International Journal of Advanced Science and Technology
Vol. 29, No.02, (2020), pp. 2092-2098

New Metrics for Steganography Algorithm Quality

Qasim Mohammed Hussein
Tikrit University, College of Petroleum & Minerals Engineering
Salah Adin, Iraq,

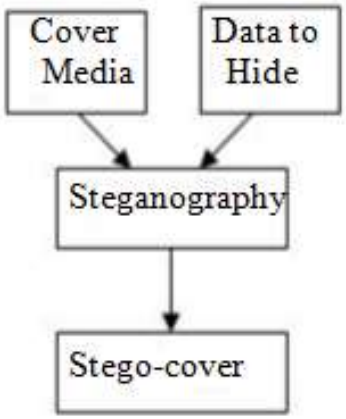
Abstract

Steganography is used to invisible secret communication between parties by hiding the existence of the secret information changing inside other media. Many techniques and approaches are proposed by the researchers. There is a necessity to use some metrics and tests to check the quality of these algorithms and approaches. Now, many tests were used to check this quality such as Mean Square Error, Peak Signal to Noise Ratio, Cross Correlation coefficient, Entropy, and the signal-to-noise ratio. These tests are taken in the consideration the number of bits that were changed in cover through the hidden process and check the quality of the stego-cover. But these tests don't take in the account the total amount of the secret data that will be hidden in the stego-cover or the locations of changed bits in the byte, or pixel, within the stego-cover. Therefore, there is a lack of precision in measuring the efficiency of the proposed algorithms. So, there is a need to use metric that meet these requirements in tests. The paper introduces two metrics to evaluate the steganography algorithms that take into account the total amount of hidden data and the number of bits affected in the stego cover by the hiding processes and their locations in the byte, or pixel.

Keywords: *Steganography, Evaluation parameters, Embedded data, Hiding process, Cover media, PSNR, MSE.*

1. Introduction

Steganography uses to invisible communication on unsecure channel, by hiding secret information inside other digital media to hide their existences [1]. The digital media may be text, image, video and audio. So, the purpose of steganography is to hide the existence of the secret message from a third party so that intruders can't detect the communication [2]. A system in steganography consists of three elements: the secret message, cover media that uses to hides the secret message within it, and the stego-cover that is the cover after embedding the secret information inside it [3], as shown in Figure 1. The embedded data in the cover must be imperceptible to the observer. This imperceptibility can be indicated by comparing the original image and its counterpart with embedded data to determine if their visual or aural are the same, or can be expressed by mathematical relationships between the original cover and the stego cover. Spatial Domain or frequency Domain based are two main categories of steganography. In the spatial domain techniques, secret data embed directly within the cover. While the frequency domain techniques use some transformations to transform the cover media and then hide the secret message within it[4].



ISSN: 2005-4238 IJAST

Figure 1. Steganography system parts

Many steganography techniques and approaches proposed by the researchers, each one has its advantages and limitations. Using LSB is a simplest and easiest method of hiding the data in covers where the data is hiding into LSBs of the cover bytes or in pixels of image. This method has small affect in the cover and the human eyes would not able to discover it [5]

To measure and evaluate the quality of steganography approaches, many metrics are used to measure the overall performance of the proposed steganography techniques. The most statistical metrics that used are Mean Square Error (MSE),Peak Signal-to-Noise Ratio (PSNR) ,Root Mean Square Error (RMSE). The histogram able to tell whether or not the image has been properly exposed, it works by examining the repetitions of each image color [6]. While structural similarity (SSIM) measures the similarity between the original and stego-covers [7]. Most of statistical metrics don't take in the account the total secret data that embedded in the cover nor the locations of bits in the byte, or pixel, which will be changed after hiding process.

This paper will introduce a new Steganography metrics to meet the previous requirement able to measure the efficiency of the proposed Steganography algorithms to hide secret information in a covers.

2. Steganography System Evaluation Parameters

To evaluate the steganography algorithms, many parameters must take in the account to hide secret data without perceptible degrading of covers multimedia qualities and provide better resistance against process of steganalysis. The main parameters are: hiding capacity, measure, and security, Figure 2[8] [9] [10].

The capacity refers to the maximum amount of data that can be embedded in the cover media, it can be represented in bits or bytes or kilobytes. Or it is the maximum number of bits that can be hidden per pixel or byte. Increasing the hidden information capacity of the cover yields more degraded in it and become more detectable. Distortion measures the ability to survive attacks aiming at modifying or removing embedded payload. The stego-images distortion should not be noticeable, imperceptible. The distortion can be measured by using many metrics such as: Mean Square Error (MSE), Root Mean Square Error (RMSE), Peak signal-to-noise Ratio (PSNR) ... etc.

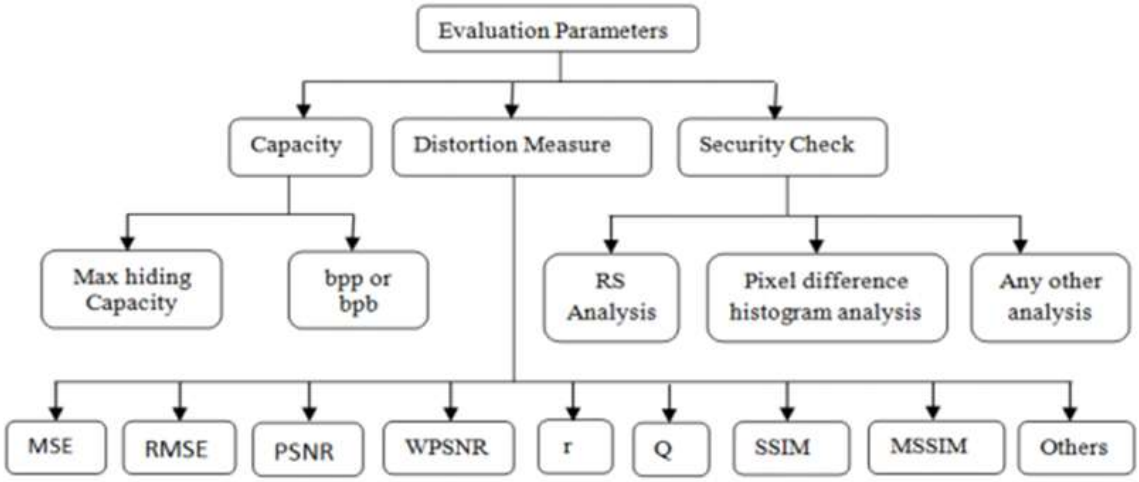


Figure 2. Steganography system evaluation parameters [9]

While the security is the resistant to various steganalytic attacks methods and robustness to the modification the stego cover can withstand before an adversary can destroy hidden data. The complexity of the algorithm may be considered as the fourth parameter. In literature no researcher has considered the algorithm complexity as an evaluation parameter.

ISSN: 2005-4238 IJAST

3. Related Works

The most common metrics that use to evaluate the distortion of steganography algorithms are [11] [12][13] [14][15]:

1. **Mean Square Error (MSE):** This metric uses to measure the distortion in the cover after hiding the data within. It by calculating the cumulative squared error between the original cover and the stego-cover. A small value of MSE means lower differences between the original cover and the stego-cover. Equation (1) is used to calculate MSE.

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(S(i,j) - C(i,j))^2 \quad \dots (1)$$

Where $S(i,j)$ represents the original cover and $C(i,j)$is stego-cover and M and N are height and width of the cover. A small value of the MSE means that the average level of difference between them is little.

2. **Peak Signal to Noise Ratio (PSNR):** This metric uses as quality measurement between the original cover and the stego-cover in the hiding process after the hidden data is embedded is better. Equation (2) is used to calculate the PSNR.

$$PSNR = 10 * log_{10}\left(\frac{R^2}{MSE}\right) \quad \dots (2)$$

Where R is the maximum value of the byte, or pixel, for 8-bit it is 255.

3. **The signal-to-noise ratio (SNR)** is used to quantify how much a signal has been distorted by a noise, the higher SNR ratio shows the less obtrusive the background noise. The SNR calculated using equation (3).

$$SNR = 10 * log_{10}\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}s(i,j)^2)}{\sum_{i=1}^{n}\sum_{j=1}^{m}(S(i,j) - C(i,j))^2} \quad \dots (3)$$

4. **Entropy**: It is a statistical measure of randomness amount that presents in the cover. It calculated using equation (4).

$$Entropy = -\sum_{i=1}^{n} P_i log P_i \quad \dots (4)$$

5. **Normalized cross correlation (NCC)**: It used to measure the degree of similarity (or dissimilarity) between the original cover and stego cover. It calculated using equation (5).

$$NCC = \frac{\sum_{I=1}^{N}\sum_{j=1}^{M}S_{i,j}C_{i,j}}{\sum_{i=1}^{N}\sum_{j=1}^{M}(S_{i,j})^2} \quad \dots (5)$$

where $S_{i,j}$ is the original cover and $C_{i,j}$ is the stego cover.

4. The proposed Distortion Metrics

In most steganography algorithms that used to embed secret data in the digital covers, many bits values of the cover media may change, 0 may be change to 1 or 1 change to 0. These changes will effect on the quality of the original cover. Therefore, many quality metrics are used to determine and assess the quality cover media through various indicators of quality, such as MSE, PSNR . . . etc.

ISSN: 2005-4238 IJAST

tests. These metrics results depend on the number of bits in the cover media whose values have been changed as a result of hiding secret data. The number of these changed bits is depending on the nature of the techniques that used in the hiding data within the cover without taking into account the size of the secret text that was hidden in the cover. On other hand, there is no standard length for the hiding text size that must be used in testing the steganography algorithms.

This paper suggests two metrics. The first takes into account the amount of total data that embedded in the digital cover and the number of cover bits that will be changed in the cover within the hiding processes. While the second, it cares on the locations of bits in the byte, pixel, that will changed as a result of the hiding process.

4.1 Affect the size of embedded data

During the hiding of secret data in the cover media, some cover bits values will change. Whenever the number of changed bits increased, the quality of the cover will be reduced and the hiding process becomes perceptible to human visually or aurally. In addition to, these increasing results anomalies give the steganalysis a sufficient statistical evidence to support the presence of hiding process. In hiding process, if the secret data is small and the cover size is much larger than hidden data, the number of changed cover bits will be too small in compared with the cover size. Therefore, the used metrics results will not reflect the efficiency of the used steganography algorithms in hiding process. In this case, therefore, these results will suggest that this steganography algorithm is efficient, but in fact, it may be a weak technique. When hiding larger number of secret data within the same cover, one can detect the weakness of this approach since the changed cover bits become more. For example, if the size of the cover is 100 bits and secrets data is 10 bits. And suppose that after hiding process, the values of 5 bits from the cover are changed. so, statistically, the ratio of bits that were changed as a result of the hiding process relative to the number of secret data bits will be 5/100=5%. If the size of the cover is 100 bits and the secret data is 20 bits and the number of bits changed is 10, the number of changed bits/ cover size bits= 10/100 = 10%. When comparing the two results, one can conclude that the first case is better efficiency than the second case. While, in fact, they have same efficiency, when taking in the account the total number of secret information that will be hidden in the cover media. They have the same ratio; 5/10= 0.5 and 10/20=0.5.

So, there is a necessity to use a metric that take in the account the total number of secret data that will be hidden in the cover media and the number of cover bits that will change as a result of hiding process. Equation (6) present Q-factor that meets the previous requirements.

$$Q - factor = \frac{number\ of\ changing\ in\ cover\ bits}{number\ of\ secret\ text\ \ bits} \quad (6)$$

A small value of the Q-factor means that difference between original and stego cover is little. The best efficient of the steganography algorithm in which the values of the Q-factor = 0. But this value is impossible since it is not possible to embed data in the cover without changing the values of its bits. A greater value of Q-factor means there is an effect on the stego cover during the hiding process. This effect may cause distortion in the cover, or it gives the steganalysis an indicator that there is an embedded data hidden within it. The value of Q-factor is $1 \leq$ Q-factor ≤ 0. For the best case Q-factor =0, while Q-factor =1 for the worst case. Steganography system is considered secure when Q-factor is close to zero. Despite our conviction, there is no a method to hiding data in any cover unless the cover bits will be changed.

4.2 The effective of bits changing position in the byte

In most steganography algorithms that used to hidden data in the digital covers, many cover bits value will change, 0 may be change to 1 or 1 change to 0. Effecting of changing cover bits on the quality of the original cover are varied from bit to other depending on their locations in their bytes. For example a change in least significant bit (LSB) of the cover has limited effect on the perception of

ISSN: 2005-4238 IJAST

the human visual or aural systems. While changing in the most significant bits (MSB) leaves a clear effect on the cover, which makes some perceptibility to the human visual or aural systems that give an indicator about the existing of embedded data. Statistically, there is difference between changing LSB value or MSB value, although the LSB and the MSB have not the same visual or aural effect on the stego cover. In other hand, most metrics does not take into account the number of bits whose values have been changed as a result of the hiding process, without significant attention to the position of the bit within the byte, dissimilarity between the two positions. Therefore, there is a need to use metric deals with bits locations within byte.For this reason, this paper suggests a k-factor metric that takes locations of the changed bits in the account. It depends on using a weight for each bit in the byte, pixel. The weights were given depending on how much impact of the changed bit on perceptibility and degrading of stego cover qualities. The LSB in the byte has the lowest weight, while the MSB has the largest weight, table (1) presents weights of each bits in the cover byte.

Table 1. Weights of bits.

Location	B_0 (LSB)	B_1	B_2	B_3	B_4	B_5	B_6	B_7(MSB)
Weight	1	2	4	8	16	32	64	128

Depending on the weights of bits, the k-factor can be calculated using equation (7).

$$k_{factor} = \frac{\sum_{i=1}^{n} wb_i}{Total\ size\ of\ secret\ data\ (in\ bits)} \quad \ldots (7)$$

Where n is the number of bits in the stego cover that their values are changed through hiding process, wb is the weight of the changed bit in stego cover byte.

The minimum value of k-factor is zero, whereas the maximum value of it for each location in the byte is $(n*2^L)/m$, where L is the number of changed bit within the cover, m is the number of secret text bits and n is the location of changed bit within the byte, $0 \leq L \leq 7$. Table 2 includes the minimum and maximum values of k-factor for each bytes location that changed through hiding process

Table 2: Minimum and maximum k-factor values for bits locations in bytes

Location	B_0 (LSB)	B_1	B_2	B_3	B_4	B_5	B_6	B_7 (MSB)
Minimum value	0	0	0	0	0	0	0	0
Maximum value	n*1	n*2	n*4	n*8	n*16	n*32	n*64	n*128

A small value of the k-factor means that there is a difficulty to perceptible the differences between the original cover and stego-cover. The steganography system is considered secure when k-factor is close to zero.

5. EXPERIMENTAL RESULTS AND DISCUSSION

Many experiments were conducted to evaluate the two proposed metrics by hiding data in images and calculated their values. Thereafter, results of the two metrics compared with the results of most common metrics that used to measure the efficiency of steganography approaches. In these experiments, the same secret text, the same cover, and the same cover bits that changed were taking in the account. Table (3) and table (4) illustrate the result of six experiments. Column 1 includes the results when changed bits are in byte LSB locations only. In column 2, only the MSB locations of bytes are changed through hiding process. While in column 3 the changed bits locations are distributed on all locations in the bytes. In table 3, the image size=128×128, secret text length=1000 bits and the number of bits that were changed after hiding process within the cover = 500 bits. Whereas, in table 4, the image size=128×128, secret text length=1000 bits and the number of bits that were changed within the cover = 250 bits.

ISSN: 2005-4238 IJAST

Table 3. Metric result where text=1000, changed bits=500

Metric	LSB changing	MSB changing	Random bits changing
MSE	0.03	500	74.36
SNR	43.2	1.05	9.33
PSNR	63.29	21.14	29.42
NCC	0.99	0.63	0.85
Q-factor	0.5	0.5	0.5
K-factor	0.5	64	13.84

Table 4: Metric result where text=1000, changed bits=250

Metric	LSB changing	MSB changing	Random bits changing
MSE	0.02	250	0.31
SNR	43.16	1.02	30.02
PSNR	66.30	24.15	53.16
NCC	1	0.63	1
Q-factor	0.25	0.25	0.25
K-factor	0.25	32	0.92

When comparing the results of the two tables, one can observe differences in the results of common metrics, although the two cases had the same effect; the number of bits within the cover that are changed equal to half of the number of secret text bits. Whereas, Q-factor metric gave the same results for both cases. Thus, the Q-factor metric gave the real effect of the steganography method that used. For k-factor, it clearly shows the effect of bits locations that had changed within the cover as a result of the hiding process. The value was very low in the case of LSB in bytes were changed. While its value was high in the case of change is in byte MSB within the cover. From the above results, it is clear that the two proposed metrics should be used together to give a real assessment for the steganography algorithm efficiency depending on the extent of change in the stego-cover bits values.

CONCLUSION

The paper suggests two metrics that are useful to evaluate the distortion of steganography algorithm efficiency in stego-cover, Q-factor and k-factor. They based on statistical information from secret text and the changing that will occur in the stego-cover data. Q-factor takes into account the amount of total secret text bits that will hide in the digital cover and the number of cover bits that will be changed through the hiding processes. While k-factor takes in the account bits locations within cover bytes whose value will be change through the hiding process. The suggested metrics provide a clearer expectation of the efficiency of the steganography algorithm in terms of its effecting on the cover data. They will improve the evaluation of steganography algorithm security and distortion evaluation. The Q-factor and k-factor can, together with the other metrics, reveal the hiding effect extent of the steganography algorithm on the stego-cover data

References

[1] Hanaa Mohsin Ahmed Salman & Halah H. Mahmoud, Review on Image Steganalysis Using INRIA Dataset,Al-Nahrain Journal of science, vol. 21, No. 4, 2018.
[2] Nada Qasim Mohammed, Qasim Mohammed Hussein, Mohammed Sh. Ahmed, Suitability of Using Julia Set Images as a Cover for Hiding Information. Al-Mansour International Conference on New Trends in Computing, Communication, and Information Technology (NTCCIT 2018). IEEE, 2018.

ISSN: 2005-4238 IJAST

[3] Hayat Shahir Al-Dmour, Enhancing Information Hiding and Segmentation for Medical Images using Novel Steganography and Clustering Fusion Techniques, Ph.D. thesis, University of Technology Sydney, 2018.
[4] Anita Pradhan, Aditya Kumar Sahu, Gandharba Swain, K. Raja Sekhar, Performance Evaluation Parameters of Image Steganography Techniques, International Conference on Research Advances in Integrated Navigation Systems (RAINS - 2016), Bangalore, India, 2016.
[5] Ramanpreet Kaur , Baljit Singh , Ishpreet Singh, A Comparative Study of Combination of Different Bit Positions In Image Steganography, International Journal of Modern Engineering Research (IJMER), Vol.2, No.5, . 2012 pp-3835-3840
[6] Erfaneh Noroozi, Salwani Bt Mohd Daud, Ali Sabouhi, Critical Evaluation on Steganography Metrics, proceedings of 2011 International Conference on Electrical Engineering and Applications, Chennai, India, 2011.
[7] A. Al-Mohammad, Steganography-based secret and reliable communications: Improving steganographic capacity and imperceptibility. PhD thesis, Brunel University, School of Information Systems, Computing and Mathematics Theses, 2010.
[8] Zinah Talaat Rashid, Security Enhancement of Image Steganography Using Embedded Integrity Features, master thesis, Faculty of Information Technology- Middle East University – Amman – Jordan, 2017,
[9] Anita Pradhan, Aditya Kumar Sahu, Gandharba Swain, K. Raja Sekhar, Performance Evaluation Parameters of Image Steganography Techniques , International Conference on Research Advances in Integrated Navigation Systems (RAINS), India, IEEE, 2016.
[10] Qasim Mohammed Hussein, Hiding Message in Color Image Using Auto Key Generator, 3rd International Conference on Advanced Computer Science Applications and Technologies, Amman, Jordan, 2014, pp:151-155.
[11] Mazhar Tayel, Hamed Shawky, Proposed Assessment Metrics for Image Steganography, International Journal on Cryptography and Information Security (IJCIS), Vol. 4, No. 1, 2014
[12] Erfaneh Noroozi , Salwani Bt Mohd Daud , Ali Sabouhi, Critical Evaluation on Steganography Metrics, Proceedings of 2011 International Conference on Electrical Engineering and Applications (EEA 2011).
[13] Raihan Sabirah Sabri, Roshidi Dini, Aida Mustapha, Analysis Review on Performance Metrics for Extraction Schemes in Text Steganography, Indonesian Journal of Electrical Engineering and Computer Science Vol. 11, No. 2, 2018, pp.761-767.
[14] Peter Ndajah, Hisakazu Kikuchi, Masahiro Yukawa, Hidenori Watanabe, Shogo Muramatsu, An Investigation on The Quality of Denoised Images, International Journal of Circuits, Systems and Signal Processing, Volume 5, No. 4, 2011, pp 423-434.
[15] Qasim Mohammed Hussein, Ahmed Saadi Abdullah, Nada Qasim Mohammed, The efficiency of color models layers at color image as cover in text hiding. Tikrit Journal of Pure Science Vol.21, No. 1, 2016, PP 130-139.

ISSN: 2005-4238 IJAST