ARTICLE

A Dynamic Steganography Method for Web Images with Average Run-Length-Coding

Jin Liu * Yiwen Zhang

College of Computer Science and Technology, National Huaqiao University, Xiamen, 361000, China

ARTICLE INFO

Article history
Received: 28 December 2020
Accepted: 28 January 2021
Published Online: 31 January 2021

Keywords:
Steganography
Web image
Covering codes
Run length coding

1. Introduction

Nowadays, flows through world wide web has dominated the vast majority of the Internet bandwidth. Thus new information redundancy could be explored from various types of web widgets (e.g., dynamic images) and complex code frames. Different from existing information hiding methods, which can be roughly divided into two categories. The first one mainly studies the traditional overt carrying covers, always commonly used static multimedia covers, such as static image [1], audio [2] and videos [3]. While the other category employed the Internet streaming media among which are voice over IP (VoIP), video streaming and even various instant messages.

Among those traditional cover media for information hiding, image is the mostly used and investigated, that’s largely due to the variety and practicability feature of the carrier. The existing steganography methods based on image can be classified into two categories. The first type takes advantages of the visual imperceptibility of human vision, such as variant least significant bit (LSB) methods [4]. The LSB of image pixels are modified according to specific embedding method for steganography, which inevitably change the statistic features of them. Differing from the first type which falls into a space domain methods, the other type of image steganography mainly use the transform domain characteristics, which transform the image data into a certain mathematical transformation [5]. After that, some acceptable feature parameters are selected to embed secret bits, then the recovered images to be shown has a well-proportioned modification on pixels, which conforms to human visual

*Corresponding Author:
Jin Liu,
College of Computer Science and Technology, National Huaqiao University, Xiamen, 361000, China;
Email: geneleo@hqu.edu.cn
characteristics. However, the above mentioned methods has a disadvantage that the steg-image (image with secret bits) is static, which draws inevitably suspicious when it is applied to a covert communication activity because of its unexpected behaviors for image transmission.

When it comes to the dynamic elements on web pages, the image content contains many additional redundant information besides traditional image contents, such as position parameter, web programming script elements for image and constantly changing image flows for peculiar purposes \[9\]. Html space coding \[10\] is the first method described for secret bits hiding, then other html elements has been exploited after it, such as different color methods \[9\], the alternative writing style redundancies in html \[9\], or even the showing text content for users based on previously proposed text steganography \[10\]. Although the existing methods as above mentioned for images and html characters, the characteristics of image features in html environment has been seldom discussed, which is a promising direction for image steganography. That contains many image transmission redundancies, image coding redundancies in html and image activity characteristics, which will be discussed as follows.

Sections 2 talks about the principle of covering code to improve the imperceptibility of embedding, and depicts the coding rules of run length code. Then section 3 gives a detailed discussion about our proposed web image steganography method. The experiment test and analysis are described in section 4. At last, section 5 gives a conclusion of our research.

2. Covering Codes and Run Length Coding

In information hiding field, covering codes are commonly used to embed secret bits into the covers or extract recovering bits from them. Matrix coding is the most widely used covering codes, which are always used for improving the embedding efficiency and imperceptibility of steganography. \(\text{Cov}(R, N, n)\) denotes that \(n\) bits of secret bits can be embedded into \(N\) bits of cover by modifying \(R\) bits at most \[11\]. Therefore, more cover bits can be introduced to reduce the distortion it brings after the embedding procedure compared with the direct embedding, the bit change rate can be decreased accordingly. The key point here is to make use of more redundancy information of the image cover in html to increase the cover spaces, that is just what the proposed method brings. We adopt the Hamming coding method \(\text{Cov}(1, 2^n - 1, n)\) as an example, the parameter \(n\) can be adjusted according to practical embedding secret bits. Then the embedding efficiency (E) can be formulated as follows:

\[
E = \frac{m}{R} = \frac{m \cdot 2^m}{\sum_{i=1}^{R} iC_i^m - 1}
\]

Where \(m\) represents the number of secret bits, and \(\bar{R}\) denotes the average bit change of the embedding algorithm.

Run length codes were previously used in compressing algorithms, which could also be utilized in steganography algorithms to enhance the security characteristics for image embedding. The length of continuous emerged 0 or 1 bits represents run length, whose value ranges from 1 to \(m\). Given an image with a size of \(p \times q\), combined with the embedding method the average run length of 0 or 1 of a given image varies from 1 to \(N\), where \(N\) represents the total pixels for an image. Then the various values of run length of 0 and 1 bits could be calculated and used for embedding secret bits. More specifically, the different adjacent run length, such as 3 or 4 categories, are used for coding in the embedding algorithm. Together with our previously proposed multi-ary coding method\[12\], this redundancies could be encoded as the carrier bits for embedding. That is, the embedding capacity for a given size \(l\) of run length with value \(v\), the average capacity is \(\log_v l\). As a result the total image has a capacity of \(\sum_{l=1}^{p\times q} \log_v l\). The average embedding capacity and bit change rate depends on the average run length code of the entire image. By using the covering codes, and the intentional designed embedding balance between the two types of 0 and 1 run length, the average distortion it brings is largely suppressed.

3. The Proposed Web Image Steganography Method

According to the particular characteristics of web environment, the image contents on web html pages are largely different from those traditional counterparts. At first, the images size on web pages depends on the specific application we use. And the number of images varies from one site to another, even on a same site, the image numbers and size may be changing following with the users’ behaviors and network status. The second is that the image formats may conform to the html requirements, such as size limitation for a quick response or requirements regarding the color. Generally speaking, the dynamically changed image contents may require the embedding method’s flexibility and universality. On the other hand, these abundant dynamic information gives us more opportunity to use its redundancy information.

As discussed above, the web image contents and side information can be utilized as redundant information, then it is encoded for steganography. Then combined with
the average run length code of image contents, which are used for the cover bits for the covering codes and multi-ary methods, the imperceptibility and security could be increased. Figure 1 gives a brief description of the embedding process our proposed average run-length-code steganography.

### 3.1 The Embedding Procedure

The web information on Internet contains various covers, here we make use of the web programming redundancy codes [13], which constitute the cover bits to be modified in the covering codes. The proposed method uses modified matrix encoding techniques COV(2, w, y) Hamming Code [14] modifying maximum 2bits. In the Hamming Codes, \( h \), denotes a \( w \times 1 \) vector of \( H=(h_1, h_2, ..., h_{2^w-1}) \) binary Hamming Matrix. The embedding distortion for the image pixels could be decreased by using more covering bits. In our method, the covering bits consists of three types. That is, the html page coding redundancies, the dynamic image information and the inherent image content redundancies (such as space domain and frequency domain).

Therefore, after the extract bits of the first step combined with covering codes, the average run-length-codes are counted. For the run-length-codes sizes range from \( l_0 \) to \( l_{p\times q} \), the occurrence number of each code denote as \( l_i \), then the average length of run-length can be calculated as 

\[
L_{ai} = \frac{1}{m} \sum_{i=0}^{m-1} l_i 
\]

The value of \( l_i \) is treated as the center value of embedding, the run-length greater than this value represents secret bit 0 or 1, while the lower one means 1 or 1, respectively. Thus, secret bits are embedded accordingly.

Figure 2 shows how the run-length cover bits are embedded according to the secret message bits. In addition to this, the image position in html pages also contains too many redundant information, which can also extend to any other objects embedded in the page, such as audio, video and also paragraphs. Under that circumstances, this kind of images include the website logo, floating image, fixed image, popup image and hidden image. Also, during the login process, a login page often receive some kinds of dynamic some verification image, which is often changed several times in one login process. As a result the verification images is an appropriate cover for our steganography method. After the embedding procedure, our previously proposed multi-ary coding method is used to enlarge the space of embedding key, thus enhance the security of covert communication.
3.2 Secret Bits Extracting

The extracting of secret bits is relatively much easier compared to the embedding procedure. What we need are the received steg-images, the average run-length code list L and the steganography parameters computed from the decoding algorithm. The extracting procedure can be briefly described in Figure 3 as follows, where the covering codes is processed using the Hamming codes retrieving methods.

4. Performance Analysis

To evaluate our proposed method, the embedding capacity and imperceptibility are mainly discussed here. To analyze the received page from the receiver, the covert information should be divided into two type as explained above, the first is page redundant information, which could only change the page size or page html elements. According to the human visual system, this kind of statistic information cannot draw people’s attention.

Moreover, the more significant feature comes from the embedding procedure of image content. However, after the introduce of covering codes and multi-ary coding into our embedding algorithm, the bit change rate (also could be regarded as imperceptibility) is decreased and the embedding security is largely increased, when compared with traditional space domain methods LSB, frequency domain DCT steganography. Table 1 shows the results of the comparison, which shows that the proposed method has a better security and low bit change rate. Given that we have an \( p \times q \) image which utilizes \( \text{Cov}(1, 2^n-1, n) \) covering codes and multi-ary coding, where \( s \) denotes the number of images in a web page.

5. Conclusion and Future Work

This paper proposes a new steganography method for web images. The main contribution is that we utilize the redundant information from the characteristics of web page and the average run-length codes of dynamic images. According to the performance analysis, the proposed method offers a comparable effects for practical use in the covert communication environment. Our future work will focus on the image features on web flows, such as the transcoding, the potential image compression and so on.

6. Acknowledgments

This work is supported in part by the First Batch of Youth Innovation Fund Projects in 2020 under Grant No.35022202006012 and the Experimental Teaching Reform Project of National Huaqiao University under Grant No.SY2019L013.

References


