Signal Processing 146 (2018) 99–111

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

On the fault-tolerant performance for a class of robust image steganography

Yi Zhang[a,b], Chuan Qin[c], Weiming Zhang[d], Fenlin Liu[a,b], Xiangyang Luo[a,b,*]

[a] State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China
[b] Zhengzhou Science and Technology Institute, Zhengzhou 450001, China
[c] School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China
[d] School of Information Science and Technology, University of Science and Technology of China, Hefei 230026, China

ARTICLE INFO

Article history:
Received 21 September 2017
Revised 4 December 2017
Accepted 1 January 2018
Available online 6 January 2018

Keywords:
Robust steganography
JPEG compression resistant
Statistical detection resistant
Fault-tolerant performance
RS codes
STC codes

ABSTRACT

The mainstream adaptive steganography algorithms often cannot transmit secret messages correctly when stego images suffer from JPEG compression. In this respect, researchers proposed a series of robust adaptive steganography methods based on the framework of "Compression-resistant Domain Constructing + RS-STC Codes" in previous studies. However, these methods leave behind the fault tolerance analysis, resulting in potential mistakes in extracted messages, which brings uncertainty to practical application. To solve this problem, an error model based on burst errors and STCs decoding damage is given in this manuscript, utilizing the burst error model based on Poisson distribution. Then the model is verified using the hypothesis test problem judged by the χ^2 test method. Based on the proposed model, the error conditions of received stego sequence are depicted, and the fault-tolerant performance of the robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes" is deduced, that is, the probability lower bound for RS-STCs decoding to correctly extract embedded messages. Experiments demonstrate that the practical fault-tolerant results of previous robust steganography methods consist with the theoretical derivation results, which provides a theory support for coding parameter selection and message extraction integrity to the robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes".

1. Introduction

With the rapid development of network multimedia technology and intelligent mobile devices, using intelligent mobile devices to capture, process, and transmit images, as a convenient and efficient way of communication in digital age [1], has been widely applied. Digital images transmitted by intelligent mobile devices have become an important potential carrier for covert communication. The image steganography technology based on intelligent mobile devices can effectively realize the secure transmission of secret messages, so as to meet the demand for convenient and secure communication, which is expected to become a new research hotspot in the field of information hiding. In most cases, the JPEG compression is performed on digital images before the transmission between mobile devices, such as sending images though FaceBook, Twitter, WeChat and many other applications, due to the constraints of the network traffic, bandwidth and intelligent device processing capacity, resulting in images degradation and loss of information [2]. For this reason, image steganography technology applied to mobile terminals need to ensure not only a high detection resistance of stego images [3], but also a strong JPEG compression resistance of embedded messages [4].

Because of the high detection resistance, adaptive steganography proposed in recent years has become a hotspot in the field of information hiding. Utilizing the structure of "Distortion function + STC Codes", many adaptive steganography algorithms have been proposed, such as HUGO (Highly Undetectable steGO) steganography [5], WOW (Wavelet Obtained Weights) steganography [6], J-UNIWARD (JPEG UNIversal Wavelet Relative Distortion) steganography [7], and so on [8,9]. In the above structure, the distortion function is used to calculate the embedding distortion in different locations of the cover images, and the syndrome-trellis codes (STCs) [10] is used to select the modifying positions adaptively according to the calculated distortion and embed messages with minimum embedding distortion. Although the existing typical steganography algorithms have a good detection resistant performance, however, these algorithms often do not take into account the condition that

* Corresponding author.
E-mail address: xiangyangluo@126.com (X. Luo).

https://doi.org/10.1016/j.sigpro.2018.01.011
0165-1684/

stego images being attacked when the transmission channel is exposed, thus cannot transmit secret messages correctly and realize covert communication successfully when stego images suffering from JPEG compression.

In contrast, robust watermarking technology is more concerned about the robustness of the embedded information against various attacks [11,12], and JPEG compression resistant capability is usually one of the important indicators of the algorithm performance [13]. By constructing robust embedding domains, such as the wavelet domain [14], contourlet domain [15,16], and other transform domains [17], various robust watermarking algorithms [18–20] have been proposed and have a strong robustness against JPEG compression, additive Gaussian noise, median filtering, and many other attacks. However, the embedding capacity of robust watermark might be limited considering the visual quality of watermarked images [21], which is usually insufficient for covert communication and designing steganography algorithms. Furthermore, the robust watermark usually do not need to ensure the completely correct extraction of embedded messages after JPEG compression, and do not have to consider the resistance against statistical detection of the watermarked images.

In order to overcome the shortcomings of existing algorithms which cannot take into account the JPEG compression and detection resistance at the same time, in our previous work, the robust adaptive steganography technology resisting both JPEG compression and statistical detection is first proposed and realized in [23]. Then, a framework of adaptive steganography resisting JPEG compression and detection [22] is proposed, which is the framework of "Compression-resistant Domain Constructing + RS-STC Codes". To obtain a good resistant performance against statistical detection, this framework utilizes the "Distortion function + STC Codes" structure of adaptive steganography. In addition, the framework combines many kinds of compression resistant methods such as the construction of robust embedding domain, and the error correction codes, to obtain a strong resistance against JPEG compression. Based on this framework, there are a total of three robust steganography methods by now, that is, the DCRAS (DCT Coefficients Relationship based Adaptive Steganography) method [23], FRAS (Feature Regions based Adaptive Steganography) method [24], and DMAS (Dither Modulation based Adaptive Steganography) method [25]. These methods inherit the good detection resistant capability of adaptive steganography, and integrate the advantages of robust watermark, thus holding both the compression and detection resistant ability [26].

However, the DCRAS, FRAS, and DMAS methods mainly focus on the robustness of message embedding domains against JPEG compression and detection, while leaving behind the analysis of message extraction integrity after JPEG compression, which brings uncertainty to the practical application. Thus, the errors in the stego elements caused by JPEG compression needs to be analyzed, and the error correction performance of RS-STC codes needs to be deduced specifically since the STCs decoding damage will interfere the performance of RS codes and cause more error bits in the extracted messages after JPEG compression [27]. In this manuscript, we will firstly model the residuals caused by JPEG compression in the stego sequence using Poisson distribution. Then the error condition of the receiving sequence can be depicted and the fault-tolerant performance of robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes" can be given out, so as to obtain a theoretical guidance for the coding parameter selection. Therefore, a theory support can be provided for the message extraction integrity of the robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes".

The rest of the manuscript is organized as follows. Section 2 introduces the Preliminary works. Section 3 gives the error model based on burst errors and STCs decoding damage, and the rationality of the model is verified by experiments. In Section 4, the fault-tolerant performance of robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes" is given. In Section 5, the validity and rationality of the conclusion on fault-tolerant performance analysis are verified by experiments, and the recommended RS coding parameters under different conditions are given. Finally, Section 6 summarizes the full manuscript.

2. Preliminary works

In order to analyze the fault-tolerant performance of the robust steganography methods based on the "Compression-resistant Domain Constructing + RS-STC Codes", this section mainly introduces preliminary works in the following two aspects: a class of robust image steganography methods based on the "Compression-resistant Domain Constructing + RS-STC Codes" and the error correction performance of RS codes on a bursty-noise channel.

2.1. A class of robust image steganography methods

To overcome the shortcomings of the existing algorithms which cannot take into account the JPEG compression and detection resistant performance at the same time, the steganography framework resisting both JPEG compression and statistical detection is proposed in [22] in our previous works, which is the "Compression-resistant Domain Constructing + RS-STC Codes" framework and is shown in Fig. 1. The embedded part of the framework mainly includes four parts: constructing compression-resistant domain, designing embedding cost calculation methods, error correction coding and the minimize embedding costs coding. The extraction part mainly includes extracting stego elements, minimize embedding costs decoding, and error correction decoding.

In the framework, the pixels (or coefficients), regions, or relative relationships that are not affected or less affected by JPEG compression are constructed as the embedding domain at first. By extracting cover elements from the compression resistant embedding domain, and designing the appropriate modifying methods, messages embedded in stego images can be saved or recovered by error correction coding after JPEG compression in a high probability. Then utilizing the advantage that distortion functions of adaptive steganography measuring the detection resistant capability of cover elements, the embedding costs calculation methods are designed corresponding to the embedding domain. In the third part, the secret messages are encoded with error correction codes in order to improve the correct rates of extracted messages. At last, the minimize embedding costs coding is utilized to embed encoded secret messages into the cover elements with minimum distortion. This framework combines the good detection resistant capability of adaptive steganography, and inherits the strong JPEG compression resistance of robust watermarking technology, thus ensuring the compression and detection performance at the same time. Based on this framework, different steganography methods resisting compression and detection can be designed for different application scenarios.

Based on the above framework, there are three compression and detection resistant steganography methods in all, that is, the DCT coefficients relationship based steganography method (DCRAS) [23], the feature region based steganography method (FRAS) [24], and the dither modulation based steganography method (DMAS) [25]. The specific ideas of the three methods are as follows:

- **DCRAS method** [23]: In this method, the compression resistant relationship between DCT coefficients are utilized to construct the message embedding domain, and the changing method of

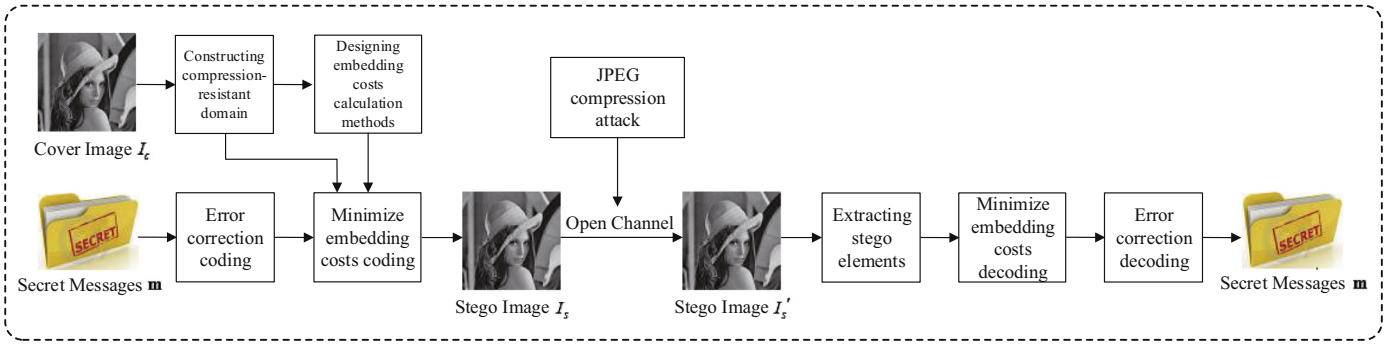


Fig. 1. Framework of "Compression-resistant Domain Constructing + RS-STC Codes".

DCT coefficients is designed in order to minimum the embedding distortion while guaranteeing the JPEG compression resistant performance. The costs function in J-UNIWARD steganography are utilized and improved to measure the cover elements' detection resistant capability. The RS codes is also introduced and combined with STC codes to embed messages with minimal costs and achieve a good compression and detection resistant performance at the same time.

- **FRAS method** [24]: In this method, based on the message embedding domain designed in DCRAS, the great compression resistant ability of Harris-Laplacian feature is utilized to construct the compression maintainable, complex and difficult to model image regions, achieving the balance between the JPEG compression and detection resistant properties of cover elements. Thus, a feature extraction algorithm resisting JPEG compression and a feature region selection algorithm minimizing the embedding costs and maximizing the feature region distribution are proposed. With the help of embedding cost function improved from J-UNIWARD steganography, RS coding and STC codes, this method can maintain the good JPEG compression resistant ability, while improving the detection resistant performance.
- **DMAS method** [25]: In this method, the correspondence between JPEG compression quantization tables and the coefficients' changes caused by compression are utilized to improve the performance. The embedding domain are designed utilizing DCT coefficients based dither modulation methods according to the quantization tables of JPEG compression, and an adaptive dither modulation algorithm based on quantization tables are proposed. Then, an embedding cost function based on side information are proposed, and the RS coding is used combined with STCs to embed messages with minimal costs. This method not only keeps the good JPEG compression resistant performance, but also has a better detection resistant performance and a higher efficiency.

From the basic ideas of the above robust steganography methods, it can be concluded that the robustness of secret messages against JPEG compression is mainly realized by constructing the compression-resistant domain and the "RS-STC Codes" based message embedding. In addition to the robustness against JPEG compression of message embedding domain which has been discussed in [23–25], the fault tolerance of coding methods also affects the message extraction integrity after JPEG compression. Therefore, it is necessary to analyze the fault-tolerant performance of the methods based on "Compression-resistant Domain Constructing + RS-STC Codes", remove the potential mistakes in the extracted messages, and put forward these methods into practical uses.

2.2. Error correction performance of RS codes on the bursty-noise channel

In the field of channel coding, RS (Reed-Solomon) codes [28] is a kind of BCH codes with strong error correction capability. It was first constructed using the MS (Mattson-Solomon) polynomial by Reed and Solomon in 1960, which could be defined as follows.

Definition 1 ([28])**.** Set q to a prime power and $q \neq 2$, the BCH codes which the roots of the code symbol and the codes' generation polynomial $g(x)$ are all derived from the finite field $GF(q)$ are called Reed-Solomon codes, or RS codes for short.

Researchers point out in [28] that RS codes is a kind of forward error correction channel coding. In (n^*, k^*) RS codes (where n^* is the code length, and k^* is the message length), the input messages is divided into groups of $m^* k^*$ bits, and each group includes k^* symbols which consist of m^* bits. The error correction capability of (n^*, k^*) RS codes is $t = (n^* - k^*)/2$, and the minimum code distance is $d = 2t - 1$. Since the RS code is a class of codes with the largest minimum distance, it has a strong ability to correct errors [24]. The RS codes in the finite field $GF(q)$ with error correction ability t can correct up to t erroneous symbols whose length is short than m^* and a burst error whose length is short than m^*t, if the locations of the error symbols are unknown in advance. If the RS codes are used to correct a burst error, it can correct up to the length of $(t-1)m^* + 1$ in $GF(q)$. Therefore, the measurement Z of RS codes' optimality is represented as follows:

$$Z = \frac{2(t-1)m^* + 2}{2tm^*}, \tag{1}$$

where Z is close to 1 when t is larger, thus RS codes is a near optimal code.

Especially, it has been pointed out in [29] that the bursty noise could be defined to be as the background Gaussian noise plus burst noise, where burst noise is defined to be a series of finite-duration Gaussian-noise pulses with fixed duration and Poisson occurrence times, which can be expressed as follows:

$$n_d(t) = n_g(t) + n_b(t), \tag{2}$$

$$n_b(t) = a(t) \sum_{t=-\infty}^{\infty} \prod \left(\frac{t - t_i - d/2}{d} \right), \tag{3}$$

where $n_g(t)$ is the background Gaussian-noise component and $n_b(t)$ is the burst-noise component. The combination of the background Gaussian noise and burst noise is referred to as bursty noise. In addition, $a(t)$ denotes a sample function from a delta-correlated Gaussian stochastic process with zero mean and double-sided power spectral density (PSD) $N_b/2$, thus $a(t) = \left(1/\sqrt{2\pi}\sigma\right) \exp\left(-t^2/2\sigma^2\right)$, and $\{t_i\}$ denotes a set of Poisson points with average rate v in Eq. (3). $\Pi(t/d)$ is defined to be a unit-amplitude pulse of width d centered at $t = 0$, and t_i is the time

at which the burst begins, thus $t_i + d/2$ is the center time of the burst that begins at t_i with width d.

On the above bursty-noise model, over a time interval I of duration T_s, let ξ_g denotes the event when no noise burst overlaps I, and let ξ_b denotes the event when a noise burst completely overlaps I, for (n^*, k^*) RS codes with code length T_{RS}. For a bursty-noise channel with ideal characteristics that the noise burst locations $\{t_i\}$ are restricted to lie on code symbol boundaries at the receiver, if $vd \ll 1$, $d \gg T_{RS}$, and let ε denote a binary symbol error in the received codeword, the binary symbol error probability after decoding can be given by Eq. (4), according to [29].

$$\begin{aligned} P(\varepsilon) &\approx P(\varepsilon|\xi_g)P(\xi_g) + P(\varepsilon|\xi_b)P(\xi_b) \\ &= Q\left(\sqrt{\frac{2}{1-\Psi}\left(\frac{E_s}{N_t}\right)}\right)P(\xi_g) \\ &\quad + Q\left(\sqrt{\frac{2}{1-\Psi+\Psi/(vd)}\left(\frac{E_s}{N_t}\right)}\right)P(\xi_b) \\ &= Q\left(\sqrt{\frac{2}{1-\Psi}\left(\frac{E_b}{N_t}\right)\left(\frac{k^*}{n^*}\right)}\right)P(\xi_g) \\ &\quad + Q\left(\sqrt{\frac{2}{1-\Psi+\Psi/(vd)}\left(\frac{E_b}{N_t}\right)\left(\frac{k^*}{n^*}\right)}\right)P(\xi_b), \end{aligned} \tag{4}$$

where $Q(x) = 1/\sqrt{2\pi}\int_x^\infty e^{-\lambda^2/2}\,d\lambda$, E_b is the signal power, E_s is the received symbol energy, and $P(\xi_g)$, $P(\xi_b)$ denote the probability for events ξ_g and ξ_b to occur separately. $S_g(f)$, $S_b(f)$ are the double-sided PSD for Gaussian noise and burst noise separately. $\Psi = S_b(f)/\left(S_g(f) + S_b(f)\right)$ is the fraction of bursty noise, and $N_t = 2 \times \left(S_g(f) + S_b(f)\right)$ is the double-sided PSD for bursty noise.

Utilizing the bursty noise model based on Poisson distribution, the error correction performance of RS codes on the bursty-noise channel can be deduced from channel parameters and the signal power in the field of communication. Learning from the above ideas, the effects of JPEG compression and STCs decoding to the stego sequence in a certain stego image can be considered as a kind of bursty noise, since the similarity between burst errors and the errors in stego sequences caused by JPEG compression and STCs decoding damage. Then, an error model of JPEG compression and STC decoding based on burst errors can be established, and the error condition of received stego sequence caused by JPEG compression and STCs decoding can be depicted. Thus the fault-tolerant performance, which is also a theory support of message extraction integrity, can be given for robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes" .

3. Error model based on burst errors and STCs decoding damage

In order to analyze the fault-tolerant performance of methods based on "Compression-resistant Domain Constructing + RS-STC Codes" against JPEG compression, the residual model of JPEG compression based on burst errors is proposed in this section, and the STCs decoding damage are also taken into account to enrich this model. Thus, an error model based on burst errors and STCs decoding damage is proposed. In addition, the model is verified using the hypothesis test problem judged by the χ^2 test method. In the following parts of this section, the residual model of JPEG compression and the error model considering STCs decoding damage are described in detail separately.

3.1. Residual model of JPEG compression based on burst errors

Utilizing the similarity between DCT coefficients' residuals caused by JPEG compression and burst errors, the Poisson distribution is used to depict the errors in stego sequences caused by JPEG compression. Thus a residual model of JPEG compression based on burst errors is proposed in this section, and the validity and rationality of the model is verified by experiments.

3.1.1. Model establishment

The JPEG compression is mainly realized by block DCT transformation, quantization, entropy coding and so on. Since the DCT transform is a nonlinear transformation, the low frequency part of the transformed data is concentrated in the upper left corner of the block, and the high frequency part is concentrated in the lower right corner of the block. For the reason that a large number of non-zero coefficients are located in the upper left corner of the block, the non-zero residuals between JPEG-compressed DCT coefficients and original DCT coefficients are usually concentrated. Therefore, the burst error model consists with the truth of concentrated residuals caused by JPEG compression. In this section, the influence of JPEG compression on stego sequences is analyzed by establishing the JPEG residual error model based on burst error.

Let $\mathbf{c}$, $\mathbf{c}'$ denote the stego sequences before and after JPEG compression separately, and $\mathbf{\Delta}_J = \mathbf{c} - \mathbf{c}'$ is the residuals of stego sequences caused by JPEG compression. Since the Poisson distribution is often used to depict burst errors, we can define the errors in stego sequences caused by JPEG compression to be a series of Poisson points $\{t_i\}$ with average rate v and average length d, then the number of non-zero residual elements n_l in each successive stego sequences of length l obey the Poisson distribution with parameter $\lambda = vd$, whose probability density function can be expressed as follows,

$$P(n_l = k) = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, \cdots, \tag{5}$$

where λ is the rates of non-zero residual elements in each successive stego sequences of length l.

3.1.2. Model validation

To test the rationality of the above JPEG compression residual model, consider the following hypothesis test problem:

$$\begin{cases} H_0 : F(n_l) = F_0(x) \\ H_1 : F(n_l) \neq F_0(x) \end{cases}, \tag{6}$$

where $F_0(x)$ is the distribution function of the Poisson distribution $P(n_l = k) = \left(\lambda^k/(k!)\right)e^{-\lambda}, k = 0, 1, \cdots,$ and $\lambda > 0$ is the parameter in the Poisson distribution.

The χ^2 test method is utilized to judge H_0. The validity and rationality of the model are verified by fitting goodness testing of compression residuals from the aspects of a single image testing and multiple images testing. The RS coding parameters used in the experiments is (31,15), and the randomly generated binary sequences are used as secret messages. The parameters of DCRAS [23], FRAS [24], and DMAS [25] methods are shown in Table 1. ('\' means that there is no this kind of parameter in the corresponding method.)

(1) Fitting goodness testing for a single image

Taking the "Lena" image with quality factor 75 as an example, the DCRAS method is used to embed messages at 0.1bpnzAC payload, and the corresponding stego image is generated and compressed with quality factor 55. Set the length of the observation l to 300, then the changes in the stego sequences before and after JPEG compression are recorded, and the numbers of error elements n_l in each successive stego sequences of length l are shown in Table 2.

Table 1
Parameters setting in DCRAS, FRAS, and DMAS methods.

Parameters	Methods		
	DCRAS [23]	FRAS [24]	DMAS [25]
Maximum distortion	$wetconst = 10^8$	$wetconst = 10^8$	$wetconst = 10^8$
Iterations T_{step}	$T_{step} = 3$	\	\
Population size	\	$N_g = 100$	\
Iterations N_i	\	$N_i = 20$	\
Quantify tables	\	\	T_{65}, T_{75}, T_{85} are quantify tables corresponding to quality factors 65, 75, and 85.

Table 2
Errors in each successive stego sequences of length l.

Error element number n_l	Statistical frequency v_n	Theoretical frequency $n_p p_{n_l}$
0	188	187.67
1	15	15.64
2	1	0.67

In Table 2, the first column indicates the errors in each successive stego sequences of length l, and the second column indicates the occurrence numbers v_n of the event that n_l errors occur in successive l stego elements, that is, the statistical frequency. By the maximum likelihood estimation method, the estimated value of the parameters λ is:

$$\hat{\lambda} = \frac{1}{n_p}\sum_{n_l} n_l v_n \approx 0.083, \tag{7}$$

where $n_p = \lfloor n_c/l \rfloor$, n_c is the length of stego sequences, and $\lfloor \bullet \rfloor$ denotes taking integers downwardly.

The values of the third column in Table 2 are the theoretical frequency $n_p p_{n_l}$, which are calculated by p_{n_l} from the Poisson distribution.

$$p_{n_l} = \frac{\left(\hat{\lambda}\right)^{n_l}}{n_l!} e^{-\hat{\lambda}}, n_{p_l} = 0, 1, 2, \cdots, \tag{8}$$

Then the statistic η whose limit distribution is subject to $\chi^2_{(m_l-1-1)}$ is established and considered as the fitting goodness test result by the following expression:

$$\eta = \sum_{n_l=1}^{m_l} \frac{\left(v_n - n_p p_{n_l}\right)^2}{n_p p_{n_l}} = \sum_{n_l=1}^{m_l} \frac{{v_n}^2}{n_p p_{n_l}} - n_p, \tag{9}$$

where m_l is the number of groups in Table 2, then it can be calculated $\eta = 0.19$ from Table 2.

Given the significance level $\alpha_l = 0.05$, since the degrees of freedom $m_l - 1 - 1 = 1$, the critical value $\chi^2_1(0.05)$ is 3.841. Because of $\eta = 0.19 < 3.841$, H_0 cannot be denied. Therefore, we can consider the above residuals caused by JPEG compression follow the Poisson process with parameter $\hat{\lambda} = 0.083$ in practical application.

(2) Fitting goodness testing for multiple images

In addition to the "Lena" image, we test more images in the BOSSbase-1.01 image database.[1] Randomly select 1000 cover images generated by JPEG compression of quality factor 75, and utilize the DCRAS, FRAS, and DMAS methods to generate the corresponding stego images at 0.1bpnzAC payload. Then compress the stego images with quality factors of 55, 65, 75, and 85, and the JPEG compression residuals of stego sequences before and after the compression are recorded. Set the observation length l to be 300, and the numbers of residual groups m_l for residual sequences corresponding to three methods are set to 3, 5, and 5 separately according to actual conditions. Repeat the above mentioned fitting goodness test for a single image, and calculate the inspection value η of each image. The fitting goodness test results η between residual sequences caused by JPEG compressing stego images and the Poisson sequence are shown in Fig. 2(a)–(c).

In the boxplots drawn in Fig. 2 (a)–(c), the horizontal axis represents the quality factor Q of JPEG compression, and the vertical axis represents the fitting goodness test results η. The upper and lower limit of the rectangle are the upper and lower quartiles of test results separately, and the difference between the upper and lower quartile is the quartile difference IQR. The red line in the rectangle is the median of η. The two black horizontal lines at $Q_3 + 1.5IQR$ and $Q_1 - 1.5IQR$ are the cut-off points for abnormal values, known as the internal limit. The data outside the internal limits is outliers and is represented by the red '+'. Since $m_l = 3, 5, 5,$ the degrees of freedom for the χ^2 distribution are 1,3, and 3 respectively. The critical values $\chi^2_1(0.05) = 3.841$ and $\chi^2_3(0.05) = 7.815$ are shown in the blue dotted line in Fig. 2.

For Fig. 2 (a), since the DCRAS method utilizes the robustness of DCT coefficients' relationship against JPEG compression, the number of non-zero residuals caused by JPEG compression to the embedding domain is small, and the inspection values η of quality factor 75 is the minimum since this quality factor is the same as the cover images'. While in FRAS method, the feature regions located in complex areas are taken into account based on the embedding domain constructed in DCRAS method to improve the detection resistant ability, thus the non-zero residuals caused by JPEG compression in this embedding domain are more than that in DCRAS methods, and the inspection values η in Fig. 2 (b) are a little bit higher than that in Fig. 2 (a). Different from the DCRAS and FRAS methods, the DMAS method use the dither modulation method to embed messages, therefore, the distribution of inspection values η in Fig. 2 (c) is a little bit different from that in Fig. 2 (a) and (b). Because most of the fitting goodness test results are less than the critical values, it can be considered that the residual sequences caused by JPEG compressing the stego images generated by the DCRAS, FRAS and DMAS methods follow the Poisson distribution.

In addition, similar results can be obtained from the fitting goodness test at other payloads, which demonstrates that the proposed residual model is valid and reasonable. By establishing the residual model of JPEG compression based on burst errors, the influence of JPEG compression on stego sequences can be deterministically measured, thus establishing the basis for the fault-tolerant performance analysis of robust steganography methods based on "Compression-resistant Domain Constructing + RS-STC code".

3.2. Error model considering STCs decoding damage

As the error diffusion phenomenon occurs in the STCs decoding [27] when stego sequences are damaged, the effects on message

[1] Proposed by P. Bas, T. Filler, T. Pevný in ICASSP 2013, available: http://agents.fel.cvut.cz/stegodata/.

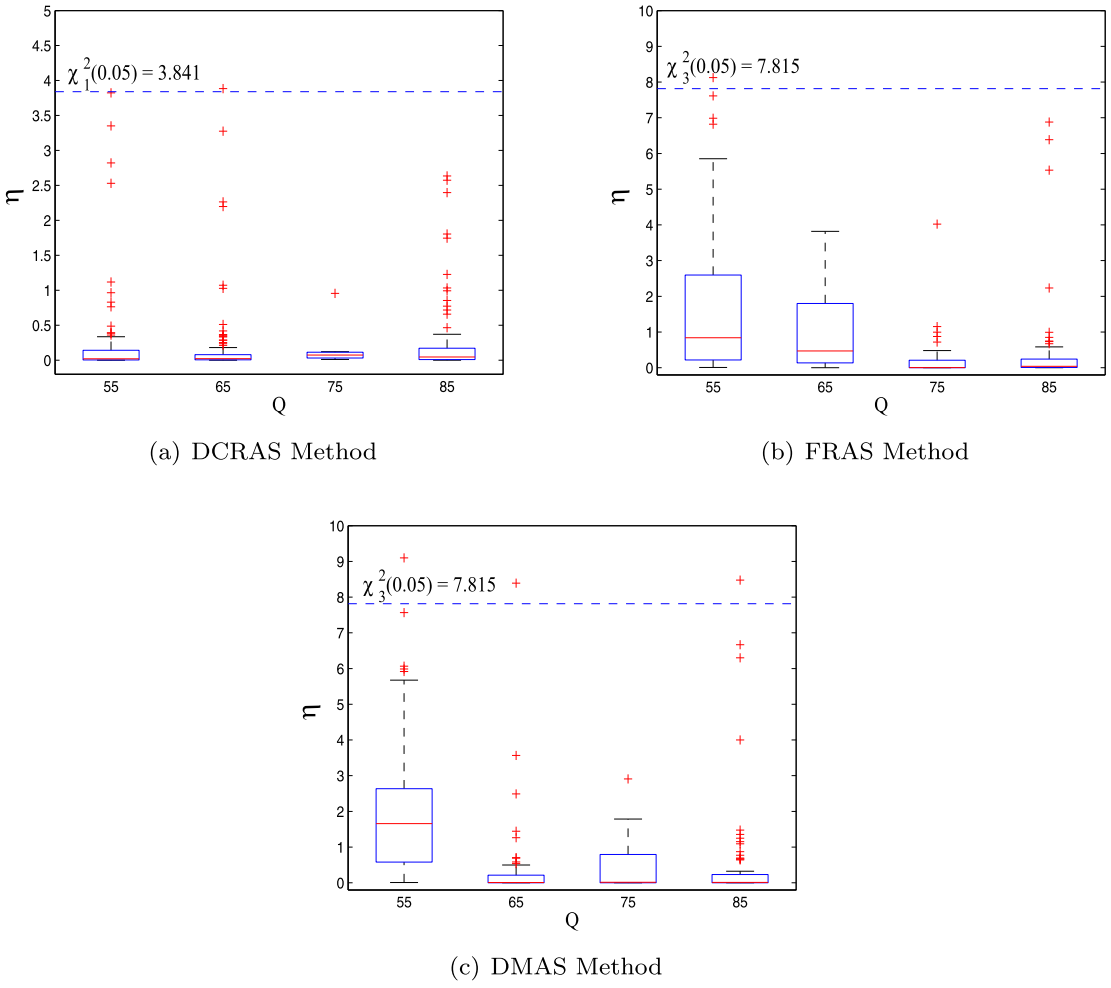


Fig. 2. Fitting goodness test results between residual sequences and Poisson distribution. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

sequences of JPEG compression and STCs decoding are further analyzed in this section. Combined with the damage gain of the STC codes, the error model considering STCs decoding damage is proposed and tested by experiments.

3.2.1. Error model establishment

The STCs coding method that minimizes embedded costs is a matrix coding method. In order to solve the problem of damage diffusion in matrix coding, the relationship between the damage rates of secret messages and that of stego sequences is analyzed in [27], and the damage gain function of matrix coding is obtained, which proves that the matrix coding is damage spreading. According to the analysis, when stego sequences are damaged, the damage rates σ_m of secret messages in matrix coding method can be expressed as follows:

$$\sigma_m = 1 - \sum_{i=1}^{m} (1-\delta)^{\gamma_i}/m, \tag{10}$$

where γ_i denots the number of 1 in the ith row (also known as the row weight), δ denotes the damage rates of stego sequences, and m is the length of secret messages. Since $\gamma_i \geq 1$, $\sigma_m > \delta$ which means the matrix coding can cause damage spreading. If the ratio between the damage rate of the secret messages and that of the stego sequences is defined as the damage gain, denoted by $\zeta = \sigma_m/\delta$, and the STCs coding has a fixed row weight $\gamma_i = \gamma$ for each row, the damage gain function can be expressed as follows:

$$\zeta_{STCs} \approx \left[1-(1-\delta)^{\gamma}\right]/\delta. \tag{11}$$

Combined with the JPEG compression residual model based on burst errors, the damage rates δ of stego sequences caused by JPEG compression is equal to vd, the damage gain ζ_{STCs} is approximately equal to $\left[1-(1-vd)^{\gamma}\right]/(vd)$, and the damage rates of secret messages σ_m is approximately equal to $1-(1-vd)^{\gamma}$.

Based on the above STCs decoding damage analysis, the JPEG compression residual model based on burst errors in Section 3.1 can be enriched. If we define the errors in stego sequences caused by JPEG compression and STCs decoding to be a series of Poisson points $\{t_i\}$ with average rate v_δ and average length d_δ, then the number of non-zero residual elements n_l in each successive stego sequences of length l obey the Poisson distribution with parameter $\lambda_\delta = v_\delta d_\delta$, whose probability density function can be expressed as follows,

$$P(n_l = k) = \frac{{\lambda_\delta}^k}{k!} e^{-\lambda_\delta}, k = 0, 1, \cdots, \tag{12}$$

where $d_\delta = d\zeta_{STCs} \approx \left[1-(1-vd)^{\gamma}\right]/v$ is the average length of burst errors, and $\lambda_\delta = \sigma_m = vd_\delta = 1-(1-vd)^{r}$ is the error rates in each successive stego sequences of length l.

3.2.2. Model validation

In order to test the rationality of the above error model considering STCs decoding damage, consider the following hypothesis

test problem:

$$\begin{cases} H_0 : F(n_l) = F_0(x) \\ H_1 : F(n_l) \neq F_0(x) \end{cases}, \tag{13}$$

where $F_0(x)$ is the distribution function of the Poisson distribution $P(n_l = k) = \left(\lambda_\delta{}^k/(k!)\right)e^{-\lambda_\delta}, k = 0, 1, \cdots,$ and $\lambda_\delta > 0$ is the parameter in the Poisson distribution.

Similar to the verification process in Section 3.1.2, the χ^2 test method is utilized to judge H_0. The RS coding parameters used in the experiments is (31,15), and the randomly generated binary sequences are used as secret messages. The parameters of DCRAS [23], FRAS [24], and DMAS [25] methods are shown in Table 1. ('\' means that there is no this kind of parameter in the corresponding method.)

Randomly select 1000 cover images generated by JPEG compression of quality factor 75, and utilize the DCRAS, FRAS, and DMAS methods to generate the corresponding stego images under 0.1bpnzAC payload. Then compress stego images with quality factors of 55, 65, 75, and 85, and extract embedded messages and perform STCs decoding. The message extraction errors caused by JPEG compression and STCs decoding are recorded. Set the observation length l to be 300, and the group numbers m_l for extracted message error sequences corresponding to three methods are all set to 5 according to actual conditions. Repeat the fitting goodness test for a single image in Section 3.1.2, and calculate the inspection value η of each image. The fitting goodness test results η between the error sequences caused by JPEG compression and STCs decoding and the Poisson sequence are shown in Fig. 3(a)–(c).

In the boxplots drawn in Fig. 3 (a)–(c), the horizontal axis represents the quality factor Q of JPEG compression, and the vertical axis represents the fitting goodness test results η. The upper and lower limit of the rectangle are the upper and lower quartiles of test results separately, and the difference between the upper and lower quartile is the quartile difference IQR. The red line in the rectangle is the median of η. The two black horizontal lines at $Q_3 + 1.5IQR$ and $Q_1 - 1.5IQR$ are the cut-off points for abnormal values, known as the internal limit. The data outside the internal limits is outliers and is represented by the red '+'. Since $m_l = 5$, the degrees of freedom for the χ^2 distribution are all equal to 3, and the critical values of three methods $\chi^2_3(0.05)$ is 7.815, as shown in the blue dotted line in Fig. 3.

Similar with Fig. 2, since the quality factor 75 is the same as that of cover images, the inspection values η in Fig. 3 (a) at quality factor 75 is the minimum. Because of the STCs decoding damage, the distribution of inspection values η in Fig. 3 (a) is different from that in Fig. 2 (a). In FRAS method, since the embedding domains are more concentrated than that of DMAS methods, the error sequences caused by STCs decoding damage distribute more concentrated as well. Thus, the distribution of inspection values η in Fig. 3 (b) is more uniform than that in Fig. 3 (a). Different from the DCRAS and FRAS methods, the DMAS method uses the dither modulation method to construct the JPEG compression resistant embedding domain, thus the distribution of error sequences caused by JPEG compression and STCs decoding damage is also a little bit different, and the distribution of inspection values η is shown in Fig. 3 (c). Because most of the fitting goodness test results are less than critical values, it can be considered that the error sequences of extracted messages caused by JPEG compression and STCs decoding in the three robust steganography methods follow the Poisson distribution.

In addition, similar results can be obtained from the fitting goodness test at other payloads, which demonstrates that the proposed error model is valid and reasonable. Based on the error model enriched by STCs decoding damage, the errors caused by JPEG compression and STCs decoding to message sequences can be deterministically measured, thus a basis can be provided for analyzing the fault-tolerant performance of robust steganography methods based on "Compression-resistant Domain Constructing + RS-STC Code".

4. Fault-tolerant performance of robust steganography methods

With the help of the proposed error model, the error rates of extracted messages after JPEG compression and STCs decoding can be expressed approximately as λ_δ which denotes the error rates in each successive stego sequences of length l. Considering the error correction performance of RS codes under a certain error rates, the probability lower bound for RS-STCs decoding to correctly extract embedded messages can be deduced, which provides a theory support for message extraction integrity to robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes".

For robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes", which utilizes the (n^*, k^*) RS codes with m^* bits in each symbol, let w_i denote the event corresponding to i binary symbol errors in a codeword of length n^* after STCs decoding, and let W_j denote the event corresponding to j code symbol errors in a received codeword after STCs decoding. If event W_j occurs, then the received code word must contain $j \le i \le m^* j$ binary symbol errors. If $i = j$, then each code symbol error must result from a single binary symbol error among its component binary symbols. Similarly, if $i = m^* j$ then each code symbol error must result from exactly m^* binary symbol errors. Let ε denote a binary symbol error occurs in a received codeword after STCs decoding, then probability of ε can be expressed as follows.

$$P(\varepsilon) = \sum_{j=0}^{n^*} \sum_{i=0}^{m^* n^*} P\left(\varepsilon, w_i, W_j\right). \tag{14}$$

Clearly, $P\left(\varepsilon, w_i, W_j\right) = 0$ for $i < j$ or $i > m^* j$.

Since the (n^*, k^*) RS codes' error correction ability is $t = (n^* - k^*)/2$, it can correct up to t erroneous symbols whose length is short than m^* and a burst error whose length is short than $m^* t$, if the locations of the error symbols are not know in advance. Let ξ_1 denote that more than t erroneous symbols occur in the received codeword with length n^* after STCs decoding, and let ξ_2 denote that a burst error whose length is more than $m^* t$ occurs in the received codeword with length n^* after STCs decoding. Ignore the situation that the length of some burst errors is more than m^* when more than one burst error occur (according to the practical statistical results, this kind of events almost never happen), then $P(\varepsilon)$ satisfies the following formula:

$$P(\varepsilon) < P(\xi_1) + P(\xi_2). \tag{15}$$

From the above formula, the upper bound of RS decoding error in RS codes with length n^* after JPEG compression and STCs decoding can be estimated. Specifically, the probability of event ξ_1 that more than t burst erroneous symbols occur in RS codes of length n^* after STCs decoding can be expressed by the following expression, with the help of conditional total probability formula,

$$\begin{aligned} P(\xi_1) &= \sum_{j=t+1}^{n^*} \sum_{i=j}^{m^* j} P\left(\varepsilon, w_i, W_j\right) \\ &= \sum_{j=t+1}^{n^*} \sum_{i=j}^{m^* j} \left[P\left(\varepsilon \middle| w_i, W_j\right) P\left(W_j | w_i\right) P(w_i)\right], \end{aligned} \tag{16}$$

where $P(\varepsilon|w_i, W_j.)$ denotes the decoding error probability when there are i erroneous bits and j erroneous code symbols in a received codeword after STCs decoding, which can be calculated by

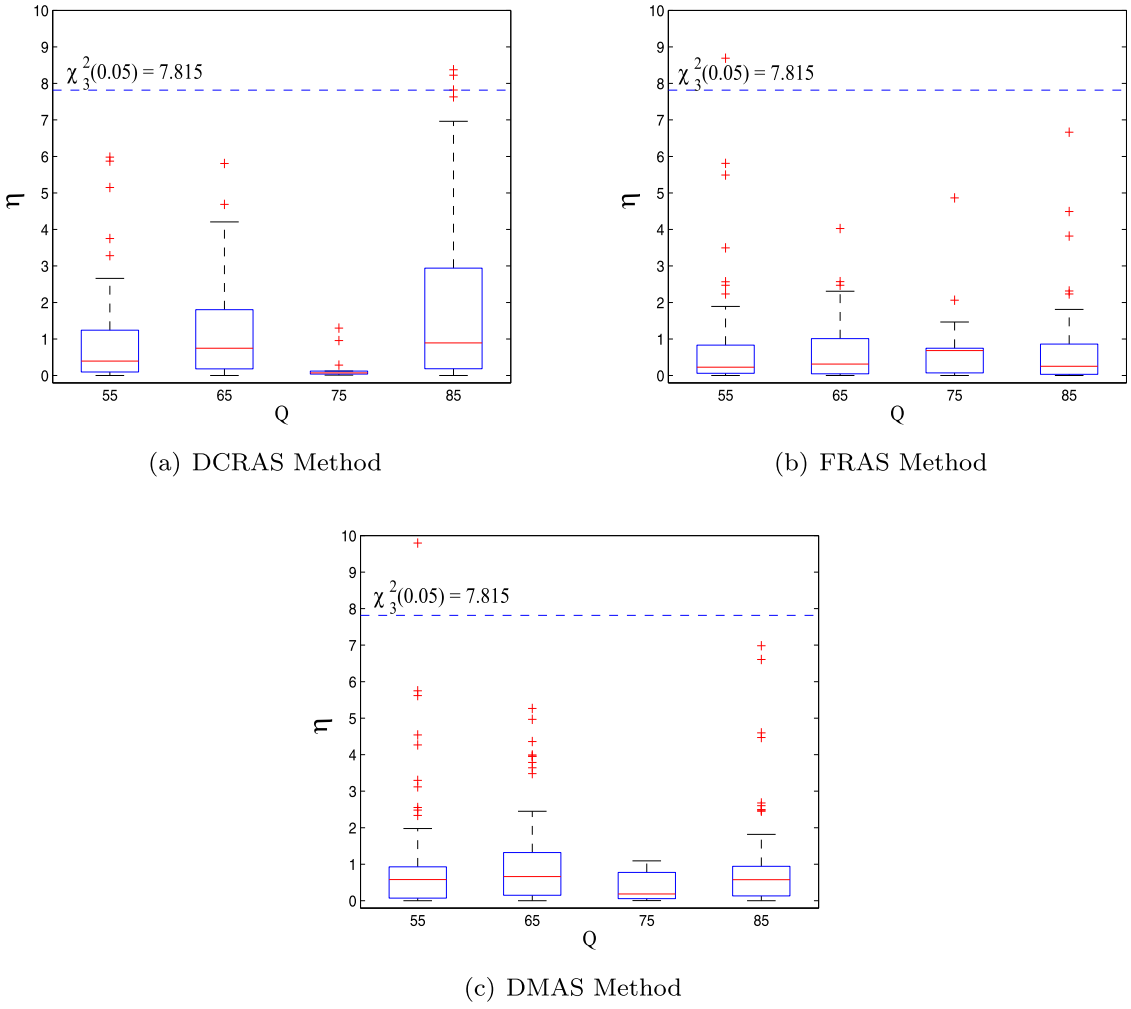


Fig. 3. Fitting goodness test results between error sequences and Poisson distribution. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

the following formula:

$$P\left(\varepsilon \middle| w_i, W_j\right) = \frac{i}{m^* n^*}. \tag{17}$$

Suppose that every erroneous bit is independent of one another in the received codes sequences, then $P(w_i)$ can be expressed as following:

$$P(w_i) = C^i_{m^* n^*} P^i(\varepsilon_d)[1 - P(\varepsilon_d)]^{m^* n^* - i}, \tag{18}$$

where $P(\varepsilon_d)$ is the error rates in the received codes sequences after JPEG compression and STCs decoding, and $P(\varepsilon_d) \approx \lambda_\delta$.

In addition, the following equation can be derived from [18],

$$P\left(W_j | w_i\right) = \frac{C^j_{n^*}}{C^i_{m^* n^*}} \sum_{q=0}^{r} \left[(-1)^q C^q_j C^i_{m(j-q)}\right], \tag{19}$$

where $r = \lfloor j - i/m^* \rfloor$, and $\lfloor \bullet \rfloor$ denotes taking integers downwardly.

Suppose that every erroneous bit is independent of one another in the received codes sequences. Then the probability of event ξ_2, that a burst error whose length is more than $m^* t$ occurs in the received codes of length n^* after STCs decoding can be calculated by the following expression,

$$P(\xi_2) = \sum_{q=m^* t}^{m^* n^*} \left[(m^* n^* - q + 1) P^q(\varepsilon_d)(1 - P(\varepsilon_d))^{m^* n^* - q}\right], \tag{20}$$

where $P(\varepsilon_d)$ is the error rates in the received codes sequences after JPEG compression and STCs decoding, and $P(\varepsilon_d) \approx \lambda_\delta$.

Utilizing the error model based on burst errors and STCs decoding damage, the parameter λ_δ, $P(\xi_1)$, and $P(\xi_2)$ can be calculated. Then the upper bound of decoding error $P(\varepsilon)$ in the received code of length n^* can also be determined. Let n_p denote the length of messages, then the total group number n_r is $\lceil n_p / k^* \rceil$, and the total length of encode sequences n_{rs} is $\lceil n_p / k^* \rceil n^* m^*$, where $\lceil \bullet \rceil$ denotes the operation of taking integers upwardly. The message correct extraction probability $P(\zeta)$ after JPEG compressed and STCs decoding can be expressed as follows.

$$P(\zeta) = (1 - P(\varepsilon))^{n_r}. \tag{21}$$

Plug the upper bound of $P(\varepsilon)$ in Eq. (15) into Eq. (21), then

$$P(\zeta) > (1 - P(\xi_1) - P(\xi_2))^{n_r}. \tag{22}$$

From formula (22), the error correction performance under different RS encoding parameters for robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes" can be deduced. Suppose the group numbers of RS codes n_r is 100, then under different error rates $P(\varepsilon_d)$ and encoding parameters, the probability lower bounds $P(\zeta)$ of message correct extraction after JPEG compression and STCs decoding are draw in Fig. 4.

Specifically, because of the length limitation of secret messages in practical application, we can assume that secret messages are correctly extracted by RS decoding in the condition of $n_p(1 -$

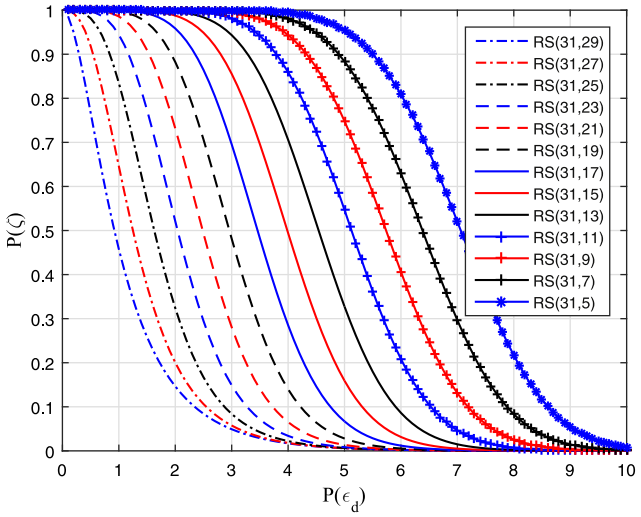


Fig. 4. Theoretical lower bounds of RS codes' decoding performance.

$P(\zeta)) < 1$. Since $n_p(1-P(\zeta)) < n_p(1-(1-P(\xi_1)-P(\xi_2))^{n_r})$, the probability $P(\zeta)$ of messages correct extraction by RS decoding can be calculated by the following expression:

$$P(\zeta) = \begin{cases} (1-P(\varepsilon))^{n_r}, & n_p(1-(1-P(\xi_1)-P(\xi_2))^{n_r}) \geq 1 \\ 1, & n_p(1-(1-P(\xi_1)-P(\xi_2))^{n_r}) < 1 \end{cases} \quad (23)$$

In conclusion, the error model based on burst errors and STCs decoding damage is utilized in this section, combined with the error correction theory of RS codes on a bursty-noise channel, to analyze the fault-tolerant performance of the robust steganography methods, and the probability lower bound for decoding methods to correctly extract embedded messages is deduced. Thus, a theory support for message extraction integrity is provided to the robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes".

5. Experimental verification and parameter recommendation

In this section, the validity and rationality of the fault-tolerant performance analysis are verified by experiments, and the recommended RS encoding parameters are given for different conditions.

5.1. Experiments settings

In order to verify the analysis results of the fault-tolerant performance, experiments are conducted in the condition of Matlab R2015a. The original images of the experiments come from BOSSbase-1.01 image database (10,000 grayscale images), and involve a wide range of image content, including natural scenery, artificial facilities, human portraits, and so on. Then the original images are compressed under quality factors of 65, 75 and 85 to generate alternative cover sets ($10,000 \times 3$ images). Randomly select 2000 images separately under different quality factors, and obtain 2000×3 cover images. Under payloads ranging from 0.01bpnzAC to 0.1bpnzAC, 4 groups of RS coding parameters, and 3 different quality factors, generate the stego images utilizing the DCRAS [23], FRAS [24], and DMAS [25] steganography methods separately, thus obtaining $2000 \times 3 \times 10 \times 4 \times 3 \times 3 = 2,160,000$ stego images. The parameter settings of steganography methods are shown in Table 1, and the experiment parameter settings are shown in Table 3.

Table 3
Experiment parameter settings.

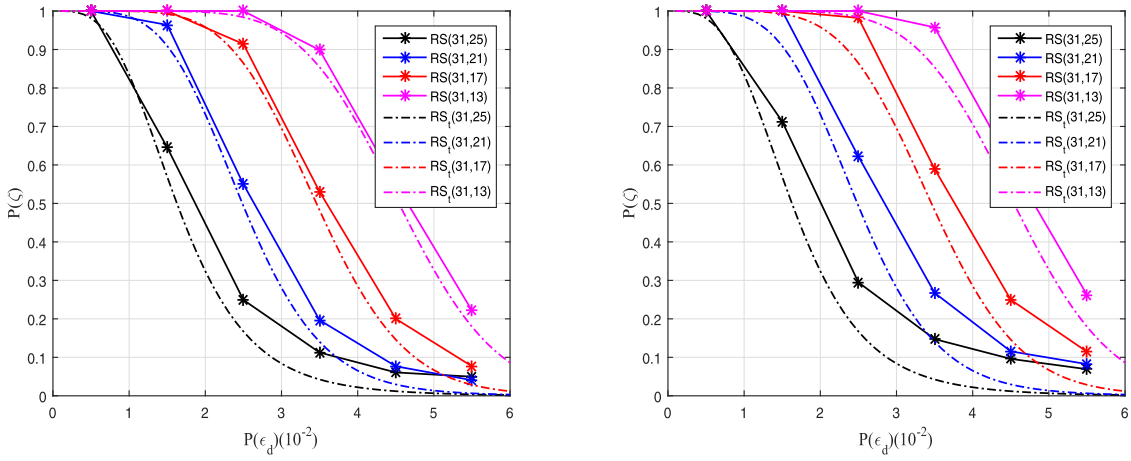
Experiment parameters	Experiment settings
Image source	BOSSbase 1.01 image database (10,000 images)
Image size	512×512
Image color	Grayscale
Quality factors	65/ 75/ 85
Alternative cover sets	$10,000 \times 3$
Number of cover images	2000 (Randomly selected) $\times$ 3
Payloads	0.01, 0.02, ..., 0.1 bpnzAC
Secret messages	Randomly generated binary sequences
RS coding parameters	(31,25)/ (31,21)/ (31,17)/ (31,13)
Steganography methods	DCRAS [23]/ FRAS [24]/ DMAS [25]
Number of stego images	$2000 \times 3 \times 10 \times 4 \times 3 \times 3$

5.2. Fault-tolerant performance verification

In this section, the rates of stego images from which embedded messages can be completely correctly extracted after JPEG compression are calculated for the DCRAS, FRAS, and DMAS methods, and compared with theoretical values, that is, the probability lower bounds for decoding methods to extract massages correctly, in order to verify the conclusion in Section 4.

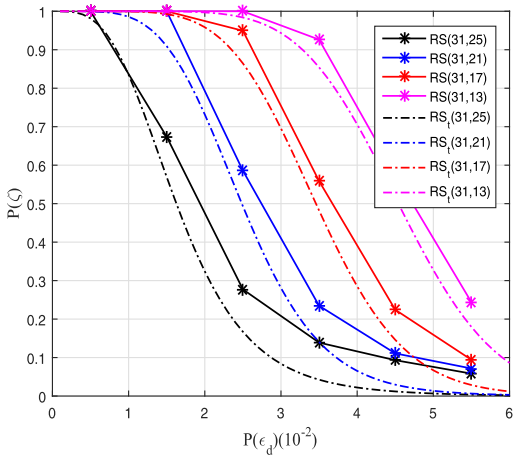
For the 2,160,000 stego images generated by different payloads, RS coding parameters, robust steganography methods shown in Table 3, conduct JPEG compression with quality factors 65, 75 and 85 according to the quality factors used to generate the corresponding stego images, thus obtain a total of 2,160,000 compressed stego images. For each robust steganography methods, extract embedded messages from the corresponding compressed images, and calculate the error rates $P(\varepsilon_d)$ of extracted messages after JPEG compression and STCs decoding, then divide stego images according to $P(\varepsilon_d)$ by the interval of 1×10^{-2}. In each interval, calculate the rates of stego images from which the embedded messages can be correctly extracted p_c under different RS encoding parameters. The relationship between p_c and message error rates $P(\varepsilon_d)$ under different RS coding parameters is shown as the solid line in Fig. 5. Meanwhile, the lower bounds of message correct extraction probability $P(\zeta)$ by RS decoding under different error rates $P(\varepsilon_d)$ are shown as the dotted line in Fig. 5.

In Fig. 5, there are a total of 6 intervals of error rates $P(\varepsilon_d)$ in every groups of experiments according to the statistical results. In each interval, the ratio between stego images from which embedded messages can be correctly extracted and all stego images p_c is higher than the lower bound of message correct extraction probability $P(\zeta)$. For example, in Fig. 5 (a), there are 47,613 images in which the error rates $P(\varepsilon_d)$ of extracted messages after JPEG compression and STCs decoding is between 2.1×10^{-2}–3.0×10^{-2}, and the number of stego images from which messages can be correctly extracted is 11,903, then the ratio p_c of this interval is approximately equal to 0.25 which is higher than $P(\zeta) = 0.1672$ in this interval. From Fig. 5 (a)–(c), it can be concluded that the rates p_c of stego images from which messages can be extracted correctly after JPEG compression are above the lower bounds of message correct extraction probability $P(\zeta)$, under different payloads, quality factors of JPEG compression, and RS encoding parameters for the three robust steganography methods. Therefore, it proves that the practical fault-tolerant results of steganography methods based on "Compression-resistant Domain Constructing + RS-STC Codes" consists with the theoretical derivation results. In addition, the experimental results show that the proposed error model can depict the errors in stego images caused by JPEG compression and STCs decoding reasonably, and verifies the validity and rationality of the error model and fault-tolerant performance analysis.



(a) DCRAS Method

(b) FRAS Method



(c) DMAS Method

Fig. 5. Rates of message correct extraction.

5.3. Determination of RS encoding parameters

In this section, the error conditions in stego elements caused by JPEG compression and STCs decoding are presented and analyzed. The recommend RS encoding parameters are given in the condition of different error rates $P(\varepsilon_d)$, utilizing the conclusion deduced by the fault-tolerant performance analysis, in order to guarantee the integrity of message extraction after JPEG compression.

The content of this section can be devided into three parts: (1) errors in stego sequences caused by JPEG compression, (2) errors in stego sequences caused by JPEG compression and STCs decoding, and (3) RS encoding parameter recommendation based on message extraction integrity.

5.3.1. Errors in stego sequences caused by JPEG compression

For the 2000 randomly selected cover images of quality factors 65, 75, and 85 shown in Table 3, conduct JPEG recompression with quality factors 65, 75 and 85, respectively. Then calculate the changing rates of DCT coefficients $\delta_d = n_{cd}/n_d$, where n_{cd} is the number of changed DCT coefficients, and n_d is the number of all DCT coefficients. The average changing rates of DCT coefficients of all 2000 images, that is, the error rates of cover elements caused by JPEG compression with different quality factors λ', are shown in Table 4.

Table 4
Error rates of cover sequences caused by JPEG compression.

Quality factor Q	65	75	85
Error Rates λ'	2.21×10^{-4}	2.73×10^{-4}	8.36×10^{-4}

Table 5
Error rates of stego sequences caused by JPEG compression.

λ'	Q		
	65	75	85
λ'_1 of DCRAS [23]	1.21×10^{-5}	8.75×10^{-6}	3.38×10^{-5}
λ'_2 of FRAS [24]	7.37×10^{-4}	6.61×10^{-4}	8.11×10^{-4}
λ'_3 of DMAS [25]	1.61×10^{-5}	2.62×10^{-5}	2.94×10^{-5}

The error rates λ' in Table 4 are the average results of 2000 cover images under different quality factors. Therefore, the error rates λ' can be used as the error rates of cover sequences caused by JPEG compression determined by quality factors Q.

For the 2000 randomly selected cover images of quality factors 65, 75, and 85 shown in Table 3, repeat the recompression operation to the JPEG compression resistant domain constructed in DCRAS, FRAS, and DMAS methods, and calculate the changing rates λ_1', λ_2', and λ_3' of elements in the corresponding embedding domains caused by JPEG compression with different quality factors, which are shown in Table 5.

Table 6
Error rates caused by JPEG compression and STCs decoding (DCRAS [23]) (10^{-3}).

Q	α									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
65	2.13	2.02	2.37	1.83	4.80	5.56	6.66	5.18	7.06	6.45
75	2.32	3.80	4.47	2.50	6.01	7.05	5.76	6.14	8.73	10.88
85	1.96	5.68	10.47	11.84	15.96	16.22	20.69	24.90	30.17	30.32

Table 7
Error rates caused by JPEG compression and STCs decoding (FRAS [24]) (10^{-2}).

Q	α									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
65	0.96	1.31	1.96	1.57	1.73	2.23	2.08	2.77	3.55	6.14
75	1.03	1.59	1.76	2.21	2.80	3.96	4.65	5.04	5.33	6.52
85	2.35	2.94	3.76	3.41	4.62	5.08	5.15	5.87	6.11	6.60

Table 8
Error rates caused by JPEG compression and STCs decoding (DMAS [25]) (10^{-2}).

Q	α									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
65	0.37	0.40	0.56	0.72	0.85	1.04	1.22	1.61	2.23	2.44
75	0.43	0.49	0.51	0.88	0.96	1.30	1.99	2.44	2.57	2.72
85	0.68	0.76	0.82	1.41	1.57	1.93	2.03	2.54	2.65	2.88

Table 9
Recommended RS encoding parameters (n^*, k^*) (DCRAS [23]).

Q	λ									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
65	(31,29)	(31,29)	(31,29)	(31,29)	(31,27)	(31,27)	(31,27)	(31,27)	(31,27)	(31,27)
75	(31,29)	(31,29)	(31,29)	(31,29)	(31,27)	(31,27)	(31,27)	(31,27)	(31,25)	(31,25)
85	(31,29)	(31,27)	(31,25)	(31,25)	(31,23)	(31,21)	(31,19)	(31,17)	(31,15)	(31,15)

Table 10
Recommended RS encoding parameters (n^*, k^*) (FRAS [24]).

Q	λ									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
65	(31,25)	(31,23)	(31,21)	(31,21)	(31,21)	(31,19)	(31,19)	(31,17)	(31,13)	(31,5)
75	(31,25)	(31,21)	(31,21)	(31,19)	(31,17)	(31,13)	(31,11)	(31,9)	(31,7)	(31,5)
85	(31,19)	(31,15)	(31,13)	(31,13)	(31,11)	(31,9)	(31,9)	(31,5)	(31,5)	(31,3)

Table 11
Recommended RS encoding parameters (n^*, k^*) (DMAS [25]).

Q	λ									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
65	(31,29)	(31,29)	(31,29)	(31,27)	(31,27)	(31,25)	(31,23)	(31,21)	(31,19)	(31,19)
75	(31,29)	(31,29)	(31,29)	(31,27)	(31,25)	(31,23)	(31,19)	(31,19)	(31,17)	(31,17)
85	(31,27)	(31,27)	(31,27)	(31,23)	(31,21)	(31,21)	(31,19)	(31,17)	(31,17)	(31,17)

The error rates in Table 5 is the 2000 images' average results of the JPEG compression resistant domain constructed in DCRAS, FRAS, and DMAS methods. Therefore, the error rates can be used as the error rates of element sequences in constructed embedding domains caused by JPEG compression under different quality factors Q, and can be used as the approximate value of λ in Section 3.1.1 as well.

5.3.2. Errors in stego sequences caused by JPEG compression and STCs decoding

For the 2,160,000 stego images generated by different payloads, RS coding parameters, robust steganography methods shown in Table 3, conduct JPEG compression with quality factors 65, 75 and 85 according to quality factors used to generate the corresponding stego images, thus obtain a total of 2,160,000 compressed stego images. Then extract embedded messages by STCs decoding (without RS decoding) and calculate the error rates of extracted messages $r_{error} = n_{error}/n_m$, where n_{error} is the number of error bits, and n_m is message length. The average error rates R_{error} of extracted messages in different robust steganography methods with different quality factors and payloads, that is, the error rates of extracted secret messages caused by JPEG compression and STCs decoding λ'_{δ}, which are shown in Tables 6–8.

The error rates λ'_{δ} of extracted messages in Table 6–8 is the average error rates calculated from the DCRAS, FRAS, and DMAS methods under different quality factors and payloads. Therefore,

the above results can be used as the error rates of message sequences determined by JPEG compression resistant domains, quality factor Q, and payloads α, and also can be used as the approximate value of $P(\varepsilon_d)$.

5.3.3. RS Encoding parameter recommendation based on message extraction integrity

From the error conditions caused by JPEG compression and STCs decoding, the bit error rates $P(\varepsilon_d)$ in received message sequences affected by steganography methods, payloads, and quality factors of JPEG compression can be evaluated. According to Eq. (22) and Fig. 4, the recommended RS encoding parameters can be determined when the probability of message correct extraction $P(\zeta)$ is given, in order to guarantee the correct extraction integrity of embedded secret messages.

Use the error rates λ'_{δ} in Tables 6–8 as the evaluates of the error rates in received messages $P(\varepsilon_d)$. Given the probability of message correct extraction $P(\zeta) > 0.99$, the recommended RS encoding parameters can be determined for different steganography methods, different payloads, and different quality factors, as shown in Table 9–11. (The code length of RS codes n^* is set to 31.)

In Table 9–11, the recommended RS encoding parameters are given for different steganography methods, payloads, and quality factors according to the derived results, thus providing a theory support to the robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes" for message extraction integrity.

6. Conclusion

In order to solve the problem of message extraction integrity in current robust adaptive steganography based on "Compression-resistant Domain Constructing + RS-STC Codes", an error model based on burst errors and STCs decoding damage is established to model the errors caused by JPEG compression and STCs decoding. Based on the error model, the fault-tolerant performance of the robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes" is deduced, that is, the probability lower bound for decoding methods to correctly extract embedded messages. Furthermore, the recommended encoding parameters are given in different conditions. Experimental results under different payloads and RS encoding parameters demonstrate that the fault-tolerant performance of the DCRAS, FRAS, and DMAS methods accords with the deduced theoretical bounds, thus verifying the validity and rationality of the error model and fault-tolerant performance analysis. Consequently, a theory support for the integrity of message extraction is provided to the robust steganography based on "Compression-resistant Domain Constructing + RS-STC Codes". In future research, we will focus on the efficiency improvement of coding methods and robust steganography methods.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No. U1636201, U1636219, 61572052, 61672354, and 61772549), the National Key Research and Development Program of China (No. 2016YFB0801303, and 2016QY01W0105), the Key Science and Technology Research and Development Program of Henan Province of China (No. 162102210032).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sigpro.2018.01.011.

References

[1] N.D.W. Cahyani, N.H.A. Rahman, W.B. Glisson, K.K.R. Choo, The role of mobile forensics in terrorism investigations involving the use of cloud storage service and communication apps, Mobile Netw. Appl. 22 (2) (2017) 240–254, doi:10.1007/s11036-016-0791-8.

[2] J. Fridrich, T. Pevný, J. Kodovský, Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities, in: Proc. The 9th ACM Workshop on Multimedia and Security, 2007, pp. 3–14, doi:10.1145/1288869.1288872.

[3] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications, Cambridge University Press, 2009, doi:10.1109/MSP.2011.941841.

[4] M. Amini, M.O. Ahmad, M.N.S. Swamy, A new locally optimum watermark detection using vector-based hidden markov model in wavelet domain, Signal Process 137 (2017) 213–222, doi:10.1016/j.sigpro.2017.01.019.

[5] T. Pevný, T. Filler, P. Bas, Using high-dimensional image models to perform highly undetectable steganography, in: Proc. The 12th ACM International Workshop on Information Hiding, 2010, pp. 161–177, doi:10.1007/978-3-642-16435-4_13.

[6] V. Holub, J. Fridrich, Designing steganographic distortion using directional filters, in: Proc. The IEEE Workshop on Information Forensic and Security, 2012, pp. 234–239, doi:10.1109/WIFS.2012.6412655.

[7] V. Holub, J. Fridrich, Digital steganography using universal distortion function, in: Proc. The ACM Workshop on Information Hiding and Multimedia Security, 2013, pp. 59–68, doi:10.1145/2482513.2482514.

[8] T. Filler, J. Fridrich, Design of adaptive steganographic schemes for digital images, in: Proc. SPIE - Electronic Imaging, Media Watermarking, Security and Forensics of Multimedia XIII, 7880, 2011, pp. 1–14, doi:10.1117/12.872192.

[9] L. Guo, J. Ni, Y. Shi, An efficient JPEG steganographic scheme using uniform embedding, in: Proc. The 4th IEEE International Workshop on Information Forensics and Security, 2012, pp. 2–5, doi:10.1109/WIFS.2012.6412644.

[10] T. Filler, J. Judas, J. Fridrich, Minimizing additive distortion in steganography using syndrome-trellis codes, IEEE Trans. Inf. Foren Sec. 6 (3) (2011) 920–935, doi:10.1109/TIFS.2011.2134094.

[11] C. Deng, X. Gao, X. Li, D. Tao, A local tchebichef moments-based robust image watermarking, Signal Process 89 (8) (2009) 1531–1539, doi:10.1016/j.sigpro.2009.02.005.

[12] M. Amini, M.O. Ahmad, M.N.S. Swamy, Digital watermark extraction in wavelet domain using hidden markov model, Multimed. Tools Appl. 75 (21) (2016) 3731–3749, doi:10.1007/s11042-016-3975-0.

[13] C. Wang, X. Wang, Z. Xia, Geometrically invariant image watermarking based on fast radial harmonic fourier moments, Signal Process-Image 45 (C) (2016) 10–23, doi:10.1016/j.image.2016.03.007.

[14] M. Amini, M.O. Ahmad, M.N.S. Swamy, A robust multibit multiplicative watermark decoder using vector-based hidden markov model in wavelet domain, IEEE Trans. Circuits Syst Video Technol. 99 (2016) 1–12, doi:10.1109/TCSVT.2016.2607299.

[15] H. Sadreazami, M.O. Ahmad, M.N.S. Swamy, Optimum multiplicative watermark detector in contourlet domain using the normal inverse gaussian distribution, in: Proc. The IEEE International Symposium on Circuits and Systems, 2015, pp. 1050–1053, doi:10.1109/ISCAS.2015.7168817.

[16] H. Sadreazami, M.O. Ahmad, M.N.S. Swamy, Multiplicative watermark decoder in contourlet domain using the normal inverse gaussian distribution, IEEE Trans. Multimed. 18 (2) (2016) 196–207, doi:10.1109/TMM.2015.2508147.

[17] J.S. Tsai, W.B. Huang, Y.H. Kuo, M.F. Horng, Joint robustness and security enhancement for feature-based image watermarking using invariant feature regions, Signal Process 92 (6) (2012) 1431–1445, doi:10.1016/j.sigpro.2011.11.033.

[18] H. Sadreazami, M.O. Ahmad, M.N.S. Swamy, A robust quantization-based image watermarking scheme in the wavelet-based contourlet domain, Comput. Electr. Eng. (2016) 1–4, doi:10.1109/CCECE.2016.7726727.

[19] M. Amini, M.O. Ahmad, M.N.S. Swamy, A new blind wavelet domain watermark detector using hidden markov model, in: Proc. The IEEE International Symposium on Circuits and Systems, 2014, pp. 2285–2288, doi:10.1109/ISCAS.2014.6865627.

[20] M. Amini, H. Sadreazami, M.O. Ahmad, M.N.S. Swamy, A blind multiplicative watermark detector using vector-based hidden markov model, in: Proc. The IEEE Midwest Symposium on Circuits and Systems, 2017.

[21] Z. Shao, Y. Shang, Y. Zhang, X. Liu, G. Guo, Robust watermarking using orthogonal fourier-mellin moments and chaotic map for double images, Signal Process 120 (2016) 522–531, doi:10.1016/j.sigpro.2015.10.005.

[22] Y. Zhang, X. Luo, C. Yang, D. Ye, F. Liu, A framework of adaptive steganography resisting JPEG compression and detection, Secur. Commun. Netw. 9 (15) (2016) 2957–2971, doi:10.1002/sec.1502.

[23] Y. Zhang, X. Luo, C. Yang, D. Ye, F. Liu, A JPEG-compression resistant adaptive steganography based on relative relationship between DCT coefficients, in: Proc. The IEEE 10th International Conference on Availability, Reliability and Security, 2015, pp. 461–466, doi:10.1109/ARES.2015.53.

[24] Y. Zhang, X. Luo, C. Yang, F. Liu, Joint JPEG compression and detection resistant performance enhancement for adaptive steganography using feature regions selection, Multimed. Tools Appl. 76 (3) (2017) 3649–3668, doi:10.1007/s11042-016-3914-0.

[25] Y. Zhang, X. Zhu, C. Qin, C. Yang, X. Luo, Dither modulation based adaptive steganography resisting JPEG compression and statistic detection, Multimed Tools Appl (3) (2017) 1–23, doi:10.1007/s11042-017-4506-3.

[26] J. Mao, W. Sheng, Y. Hu, G. Xiao, Z. Qu, X. Niu, L. Zhu, Research on watermarking payload under the condition of keeping JPEG image transparency, Multimed. Tools Appl. 76 (6) (2017) 8423–8448, doi:10.1007/s11042-016-3477-0.
[27] W. Liu, G. Liu, Y. Dai, Damage-resistance matrix embedding framework: the contradiction between robustness and embedding efficiency, Secur. Commun. Netw. 8 (9) (2014) 1636–1647, doi:10.1002/sec.1111.
[28] F.J. MacWilliams, N.J.A. Sloane, The Theory of Rrror Correcting Codes, Elsevier, 1977.
[29] W.J. Ebel, W.H. Tranter, The performance of reed-solomon codes on a bursty-noise channel, IEEE Trans. Commun 43 (234) (1995) 298–306, doi:10.1109/26.380048.