Математические структуры и моделирование 2020. № 4(56). С. 103–111

УДК 004.932.2
DOI 10.24147/2222-8772.2020.4.103-111

СТЕГОАНАЛИЗ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ С НИЗКИМ ЗАПОЛНЕНИЕМ СТЕГОКОНТЕЙНЕРА С ИСПОЛЬЗОВАНИЕМ ПРОГРАММНОГО КОМПЛЕКСА

А.К. Гуц
д.ф.-м.н., профессор, e-mail: aguts@mail.ru
Д.Э. Вильховский
ассистент кафедры информационной безопасности, e-mail: vilkhovskiy@gmail.com

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. В статье представлен программный комплекс, использование которого позволяет проводить стегоанализ цветных изображений на предмет выявления и извлечения скрытых сообщений в виде LSB-вставок, а также вставок, выполненных методом Коха–Жао. Комплекс объединяет ранее разработанные авторами алгоритмы и эффективен даже при малом заполнении стегоконтейнера (10–30 %). Приводятся результаты компьютерного эксперимента и сравнение эффективности предложенного комплекса с эффективностью других моделей

Ключевые слова: стегоанализ, стеганографический анализ, анализ стегоконтейнера, обнаружение LSB-вставки, обнаружение DCT-вставки, обнаружение вставки Коха–Жао.

Введение

Несмотря на развитие стеганографии и появление новых методов и алгоритмов методы замены наименее значащих бит (LSB), работающие с пространственной областью изображений, и метод Коха–Жао с использованием дискретного косинусного преобразования (DCT), работающий с частотной областью, до сих пор активно применяются для сокрытия данных. Следовательно, применение методов стегоанализа, направленных на обнаружение этих видов вставок, является по-прежнему актуальным.

Более того, злоумышленники, стремясь обойти существующие алгоритмы стегоатак, уменьшают объём встраиваемого сообщения, т. е. используют меньшее заполнение стегоконтейнера, что повышает уровень незаметности вставок. Следовательно, актуальной является разработка таких методов стегоанализа, которые могли бы обнаруживать изображения со встроенными сообщениями даже при низкоуровневой нагрузке.

Для решения данной задачи стеганографического анализа авторами разработан специальный программный комплекс на основе алгоритмов, обладающих высокой чувствительностью к стеговставкам при работе с цветными изображениями с малым заполнением стегоконтейнера. Объединение данных алгоритмов

в комплекс позволяет эффективно решать проблему обнаружения LSB-вставок и вставок, выполненных методом Коха-Жао.

С целью унификации трактовок в таблице 1 приведён ряд используемых в статье сокращений, заимствованных из англоязычных исследований и укоренившихся в отечественной терминологии.

Таблица 1. Перечень терминов, используемых в статье

Термин	Расшифровка на английском языке	Значение на русском языке
DCT	Discrete Cosine Transform	Дискретное косинусное преобразование
LSB	Least Significant Bit	Наименее значащий бит
TP	True positive	Истинно положительный
TN	True negative	Истинно отрицательный
FP	False positive	Ложно-положительный
FN	False negative	Ложно-отрицательный

1. Общая характеристика разработанного программного комплекса

Представляемый программный комплекс – это программное обеспечение, разработанное в виде веб-приложения с микросервисной архитектурой. Взаимодействие между клиентом и сервером осуществляется через GET/POST-запросы. Общая схема работы программного комплекса представлена на рисунке 1.

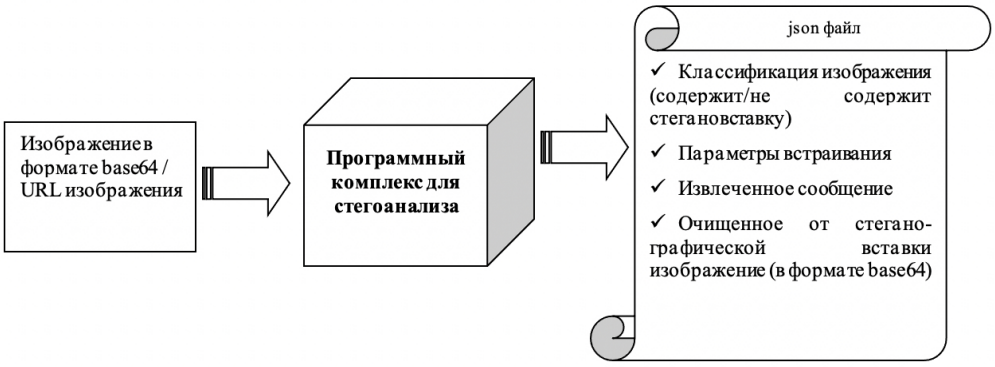

Рис. 1. Общая схема работы программного комплекса

Таким образом, для проведения стегоатаки можно либо загрузить анализируемое изображение непосредственно в программу, либо указать адрес его

размещения в сети. В результате программа возвращает не только констатацию факта наличия или отсутствия встроенного сообщения, но и параметры встраивания, а также извлечённое сообщение и изображение, очищенное от вставки.

Программный комплекс легко интегрируется с системой документооборота организации. Интеграция реализуется через хуки приложения, вызов которых осуществляется при загрузке файла на сервер или при добавлении или обновлении таблиц базы данных.

2. Основные компоненты программного комплекса

2.1. Первая компонента – алгоритм обнаружения стеговставок, выполненных методом LSB-замены

Алгоритм позволяет выявить скрытые сообщения, встраивание которых выполнено методом замены наименее значащих битов (LSB), при низком уровне заполнения стегоконтейнера минимальная величина встраиваемой нагрузки, при которой данный алгоритм всё ещё показывает высокую эффективность составляет 10 %. Основан на анализе младших слоев цветного цифрового изображения с применением алгоритма таксономии и метода анализа иерархий.

Использование метода анализа иерархий, описанного в работе [3], позволило существенно усовершенствовать модели, предложенные в работах [9, 10], и представить алгоритм в новой, более эффективной модификации.

Для анализа изображения на предмет наличия LSB-вставок и определения области встраивания проводится иерархический анализ трёх младших слоев — нулевого и двух вышележащих слоев. Пиксели каждого из трёх слоев анализируются на основе предложенных критериев и весовых коэффициентов, представленных в таблице 2. Иерархическая взаимосвязь младших битов, а также младших битов и битов более высокого порядка и её особенности позволяют классифицировать изображение как чистое или имеющее вставки.

Таблица 2. Критерии принятия решения и их весовые коэффициенты

Критерии	Весовые коэффициенты
Сравнение значения бита и значений соседних по сторонам битов	$r_1 = \dfrac{nk}{nk+k+1}$
Сравнение значения бита и значений соседних по углам битов	$r_2 = \dfrac{k}{nk+k+1}$
Сравнение значения бита и значений окружающих битов	$r_3 = \dfrac{1}{nk+k+1}$

Более того, на основе анализа поведения битов всех трёх слоев алгоритм способен не только сделать вывод о наличии или отсутствии замены младших битов, но и сформировать список пикселей со встроенным сообщением, что позволяет не только классифицировать изображение как стегоконтейнер, но и

определить область встраивания с высокой степенью точности.

Далее предложенный алгоритм был усовершенствован для возможности его применения в областях градиентной заливки, что достигается в том числе предварительной обработкой изображения.

Более детальное описание алгоритма представлено в работе [1].

2.2. Вторая компонента — алгоритм обнаружения стеговставок, выполненных методом Коха–Жао

Алгоритм позволяет выявлять стеганографические вставки, встраивание которых выполнено методом Коха–Жао, описанным в работе [7]. Для выявления блоков встраивания алгоритм использует анализ среднечастотных коэффициентов дискретного косинусного преобразования (далее DCT-коэффициентов) в трёх видах последовательности их величин. Необходимым условием успешного анализ является определение правильного порогового значения.

Анализ последовательности величин DCT-коэффициентов основан на том факте, что встраивание изменяет как минимум одну из последовательностей, что на гистограмме изображения будет иметь вид некой ступени, имеющей определённую высоту, равную величине порога.

В данную компоненту программного комплекса встроен модуль поиска ступенчатых изменений, что позволяет последовательно проходить по всем блокам изображения в поисках блоков встраивания.

В результате анализа последовательности и численного дифференцирования с использованием разностных схем для определения границ ступеней гистограммы формируются матрицы данных коэффициентов дискретного косинусного преобразования.

Таким образом, выявление факта изменения одной из последовательностей означает, что исследуемое изображение содержит встроенное сообщение. Более того, высокие пики на гистограмме зависимостей позволяют достоверно определить границы этого сообщения, а также осуществить его извлечение с минимальными потерями.

Более детальное описание алгоритма представлено в работе [2].

3. Эффективность программного комплекса

Эффективность предложенного программного комплекса тестировалась при помощи компьютерного эксперимента с использованием базы изображений BossBase и базы изображений праздников INRIA. Выбор данных баз изображений обусловлен, в первую очередь, популярностью их использования в других отечественных и зарубежных исследованиях, что позволило в дальнейшем получить сопоставимые данные по чувствительности предложенного программного комплекса и некоторых других моделей стегоанализа.

Тестирование проводилось с различными уровнями нагрузки (уровнями заполнения стегоконтейнера): 30 %, 25 %, 20 %, 15 %, и 10 %. Встраивание производилось методом LSB-замены и методом Коха–Жао (далее — DCT-вставки).

Встраивание методом LSB-замены осуществлялось в синюю компоненту. Результаты проведённого эксперимента представлены в таблицах 3 и 4.

Таблица 3. Результаты стегоатак, проведённых при помощи программного комплекса с использованием базы изображений BossBase, %

Тестируемая нагрузка	Обнаружение LSB-вставок				Обнаружение DCT-вставок			
	TP	FN	TN	FP	TP	FN	TN	FP
30 %	91,15	8,85	79,12	20,88	84,73	15,27	76,53	23,47
25 %	87,91	12,09			79,11	20,89		
20 %	79,87	20,13			74,64	25,36		
15 %	72,31	27,69			68,12	31,88		
10 %	63,12	36,88			63,88	36,12		

В самом общем виде можно сделать вывод, что оба алгоритма показывают лучшую чувствительность при работе с изображениями, содержащими скрытые сообщения, чем с чистыми изображениями. Однако полученные величины погрешностей при классификации чистых изображений находятся в пределах допустимых значений.

Таблица 4. Результаты стегоатак, проведённых при помощи программного комплекса с использованием базы изображений INRIA, %

Тестируемая нагрузка	Обнаружение LSB-вставок				Обнаружение DCT-вставок			
	TP	FN	TN	FP	TP	FN	TN	FP
30 %	90,36	9,64	78,31	21,69	86,3	13,7	79,15	20,85
25 %	86,89	13,11			80,26	19,74		
20 %	79,6	20,4			76,01	23,99		
15 %	72,03	27,97			69,94	30,06		
10 %	62,56	37,44			65,04	34,96		

Сводные данные по эффективности каждого из алгоритмов программного комплекса представлены в таблице 5.

Таблица 5. Эффективность программного комплекса, %

Тестируемая нагрузка	BossBase		INRIA	
	Обнаружение LSB-вставок	Обнаружение DCT-вставок	Обнаружение LSB-вставок	Обнаружение DCT-вставок
30 %	91,15	84,73	90,36	86,3
25 %	87,91	79,11	86,89	80,26
20 %	79,87	74,64	79,6	76,01
15 %	72,31	68,12	72,03	69,94

10 %	63,12	63,88	62,56	65,04

Данные, полученные в результате проведённого компьютерного эксперимента, свидетельствуют о высокой результативности предложенного программного комплекса. Ожидаемо, максимальная эффективность достигается при более высоких уровнях заполнения стегоконтейнера и снижается по мере уменьшения размера встраиваемого сообщения: 91,15 % и 86,3 % против 63,12 % и 65,04 % для LSB-вставок и DCT-вставок соответственно.

Также отмечается, что алгоритмы, используемые в программном комплексе, показывают некоторую чувствительность к базам изображений. Так, программный комплекс показывает лучшую результативность в обнаружении вставок, выполненных методом LSB-замены, при работе с изображениями из BossBase, по сравнению с изображениями из базы INRIA, тогда как алгоритм выявления сообщений, скрытых методом Коха–Жао, лучше работает с базой изображений праздников INRIA, чем с BossBase. Возможно, последний факт связан с тем, что база изображений праздников INRIA содержит изображения в формате JPEG, что соответствует принципам DCT, на основе которых работает алгоритм обнаружения вставок Коха–Жао, в результате чего наблюдается снижение ошибок классификации.

Для получения более полной картины о преимуществах и целесообразности использования предлагаемого нами программного комплекса для обнаружения LSB-вставок, сравним его результативность с результативностью некоторых методов, предложенных в исследованиях других авторов. В качестве базы сравнения нами рассмотрены ансамблевый SW-стегоанализ [4, 5], SVM-PSO-стегоанализ JPEG изображений [8, 11] и компактный метод стегоанализа [6]. Выбор данных исследований продиктован наличием чётких количественных выводов об эффективности обнаружения стеганографических вставок при низком уровне заполнения стегоконтейнера (10 или 25 %), что в полной мере соответствует объёму встраиваемой нагрузки, таргетируемой представленным программным комплексом. Сравнительная эффективность обнаружения представлена в таблице 6, сравнение происходит по наилучшим значениям работы программного комплекса.

Таблица 6. Сравнительная эффективность обнаружения

Встраиваемая нагрузка	Предлагаемый программный комплекс	Ансамблевый SW-стегоанализ	SVM-PSO стегоанализ JPEG изображений	Компактный метод стегоанализа
25%	87,91	86,77	–	74,53
10%	65,04	–	51,07	–

На основании данных, представленных в таблице 6, можно сделать вывод, что алгоритмы, реализованные в разработанном программном комплексе, по

своей эффективности обладают существенным превосходством по сравнению с другими рассматриваемыми алгоритмами. При этом, если при 25 % заполнении стегоконтейнера преимущество представленных нами алгоритмов по сравнению с моделями, использующими машинное обучения невелико и составляет чуть более 1 % (хотя и показывает лучшие значения по сравнению с компактным методом), то при 10 % заполнении стегоконтейнера алгоритмы предложенного программного комплекса выдают значительную разницу в результативности (65,04 % против 51,07 %)

Заключение

На основе разработанных алгоритмов обнаружения LSB-вставок с использованием метода анализа иерархий и таксономии, а также алгоритма обнаружения вставок, выполненных методом Коха–Жао, использующего дискретное косинусное преобразование, был создан программный комплекс. Тестирование его на двух базах показало, что данный комплекс позволяет обнаруживать (в наилучших значениях) до 91,15 % изображений, содержащих LSB-вставки и до 86,3 % изображений, содержащих вставки Коха–Жао. Ошибки классификации стего-изображений составляют, соответственно, 8,85 % и 13,7 %, в то время как ошибки классификации чистых изображений составляют, соответственно, 20,88 % и 20,85 % (в наименьших значениях), что в целом является допустимым уровнем.

Таким образом, делается вывод о высокой эффективности и целесообразности применения данного программного комплекса при проведении стеганографического анализа цветных изображений с малым заполнением стегоконтейнера (малой нагрузкой).

Литература

1. Белим С.В., Вильховский Д.Э. Алгоритм выявления стеганографических вставок типа LSB-замещения на основе анализа слоя младших битов // Информатика и системы управления. 2017. № 4 (54). С. 3—11.
2. Белим С.В., Вильховский Д.Э. Стеганоанализ алгоритма Коха–Жао // Математические структуры и моделирование. 2018. № 4(48). С. 113—119.
3. Amritha P.P., Sreedivya-Muraleedharan M., Rajeev K., Sethumadhavan M. Steganalysis of LSB Using Energy Function // Advances in Intelligent Systems and Computing. 2016. No. 384. P. 549–558.
4. Chaeikar A. Ensemble SW image steganalysis: A low dimension method for LSBR detection // Signal Process Image Commun. 2019. No. 70. P. 233–245.
5. Chaeikar S.S., Ahmadi A. SW: A blind LSBR image steganalysis technique // In Proceedings of the 10th International Conference on Computer Modeling and Simulation, Sydney Australia, 8 January 2018. P. 14-–18.
6. Juarez-Sandoval O., Cedillo-Hernandez M., Sanchez-Perez G. et all. Compact Image Steganalysis for LSB-Matching Steganography // 5th International Workshop on Biometrics and Forensics (IWBF). 2017. P. 1–6.

7. Koch E. Towards robust and hidden image copyright labeling // IEEE Workshop on Nonlinear Signal and Image Processing. 1995; P. 452—455.
8. Kumar U.P., Shankar D.D. Blind Steganalysis for JPEG Image susing SVM and SVM-PSO Classifiers // International Journal of Innovative Technology and Exploring Engineering (IJITEE). 2019, V. 8; P. 1239—1246.
9. Ojala T., Pietikainen M., Harwood D. A Comparative Study of Texture Measures with Classification Based on Feature Distributions // Pattern Recognition, 1996. No. 29. P. 51—59.
10. Saaty T.L. Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors // The Analytic Hierarchy/Network Process. Review of the Royal Spanish Academy of Sciences. Series A, Mathematics. 2008. No. 102(2). P. 251-–318.
11. Shankar D.D., Azhakath A.S. Minor blind feature based Steganalysis for calibrated JPEG images with cross validation and classification using SVM and SVM-PSO // Multimedia Tools and Applications. 2020. DOI: 10.1007/s11042-020-09820-7.

STEGANALYSIS OF LOW STEGO-PAYLOAD COLOR IMAGES USING SOFTWARE COMPLEX

A.K. Guts
Dr.Sc. (Phys.-Math.), Professor, e-mail: aguts@mail.ru
D.E. Vilkhovskiy
Assistant of the Department of Information Security, e-mail: vilkhovskiy@gmail.com

Dostoevsky Omsk State University, Omsk, Russia

Abstract. In the paper, a new software package is presented as a tool that allows to conduct steganalysis of color images and reveal embedded messages such as LSB inserts or Koha-Zhao inserts. The complex combines two algorithms previously developed by the authors and is highly effective even for small stegocontainer payload (10–30 %). The computer experiment results and a comparison of the software package and some other models performance are presented and discussed.".

Keywords: steganalysis, steganographic analysis, stegocontainer analysis, LSB-insert detection, DCT-insert detection, Koch-Zhao insert detection..

References

1. Belim S.V. and Vil'khovskii D.E. Algoritm vyyavleniya steganograficheskikh vstavok tipa LSB-zameshcheniya na osnove analiza sloya mladshikh bitov. Informatika i sistemy upravleniya, 2017, no. 4(54), pp. 3—11. (in Russian)
2. Belim S.V., and Vil'khovskii D.E. Steganoanaliz algoritma Kokha–Zhao. Matematicheskie struktury i modelirovanie, 2018, no. 4(48), pp. 113—119. (in Russian)
3. Amritha P.P., Sreedivya-Muraleedharan M., Rajeev K., and Sethumadhavan M. Steganalysis of LSB Using Energy Function. Advances in Intelligent Systems and Computing, 2016, no. 384, pp. 549–558.

4. Chaeikar A. Ensemble SW image steganalysis: A low dimension method for LSBR detection. Signal Process Image Commun., 2019, no. 70, pp. 233–245.
5. Chaeikar S.S. and Ahmadi A. SW: A blind LSBR image steganalysis technique. In Proceedings of the 10th International Conference on Computer Modeling and Simulation, Sydney Australia, 8 January 2018, pp. 14-–18.
6. Juarez-Sandoval O., Cedillo-Hernandez M., Sanchez-Perez G. et all. Compact Image Steganalysis for LSB-Matching Steganography. 5th International Workshop on Biometrics and Forensics (IWBF), 2017, pp. 1–6.
7. Koch E. Towards robust and hidden image copyright labeling. IEEE Workshop on Nonlinear Signal and Image Processing, 1995, pp. 452—455.
8. Kumar U.P. and Shankar D.D. Blind Steganalysis for JPEG Image susing SVM and SVM-PSO Classifiers. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019, vol. 8, pp. 1239—1246.
9. Ojala T., Pietikainen M., and Harwood D. A Comparative Study of Texture Measures with Classification Based on Feature Distributions, Pattern Recognition, 1996, no. 29, pp. 51—59.
10. Saaty T.L. Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors. The Analytic Hierarchy/Network Process, Review of the Royal Spanish Academy of Sciences, Series A, Mathematics, 2008, no. 102(2), pp. 251-–318.
11. Shankar D.D. and Azhakath A.S. Minor blind feature based Steganalysis for calibrated JPEG images with cross validation and classification using SVM and SVM-PSO. Multimedia Tools and Applications, 2020. DOI: 10.1007/s11042-020-09820-7.

Дата поступления в редакцию: 23.11.2020