2010 Second International Workshop on Education Technology and Computer Science

Selecting Cover for Image Steganography by Correlation Coefficient

Yifeng Sun, Fenlin Liu
Zhengzhou Information Science and Technology Institute
Zhengzhou, China
yfsun001@163.com

***Abstract*—The security of steganographic system is improved by selecting cover. The cover data are modeled as Gauss-Markov process, where the correlation coefficient of two arbitrary data elements is the exponent of correlation parameter. The KL divergence and Bhattacharyya distance of Spread Spectrum steganographic system increase with the correlation parameter. Thus the cover with smaller correlation parameter is selected to improve security. For image spatial domain steganography, the correlation parameter of image data is calculated by a special exponential model of correlation coefficients and the least squares estimator. Experiments show that the cover selection method is efficient on improving the security of image steganoagraphy.**

Keywords-Steganography; Security; Steganalysis; information hiding; Image Processing

I. INTRODUCTION

Steganography is to hide the presence of communication by embedding messages in innocuous cover. There have been a number of practical steganography methods. But with the development of steganalysis techniques which detect the presence of steganography by statistical analysis, current steganography methods face the challenge of security.

This results to the research on steganography security. Cachin [1] defined the security of a steganographic system by Kullback-Leibler(KL) divergence (also known as relative entropy) from a information theoretic point of view. Sullivian[2] modeled cover data as a Markov chain and utilized the divergence of empirical matrices to measure the security of steganography. Korzhik [3] used Bhattacharyya distance as the gauge of steganography security. Based on the benchmark of steganography security, scholars investigate the methods of improving the security of steganographic system. Wang [4] discussed the construction of perfectly secure steganographic system under the *N*-dimensional Gaussian distribution cover. The literature [5-7] proposed embedding methods with statistical restoration of the first-order and second-order statistics. The literature [8-11] proposed embedding methods based on linear and nonlinear codes which can decrease statistical change.

In steganography, the sender can also choose a cover image that results in the least detectable stego image to improve the security of steganographic system. Kharrazi[12] investigated the problem by experiments, but lacks theoretical analysis. In this paper, we propose a cover selection method based on correlation coefficient. We model cover data as Gauss-Markov process and find that the KL divergence and Bhattacharyya distance of Spread Spectrum steganographic system increase with the correlation coefficients of cover data. The correlation coefficient of two arbitrary data elements in Gauss-Markov process is the exponent of correlation parameter. Thus the cover with smaller correlation parameter is better. For spatial domain image steganography, the correlation coefficient of image pixel can be modeled by a special exponent of correlation parameter. The correlation parameter of image pixel is calculated by the least squares estimator. Experiments show that the cover selection method is efficient on improving the security of steganoagraphy.

II. IMPROVING STEGANOGRAPHY SECURITY BASED ON COVER SELECTION

In general, one would argue that the "better" the embedding technique, the less likely that the stego would be detected by the attacker. But other than the choice of the embedding technique, the sender has the freedom to choose any cover for the embedding process. If he chooses the cover that results in the least detectable stego, the secret communication are also successful. We select cover based on the correlation characteristic in cover data.

A. Principle

We model cover data as a realization of a Gauss stationary process $\mathbf{C} = (C_1, C_2, \cdots, C_n)$ where n is the number of cover data elements. Denote the joint probability distribution of $\mathbf{C}$ as $P_{\mathbf{C}} = N(\mathbf{m}, \mathbf{R}_{\mathbf{C}})$, where $\mathbf{m} = (m, m, \cdots, m)'$ is mean vector, and $\mathbf{R}_{\mathbf{C}}$ is the covariance matrix. Here restrict

$$\mathbf{R}_{\mathbf{C}} = \sigma_c^2 \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{bmatrix} \tag{1}$$

where σ_c^2 is the variance of marginal distribution, and ρ is the parameter which reflects the degree of cover data dependency (correlation characteristic). $\rho \in [0,1)$. The correlation coefficient between two adjacent elements C_i and C_{i+1} is ρ, and the correlation coefficient between two arbitrary elements

This work is supported by Natural Science Foundation of China (60970141, 60902102)

978-0-7695-3987-4/10 $26.00 © 2010 IEEE
DOI 10.1109/ETCS.2010.33

C_i and C_j is equal to $\rho^{|j-i|}$, where $1 \le i, j \le n$. The signal with the above Gauss distribution is also named as Gauss-Markov signal [13]. In this signal model, the correlation coefficient between two arbitrary elements is the exponent of the parameter ρ. So the correlation characteristic in data is easily measured by the single parameter ρ .The bigger ρ, the stronger correlation. Here ρ is named as correlation parameter.

We assume that the steganography adopts Spread Spectrum embedding. SSIS [14] and Cox' method [15] are the representative of SS embedding. We assume the sender embeds message into each cover elements, that is to say, with maximum embedding rate. Under the above assumption, Spread Spectrum steganography equals to adding Gauss white noise $\mathbf{W} = (W_1, W_2 \cdots, W_n)$ to cover $\mathbf{C} = (C_1, C_2, \cdots, C_n)$. The stego is denoted as $\mathbf{S} = \mathbf{C} + \mathbf{W}$. Let the joint probability distribution of $\mathbf{W}$ be $P_{\mathbf{W}} = N(\mathbf{0}, \mathbf{R}_{\mathbf{W}})$,

$$\mathbf{R}_{\mathbf{W}} = \sigma_w^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2}$$

where σ_w^2 is the variance of embedding noise. The joint probability distribution distribution of $\mathbf{S}$ is $P_{\mathbf{S}} = N(\mathbf{m}, \mathbf{R}_{\mathbf{S}})$,

$$\mathbf{R}_{\mathbf{S}} = \sigma_c^2 \begin{bmatrix} 1+\sigma_w^2/\sigma_c^2 & \rho & \cdots & \rho^{n-1} \\ \rho & 1+\sigma_w^2/\sigma_c^2 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1+\sigma_w^2/\sigma_c^2 \end{bmatrix} \tag{3}$$

Let $\Gamma = \sigma_w^2 / \sigma_c^2$. Γ reflects the intensity of embedding.

KL divergence is usually the measure of steganography security [1]. For two multivariate Gaussian distribution $P_i = N(\mathbf{m}_i, \mathbf{R}_i), i = 1,2$, KL divergence between P_1 and P_2 is [16]

$$D_{KL}(P_1, P_2) = \frac{1}{2} \ln \frac{\det(\mathbf{R}_2)}{\det(\mathbf{R}_1)} + \frac{1}{2} \mathrm{tr}\left\{\mathbf{R}_1[\mathbf{R}_2^{-1} - \mathbf{R}_1^{-1}]\right\} + \frac{1}{2} \mathrm{tr}\left\{\mathbf{R}_2^{-1}[\mathbf{m}_1 - \mathbf{m}_2][\mathbf{m}_1 - \mathbf{m}_2]'\right\} \tag{4}$$

where tr() represents the trace of a matrix, and det() represents the determinant of a matrix. Thus the KL divergence between $P_{\mathbf{C}}$ and $P_{\mathbf{S}}$ is

$$D_{KL}(P_{\mathbf{C}}, P_{\mathbf{S}}) = \frac{1}{2} \ln \frac{\det(\mathbf{R}_{\mathbf{S}})}{\det(\mathbf{R}_{\mathbf{C}})} + \frac{1}{2} \mathrm{tr}\left\{\mathbf{R}_{\mathbf{C}}[\mathbf{R}_{\mathbf{S}}^{-1} - \mathbf{R}_{\mathbf{C}}^{-1}]\right\} \tag{5}$$

Let σ_c^2 be 16, and n be 100. From (5), we numerically calculate KL divergence when $0 \le \rho < 1$ and $0 < \Gamma \le 1$. Fig.1 shows that the value of KL divergence varies with ρ and Γ. We find that the bigger the correlation parameter ρ, KL divergence is larger. The security of steganography is poor in the sense of KL divergence.

So we think that the correlation characteristic inherent in multimedia cover data has an important effect on the security of steganography. The correlation characteristic in cover data is generally destroyed by steganography. Many steganalysis methods, such as SPA[17] and Shi's method[18], exploit the fact. Thus if the sender selects the cover of weaker correlation which corresponds to smaller ρ, the above destruction would be not distinct, and the security of steganography would be improved.

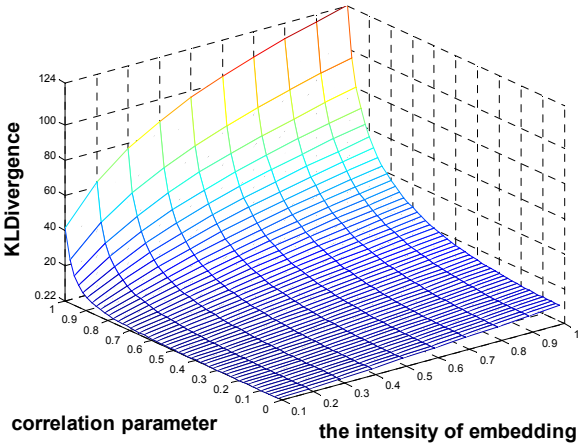

Fig1. KL divergence with ρ and Γ, under $\sigma_c^2 = 16$ and $n = 100$

B. Bhattacharyya Distance as the security measure

Bhattacharyya Distance (BD) can also be used as the measure of steganography security[3]. BD between $P_{\mathbf{C}}$ and $P_{\mathbf{S}}$ is denoted as $D_B(P_{\mathbf{C}}, P_{\mathbf{S}})$. If $D_B(P_{\mathbf{C}}, P_{\mathbf{S}}) = 0$, the steganography is almost secure. If $D_B(P_{\mathbf{C}}, P_{\mathbf{S}}) = \varepsilon$, it is ε secure.

Proposition1: $D_B(P_{\mathbf{C}}, P_{\mathbf{S}})$ is monotonously increasing function of ρ when $\rho \in [0,1)$ under $P_{\mathbf{C}} = N(\mathbf{m}, \mathbf{R}_{\mathbf{C}})$ and $P_{\mathbf{S}} = N(\mathbf{m}, \mathbf{R}_{\mathbf{S}})$.

Proof:

According to [16],

$$D_B(P_{\mathbf{C}}, P_{\mathbf{S}}) = \frac{1}{2} \ln\left(\frac{\det(\mathbf{R})}{\sqrt{\det(\mathbf{R}_{\mathbf{C}}) \cdot \det(\mathbf{R}_{\mathbf{S}})}}\right) \tag{7}$$

$$\mathbf{R} = \frac{\mathbf{R}_{\mathbf{C}} + \mathbf{R}_{\mathbf{S}}}{2} = \sigma_c^2 \begin{bmatrix} 1+\sigma_w^2/2\sigma_c^2 & \rho & \cdots & \rho^{n-1} \\ \rho & 1+\sigma_w^2/2\sigma_c^2 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1+\sigma_w^2/2\sigma_c^2 \end{bmatrix} \tag{8}$$

From [3], the determinant

$$a_n = \begin{vmatrix} d & \rho & \cdots & \rho^{n-1} \\ \rho & d & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & d \end{vmatrix} \tag{9}$$

has

$$a_n \approx d(d + (d-2)\rho^2)^{n-1} \quad \text{if } d \approx 1. \tag{10}$$

We can think that σ_w^2/σ_c^2 and $\sigma_w^2/2\sigma_c^2$ approximate to zero due to the requirements of imperception for steganography. Thus $1+\sigma_w^2/\sigma_c^2 \approx 1$ and $1+\sigma_w^2/2\sigma_c^2 \approx 1$. So

$$\det(\mathbf{R_S}) \approx \sigma_c^{2n}(1+\sigma_w^2/\sigma_c^2)(1+\sigma_w^2/\sigma_c^2+(\sigma_w^2/\sigma_c^2-1)\rho^2)^{n-1} \quad (11)$$

$$\det(\mathbf{R}) \approx \sigma_c^{2n}(1+\sigma_w^2/2\sigma_c^2)(1+\sigma_w^2/2\sigma_c^2+(\sigma_w^2/2\sigma_c^2-1)\rho^2)^{n-1} \quad (12)$$

$$\det(\mathbf{R_C}) = \sigma_c^{2n}\cdot c_n = \sigma_c^{2n}(1-\rho^2)^{n-1} \quad (13)$$

$$\frac{dD_B(P_\mathbf{C},P_\mathbf{S})}{d\rho} = \frac{(n-1)\rho(1+\rho^2)\sigma_w^4}{(1-\rho^2)[2(1-\rho^2)\sigma_c^2+(1+\rho^2)\sigma_w^2)][(1-\rho^2)\sigma_c^2+(1+\rho^2)\sigma_w^2)]} \quad (14)$$

When $\rho\in[0,1)$, $\frac{dD_B(P_\mathbf{C},P_\mathbf{S})}{d\rho}>0$. □

From Proposition1, the smaller ρ, the smaller $D_B(P_\mathbf{C},P_\mathbf{S})$ when $\rho\in[0,1)$. That is to say that the security of steganography will be improved if we select the cover whose ρ is smaller.

III. Cover Image Selection Method

We select the "better" cover according to the value of the correlation parameter ρ. In the following, we discuss only the problem of selecting the "better" cover image for spatial domain image steganography, such as SSIS [14] and LSB matching [19]. LSB matching can be seen as a special Spread Spectrum embedding. Spatial domain image steganography embeds data directly into pixel data, such as the intensity value of pixel. We should estimate ρ from pixel data.

In some sense, pixel data in an image should also be a realization of Gauss-Markov process after two dimensional pixel data is converted to one dimension. But the conversion process should be complicated. Thus it is not easy to estimate the correlation parameter ρ by Maximum Likelihood estimator.

Assume that an image I is stationary. The correlation coefficient between two arbitrary pixels $I(x,y)$ and $I(x+\Delta x,y+\Delta y)$ can be denoted as $r(\Delta x,\Delta y)$, and the special exponential model

$$r(\Delta x,\Delta y) = \rho^{\Delta r} = \rho^{[k_1(\Delta x)^2+k_2(\Delta y)^2]^{1/2}} \quad (15)$$

can be adopted. Here k_1 and k_2 are the parameters which reflect the difference between horizontal correlation and vertical correlation in an image. In (15), the correlation characteristic in an image is also described by the correlation parameter ρ.

For an image I which has $M\times N$ pixels, $r(\Delta x,\Delta y)$ can be estimated by the expression

$$\bar{r}(\Delta x,\Delta y) = \frac{\sum_{x_m=1}^{M-\Delta x}\sum_{x_n=1}^{N-\Delta y}[I(x_m,y_m)-\tilde{I}][I(x_m+\Delta x,y_m+\Delta y)-\tilde{I}]}{\sum_{x_m=1}^{M-\Delta x}\sum_{x_n=1}^{N-\Delta y}[I(x_m,y_m)-\tilde{I}]^2} \quad (16)$$

$$\tilde{I} = \frac{\sum_{x_m=1}^{M}\sum_{x_n=1}^{N} I(x_m,y_m)}{M\times N} \quad (17)$$

For each $(\Delta x,\Delta y)$, $1\le\Delta x\le M-1$, $1\le\Delta y\le N-1$, we can get the corresponding $\bar{r}(\Delta x,\Delta y)$. The estimation value of correlation parameter ρ is calculated by least squares estimator

$$\bar{\rho} = \arg\min_{\rho}\sum_{\Delta x=1}^{M-1}\sum_{\Delta y=1}^{N-1}(\rho^{[k_1(\Delta x)^2+k_2(\Delta y)^2]^{1/2}}-\bar{r}(\Delta x,\Delta y))^2 . \quad (18)$$

Suppose that we have an image gallery for spatial domain image steganography, we calculate the ρ value of each image in the gallery. In order to improve the security, we should select the image which has the smallest ρ as cover.

IV. Experiments

In the experiments, the cover database consists of 3048 images which were downloaded from USDA NRCS Photo Gallery [20]. The images are very high resolution TIF files (mostly 2100×1500) and appear to be scanned from a variety of paper and film sources. For testing, the cover images were resampled to 614×418 and converted to grayscale(The tool used is Advanced Batch Converter 3.8.20, and the selected interpolation filter is bilinear). They are regarded as seeds to generate 2 groups of 3048 stego images of SSIS[14] and LSB matching[19]. The max embedding rate is adopted.

Cover images and the corresponding stego images are divided into 2 groups according the ρ value of cover images. If $\rho>0.9$, the cover image and stego image belong to the big correlation group. If $\rho\le0.9$, they belong to the small correlation group. We test the detectability of two groups using the detector in [21] due to its better performance.

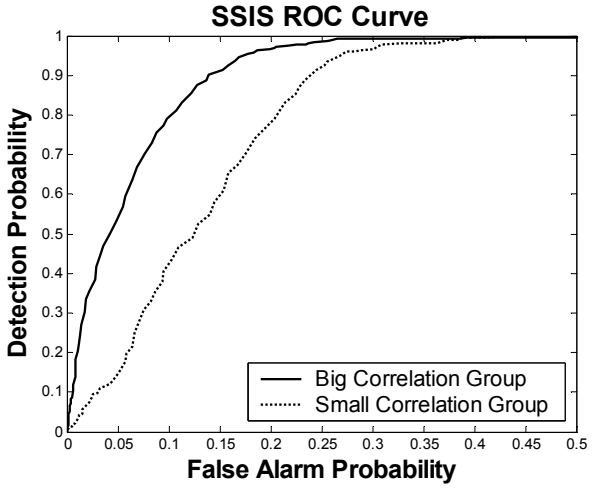

Fig2. ROC curves for detecting SSIS

Fig.2 shows the receiver operating characteristic (ROC) curves for SSIS. The ROC curves show how the detection probability (the fraction of the stego images that are correctly classified) and the false alarm probability (the fraction of the cover images that are misclassified as stego images) vary as

detection threshold is varied. The lower the ROC curve is, the difficult the detection of steganography is. From Fig.2, the ROC curve of small correlation group is lower than that of big correlation group. It validates that in detecting SSIS, the small correlation group are more difficult than the big correlation group.

Fig.3 shows ROC curves of big correlation group and small correlation group in detecting LSB matching. The ROC curve of small correlation group is lower than that of big correlation group. It validates that in detecting LSB matching, the small correlation group are more difficult than the big correlation group. Comparing Fig.2 with Fig.3, we also find that detecting LSB matching is more difficult than SSIS.

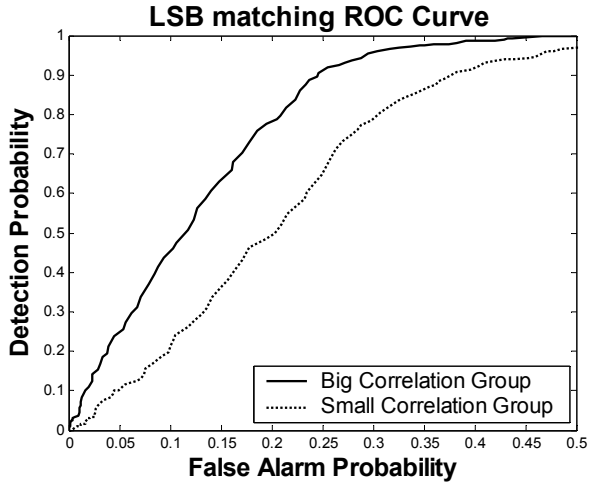

Fig3. ROC curves for detecting LSB matching

V. CONCLUSIONS

Other than the choice of the embedding technique, the sender can choose "better" cover to improve the security of steganography. We find that the covers whose data have smaller correlation characteristic are "better". The point of view is obtained from Gauss-Markov cover model and spread spectrum embedding model. Note that Gauss-Markov cover model still has a gap with the practical multimedia data, for example digital image. The effect of the gap on the security of steganography is still a question. But in the experiments, for image spatial domain steganography, the images that have smaller correlation parameters are more difficult in discriminating cover with stego than the images that have larger correlation parameters.

REFERENCES

[1] C. Cachin, "An information-theoretic model for steganography,". in IH 98, LNCS 1525, Heidelberg: Springer-Verlag,1998, pp. 306-318.

[2] K. Sullivan, U. Madhow, and S. Chandrasekaran, et al., "Steganalysis for Markov cover data with applications to images," IEEE Trans. on information forensics and security, vol. 1, no. 2, pp. 275-287, June 2006.

[3] V. Korzhik, H. Imai, and J. Shikata, etc., "On the use of Bhattacharyya Distance as a measure of the detectability of steganographic systems,". Transactionss on DHMS III, LNCS 4920, Heidelberg: Springer-Verlag, 2008, pp.23-32.

[4] Y. Wang and P. Moulin, "Steganalysis of block-structured stegotext," in Security, Steganography, and Watermarking of Multimedia Contents VI, vol.5306, SPIE-IS&T, 2004, pp.477-488.

[5] J. Egger, R. Bäuml, and B. Girod, "A communications approach to image steganography," in Security, Steganography, and Watermarking of Multimedia Contents IV, vol. 4675, SPIE-IS&T, 2002, pp. 26-37.

[6] P. Sallee, "Mode-based steganography," in IWDW 2003, LNCS 2939, Heidelberg: Springer-Verlag, 2004, pp.154-167.

[7] A. Sarkar, K. Solanki, and U. Madhow, etc., "Secure stegnography: statistical restoration of the second order dependencies for improved security", in IEEE ICASSP 2007, pp.II-277- II-280.

[8] J. Fridrich, and D. Soukal, "Matrix embedding for large payloads," IEEE Trans. on Information Forensics and Security, vol.1, no.3, pp.390-395, September 2006.

[9] J. Fridrich, M. Goljan, and D. Soukal, "Wet paper codes with improved embedding efficiency," IEEE Trans. on Information Security and Forensics, vol. no.1, pp. 102-110, March 2006.

[10] X. Zhang and S. Wang, "Dynamical runing coding in digital steganography," IEEE Signal Proceesing Letters, vol. 13, no.3, pp.165-168, 2006.

[11] W. Zhang, X. Zhang, and S. Wang, "Maximizing steganographic embedding efficiency by combining Hamming codes and wet paper codes," in IH 08, LNCS 5284, Heidelberg: Springer-Verlag, 2008, pp.60–71.

[12] M. Kharrazi, H. T. Sencar, and N. Memon, "Cover selection for steganographic embedding," in IEEE ICIP 2006, pp.117-120.

[13] Y. Sung, L. Tong, and H. Poor, "Neyman-Pearson detection of Gauss-Markov signals in noise: closed-form error exponent and properties," IEEE Trans. on Information Theory, vol. 52, no.4, pp. 1354-1365, 2006.

[14] L. M. Marvel, C. G. Boncelet, and C. T. Retter, "Spread spectrum image steganography," IEEE Trans. on Image Processing, vol.8, no.8, pp. 1075-1083. 1999.

[15] I. Cox, J. Kilian, and F. Leighton, etc., "Secure spread spectrum watermarking for multimedia," IEEE Trans. on Image Processing, vol.6, no.12, pp.1673-1687,1997.

[16] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," IEEE Trans. on Communication Technology, vol.com-15, no.1, pp.52-60, 1967.

[17] S. Dumitrescu, X. Wu, and Z. Wang, "Detection of LSB steganography via sample pair analysis," IEEE Trans. on Signal Processing, vol.51, no.7, pp.1995-2007, 2003.

[18] Y. Shi, C. Chen, and W. Chen. A Markov process based approach to effective attacking JPEG steganography. in IH 06, LNCS 4437, Heidelberg: Springer-Verlag, 2007, pp.249–264.

[19] A. Ker, "Improved detection of LSB steganography in grayscale images." In IH 04. LNCS 3200, Heidelberg: Springer-Verlag, 2004, pp. 97–115.

[20] USDA NRCS Photo Gallery , http://photogallery.nrcs.nsda.gov

[21] Y. Sun, F. Liu, and B. Liu, etc., "Steganalysis based on difference image," in IWDW 2008, LNCS 5450, Heidelberg: Springer-Verlag, 2009, pp.184-198.