АЛГОРИТЪМ ЗА КОЛОННА LSB МОДИФИКАЦИЯ С ПОСТАВЯНЕ НА МАРКЕР*

ХРИСТО И. ПАРАСКЕВОВ, ПЛАМЕН А. СИМЕОНОВ

ALGORITHM FOR COLUMN LSB MODIFICATION WITH PLACING A TAG

HRISTO I. PARASKEVOV, PLAMEN A. SIMEONOV

ABSTRACT: *Computer steganography is one of the areas of science steganologiya through which you can send and receive secret messages without having to cause suspicion on the watcher. The sender uses steganography key deployment secret message using a specific steganography algorithm in multimedia container. The recipient applied "reverse" steganography algorithm together with steganography key, extract the secret message.*

One of the most important requirements for stego system is resistant to attacks, including a security retrieving the secret message. This requires seeking solutions against active attacks.

KEYWORDS: *information security, steganography.*

I. Въведение

При анализа на методите в компютърната стеганография има два подхода. Първият е според използваните контейнери, както това е направено в [1, 2], а вторият – според метода за скриване на съобщенията [3]. Той е и най-предпочитан в стеганографската изследователската общност.

Според използваният принцип на вграждане на скритите съобщения, методите на компютърната стеганография в мултимедийни файлове се разделят на няколко основни класа:

- Методи в пространствената област (Spatial Domain) - LSB (Least Significant Bit) и BPCS (Bit Plane Complexity Segmentation) [4].
- Методи в честотната област (Transform Domain) - използват тригонометрични преобразования (discrete cosine transformation, DCT) или DWT (discrete wavelet transformation) [4].
- Стеганография с разпръснато вграждане (Spread Spectrum)–при тази техника скритото съобщение се разпръсква по цялия носещ файл [5].
- Статистически методи - Patchwork – кодиране в 2 отделни области от пиксели като се увеличава/намалява интензитета на осветеността им.
- Деформиращи методи - извършване на промени в псевдослучайно избрани пиксели от контейнера и се сравнява с оригиналния файл за разлики.
- Изграждащи методи - към съобщението се добавя голям по обем текст за прикриване на скритата комуникация.
- Хибридни методи - спрямо характеристиките на изображението се взема решение какви промени могат да бъдат направени.

* Настоящата статия е финансирана от Фонд „Научни изследвания” към Шуменския университет „Епископ Константин Преславски“ по проект № РД-08-119/2016 г.

Според използваният контейнер, методите на компютърната стеганография се разделят на три основни вида:

- в графични изображения;
- в аудиофайлове;
- във видеофайлове.

Най-често използвания мултимедиен контейнер в компютърната стеганография е графичния файл поради масовото му използване и сравнително по-лесната му обработка. Съществуват различни методи, алгоритми и подходи, използващи един или повече контейнери за вмъкване на съобщението. В [6] са дадени някои съображения при избор на графичен контейнер:

- да не се използват общоизвестни изображения (например като „Джоконда”);
- да не се използват изображения, които са получени при конвертиране от един в друг формат;
- да се използват изображения, създадени с фотоапарат или скенер, а не с графичен редактор;
- да не се прави избор на контейнер с голям размер;
- в контейнера да не е добавен по друг начин шум;
- отсъствие на плавни преходи и монотонни области;
- да се използват пъстри, „шарени” изображения;
- наличие на много пиксели, с цвят който слабо се различава от човешкото око, например зеления и жълтия.

II. Стегоатаки и устойчивост

Основната цел на стеганографията е да се постигне максимална незабележимост на промяната след вмъкването на скритото съобщение в носещия файл. При осъществяването на тази цел се оформят няколко задачи, някои от които са: какъв да е формата, големината и броя на носещите файлове, какъв стегоалгоритъм да се използва, по какъв начин да се увеличи надеждността на извличането на тайното съобщение при нарочна атака към стего файла и други.

Всеки опит да се открие, извлече и унищожи внедреното съобщение от 3-ти лица се нарича „стего атака“.

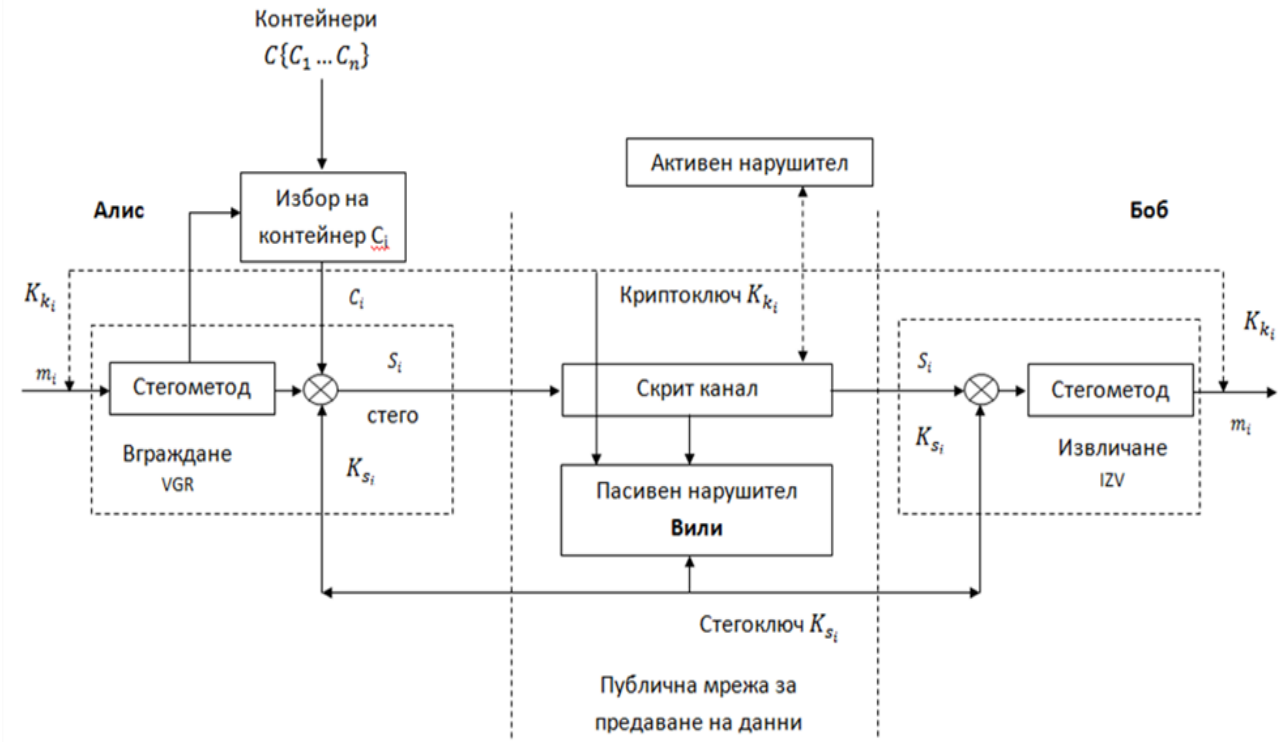
Предпазването на съобщението от атака, означава да се направи устойчиво на външна намеса, за която може да се предприеме противодействие за конкретни действия. Устойчивостта се изразява в това, дали след атака от 3-то лице ще може да се открие наличието на съобщението, дали може да се извлече информацията, дали може да се добави информация наречена „шум“, която да помогне за разобличаването съобщението или дали може да се повреди и унищожи съобщението за да се прекъсне контакта между подателя и получателя на съобщението.

Съгласно фиг. 1 [7] възможните атаки, които могат да се приложат към стего-комуникацията са пасивни и активни.

Към пасивните атаки се отнасят тези, които не правят корекции по графичната матрица от пиксели и имат за цел само да прослушват канала за връзка, да анализират в последствие трафика, но без намеса в него.

Към активните атаки се отнасят, тези които правят директно промени по графичната матрица от пиксели – завъртане по градус (rotate), завъртане огледално (flip) по хоризонталата или вертикалата на изображението, инвертиране на всички най-младши битове в изображението, изрязване на част от изображението (crop), добавяне на шум в изображението, добавяне на различни цветови филтри, изостряне на един или повече

цветове и други. С всички тези промени биха се очертали дефекти, които могат да разобличат скритото съобщение.



Фиг. 1 Модел на стегосистема

Атака Rotate – това е завъртане на изображението, което лесно може да се приложи с помощта на програма за разглеждане на снимки. При ротацията на 90, 180 или 270 градуса няма да се внесат никакви промени в стойностите на матрицата от пиксели на изображението, а само ще я запази завъртяна.

Атака Flip – може да се приложи в два варианта: завъртане по хоризонталата и завъртане по вертикалата с помощта на програма за разглеждане на изображения. При тази атака отново няма пряка намеса в стойностите на матрицата от пиксели на изображението, но отново скритото съобщение ще се изгуби, поради промяна в позициите на пикселите.

Атака Rotate and Flip – възможно е да се комбинират двете атаки с цел максимално затрудняване на извличането на съобщението.

Атака LSB Inversion – това е вид стерилизираща процедура, която обръща стойностите в най-малкия бит в противоположните им. Ако е 0 става 1, ако е 1 става 0. Това разваля съобщението, където и да се намира то в графичния файл.

Атака Crop – представлява изрязване на част от изображението. При този вид интервенция към стегофайла не ясно коя част от изображението ще остане и коя ще бъде изрязана. Какъвто и да е алгоритъма за вмъкване след Crop атака е доста трудно да се търси скритото съобщение.

За постигане на висока степен на устойчивост по отношение на извличане на скритото съобщение, може да се добави маркер в стегофайла, чрез който да се установи каква е била атаката върху него. Позицията на маркера и неговия размер се подбират спрямо атаките, на които е необходимо да се противодейства.

III. Колонна LSB модификация с поставяне на маркер

Колонна модификация на алгоритъма на LSB метода представлява корекция в четенето на матрицата от пиксели на изображението. Посоката на четене е колона по колона, а не ред по ред.

Идеята, е че тази промяна в посоката може да затрудни опитите на 3-ти лица за атака. Матрицата от пиксели не се променя, но редът на запис на съобщението се променя. Голяма е вероятността всеки софтуер да следва класическото четене и да търси или анализира грешно, ако сме вмъкнали съобщението с колонната модификация.

На база на предложената модификация е разработен алгоритъм, с цел устойчивост при извличане на скрито съобщение след атака Rotate.

Алгоритъмът може да работи с BMP и PNG контейнери, без ограничение в размера. Препоръчително е големината на носещият файл да не надвишава 500 KB.

Тъй като алгоритъмът няма блок за предварителна компресия максималния размер на информацията, която ще се вгради в изображението се определя в зависимост от размера на носещия файл минус хедърната информация делено на осем. Размерът на стегофайла трябва да е еднакъв с този на носещия файл.

IV. Експериментални резултати

За програмната реализация на предложените алгоритми е избран езика Python. За контейнери са използвани създадена база от данни с обем от 101 изображения във формат BMP и 101 изображения във формат PNG, изтеглени от Интернет с почти еднакъв размер, вариращ между 150-151KB. Освен това е използвана и предпочитаната при графична обработка база от изображения, достъпна от [8].

Генерирани са две съобщения е с размер 18816 и 1837 байта от LoremIpsum [9], наречени условно съответно „голямо” и „малко” съобщение. Голямото съобщение има за цел да запълни почти до край капацитета на контейнера на картинката и да се изследва новия файл за визуални дефекти и промени в цвета на изображението.

Внедряването на съобщението е извършено в нормалната позиция на всяко изображение. След приключването на вмъкването всички изображения са атакувани с rotate на 90, 180 или 270 градуса на случаен принцип. При извличането на секретното съобщение, всяко изображение е върнато в нормална позиция от 0 градуса и е успешно извлечено без загуби.

Създадени са нови стегофайлове в BMP и PNG формат, съдържащи вмъкнатото съобщение и заедно с оригиналните ще бъдат оценени резултатите от експеримента.

При оценяване на резултатите се отчитат: ефективност на вграждане, визуален, хистограмен и статистически анализ.

Ефективност на вграждане E_e е съотношението на размера на най-голямото съобщение, което може да се вгради в контейнера, към размера на контейнера [7].

$$E_e = \frac{V_{съоб.}}{V_к} \quad (1)$$

където:

$V_{съоб.}$-обем на скритото съобщение в KB или MB

V_k - обем на контейнера в KB или MB.

Коефициент на ефиктивност на вграждане

В таблица 1 е показана извадка от резултатите при експериментите с двете съобщения.

Таблица 1 Ефективност на вграждане

Ефективност на вграждане	
Large msg	Small msg
0,124075466	0,012149138
0,124865661	0,012226512
0,123930864	0,012134979
0,124865661	0,012226512
0,124484127	0,012189153
0,124811013	0,012221161
0,124870631	0,012226999
0,124669647	0,012207319
0,12441254	0,012182144
0,124865661	0,012226512
0,124669647	0,012207319
0,124513773	0,012192056
0,124513773	0,012192056
0,124669647	0,012207319
0,124607746	0,012201258
0,124865661	0,012226512
0,124129471	0,012154427
0,124225324	0,012163812
0,124484127	0,012189153
0,023933411	0,002343496

Експериментално се доказва, че поради получената 0,12 ефективност на вграждане, алгоритъмът на базата на LSB с колонна модификация удовлетворява изцяло изискванията на класическия метод LSB.

Визуален анализ

На следващите фигури (фиг. 2 и 3) е показан визуалния анализ на изображения във формат BMP и PNG с вмъкнато голямо и малко съобщение.

Фиг. 2 а) оригинал BMP

Фиг. 2 б) оригинал PNG

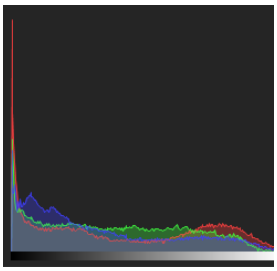
Фиг. 3. а) с голям текст BMP

Фиг. 3. б) с голям текст PNG

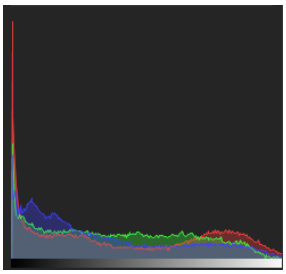
От експеримента може да се направи извод, че и в двата формата не се забелязват никакви визуални разлики след прилагането на алгоритъма.

Хистограмен анализ

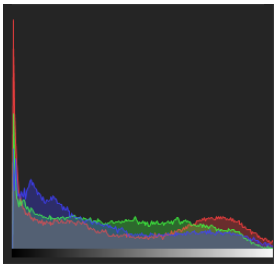
На следващите фигури (фиг. 4 и 5) е показан хистограмен анализ на изображения във формат BMP и PNG с вмъкнато голямо и малко съобщение.



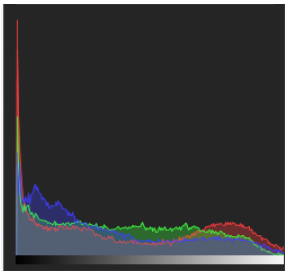
Фиг. 4 а) оригинал BMP



Фиг. 4 б) оригинал PNG



Фиг. 5. а) с голям текст BMP



Фиг. 5. б) с голям текст PNG

Не се забелязва разлика в хистограмите на оригиналните и стегофайлове, което показва равномерно изменение на стойностите при прилагане на алгоритъма.

Статистически анализ

За анализ се използва програмата е Chi-square, чрез която се изчислява вероятността дали има вмъкнато съобщение във файл (фиг. 6).

Фиг. 6. От горе на долу - BMP изображение оригинал, с голямо, с малко съобщение

Chi-square експеримента показва едни и същи резултати в оригиналния и стегофайла, което доказва подходящия избор на алгоритъм.

V. Заключение

На база направените анализи, изводи, разработки и експерименти е представен алгоритъм за внедряване покриващ изискванията на LSB метода и отговарящ на необходимото му качество, алгоритъм за извличане обратно на внедреното съобщение от стегофайл и алгоритъм устойчив на завъртане на графичната матрица, локализиращ и конвертиращ я обратно в нормалната и позиция. По този начин скритото съобщение дори и след прилагане на атака ротация, с помощта на предложения алгоритъм ще бъде намерено, прочетено правилно и извлечено обратно.

ЛИТЕРАТУРА

1. **Morkel,** T., J.Eloff and M. Olivier. An overview of image steganography. Proceedings of the 5th Annual Information Security South Africa Conference (ISSA 2005).http://mo.co.za/open/stegoverview.pdf.
2. **Sueyoshi,T.** and G.Tadiparthi.Steganography for e-Business: An Offensive Use of Information Security.Asia Pacific Management Review (2004) 9(5), pp.943-968.
3. **Almohammad, A**. Steganography-Based Secret andReliable Communications: ImprovingSteganographic Capacity andImperceptibility.PhD thesis.Brunel University,August, 2010.
4. **Аграновский**, А., А. Балакин , В. Грибунин и С. Сапожников. Стеганография, цифровые водяные знаки и стеганоанализ. Москва: Вузовская книга ,2009. ISBN 978-5-9502-0401-2.
5. **Marvel**, L.M., Boncelet Jr., C.G. & Retter, C.Spread Spectrum Steganography. IEEE Transactions on image processing, 8:08, 1999.
6. **Чиркова**,С.,Б.Бородин.Классификация критериев выбора контейнера для LSB-метода. [онлайн].[прегледан3юни 2012]. .http://network-journal.mpei.ac.ru/cgi-bin/main.pl?l=ru&n=10&pa=15&ar=4.
7. **Станев**, Ст. Стеганологична защита на информацията. Шумен: Университетско издателство „Еп. К. Преславски“, 2013. ISBN: 978-954-577-825-4.
8. http://sipi.usc.edu/database/database.php?volume=misc
9. http://bg.lipsum.com/