

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed PhD Students: Maisa'a A. A. Khohder
Computer Science Department University of Technology

Abstract:

Steganography is known as the art and technique of text messages concealed in such a manner that one apart from the meant recipient defines the message presence. Because of this reason it has become almost essential to conserve the systems free from various internet attacks and take occupy steps to remove any danger. Linguistic steganography, a kind of steganography is famous as a set of techniques and manners that allow the hiding of any digital data inside scripts based on some linguistic knowledge.

Linguistic steganography is of various shapes and has varied applications. This research offers in a new method to use two levels to hide, the first level hides by embedding and addition but the second level hides by injection. The first level embeds a secret message one bit in the LSB in the FFT with the addition of one kashida. Using DRLR is RNG to find position to hide within the text. The second level is the injection of random single or double quotation within the text.

This paper presents a linguistic steganography for scripts written in Arabic language, using kashida and Fast Fourier Transform on the basis of using new technique entitled Dynamic Random Linear Regression as location to hide secret message. The proposed approach is an attempt to present a transform linguistic steganography using levels for hiding to improve implementation of kashida, and improve the security of the secret message by using Dynamic Random Linear Regression.

This method is tested in terms of security and capacity, transparency, and robustness and this is way better than previous methods. The proposed algorithm has ideal steganographic properties.

Keywords: Arabic script, Linguistic Steganography, Dynamic Random Linear Regression, Kashida,quotation, Transform Based

I. INTRODUCTION

Linguistic steganography focuses on applying changes to a cover text so as to embed secret message, in a way that the changes do not cause any unnatural or ungrammatical text. According to cover text, linguistic steganography can be classified in natural language steganography into two groups: generation based such as in [1], or transformation based such as in [2]. Natural language steganography, as depicted in Figure (1).

1- Natural Language Steganography Methods

The public prepared manner for data hiding within the natural language scripts

can be classified into two group.

- **The first group** which is called script steganography. This manner is based on several lines, spaces, characters, or any other shapes of the offered message as shown below [3].

- 1- Line-shift coding.
- 2- Word-shift coding.
- 3- Feature coding.

- **The second group** which is called linguistic steganography is based on Linguistically modified covering script in order to encode the message. This is manually retyping the scripts using linguistic conversion such as synonym exchange, syntactic exchange (paraphrasing) or semantic conversion as follows [3].

- 1- Using probabilistic context-free grammars to generate cover text.
- 2- Synonym substitutions.
- 3- Syntactic transformations.
- 4- Semantic transformations.
- 5- Hybrid techniques [3].

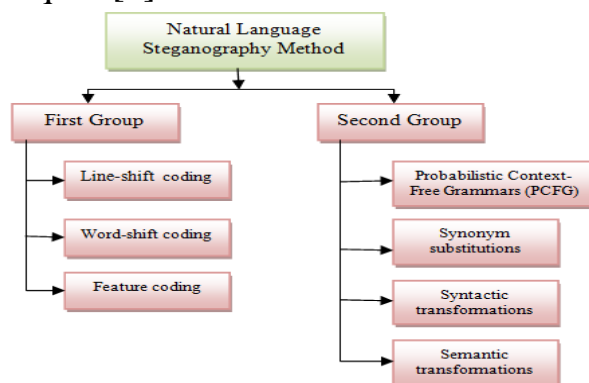


Figure (1): The Natural Language Steganography Method.

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

2- Hierarchical Structure Of Word Documents:

"Electronic documents have many advantages, including compact and lossless storage, easy maintenances, efficient retrieval and fast transmission. One of the major advantages of electronic documents is that an electronic document can have an explicit structure; it can be partitioned into: *a hierarchy of physical components*, such as pages, columns, paragraphs, text lines, words, tables, figures, halftones, etc.; *a hierarchy of logical components*, such as titles, authors, affiliations, abstracts, sections, etc.; or both. This structural information can be very useful in indexing and retrieving the information contained in the document" [4], [5].

Document understanding modules, "such as Optical Character Recognition (OCR) and graphics recognition modules, can also be selectively applied to the structural components of document images. Physical layout and logical structure analysis of document images is a crucial stage in a document image analysis system [4]. Documents in Word have a hierarchical structure. Different types of properties apply to different units in the hierarchy", as shown in figure (2) [5]:

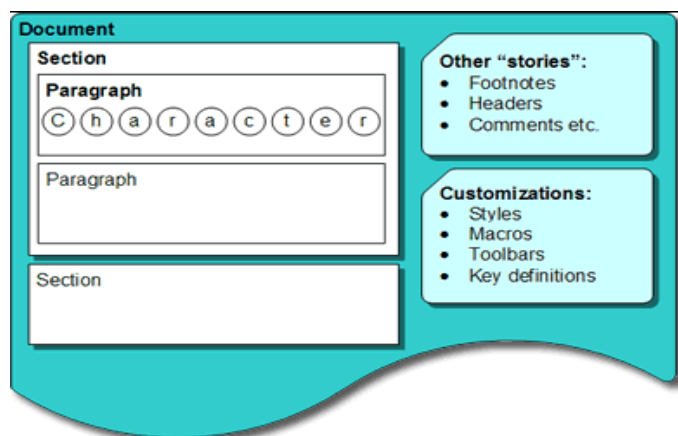


Figure (2): The Hierarchical Structure of Word Document [5].

In the present work, layers of steganography technique are proposed for script written in Arabic language using Fast Fourier Transform (FFT) and kashida. The proposed approach uses Dynamic Random Linear Regression (DRLR) to generate random location, to embed the secret message bits using FFT and kashida as a first layer followed by adding single or double quotation randomly as second layer. The proposed algorithm optimizes steganography properties such as transparency, robustness, and security of the secret message for Arabic script based secure communication.

The other sections of the paper are structured as follows: Section II presents the literature review for kashida based linguistics steganography

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

and explains fundamentals use of proposed system. Section III explains algorithm for proposed system, results and discussions are given in section V, and section IV deals with the conclusions.

II. LITERATURE REVIEW AND FUNDAMENTAL USED IN PROPOSED SYSTEM

A. *Related Work*

Kashida is an Arabic redundant character which is used to justify the text, without affecting the meaning of words. Researchers suggested using one kashida as bit zero, and two kashidas as bit one, or vice versa.

In 2007, A. Gutub, and M. Fattani, introduced a novel Arabic text steganography technique for Arabic script using letter points and kashida. The technique hides secret information as bits in Arabic letters (cover) by using kashida and points of letters. The technique considers un-point Arabic letters followed by a kashida if the secret bit is (0), and point Arabic letters followed by kashida if secret bit is (1).

Their technique enhanced robustness and security but might have some limitations with capacity of the cover media if the number of secret bits of the secret information is large. This steganography technique is found to be suitable for other languages having similar script to Arabic for example Persian and Urdu [6].

In 2009, A. H. Fahd, et al, introduced improving security, and capacity for Arabic text steganography using kashida. The approach hides secret information as bits within Arabic letters (cover) by using kashida using three scenarios. The approach discussed maximum number of kashida letters that can be added to the Arabic cover word. Also the researchers evaluated the number of hidden bits that can be embedded in the carrier file and compared the results with diacritics, and kashida methods [7].

In 2009, M. Hassan Shirali-Shahreza and Mohammad Shirali-Shahreza, introduced a novel manner to conceal data in Persian (Farsi) and Arabic languages. In Unicode scale, to conceal script two characters of 'Ya' (ي) and 'Kaf' (ك) used. The two characters of (ي) and (ك) have themselves shape however various codes which they utilized at the start or in the center of words. The major aim in this manner is perception translucence. It has a stellar perception translucence which causes the stego-text which the user seen alright like the main text [8].

In 2010, Adnan Abdul-Aziz Gutub, and Manal Mohammad Fattani, introduced an enhanced Arabic script steganography technique for Arabic text using hiding secure data bits into messages making uses two set points. Hiding a secret message in cover text, where the pointed character is used for expansion convene if secret bit 'one' and the un-

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

pointed character is used for expansion convene 'zero'. That manner enhances security, capacity, transparency and robustness for Arabic texts based secure communication [9].

In 2010, A. Ali and F. Moayad, introduced Arabic text steganography technique for Arabic script using kashida with Huffman code. The approach hides secret information as bits within Arabic letters (cover) by using extension character (kashida), and compresses the stego file using Huffman code. The technique considers absence of kashida if the secret bit is (0) and one kashida if secret bit is (1) after any connected letters. Also their technique is applied to other Arabic script than based on secure communication in different document formats [10].

B. Fast Fourier Transform and its Inverse

The mathematical formula to Fourier Transform of a time domain function $f(x)$, for

real numbers x and y is [10]:

$$F(y) = \int_{-\infty}^{+\infty} f(x) \exp[-i2\pi xy] dx \quad \dots\dots (1)$$

And the mathematical formula to its inverse is [11]:

$$f(x) = \int_{-\infty}^{+\infty} F(y) \exp[i2\pi xy] dy \quad \dots\dots (2)$$

$f(x)$ = Time domain function

$F(y)$ = Frequency domain function

x = Argument with units of time

y = Argument with units of frequency

e = Base of natural logarithms

i = Imaginary unit ($i^2 = -1$).

C. Linear Regration (LR)

Linear regression attempts to model the relationship between two variables X , and Y , by fitting a linear equation to observed data, such as[12]:

$$Y = a + b X_i \quad \dots\dots(3)$$

Where

X = The explanatory variable

Y = The dependent variable

b = The slope of the line

a = The value of y when $x = 0$.

D. Dynamic Radom Linear Regression (DRLR)

It is a new technique to generate a set of random positions X_i $i = 1.2....N$ by using this equation (4) depicted in Figure (3) for the position DRLR.

$$X_i = a + bX_{i-1} \quad \dots\dots(4)$$

Where

N = The size of generated random positions

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

- X_{i-1} = The explanatory variable
- X_i = The dependent variable
- b = The slope of the line
- a = The value of X_i when $X_{i-1} = 0$.

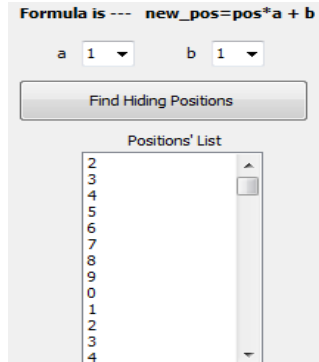


Figure (3): The position of DRLR.

III. PROPOSED SYSTEM

A. Idea for proposed system

The proposed approach main idea as depicted in Figure (4) is the embedding, and while in Figure (5) the extraction, is to use DRLR as generated random location, to add random kashida characters to the rest of Arabic word scripts as a second layer, where the first layer injects the secret message bits in the inverse FFT (LSB of (real (FFT) of selected Arabic script word)), and then one kashida character is applied. The first addition of kashida is for the hiding process of the secret information, while the second addition of the single or double quotation is for confusion purpose by insuring security of the secret message.

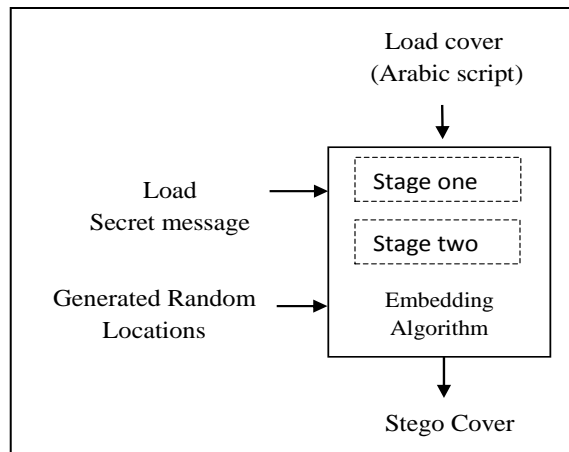


Figure (4): The proposed hiding process

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

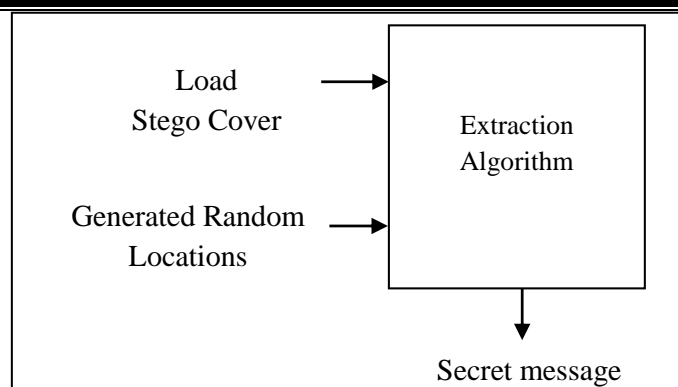


Figure (5): The proposed extraction process.

B. Embedding Process

Embedding Algorithm:

Input: secret message, seed, a, b, N, a set of Arabic scripts.

Output: stego-cover.

Process:

Step1. Secret message binarization: The secret message is hidden in form of (0) s, and (1) s, which represent (64) bit Unicode of each character using the hexadecimal representation. N, is the total number of secret message bits. Figure (6) presents the binarization process to secret message. Figure (7) is a simple example of applying binarization process to secret message.

Step2. Generate Random positions: The process of generated random positions, using DRLR, starts by using secret key (seed) to generate sequence of random values C_i , where $0 \leq c_i < 32$. The values C_i , represent offset of Arabic script words to start the embedding process. The total number of Generate Random positions is (N), where N, is the total number of secret message bits.

Step3. Cover selection: Select Arabic script (cover) that can hold input secret message bits.

Step4. Do while not end of Arabic script words

Step5. Embedding layer one: For each secret message bit and Generate Random positions do

Step6. Use C_i value as offset to next word to embed the secret message bit, into inverse FFT (LSB (real (FFT (select Arabic script word))))), then apply one kashida if the secret message bit is one or if the secret message bit is zero.

Step7. End of For.

Step8. Else

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

Step9.Embedding layer two: add single or double quotation randomly to the rest Arabic script words

Step10. End of Do.

Step11. End

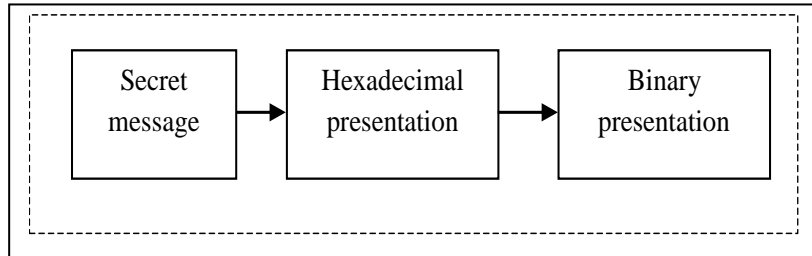


Figure (6): Secret Message binarization.

Secret message	ما نفضت الندى عن ذرى العشب فيه
Hexadecimal representation	EFC72020F2BAEBAA2020C7FBF2CFE92020DFF22020D0D1E92020C7FBDFBDA92020BAFDF3
Binary representation	11101111110001110010000000100000111100 10101110101110101110101010001000000010 00001100011111111011111100101100111111 1010010010000000100000110111111110010 00100000001000001101000011010001111010 01001000000010000011000111111110111101 11111011110110101001001000000010000010 1110101111110111110011

Figure (7): Secret message binarization example.

C. Extraction Process

Extraction Algorithm:

Input: secret message, seed, a, b, N, stego cover.

Output: secret message.

Process:

- Step1. Generate Random positions: The process of generated random positions, using DRLR, starts by using secret key (seed) to generate sequence of random values C_i , where $0 \leq c_i < 63$. The values C_i , represents offset of Arabic script words (stego-cover) to start the extraction process.
- Step2. Loading: Load stego-cover, and Generate Random positions.
- Step3. For each Generate Random Positions do
- Step4. Use C_i value as offset to next word to extract the secret message bit, from LSB of select Arabic script word (stego-cover).
- Step5. End of For.

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

Step6. Convert each eight bits into one letter the result is the secret message.End.

IV. RESULTS AND DISCUSSION

This section discusses cases to ensure the proposed technique security:

Case one: An example of result for applying the proposed technique using embedding layer one, as depicted in Figure (8).

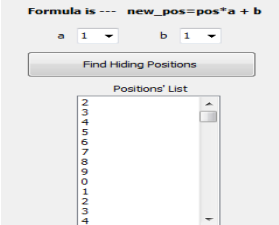
Cover	إن الطفولة هي المفتاح إلى شخصية بدر شاكر السياب وإلى أدبه وكل محاولة لفهم إنسانا وأديبا خارج دائرة الطفولة هي محاولة ناقصة. لقد تشبث بلفظة الطفل والطفولة ودلالاتهما تشبثا يعكس عمق معاناته وفرط حساسيته تجاه مرحلة قاسية من عمره رافقته آثارها حتى الكبر. لقد كانت بواكير قصائده تتضمن جزءا كبيرا من صور طفولته في جيور حيث اليتيم والحرمان العاطفي ثم الافتراق عن الأهل والحقيقة أن قصائده الأولى كانت عاطفية ذاتية مباشرة إذ كانت كل صورها في محيط حياة الشاعر وقريته ومعاناته الشخصية لكنه بعد قراءته لأشعار الغربيين وتوسع أفق رؤيته الفكرية جاء توظيفه للطفولة ببنية تعبيرية عالية التي وضعت مع مجالات توظيفات أخرى في مصافي كبار الشعراء الذين يتحول في نتائجهم كل معنى أو فكرة إلى رموز ودلالات كبرى وعميقة. تراوحت دلالة الطفولة في شعر السياب بالإيجابية دائما حتى في تلك الصور التي أظهر فيها الطفل والطفولة
Secret message	كان يومى كعام تعد المسرة فيه
DRLR	
Stego-cover FFT	إن الطفولة هي المفتاح إلى شخصية بدر شاكر السياب وإلى أدبه وكل محاول لفهم إنسانا وأديبا خارج دائرة الطفولة هي محاولة ناقصة. لقد تشبث بلفظة الطفل والطفولة ودلالاتهما تشبثا يعكس عمق معاناته وفرط حساسيته تجاه مرحلة قاسية من عمره رافقته آثارها حتى الكبر. لقد كانت بواكير قصائده تتضمن جزءا كبيرا من صور طفولته في جيور حيث اليتيم والحرمان العاطفي ثم الافتراق عن الأهل والحقيقة أن قصائده الأولى كانت عاطفية ذاتية مباشرة إذ كانت كل صورها في محيط حياة الشاعر وقريته ومعاناته الشخصية لكنه بعد قراءته لأشعار الغربيين وتوسع أفق رؤيته الفكرية جاء توظيفه للطفولة ببنية تعبيرية عالية التي وضعت مع مجالات توظيفات أخرى في مصافي كبار الشعراء الذين يتحول في نتائجهم كل معنى أو فكرة إلى رموز ودلالات كبرى وعميقة. تراوحت دلالة الطفولة في شعر السياب بالإيجابية دائما حتى في تلك الصور التي أظهر فيها الطفل
Stego-cover	إن الطفولة هي المفتاح إلى شخصية بدر شاكر السياب وإلى أدبه وكل محاولة لفهم إنسانا وأديبا خارج دائرة الطفولة هي محاولة ناقصة. لقد تشبث بلفظة الطفل والطفولة ودلالاتهما تشبثا يعكس عمق معاناته وفرط حساسيته تجاه مرحلة قاسية من عمره رافقته آثارها حتى الكبر. لقد كانت بواكير قصائده تتضمن جزءا كبيرا من صور طفولته في جيور حيث اليتيم والحرمان العاطفي ثم الافتراق عن الأهل والحقيقة أن قصائده الأولى كانت عاطفية ذاتية مباشرة إذ كانت كل صورها في محيط حياة الشاعر وقريته ومعاناته الشخصية لكنه بعد قراءته لأشعار الغربيين وتوسع أفق رؤيته الفكرية جاء توظيفه للطفولة ببنية تعبيرية عالية التي وضعت مع مجالات توظيفات أخرى في مصافي كبار الشعراء الذين يتحول في نتائجهم كل معنى أو فكرة إلى رموز ودلالات كبرى وعميقة. تراوحت دلالة الطفولة في شعر السياب بالإيجابية دائما حتى في تلك الصور التي أظهر فيها الطفل

Figure (8): The proposed technique example of embedding layer one.

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

It can be concluded from case one that it is visually easy to find the locations of secret message that is embedded in stego-cover.

Case two: An example of applying the proposed technique using embedding layer one and applying the proposed technique (embedding layer one and layer two) is depict in Figure (9) using the same secret message.

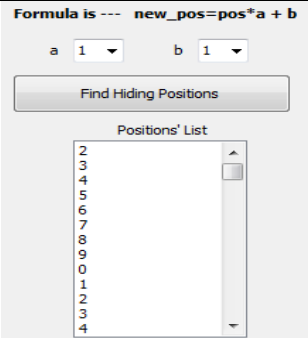
Cover	يعرف بعضهم الاغتراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالا وثيقا فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانا مغتربا على وفق الخيارات المتاحة أمامه. وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما تسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. وليقين الشاعر بأن معركته خاسرة لانعدام تكافؤ قوى الصراع فقد وجد نفسه غريبا في محيط قاس بعد أن وعى المسألة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في اليور الثورية الجماعية حيناً آخر وقد ينكفي على نفسه لانذا بها هاربا من الواقع ومعتر لا المجتمع
Secret message	جنتها والضحي يزرع الشمس
DRLR	
Stego-cover FFT	به صيرورية تتكون يعرف بعضهم الاغتراب بأنه عم من ثلاث مراحل متصلة اتصالا وثيقا فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانا مغتربا على وفق الخيارات المتاحة أمامه. وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما تسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. وليقين الشاعر بأن معركته خاسرة لانعدام تكافؤ قوى الصراع فقد وجد نفسه غريبا في محيط قاس بعد أن وعى المسألة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في اليور الثورية الجماعية حيناً آخر وقد ينكفي على نفسه لانذا بها هاربا من الواقع ومعتر لا المجتمع ولم
Stego-cover Layer one	يعرف بعضهم الاغتراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالا وثيقا فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانا مغتربا على وفق الخيارات المتاحة أمامه. وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما تسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. وليقين الشاعر بأن معركته خاسرة لانعدام تكافؤ قوى الصراع فقد وجد نفسه غريبا في محيط قاس بعد أن وعى المسألة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في اليور الثورية الجماعية حيناً آخر وقد ينكفي على نفسه لانذا بها هاربا من الواقع ومعتر لا المجتمع ولم
Stego-cover Layer two	يعرف بعضهم الاغتراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالا وثيقا فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانا مغتربا على وفق الخيارات المتاحة أمامه. وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما تسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. وليقين الشاعر بأن معركته خاسرة لانعدام تكافؤ قوى الصراع فقد وجد نفسه غريبا في محيط قاس بعد أن وعى المسألة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في اليور الثورية الجماعية حيناً آخر وقد ينكفي على نفسه لانذا بها هاربا من الواقع ومعتر لا المجتمع ولم

Figure (9): The proposed technique example of embedding.

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

It can be concluded from case two that it is visually not easy to find the locations of secret message that is embedded in stego-cover.

Case three: An example result of applying the proposed technique using embedding layer one. The steganography has no change, this state indicates to robustness, as depicted in Figure (10).

Stego-cover scanner .PDF Layer one	<p>يعرف بعضهم الاغراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالاً وثيقاً فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانياً مغترباً على وفق الخيارات المتاحة أمامه وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما اتسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. ولقّين الشاعر بأن معركته خاسرة لعدم تكافؤ قوى الصراع فقد وجد نفسه غريباً في محيط فأس بعد أن وعى المأساة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في البؤر الثورية الجماعية حيناً آخر وقد يتكفى على نفسه لأننا بها هارباً من الواقع ومعتز لا المجتمع ولم</p>
Stego-cover .DOCX Layer one	<p>يعرف بعضهم الاغراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالاً وثيقاً فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانياً مغترباً على وفق الخيارات المتاحة أمامه وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما اتسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. ولقّين الشاعر بأن معركته خاسرة لعدم تكافؤ قوى الصراع فقد وجد نفسه غريباً في محيط فأس بعد أن وعى المأساة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في البؤر الثورية الجماعية حيناً آخر وقد يتكفى على نفسه لأننا بها هارباً من الواقع ومعتز لا المجتمع ولم</p>

Figure (10): The proposed technique example of robustness in layer one.

Case four : An example result of applying the proposed technique using embedding layer two the steganography has no change, this state indicates to robustness, as depicted in Figure (11).

Stego-cover scanner .PDF Layer two	<p>يعرف بعضهم الاغراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالاً وثيقاً فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانياً مغترباً على وفق الخيارات المتاحة أمامه وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما اتسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. ولقّين الشاعر بأن معركته خاسرة لعدم تكافؤ قوى الصراع فقد وجد نفسه غريباً في محيط فأس بعد أن وعى المأساة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في البؤر الثورية الجماعية حيناً آخر وقد يتكفى على نفسه لأننا بها هارباً من الواقع</p>
Stego-cover .DOCX Layer two	<p>يعرف بعضهم الاغراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالاً وثيقاً فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانياً مغترباً على وفق الخيارات المتاحة أمامه وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كما اتسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. ولقّين الشاعر بأن معركته خاسرة لعدم تكافؤ قوى الصراع فقد وجد نفسه غريباً في محيط فأس بعد أن وعى المأساة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في البؤر الثورية الجماعية حيناً آخر وقد يتكفى على نفسه لأننا بها هارباً من الواقع</p>

Figure (11): The proposed technique example of robustness in layer two.

Case five: In this proposed technique, when deleting all kashidas, hidden information is retained. Because the secret message is hidden in FFT in LSB and the FFT is transformed to IFFT in layer one, the secret

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

message is not known by the attacker. This technique gives high security. After apply Jaro-Winkler method, as depicted in Table (1), Table (2), Table (3), and Table (4). The Jaro-Winkler method measures distance, the similarity between two strings.

The Jaro distance is: $dj = \frac{1}{3} \left(\frac{m}{|s1|} + \frac{m}{|s2|} + \frac{m-t}{m} \right)$ where

$t = \max\{[|S1|, |S2|]/2\} - 1$. explained in table (1) below:

If the word is يتعاطفون without stego, $dj = 1/3(8/8+8/8+8-1/8) = 0.9583$

where $t = 1$

If the word is يتعاطفون stego in layer one, $dj = 1/3(9/9+9/9+9-1/9) = 0.9626$ where $t=2$

else the word is 'يتعاطفون' stego in layer two, $dj = 1/3(11/11+11/11+11-2/11) = 0.9393$

else the word is "يتعاطفون" stego in layer two, $dj = 1/3(11/11+11/11+11-2/11) = 0.9393$

Table (1): Similarity between cover and stego cover in layer one.
cover without stego

	ي	ت	ع	ا	ط	ف	و	ن
ي	1	0	0	0	0	0	0	0
ت	0	1	0	0	0	0	0	0
ع	0	0	1	0	0	0	0	0
ا	0	0	0	1	0	0	0	0
ط	0	0	0	0	1	0	0	0
ف	0	0	0	0	0	1	0	0
-	0	0	0	0	0	0	0	0
و	0	0	0	0	0	0	1	0
ن	0	0	0	0	0	0	0	1

Stego cover
Layer one

Table (2): Similarity between cover and stego cover in layer two.
cover without stego

	ي	ت	ع	ا	ط	ف	و	ن
'	0	0	0	0	0	0	0	0
ي	1	0	0	0	0	0	0	0
ت	0	1	0	0	0	0	0	0
ع	0	0	1	0	0	0	0	0
ا	0	0	0	1	0	0	0	0
ط	0	0	0	0	1	0	0	0
ف	0	0	0	0	0	1	0	0
-	0	0	0	0	0	0	0	0
و	0	0	0	0	0	0	1	0
ن	0	0	0	0	0	0	0	1
'	0	0	0	0	0	0	0	0

Stego cover
Layer two
Single Quotation

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

Table (3): Similarity between cover and stego cover in layer two.

	ي	ت	ع	ا	ظ	ف	و	ن
"	0	0	0	0	0	0	0	0
ي	1	0	0	0	0	0	0	0
ت	0	1	0	0	0	0	0	0
ع	0	0	1	0	0	0	0	0
ا	0	0	0	1	0	0	0	0
ظ	0	0	0	0	1	0	0	0
ف	0	0	0	0	0	1	0	0
-	0	0	0	0	0	0	0	0
و	0	0	0	0	0	0	1	0
ن	0	0	0	0	0	0	0	1
"	0	0	0	0	0	0	0	0

Stego cover
Layer two
Double Quotation

Table (4): The explaining hide capacity ratio in proposal algorithm.

No of cover	Secret message size (Byte)	Secret message size (KB)	Carrier file size (Byte)	Carrier file size (KB)	Average of hide capacity ratio %
1	21505	21	33792	33	0.6363 B or KB
2	21505	21	46080	45	0.5666 B or KB

Case six: This proposed technique shows very high transparency, because the secret message is not seen by human vision and is not clear to attack, especially when the text is without kashida and single-double quotation, as depicted in Figure (12).

cover	يعرف بعضهم الاغتراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالا وظيفيا فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانا معتزبا على وفق الخيارات المتاحة أمامه. وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كاتسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. وليقين الشاعر بأن معركته خاسرة لانعدام تكافؤ قوى الصراع فقد وجد نفسه غريبا في محيط قاس بعد أن وعى المسألة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في البؤر الثورية الجماعية حيناً آخر وقد ينكفي على نفسه لاندأ بها هاربا من الواقع ومعتز لا المجتمع
Stego-cover Layer one	يعرف بعضهم الاغتراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالا وظيفيا فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانا معتزبا على وفق الخيارات المتاحة أمامه. وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كاتسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. وليقين الشاعر بأن معركته خاسرة لانعدام تكافؤ قوى الصراع فقد وجد نفسه غريبا في محيط قاس بعد أن وعى المسألة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في البؤر الثورية الجماعية حيناً آخر وقد ينكفي على نفسه لاندأ بها هاربا من الواقع ومعتز لا المجتمع ولم
Stego-cover Layer two	يعرف بعضهم الاغتراب بأنه عملية صيرورية تتكون من ثلاث مراحل متصلة اتصالا وظيفيا فالمرحلة الأولى تتكون نتيجة لوضع الفرد في البناء الاجتماعي ويتدخل وعي الفرد لوضعه في تشكيل المرحلة الثانية أما المرحلة الثالثة فتعكس على تصرفه إنسانا معتزبا على وفق الخيارات المتاحة أمامه. وغربة الشعراء الرواد الأربعة لا تخرج عن نطاق المراحل المذكورة فقد ربطت المؤسسة السياسية القائمة آنذاك العراق بعجلة الاستعمار وصارت الحريات العامة والشخصية وأفقرت الشعب كاتسم الوضع الاجتماعي بسيادة القوى المحافظة وجمود التقاليد. وكان على الشاعر أن يدخل معركة الحرية والتقدم ضد القوى والمؤسسات الحاكمة والمتنفذة. وليقين الشاعر بأن معركته خاسرة لانعدام تكافؤ قوى الصراع فقد وجد نفسه غريبا في محيط قاس بعد أن وعى المسألة وكافح من أجل الخلاص دون جدوى. ومن هنا فقد تشكل في داخله رد الفعل المناسب على وفق قدرته ووعيه فقد يتابع مواجهة التحدي ومن خلال أساليب تتخذ طابع التمرد الفردي حيناً وتندمج في البؤر الثورية الجماعية حيناً آخر وقد ينكفي على نفسه لاندأ بها هاربا من الواقع

Figure (12): The proposed technique example Transparency in layer one and layer two

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

Case seven: In this proposed technique the capacity changes during hiding a secret message, because in the first state Arabic text is converted to FFT and second state is addition of the kashida in layer one, and injection is single-Double quotation in layer two. The amount of hiding data is increased in cover, because addition and injection in file carrier imply relative increase in stego cover. The equation below shows this:

Hidden Ratio = amount of hidden data / carrier file size

For example:

Hide ratio = 21KB/33 KB = 0.6363

Hide ratio = 21 KB/ 45 KB = 0.5666

V. CONCLUSION:

In this paper a new layer of Arabic language steganography is implemented using the FFT and Kashida implemented as an embedding process, and DRLR as random location generator to embed the Arabic secret message in the Arabic script. Some conclusions are presented below:

1. Applying Steganography methods to document (text) files as a cover which is written in Arabic language is difficult, because of the visual sensitivity of Arabic letters to any minor change as in case one.
2. The DRLR is fast search algorithm, which is improved to be used as means to locate random positions in the cover media (Arabic scripts) to perform the embedding operation, this position can be considered as secret key.
3. Embedding methods, usually frequency methods are harder against attack than time domain method, so using FFT and Kashida as embedding method, improves security against attack.
4. The proposed algorithm prohibits any change in carrier (Arabic script) during the transmission process since the hidden secret message does not change the cover (Arabic script) file properties such as, file size, content, and format during the transmission.
5. The proposed algorithm improves the transparency property by hiding secret message inside the Arabic script using FFT. In addition another layer of hiding is applied using Kashida.
6. The proposed algorithm improves the security property by hiding secret message inside the Arabic script using FFT and applying kashida as first layer then applying single or double quotation as second layer to the rest of Arabic script.
7. This algorithm has more capacity after hiding a secret message in cover Arabic text as the equation is:

Hidden Ratio = amount of hidden data / carrier file size

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

References:

- [1] H. M. Salman, " A Natural Language Steganography Technique for Text Hiding Using LSB's", Eng.&Tech. Vol.26,No3,2008.
- [2] X. Hu, G. Luo, Y. Lu, and L. Xiang,"A Steganography on Synonym Frequency Distribution", Advances in information Sciences and Service Sciences(AISS), Vol.5, no.10, May 2013.
- [3] R. Din, A. Samsudin, and P. Lertkrai," A Framework Components for Natural Language Steganalysis", International Journal of Computer Theory and Engineering, Vol. 4, No. 4, August 2012.
- [4] S. Mao , A. Rosenfeld , and T. Kanungo," Document Structure Analysis Algorithms: A Literature Survey", Center for Automation Research University of Maryland, College Park. Available at: www.lhncbc.nlm.nih.gov/files/archive/pub2003015.pdf
- [5] A. Monem S. Rahma , H. B. Abdul Wahab, and A. Y. Al-Noori, " Proposed Steganographic Method for Data Hiding in Microsoft Word Documents Structure", Al-Mansour Journal/ Issue (15) 2011.
- [6] A. Abdul-Aziz Gutub, and M. Mohammad Fattani," A Novel Arabic Text Steganography Method Using Letter Points and Extensions", International Journal of Computer, Information, Systems and Control Engineering Vol : 1 No:3, 2007.
- [7] A.-H. Fahd, G. Adnan, A.-K. Khalid, and H. Jameel, "Improving security and capacity for Arabic text steganography using 'Kashida 'extensions," presented at the IEEE/ACS International Conference on Computer Systems and Applications, 2009.
- [8] H. Shirali-Shahreza and Mohammad Shirali-hahreza,"Arabic/Persian Text Steganography Utilizing Similar Letters With Different Codes",The Arabian Journal for Science and Engineering, Volume 35, Number 1B, December 9, 2009.
- [9] A. Abdul-Aziz Gutub, W. Al-Alwani, and A. Bin Mahfoodh, "Improved Method of Arabic Text Steganography Using the Extension 'Kashida' Character", Bahria University Journal of Information & Communication Technology Vol. 3, Issue 1, December 2010.
- [10] A. Ali and F. Moayad, "Arabic text steganography using kashida extensions with Huffman code," Journal of Applied Sciences, vol. 10, pp. 436-439, 2010.
- [11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Michael Metcalf," Numerical-Recipes-in-C-Second-Edition.", Cambridge University Press; 2 edition (October 30, 1992).
- [12] K. H. Zou, K. Tuncali, S. G. Silverman, "Correlation and Simple Linear Regression", Published online 10.1148/radiol.2273011499 Radiology 2003. Available at: <http://www.spl.harvard.edu/.../spl.../Regression.pdf>

Steganography Techniques for Arabic Language Texts Utilizing Single-Double Quotation Using RNG

Assist.Prof. Dr. Hanaa M. Ahmed , PhD Students: Maisa'a A. A. Khohder

الملخص:

ان إخفاء المعلومات يعرف بأنه فن وتقنية إخفاء النصوص والرسائل بطريقة لا يستطيع اي شخص عدا المستلم من ادراك وجود الرسالة. لأن أصبحت هذه الطريقة ضرورية للحفاظ على الأنظمة من هجمات الانترنت المختلفة، ولها خطوات لإزالة أي خطر. ان إخفاء المعلومات اللغوي، نوع من النصوص المستخدمة في إخفاء المعلومات ويمثل مجموعة من التقنيات تعد بطريقة تسمح إخفاء أي بيانات رقمية داخل النصوص بالاستناد الى بعض المعرفة اللغوية. إخفاء المعلومات اللغوي يعمل على مختلف الأشكال والتطبيقات المتفاوتة.

هذا البحث يقدم طريقة جديدة باستخدام مستويين للاخفاء، المستوى الأول يخفي عن طريق دمج وإضافة ولكن المستوى الثاني يخفي عن طريق الحقن. يتضمن المستوى الأول إخفاء الرسالة السرية بت واحدة في LSB ضمن FFT وإضافة كاشيدة واحد. باستخدام DRLR وهي RNG للعثور على الموقع الذي تخفي الرسالة داخل النص. المستوى الثاني هو حقن عشوائي للاقتباس واحد أو مزدوج داخل النص. هذا البحث يقدم إخفاء المعلومات اللغوي للنصوص مكتوبة باللغة العربية، باستخدام كاشيدة وتحويل FFT على أساس استخدام تقنية جديدة بعنوان العشوائية الديناميكية للانحدار الخطي لإخفاء رسالة سرية. النهج المقترح هو محاولة لتقديم تحويل إخفاء المعلومات اللغوي باستخدام مستويات لإخفاء لتحسين تنفيذ الكاشيدة، وتحسين أمن الرسالة السرية باستخدام العشوائية الديناميكية للانحدار الخطي.

تم اختبار هذه الطريقة من الناحية الأمنية والقدرات، والشفافية، والمتانة، وقد وجد ان هذه الطريقة أفضل من الطرق السابقة. الخوارزمية المقترحة تمتاز بخصائص مثالية في إخفاء المعلومات.