International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 3 Issue 7, July – 2019, Pages: 47-53

Create a Secure Communication Within Finding an Optimal Cover for Exchanging Information by Using Kendall's Tau Correlation

Zainalabideen Abdual Samad
Department of Computer Science
Faculty of Education-University of Kufa
zainalabideena.alsaraf@uokufa.edu.iq

Ahmed Al-Adilee
Department of Computer Science
Faculty of Education-University of Kufa
ahmeda.aladilee@uokufa.edu.iq

***Abstract**: Hiding data is one of the most important aspects that has been occurred from the early date of humanity. there are several different modern methods that have been created to hide data specially in mobile exchanging data applications. In this paper, it has been proposed a new technique for choosing an optimal cover depending on determining an optimal rank correlation which is well-known by Kendall-tau correlation. The producer of choosing the appropriate cover has two main steps. First of all, downloading n images and calculate their Kendall's tau before and after inserting the secret message. While the second step involves dividing each image to $n \times n$ blocks and then finding the optimal Kendall's tau among them. Indeed, our attempt is to set a standard method of choosing the proper cover for secure information.*

Keywords: Communication, Cover Image, Rank Correlation, statistical inference, Mobile applications.

1 Introduction

Communication between parties it means creating a secure channel between them to exchange their vital information in a safe way. This terminology of communication can be applied in both ways cryptography, and steganography. Steganography (Hiding data) is one of the most important aspects that has been occurred from the early date of humanity. Historically, sending a secure message or information was a very serious issue in order to deliver that message safely. These reasons and other became the motivations of thinking of finding methods and techniques that could guarantee sending and receiving message safe and secure. One the oldest historical story that associate with hiding information came from a Greek age when Herodotus shaved the head of his slave and write a tattoo message on his scalp,[6-9].

In the twentieth century, the need of techniques that guarantee hiding data in a secure cover were very useful in order to send the required information to other party without interrupting of enemies or any other part that wish to discover the content of the hidden data [12].

The main key of stenography analysis is based on statistical methods which are applied on depending on the mathematical principle, i.e. the robust algorithm of stenographer is coming from the way of immune statistical technique. According to this idea, there were methods that lies on statistical meaning so that we could avoid the methods of detection for hidden information through a proper channel, which are used between parties to exchange their secure content.

On the other hand, rank correlation is one of the most powerful tools in statistical world. One of these popular rank correlations is well-known by Kendall's tau (1970). Moreover, Kendall's tau represents a measure of correlation coefficient and it is much useful in many application . The principle of Kendall's tau is not difficult but it is efficient in determining the relation between random variables. In 1970, Kendall's has shown a measure of rank correlation that corresponds to the degree of intensity of correlation coefficient using a Greek symbol τ to refer to this type of correlation. The relationship between the hiding data process and correlation is so important in order to measure the strength of encryption methods and the methods of choosing the proper covers.

This study has been divided to the following part. First part shows the most common works related to the methods of hiding data and Kendall's tau usage. In the second part, there will be our main proposed technique that associate Kendall's tau concept and calculations to choosing the proper cover for hiding the desired information.

2 Literature Review

There are tremendous researches have been applied in the field of protection data. However, in this paper the new technique which combines both concepts (mathematics and statistical inference) to gain robust constitute. Moreover, the new idea refer to one of the most important factors which is related to the image addressed by correlation. Through the meaning of statistical correlation, the correlation of image and rank correlation are equivalents. While, image correlation reflexes how much pixels are adjustment in terms of pixels value. Furthermore, this fact come over the blocks as well as the entire images.

One of the most interesting efficient studies have mentioned in [1]. It's an algorithm that have been used to reduce the size of the secret message in order to remain the quality of the cover image in the same value. Indeed, the algorithm has two main parts that deals with compression techniques and using of Fibonacci algorithm in order to embed the chosen secret message [2].

Another impressive result related to steganography and encryption is the study that have been shown a new technique of representation. They have decomposed pixel intensity value to 16 bit-planes that are suitable for embedding purposes. However, this approach tries to find an efficient way that gathers payload capacity and stego quality. Indeed, its weakness point appears in the embedding message in the higher bit-planes [2].

There are also various methods of hiding data that have been proposed to find and select the right cover. In 1996, W. Bender, D. Gruhi, N. Morimoto, A. Li have shown a steganography approach of hiding data. Also, in 2005, C.B. Adsumili, M.C.Q. Farias, S.k. Mitra, M. Carli have proposed a technique of robust error covering using hiding data in order to perform a high quality of image. Furthermore, in 2008, Sotirios P. Chatzis, Dimitrios I. Kosmopoulos, Theodora A. Varvarigou have used Markov chains in hiding data. They have a proposed a system of sequential Markov chains that have been used in data modeling techniques.

On the other hand, rank correlation techniques were widely used in various applications of real life data. One of the main fields that have employed correlation coefficient, and rank correlation approaches is risk management and computed the value at risk. In 2014, Kristofer Eriksson used Kendall's tau to measure the risk of financial management and describing of the dependence structure. Indeed, Kendall's tau have widely been demonstrated in statistical inference, dependence structure, analysis of systems throughout a very interest functions that have been in many different approaches of finance, risk measure, insurance, and etc. In particular, it has been applied in finding the risk of portfolio with respect to a well-known copula which called a t-copula, see [7-8]. Finally, it should be mentioned that using of rank correlation in hiding data may not be very popular and there are few methods that have mentioned it.

2 Preliminaries and Basic concepts

In this part, we should mention some foundations of rank correlation. Indeed, our technique depend on the essential mathematical model of Kendall's tau as a correlation coefficient. Mathematically speaking, let Ω be a set of n numbers that are ordered in a way that follows a random experiment. According to the notion of Kendall the essential step of our process is determining the correlation coefficient with respect to Kendall's tau and this step consists of separating Ω to two main sets [8]. Let the first set, denoted by A, represents the counted numbers below and greater than the selected number that belongs to Ω, while the second set is the counted numbers below and less than the same selected number. Repeating the same process with all other numbers that belong to Ω leads us to the following table

Ω	A	B
a_1	1st counted numbers below a_1 and greater than it	1st counted numbers below a_1 and less than it
a_2	2nd counted numbers below a_2 and greater than it	2nd counted numbers below a_2 and less than it
⋮	⋮	⋮
a_n	n^{th} counted numbers below a_n and greater than it	n^{th} counted numbers below a_n and less than it
	$C = \sum A$	$D = \sum B$

Then the basic formula of Kendall's tau can be written by the following well-known relation

$$\tau = \frac{C - D}{C + D} \quad (1)$$

Using a permutation formula with respect to equation (1) leads us to the following equivalent relation that we have mentioned above

$$\tau = \frac{C - D}{1/2n(n-1)} = 1 - \frac{2D}{1/2n(n-1)} \quad (2)$$

Note that $-1 \leq \tau \leq 1$, C is called concordant set, and D is called the discordant set, See[7-8].

4 The Computations of Rank Correlation Kendall's Tau (RCKT)

Our Proposed technique of choosing the optimal cover that used to hide the desired message is basically depending on choosing the optimal Kendall's tau value. This process can be divided to two main steps:

Step1: Downloading n random different images with the same sizes, but different types (jpg, png,...etc). Divide any chosen image to $n \times n$ block in order to calculate its rank correlation (Kendall's tau). The calculation involves finding the Kendall's tau value for each image before and after embedding the secret information inside the desired image.

Note that it has been mention that the rank correlation with respect to Kendall's tau before inserting the desired message is called RCKTB, while after the inserting is called RCKTA. Indeed, the decision can be made by finding the absolute value of the different between RCKTB, and RCKTA ($|\mathrm{RCKTB} - \mathrm{RCKTA}|$) which represents the optimal cover. The images that we have chosen are 10 different images that are the following:

Figure1: Baboon.jpg

Figure2: Barbara.jpg

Figure3: Boat.jpg

Figure4: Fruits.jpg

International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 3 Issue 7, July – 2019, Pages: 47-53

Figure5: Perppers.jpg

Figure6: Lena.png

Figure7: Sails.png

Figure8: Computer.png

Figure9: Watch.png

Figure10: Desktop.png

Step2: After obtaining the right cover, the sender will divide an image into $n \times n$ blocks and determine the RCKT after and before once again for each block. Mathematically speaking, calculate $RCKT_i, i = 1, \dots, n$. Then selecting the optimal block according to the same producer of choosing the optimal cover. Eventually, the optimal cover and optimal block, respectively complete the process of choosing the desired cover of the message that we wish to hide inside that cover.

The following table shows each type of image with its extensions and the values of optimal cover through the RCKTB and RCKTA.

Image Type	RCKTB	RCKTA	Optimal RCKT
jpg1	-0.74	-0.75	0.01
jpg2	0.64	0.62	0.02
jpg3	0.7561	-0.6	1.3561
jpg4	0.05	-1	1.05
jpg5	0.7561	-0.6	1.3561
png1	-0.01429	0.00001	0.0143
png2	-0.2941	0.1579	0.452
png3	0.1034	0.6471	0.5437
png4	0.4091	-1	1.4091
png5	-0.4667	0.9459	1.4126

Table1: shows the type of image and its rank correlation before and after the encryption

The calculated values of each image and the determining of the optimal cover have been applied through the MATLAB 2013 software with i5 core for images ($image_1$,$image_2$,..., $image_n$). The results occurred in **Table1** show that the optimal cover corresponds to the optimal RCKT value that we have obtain in Table1. For instance, the optimal cover in our above table is 0.01 because it is the smallest value of the absolute differences between RCKTB and RCKTA, respectively.

Further, we should mention that our choice of the image associate with another important step that we have mentioned previously which the step of finding the optimal block in side that image. Certainly, this step is nothing more than repeating the same calculations of RCKT that we have explained with respect to set of images.

Also, we should statistically note that Kendall's tau cannot be more than one or minus one, but by looking carefully at the optimal values in the Table1 shows that most values are greater than one or minus one which means that they are absolutely not suitable to be chosen as an optimal cover.

Indeed, these results based on selecting the proper secret information that has been chosen randomly. In other words, the content of different secret messages leads to choosing the optimal (suitable) cover image and each hidden data has another different cover that might be suit to another different image.

Furthermore, there are different figures below that interpret the relationships between each cover and its RCKT. Figure1 shows histogram relationships between each image and its absolute difference of RCKTB, and RCKTA. A deep look at the blocks of each histogram prove that our choice of the optimal cover lies between jpg1 and its optimal value of RCKT which is equal to 0.01. While Figure2 explain the differences between the RCKTB and RCKTA for each cover and once again if we look at the graph so it is clear that the only identical RCKT that has no intercepts lies through the value of 0.01. Third Figure interprets the relationships between covers and their optimal values as a curve of each value of RCKT and a cover.

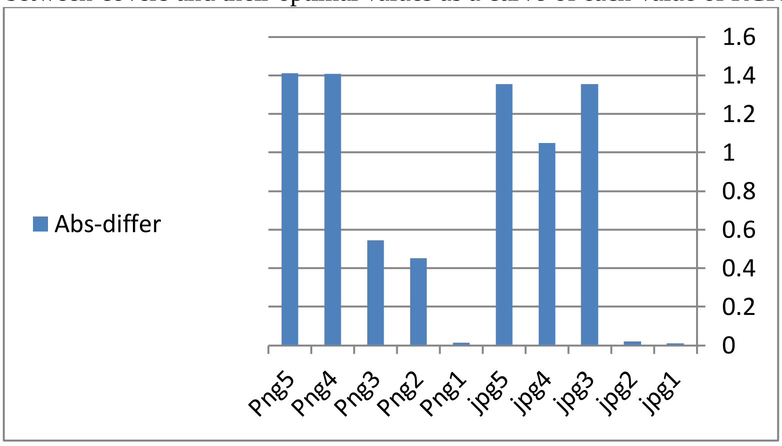

Figure 1:Histogram of the Relationshipe Between Each Image and its RCKT

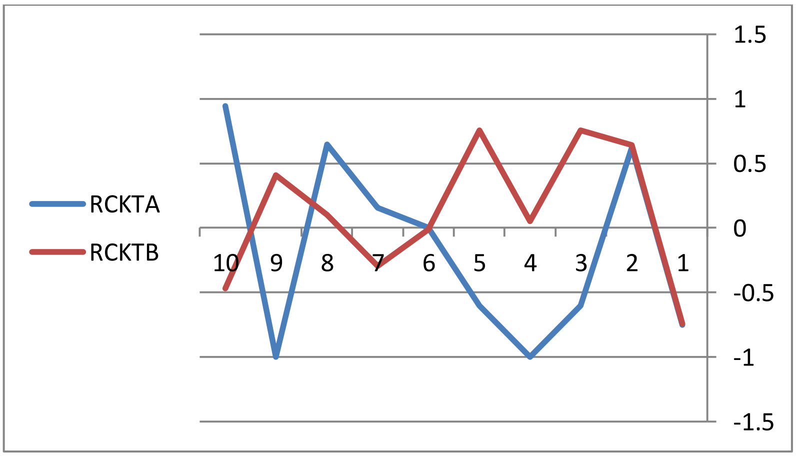

Figure 2:Diferences Between RCKTB and RCKTA

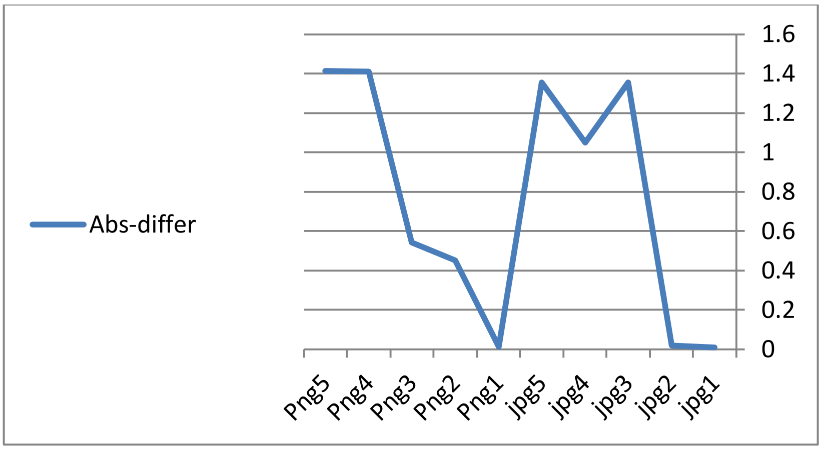

Figure 3:A Curve Relation of Each Image and its RCKT

4 Conclusion

The study shows that rank correlation (Kendall's tau) technique is a very efficient and easy way to identify the optimal cover of our secret message. The calculations of each cover led us to find identical RCKTB and RCKTA and keep their optimal value in the range of (-1,1). While some other calculations have a big difference in their optimal value and they are out of the desired range of Kendall's tau correlation. Another important fact that we have obtained from the examined cover demonstrates that optimal cover may differ according to the type of the information that one would like to hide inside a cover. Moreover, these steps and calculations have been repeated similarly inside the chosen image in order to determine the optimal block that will be used to hide the information. The efficiency of our technique as a statistical instrument is quite simple because of rank correlation theory that lies behind it, for example, Kendall's tau theory is one of the perfect theory in mathematical statistics. Indeed, this technique is totally different from any other previous techniques that have been used to locate the proper cover of the secret message, but it is robust and not complicated. Indeed, we could say that our method has an short time of execution comparing with many other methods that have been applied in the same area. Finally, there are several different aspects relevant to choosing the optimal cover that rank correlation could serve their needs perfectly and this will be our future study.

Compliance with Ethical Standards:

Conflict of interest: Author **Zainalabideen Abdual Samad**, declares that he has no conflict of interest. Author **Ahmed AL-Adilee** declares that he has no conflict of interest.

Ethical approval: This article does not contain any studies with human participants performed by any of the authors. Also, This article does not contain any studies with animals performed by any of the authors.

References

[1] Adsumili C.B., Farias M.C.Q., Mitra S.k., Carli M.: A robust error concealment technique using data hiding for image and video transmission over lossy channels, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 15, Issue: 11, 1394 – 1406, (2005).

[2] Alan A. Abdulla, Sabah A. J., Harin S.: high-capacity steganography techniques, inter.conf.royal hallway, uni. of London, 2014.

[3] Alan A. Abdulla, Harin S. Sabah A. J.: Stego Quality Enhancement by message size reduction on Fibonacci Bi1-Plane Mapping, first inter. Conference, London, 2014.

[4] Alan A. Abdulla, Harin S., and Sabah A. Jassim. "Steganography based on pixel intensity value decomposition." Mobile Multimedia/Image Processing, Security, and Applications 2014. Vol. 9120. International Society for Optics and Photonics, 2014.

[5] Aladilee A. Mohammed K.D: Copula functions and correlations personal contribution, Journal of Kerbala university, Vol. 12,issue 4, 196-205, (2014).

[6] Bender W., Gruhi D., Morimoto N., Li A.: Techniques for data hiding, IEEE, IBM Systems Journal, Vol. 35, Issue: 3.4, 313 – 336, (1996).

[7[Hogg R.V., Craig T.A.: Introduction to Mathematical Statistics, Pearson education limited, 2004.

[8] Kendall's M.G.: Rank Correlation Methods, 4th Edition, Charles Griffin, London, 1970.

[9] Kakade R., Kasar N., Kulkarni S., Kumbalpuri S., Patil S.: Image Steganography and Data hiding in QR Code, IRJET, V. 04, issue 05, 2017.

[10] Luo X., Shevchenko P. V.: The t copula with Multiple Parameters of Degrees of Freedom: Bivariate Characteristics and Application to Risk Management, Quantitative finance, vol.10, (2009).

[11] Mohammad S.S.: A new method for real time steganography", ICSP Proceedings of IEEE, 2006.

[12] Rasheed, Zainalabideen Abdual Samad: Steganography Technique for Binary Text Image, International Journal of Science and Research (IJSR) ISSN (Online), 2319-7064 (2015).

[13] Rasheed, Zainalabideen Abdual Samad: Improving Classical Fibonacci in Steganography, IJARCSSE, v.4, issue 12, (2014).

[14] Rasheed, Zainalabideen Abdual Samad, A. A. M. Baker. A Novel Method of Generating (Stream Cipher) Keys for Secure Communication, vol. 17, issue 7, 07-13, (2015).

[15] Sotirios P. Chatzis, Dimitrios I. Kosmopoulos, Theodora A. Varvarigou: Robust Sequential Data Modeling Using an Outlier Tolerant Hidden Markov Model, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, Issue: 9, 1657-1669, (2009).