Digital steganography]. Moscow, 2016, 262 p.
2. Zelenkin S.E. Model' robastnoi steganograficheskoi sistemy [Model of robust steganographic systems]. Naukoemkie tehnologii [High Tech]. Moscow, 2010, №9, pp. 84–86.
3. Fridrich J. *Steganography in Digital Media. Principles, Algorithms and Applications*. Cambridge Univ P, 2010, 462 p.
4. Winkler A. "Advances in Syndrome Coding Based on Stochastic and Deterministic Matrices for Steganography", pp. 25, 37–55, 67–77, Dissertation, Dresden, 2012.

УДК 004.023

З.К. Кабаков

Череповецкий государственный университет,

Д.В. Храмешин

Выксунский металлургический завод

ИСПОЛЬЗОВАНИЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ ОПРЕДЕЛЕНИЯ КЛЮЧЕВЫХ ПАРАМЕТРОВ В ЗАДАЧЕ СТАТИСТИЧЕСКОГО КОНТРОЛЯ МНОГОПАРАМЕТРИЧЕСКОГО ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА

Статья посвящена использованию эвристического алгоритма для быстрого выбора ключевых технологических параметров, наиболее полно определяющих качество поверхности горячего рулонного проката в условиях сильной скоррелированности исходных данных.

Генетический алгоритм, статистический контроль процессов, регрессионный анализ, эвристика, мультиколлинеарность.

The article deals with usage of heuristic algorithm, which makes swift choice of most valuable technological parameters within initial set. It evaluates synergy score of metallurgical parameters affecting surface defect known as 'skin' and compiles set of crucial variables for the needs of Statistical control process.