BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Computer Science Research
https://ojs.bilpublishing.com/index.php/jcsr

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

ABSTRACT

Web page has many redundancies, especially the dynamic html multimedia object. This paper proposes a novel method to employ the commonly used image elements on web pages. Due to the various types of image format and complexity of image contents and their position information, secret message bits could be coded to embed in these complex redundancies. Together with a specific covering code called average run-length-coding, the embedding efficiency could be reduced to a low level and the resulting capacity outperforms traditional content-based image steganography, which modifies the image data itself and causes a real image quality degradation. Our experiment result demonstrates that the proposed method has limited processing latency and high embedding capacity. What's more, this method has a low algorithm complexity and less image quality distortion compared with existing steganography methods.

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 mathematic 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*

DOI: https://doi.org/10.30564/jcsr.v3i1.2735

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 [6]. Html space coding [7] is the first method described for secret bits hiding, then other html elements has been exploited after it, such as different color methods [8], 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. *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 *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:

$$\mathrm{E} = \frac{m}{\overline{R}} = \frac{m \cdot 2^m}{\sum_{i=1}^{R} iC_{2^m-1}^{i}}$$

Where m represents the number of secret bits, and $\overline{R}$ 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_i of run length with value v_i, the average capacity is $\log_2 v_i$. As a result the total image has a capacity of $\sum_{i=1}^{p\times q} \log_2 v_i$. 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

DOI: https://doi.org/10.30564/jcsr.v3i1.2735

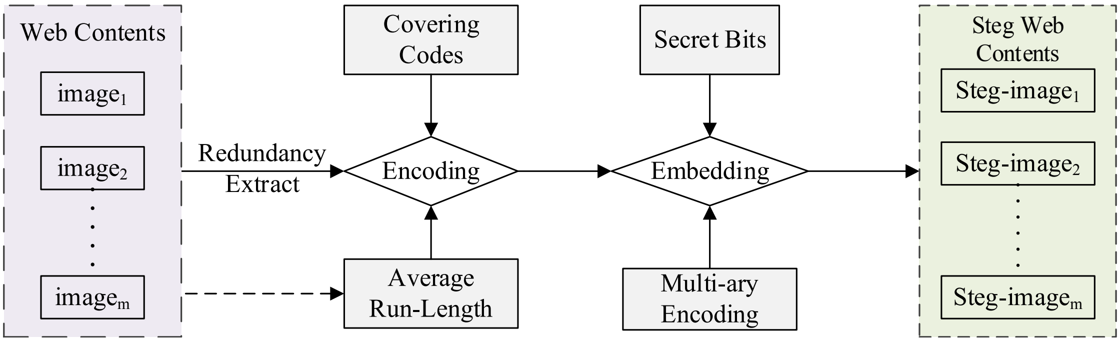

Figure 1. steganography of dynamic web image

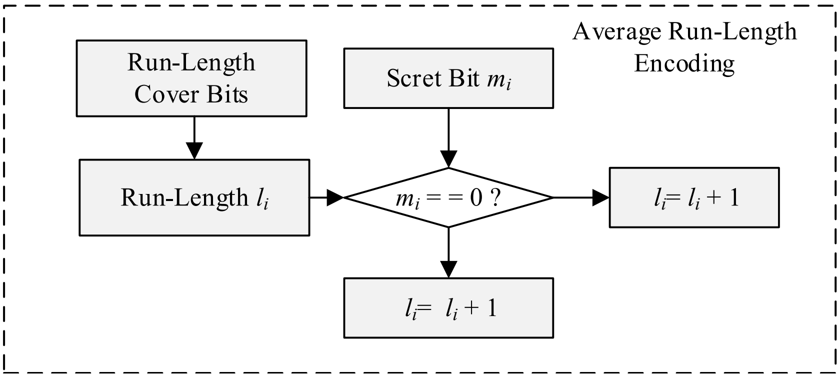

Figure 2. The embedding process of average run length codes steganography

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_i denotes a w×1 vector of H=(h_1, h_2, ..., h_{2y-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 , then the average length of run-length can be calculated as $l_{\mathrm{a_i}} = \frac{1}{m}\sum_{i=0}^{m-1} l_i$. The value of 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.

DOI: https://doi.org/10.30564/jcsr.v3i1.2735

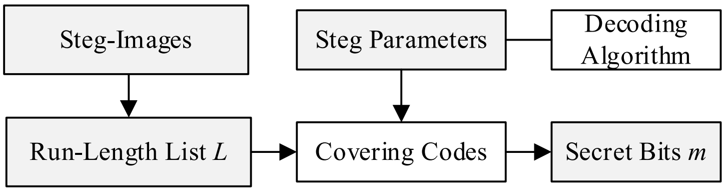

Figure 3. The secret bits extracting procedure

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 $Cov(1, 2^n\text{-}1, n)$ 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.3502Z202006012 and the Experimental Teaching Reform Project of National Huaqiao University under Grant No.SY2019L013.

References

[1] Hussain I, Zeng J, Tan S. A Survey on Deep Convolutional Neural Networks for Image Steganography and Steganalysis[J]. KSII Transactions on Internet & Information Systems, 2020, 14(3).

[2] Wu J, Chen B, Luo W, et al. Audio Steganography Based on Iterative Adversarial Attacks Against Convolutional Neural Networks[J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 2282-2294.

[3] Abdolmohammadi M, Toroghi R M, Bastanfard A.

Table 1. Steganography performance comparison (per pixel)

Steganography methods	Bit Change Rate	Secret Key space	Embedding Capacity
Average Run-Length code Steganography	$p \times q/2_n$	$>p \times q \times n+2^s$	$>p \times q \times n+2^s$
Image LSB (one bit per pixel)	0.5	$p \times q$	$p \times q$
Image DCT Coefficients	$\lfloor\frac{p}{8}\rfloor \times \lfloor\frac{q}{8}\rfloor / (p \times q)$	$\lfloor\frac{p}{8}\rfloor \times \lfloor\frac{q}{8}\rfloor$	$\lfloor\frac{p}{8}\rfloor \times \lfloor\frac{q}{8}\rfloor$

Video Steganography Using 3D Convolutional Neural Networks[C]. Mediterranean Conference on Pattern Recognition and Artificial Intelligence. Springer, Cham, 2019: 149-161.

[4] Hosam O, Ahmad M H. Hybrid design for cloud data security using combination of AES, ECC and LSB steganography[J]. International Journal of Computational Science and Engineering, 2019, 19(2): 153-161.

[5] Wang Q, Bi S. Improved method for predicting the peak signal-to-noise ratio quality of decoded images in fractal image coding[J]. Journal of Electronic Imaging, 2017, 26(1): 013024.

[6] Duan X, Guo D, Liu N, et al. A New High Capacity Image Steganography Method Combined With Image Elliptic Curve Cryptography and Deep Neural Network[J]. IEEE Access, 2020, 8: 25777-25788.

[7] Bajaj I, Aggarwal R K. RSA Secured Web Based Steganography Employing HTML Space Codes And Compression Technique[C]. 2019 International Conference on Intelligent Computing and Control Systems (ICCS). IEEE, 2019: 865-868.

[8] Liao X, Yu Y, Li B, et al. A new payload partition strategy in color image steganography[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(3): 685-696.

[9] Nolkha A, Kumar S, Dhaka V S. Image Steganography Using LSB Substitution: A Comparative Analysis on Different Color Models[M]. Smart Systems and IoT: Innovations in Computing. Springer, Singapore, 2020: 711-718.

[10] Por L Y, Delina B. Information hiding: A new approach in text steganography[C]//WSEAS international conference. Proceedings. Mathematics and computers in science and engineering. World Scientific and Engineering Academy and Society, 2008, 7.

[11] Kim C, Yang C N. Improving data hiding capacity based on hamming code[M]. Frontier and Innovation in Future Computing and Communications. Springer, Dordrecht, 2014: 697-706.

[12] Liu J, Tian H, Lu J, et al. Neighbor-index-division steganography based on QIM method for G. 723.1 speech streams[J]. Journal of Ambient Intelligence and Humanized Computing, 2016, 7(1): 139-147.

[13] Liu J, Tian H, Zhou K. Frame-bitrate-change based steganography for voice-over-IP[J]. Journal of Central South University, 2014, 21(12): 4544-4552.

[14] Tian H, Wu Y, Chang C C, et al. Steganalysis of adaptive multi-rate speech using statistical characteristics of pulse pairs[J]. Signal Processing, 2017, 134: 9-22.

DOI: https://doi.org/10.30564/jcsr.v3i1.2735