Steganography Arabic Text Based on Natural Language Process Documents

*Assist.prof. Dr. Hanaa M. Ahmed
University of Technology

**PhD Students: Maisa'a A. A. Khohder
University of Technology

Abstract: Obscurity is a main reason whereas computers can not know natural language. It have made great transaction steps trend developing instrument to morphological and syntactic analyzers for Arabic . One of the manners used in security areas is steganography. The rapid development of steganography scripts, it is a large security and confidentiality problem, it becomes necessary to find appropriate protection because of the significance, accuracy and sensitivity of the data during transmitted.

In this research is offer in a new method and to use one level to hide, this level is hiding by embedding and addition. The one level is embed a secret message twice, one bit in the LSB in the FFT and the addition of one kashida and add Single-Double Quotation in the same secret message. Using Random Singular Value Decomposition (RSVD) is NRG to find positions that are hiding within the text.

Linguistic steganography is covering all the techniques that deal with using written natural language to hide secret message. in this research presents a linguistic steganography for scripts written in Arabic language, using kashida, Single-Double Quotation and Fast Fourier Transform on the bases of using new technique entitled Random Singular Value Decomposition (RSVD) as allocation to hide secret message. The proposed approach is an attempt to present a transform linguistic steganography using one level for hiding to improve implementation of kashida and Single-Double Quotation , and improve the security of the secret message by using Random Singular Value Decomposition (RSVD). Are testing this method in terms of security and capacity, transparency, and robustness and this is way better than previous methods. The proposed algorithm ideal steganography properties.

Keyword: *Arabic text, Linguistic Steganography, random Singular Value Decomposition, Kashida, Single-Double Quotation, Transform Based*

الملخص:

الغموض هو السبب الرئيسي في عدم التعرف الى اللغات الطبيعية. وانها جعلت اكبر العمليات كخطوة باتجاه التطور وتحليل قواعد علوم الصرف في اللغة العربية. واحدة من الطرق المستخدمة في المجالات إخفاء المعلومات الأمنه. ان التطور السريع لإخفاء معلومات النصوص، بامكان السرية والامنية تواجه مشكلة كبيرة، واصبح من الضروري إيجاد حماية مناسبة بسبب أهمية ودقة وحساسية البيانات أثناء إرسالها.

هذا البحث يقدم طريقة جديدة باستخدام مستوى واحد لإخفاء المعلومات، وهذا المستوى والاخفاء يتم عن طريق دمج وإضافة. تم تضمين مستوى واحد في رسالة سرية مرتين، بت واحد في LSB وتحويل FFT وإضافة كاشيدة واحد ويضيف اقتباس الفردي والزوجي لنفس رسالة سرية. باستخدام عشوائية تفكيك القيم الفردية للصورة (RSVD) و NRG لايجاد المواقع التي يخفى داخلها النص.

إخفاء المعلومات اللغوي يغطي جميع التقنيات التي تتعامل مع استخدام كتابة اللغة الطبيعية لإخفاء رسالة سرية. في هذا البحث يقدم إخفاء المعلومات اللغوية للنصوص مكتوبة باللغة العربية، وذلك باستخدام كاشيدة واقتباس الفردي والزوجي، وتحويل FFT على أسس باستخدام تقنية جديدة بعنوان عشوائية تفكيك القيم الفردية للصورة (RSVD) في ايجاد مواقع لإخفاء رسالة سرية. الطريقة المقترح هي محاولة تحويل إخفاء المعلومات اللغوي باستخدام مستوى واحد لاخفاء لتحسين تنفيذ كاشيدة والاقتباس الفردية والزوجية ، وتحسين أمن الرسالة السرية باستخدام عشوائية تفكيك القيم الفردية للصورة (RSVD). يتم اختبار هذه الطريقة من الناحية الأمنية والقدرات، والشفافية، ومتانة، وهذا هو وسيلة أفضل من الطرق السابقة. الخوارزمية المقترحة لها خصائص في إخفاء المعلومات مثالية.

Introduction:

Linguistic steganography is focused on apply changes to a cover text so as to embed secret message, in a way that the changes do not caused any unnatural or ungrammatical text. The component of natural language steganalysis framework can be categorized into three groups [1], as depicted in Figure (1), [2]:

1- Natural Language Processing Resources and Techniques

Natural Language Processing (NLP) has accumulated a great deal of fundamental knowledge in steganalysis environment. Thus, there are four components in NLP that can be considered during the construction of

natural language steganalysis framework which are identified as:[3]

- **Corpora Resources:** "Several numbers of electronic corpora are available on Internet that has been created for NLP research on such as WordNet and VerbNet. These corpora resource are considered because of their availability and accessibility".

- **Text Paraphrasing:** "One of the challenges in natural language steganalysis is to paraphrase a text in order to detect the hidden message".

- **Natural Language Parsing:** "This task can be described as reprocessing the text sentences and reproducing new structure for the sentences".

- **Linguistic Transformations:** "There are three types of lingulistic transformation in natural language environment which are synonym substitution, syntactic transformations, and semantic transformations" [3].

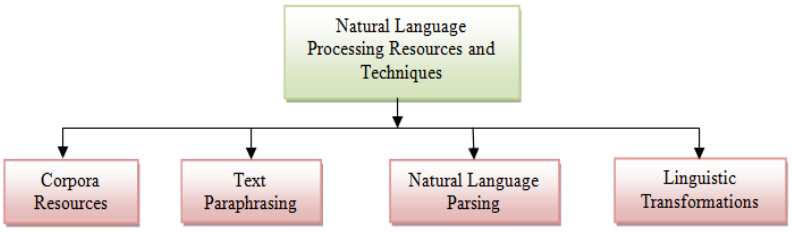

Figure (1): The component of natural language steganalysis

2- Principles of Steganography:

"Steganography involves hiding data in an overt message and doing it in such a way that it is difficult for an adversary to detect and difficult for an adversary to remove. Based on this goal, three core principles can be used to measure the effectiveness of a given steganography technique: amount of data, difficulty of detection, and difficulty of removal"[4].

- ***Amount of data:*** "suggests that the more data you can hide, the better the technique".

- ***Difficulty of detection***: "relates to how easy it is for somebody to detect that a message has been hidden. There is usually a direct relationship between how much data can be hidden and how easy it is for someone to detect it. As you increase the amount of information that is hidden in a file, you increase the chance that someone will be able to detect that there is information hidden in the file".

- ***Difficulty of removal***: "involves the principle that someone intercepting your file should not be able to remove the hidden data easily"[4].

I. LITERATURE REVIEW AND FUNDAMENTAL USED IN PROPOSED SYSTEM

Literature review

Kashida is an Arabic redundant character which is used to justify the text, without affect the meaning of words. The Researchers suggested using one kashida as bit zero, and two kashida 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 consider un-point Arabic letters followed by a kashida if the secret bit is (0), and pint Arabic letters followed by kashida if secret bit is (1).

Their technique enhanced robustness and security but might have some limitation 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 scrip to Arabic for example Persian and Urdu [5].

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, there are to conceal two characters for 'Ya' (ی) and 'Kaf' (ک). The two characters of (ی) and (ک) has itself shape however various codes which they are utilized at the start or in the center of words. The major aim in this manner is perception translucence. It have a stellar perception translucence cause the stego-text who the employee sight is aright like for the main text [6].

In 2010, Adnan Abdul-Aziz Gutub, Wael Al-Alwani, and Abdulelah BinMahfoodh, introduced an improved Arabic text steganography technique for Arabic script using kashida. The approach hides secret information as bits within Arabic letters (cover) by using extension character (kashida). The technique considers one kashida if the secret bit is (0) and two kashida if secret bit is (1) after any letter can hold it. The finishing character is embedding just after the last bit of the secret information, then the kashida as is embed randomly to the rest script in

order to enhance the security of the technique. Also their technique enhanced security, capacity and robustness for Arabic scripts based secure communication [7].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 compressed 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 applied to other than Arabic script based secure communication, with different document formats [8].

In 2013, Ammar Oden, Khaled Elleithy, Miad Faezipour, introduced an improved Arabic text steganography technique for Arabic script using variation kashida. The approach selected one of four scenarios randomly to hides secret information as bits within Arabic letters (cover) by using kashida. 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) as first scenario , and vice versa as second senior. The third senior is adding kashida after Arabic letters if the secret bit is (1) and (0) otherwise, vice versa as forth senior. Also their technique enhanced security, complexity for Arabic script based secure communication [9].

A. 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 [12]:

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

And the mathematical formula to its inverse is [10]:

$$f(x) = \int_{-\infty}^{+\infty} F(y)\exp^{[j2\pi xy]}dy \qquad \ldots\ldots\ldots (2)$$

Where:

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 ($i^2 = -1$).

B. Singular Value Decomposition (SVD)

Singular Value Decomposition technique splits given matrix into a product of orthonormal matrices and a diagonal matrix. The mathematical formula to Singular Value Decomposition is [11].

$$A = USV^T \quad \text{..... (3)}$$

$$A = [u_1\ u_2 \dots u_m] \begin{bmatrix} s_1 & 0 & \dots & 0 \\ 0 & s_2 & \dots & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & s_n \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}, \text{ ..... (4)}$$

Let A be an $m \times n$ matrix. Performing SVD to A factorizes it into a product of orthogonal matrix, diagonal matrix and another orthogonal matrix as:

$$A = USV^T, \text{ ..... (5)}$$

Where,

A: original image matrix
U: $m \times m$ product of orthogonal matrix
S: $m \times n$ diagonal matrix
V: $n \times n$ orthogonal matrix

C. Random Singular Value Decomposition (RSVD)

Is a new technique to generate a set of random position (x_i, y_j) to apply the embedding algorithm, from decomposing original image (A) using SVD, the result is B. Detect the non-zeros elements, and converges into nearest integer the results is RSVD. Figure (2) is example for the original image.

Figure (2): Original image example.

The original image A is decomposed into three matrixes: U of size mi × mj matrix, V of size mi × nj matrix, and D of size ni × nj matrix. The new array B = S*V*D as depicted in Figure (3), indicates random location from original image A.

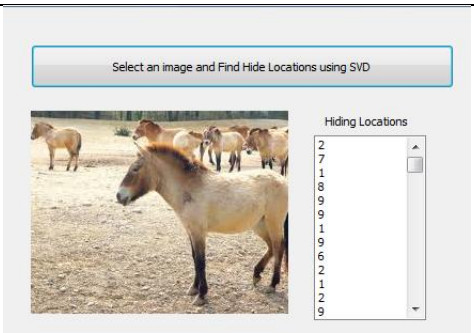

Figure (3): Random location generated using RSVD.

Algorithm RSVD:
Input: original image A
Output: random number
Process:
Step1: input A
Step2: Apply SVD algorithm to A
Step1: $B = U*S*V^T$
Step2: For i= 1 to Length of B
Step3: No= B[i];
Step4: No= absolute (No);
Step5: While (integer (No)==0)
Step6: No =No*10;
Step7: RSVD[i]=(integer (No));
Step8: Next
End of algorithm

III. PROPOSED SYSTEM

IV. IDEA FOR PROPOSED SYSTEM

The proposed approach main idea as depicted in Figure (4) the embedding, and Figure (5) the extraction, is to use RSVD as generated random offset location, to added random kashida characters and single-double quotation to the rest Arabic word texts as a first layer, where the first

layer is inject the secret message bits in the inverse FFT (LSB of (real (FFT) of selected Arabic text word))), and then apply one kashida character. The first addition of kashida is for the hiding process of the secret information, while the second addition of the first layer is inject again the secret message bit using single-double quotation is for confusion purpose to insure security of the secret message.

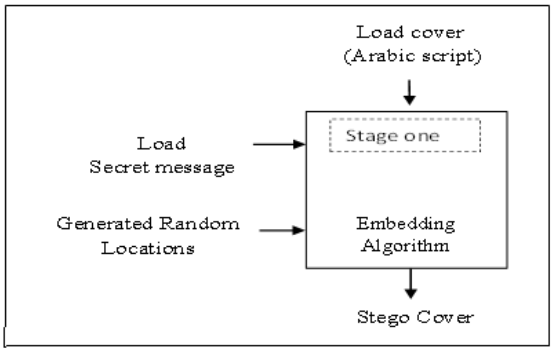

Figure (4): The proposed hiding process.

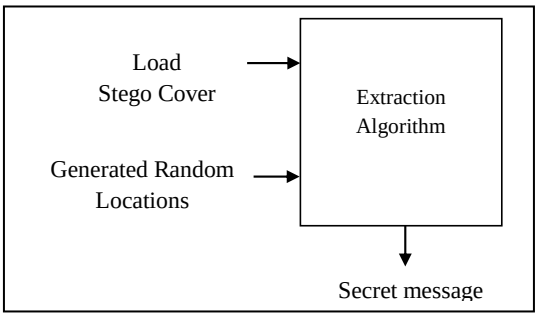

Figure (5): The proposed extraction process.

Embeding process

- The flowchart of embeding process, as shown in Figure (6).

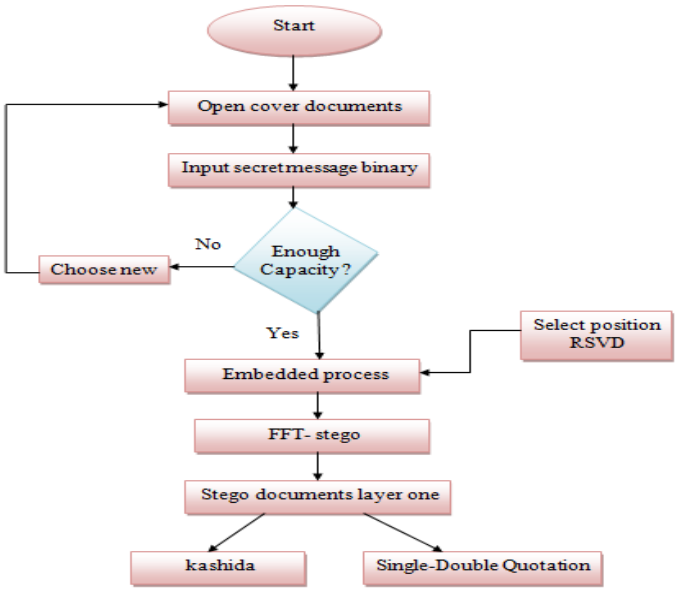

Figure (6): The flowchart of embedding process.

Embedding Algorithm:
Input: secret message, image A, set of Arabic texts.

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 (8) presents the binarization process to secret

message. Figure (7) is a simple example of applying binarization process to secret message.

Generate Random positions: The process of generated random positions, using RVSD, starts by using applying SVD algorithm to the input image (A) to generate a sequence of random values C that represents offset of Arabic text 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.

Step2. Cover selection: select Arabic text (cover) that can hold input secret message bits.
Step3. Do while not end of Arabic text words
Step4. Embedding layer one: For each secret message bit and Generate Random positions do
Step5. Use c value as offset to next word to embed the secret message bit, into inverseFFT (LSB (real (FFT (select Arabic text word)))), then apply one kashida if the secret message bit is one or if the secret message bit is zero.
Step6. End of For.
Step7. Else
Step8. Embedding layer one: add again secret message bit single-double quotation randomly to the rest Arabic text words
Step9. End of Do.
Step10. End.

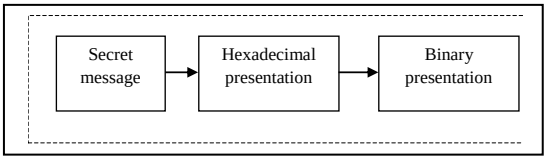

Figure (7): Secret message Binarization.

Secret Message	يا غربة الروح في دنيا من العجز
Hexadecimal Representation	20FDC72020EED1A9C92020C7FBD1E8AE2020BAF D2020CFF2FDC72020EFF22020C7FBAEADD1
Binary Representation	00100000111111011100011001000000001000 00111011101000110101001110010010010 00000010000011000111111110111101000111 10100010101110001000000010000010110101 11111101001000000010000011001111111100 10111111011100011100100000001000001110 11111111001000100000001000001100011111 11101110101110101011011101000 1

Figure (8): Secret message binarization Example.

a. Extraction Process

- The flowchart of extraction process, as shown in Figure (9).

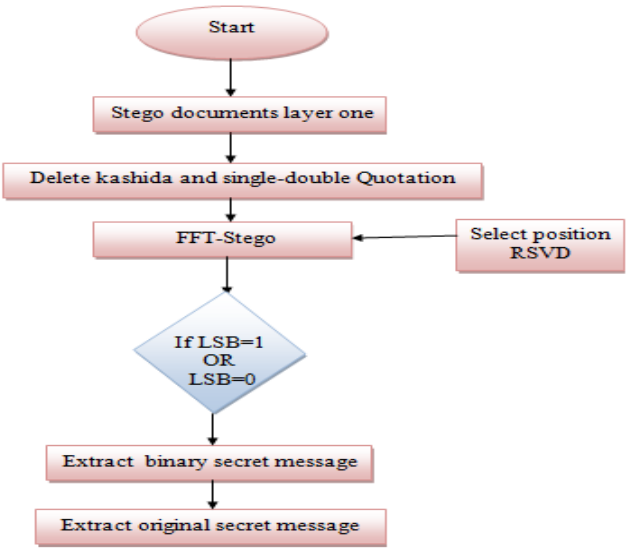

Figure (9): The flowchart of extraction process

Extraction Algorithm:

Input: secret message, image A, stego cover.
Output: secret message.
Process:

Step1. Generate Random positions: The process of generated random positions, using RSVD, start by using applying SVD algorithm to the input image (A) to generate a sequence of random values C that represents offset of Arabic text words to start the extraction process.
Step2. Loading: Load stego-cover, and Generate Random positions.
Step3. For each Generate Random Positions do
Step4. Use C value as offset to next word to extract the secret message bit, from
LSB of select Arabic text word (stego-cover).
Step5. End of For.
Step6. Converts each seven bits into one letter the result is the secret message.

End.

V. RESULTS AND DISCUSSION

In this section discuses to cases to ensure the proposed technique security:

Case one: Embedding Kashida

An example of result of applying the proposed technique using embedding layer

one, as depicted in Figure (10).

Cover	لذى السياب رموز شخصية متفردة خلقها ليلقيها على أذهننا ولتتفاعل معه من خلالها و تتأثر قله المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق المتلقي. من بين هذه الرموز الشخصية التي انفردا السياب ذكرنا رمز الموت و البعث الذي تردد في ديوانه كثيرا. و جيكورمسقط رأس الشاعر و بويب النهر الذي ذكره السياب كثيرا في ديوانه و هو يجري قرب منازل أهله وذويه والمطر أو الماء و وفيقة ابنة عم الشاعر و حبيبته التي تزوجت من
Secret message	يا غربة الروح في دنيا من الحجر
RSVD	
Stego-cover FFT	بلقيها ور شخصية متفردة خلقهنذلذى السياب ر فلهعلى أذهننا ولتتفاعل معه من خلالها و تتأثر المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق المتلقي. مم بين هذه الرموز الشخصية التي ذكرنا رمز الموت و البعث الذي ترددانفردا السياب في ديوانه كثيرا. و جيكورمسقط رأس الشاعر و بويب النراالذي ذكره السياب كثيرا في ديوامو و هو أهله وذويه والمطراأو الماء و وفيفه ةيجري قرا منار
Stego-cover Using kashida	لذى السياب رموز شخصية متفردة خلقها ليلقيها على أذهننا ولتتفاعل معه من خلالها و تتأثر قله المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق المتلقي. من بين هذه الرموز الشخصية التي انفردا السياب ذكرنا رمز الموت و البعث الذي تردد في ديوانه كثيرا. و جيكورمسقط رأس الشاعر و بويب النهر الذي ذكره السياب كثيرا في ديوانه و هو يجري قرب منازل أهله وذويه والمطر أو الماء و وفيقة ابنة عم الشاعر و حبيبته التي تزوجت

Figure (10): Proposed technique example of embedding kashida layer one.

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

Case two: **Embedding Single-Double Quotation**

An example for applying the proposed technique using embedding layer one kashida and applying the proposed technique (embedding layer one put again single or double quotation) as depict in Figure (11) using the same secret message is hide twice.

Cover	
Secret message	باعربة الروح لاشمس فانتلق
RSVD	
Stego-cover FFT	
Stego-cover Using kashida	
Stego-cover Using single-double quotation	

Figure (11): proposed technique example of embedding layer one.

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

Case three: Robustness

An example of result of applying the proposed technique using embedding layer one the steganography have no change, this state indicates to robustness in kashida or single double quotation. As depicted in figure (12).

Stego-cover scanner .PDF Layer one kashida	ذى السباب رموز شخصية مقدرة حقها ليلقيها على أدهنا وتتقاطع معه من خلالها و تشكر قله المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق المتلقي. من بين هذه الرموز الشخصية التي اقترنا السباب تذكرنا رمز الموت و البعث الذي ترند في نيوانه كثيرا. و جيكورسقط رأس الشاعر و يوبب النهر الذي ذكره السباب كثيرا في نيوانه و هو يجري قرب منازل أهله ونديه والمطر أو الماء و وفيقة ابنة عم الشاعر و حبيبته التي تزوجت من غيره ثم ماتت في سن مبكر. و القناع تقنية رمزية أخرى قد يستعير الشاعر بواسطتها شخصية تاريخية أو أدبية أو دينية و يضفي عليها من ملامحه فيتحدا لتكون ناطقة بلسانه ومعبرة عن حاله وحالة لموافقه. واستطاع السباب أن يعبر عن مشاعره الكامنة بواسطة هذا القن فجعل من شخصية أيوب و تموز و المسيح و سيزيف أقنعة في شعره ليبرز من خلالها تارة يأسه و اسيائه و ضجره من الدنيا و حيناً أمله بالشفاء و الخلاص و الإنبعاث من جديد. إنما الرمزية هي أحد الأبعاد الشعرية عند بدر شاكر السباب الشاعر الذي كان
Stego-cover .DOCX Layer one kashida	ذى السباب رموز شخصية مقدرة حقها ليلقيها على أدهنا وتتقاطع معه من خلالها و تشكر قله المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق المتلقي. من بين هذه الرموز الشخصية التي اقترنا السباب تذكرنا رمز الموت و البعث الذي ترند في نيوانه كثيرا. و جيكورسقط رأس الشاعر و يوبب النهر الذي ذكره السباب كثيرا في نيوانه و هو يجري قرب منازل أهله ونديه والمطر أو الماء و وفيقة ابنة عم الشاعر و حبيبته التي تزوجت من غيره ثم ماتت في سن مبكر. و القناع تقنية رمزية أخرى قد يستعير الشاعر بواسطتها شخصية تاريخية أو أدبية أو دينية و يضفي عليها من ملامحه فيتحدا لتكون ناطقة بلسانه ومعبرة عن حاله وحالة لموافقه. واستطاع السباب أن يعبر عن مشاعره الكامنة بواسطة هذا القن فجعل من شخصية أيوب و تموز و المسيح و سيزيف أقنعة في شعره ليبرز من خلالها تارة يأسه و اسيائه و ضجره من الدنيا و حيناً أمله بالشفاء و الخلاص و الإنبعاث من جديد. إنما الرمزية هي أحد الأبعاد الشعرية عند بدر شاكر السباب الشاعر الذي كان
Stego-cover scanner .PDF Layer one Single-double qoutation	"الماضي" فيها يبيان معا من "القروحا" البعث يتغلغل أيضاً في ثنايا قصائده التي قالها" و هو يشبه "مقاومة" حركة المد الشيوعي في "العراق" فقد اعتد فيها التوجه إلى "أسطورة" تونيس و "عشتار" وكان "يستمد" من "الأسطورتين" "وأمثالها" شعوره بأن "الغضب" ذى السباب رموز "شخصية" مقدرة حقها ليلقيها على "أدهنا" وتتقاطع معه من "خلالها" و تتأثر قله "المقدرة الكافية لكي يجعل من هذه "الرموز" "الشخصية" رموزا مشتركة ذات وقع و تأثير في "أعماق" المتلقي من بين هذه "الرموز" "الشخصية" التي اقترنا السباب تذكرنا رمز الموت و "البعث" الذي ترند في نيوانه كثيرا و هيكورسقط رأس "الشاعر" و يوبب النهر الذي ذكره السباب كثيرا في "نيوانه" و هو يجري قرب منازل أهله ونديه والمطر أو "الماء" و وفيقة ابنة عم الشاعر و "حبيبته" التي تزوجت من غيره ثم ماتت في سن مبكر و "القناع" تقنية رمزية أخرى قد "يستعير" "الشاعر بواسطتها شخصية تاريخية أو "أدبية" أو دينية و يضفي عليها من "ملامحه" فيتحدا لتكون ناطقة بلسانه ومعبرة عن حاله "وحالة" لموافقه واستطاع السباب أن يعبر عن
Stego-cover .DOCX Layer one Single-double qoutation	"الماضي" فيها يبيان معا من "القروحا" البعث يتغلغل أيضاً في ثنايا قصائده التي قالها" و هو يشبه "مقاومة" حركة المد الشيوعي في "العراق" فقد اعتد فيها التوجه إلى "أسطورة" تونيس و "عشتار" وكان "يستمد" من "الأسطورتين" "وأمثالها" شعوره بأن "الغضب" ذى السباب رموز "شخصية" مقدرة حقها ليلقيها على "أدهنا" وتتقاطع معه من "خلالها" و تتأثر قله "المقدرة الكافية لكي يجعل من هذه "الرموز" "الشخصية" رموزا مشتركة ذات وقع و تأثير في "أعماق" المتلقي من بين هذه "الرموز" "الشخصية" التي اقترنا السباب تذكرنا رمز الموت و "البعث" الذي ترند في نيوانه كثيرا و هيكورسقط رأس "الشاعر" و يوبب النهر الذي ذكره السباب كثيرا في "نيوانه" و هو يجري قرب منازل أهله ونديه والمطر أو "الماء" و وفيقة ابنة عم الشاعر و "حبيبته" التي تزوجت من غيره ثم ماتت في سن مبكر و "القناع" تقنية رمزية أخرى قد "يستعير" "الشاعر بواسطتها شخصية تاريخية أو "أدبية" أو دينية و يضفي عليها من "ملامحه" فيتحدا لتكون ناطقة بلسانه ومعبرة عن حاله "وحالة" لموافقه واستطاع السباب أن يعبر عن

Figure (12): Proposed technique example robustness in layer one.

Case four: Security

In this proposed technique, the secret message is hidden in FFT in LSB and the FFT is transformed to IFFT in layer one, the secret message is not known by the attacker. Thus where all kashidas and single double quotation in layer one are deleted, data can be retained in the hide of secret message in LSB, This technique gives high security. The Jaro-Winkler is applied method, as depicted in table (1), table (2), table (3), and table (4).

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

If the word is ليلقيها without stego, dj=1/3(7/7+7/7+7-1/7) = 0.9523

where t = 1

If the word is ليـلقيها stego in layer one, dj= 1/3(8/8+8/8+8-1/8) = 0.9583

else

If the word is ملامحه stego in layer one, dj= 1/3(6/6+6/6+6-1/6) = 0.9444

where t=2

If the word is 'ملامحه' stego in layer one, dj= 1/3(8/8+8/8+8-2/8) = 0.9166

else

If the word is الشاعر stego in layer one, dj= 1/3(6/6+6/6+6-1/6) = 0.9444

where t=2

If the word is "الشاعر " stego in layer one, dj= 1/3(8/8+8/8+8-2/8) = 0.9166

cover without stego

Stego cover Layer one kashida	ل	ي	ل	ق	ي	ه	ا
ل	1	0	0	0	0	0	0
ي	0	1	0	0	0	0	0
-	0	0	0	0	0	0	0
ل	0	0	1	0	0	0	0
ق	0	0	0	1	0	0	0
ي	0	0	0	0	1	0	0
ه	0	0	0	0	0	1	0
ا	0	0	0	0	0	0	1

Table (1): Similarity between cover and stego cover in layer one.

cover without stego

	م	ل	ا	م	ح	ه
'	0	0	0	0	0	0
م	1	0	0	0	0	0
ل	0	1	0	0	0	0
ا	0	0	1	0	0	0
م	0	0	0	1	0	0
ح	0	0	0	0	1	0
ه	0	0	0	0	0	1
'	0	0	0	0	0	0

Stego cover
Layer one
Single quotation

Table (2): Similarity between cover and stego cover in layer one.

cover without stego

	ا	ل	ش	ا	ع	ر
"	0	0	0	0	0	0
ا	1	0	0	0	0	0
ل	0	1	0	0	0	0
ش	0	0	1	0	0	0
ا	0	0	0	1	0	0
ع	0	0	0	0	1	0
ر	0	0	0	0	0	1
"	0	0	0	0	0	0

Stego cover
Layer one
Double quotation

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

No of cover	Secret messge size (Byte)	Secret messge size (KB)	Carrier file size (Byte)	Carrier file size (KB)	Average of hide capacity ratio %
1	22528	22	54272	53	0.4150 B or KB
2	22528	22	81920	80	0.275 B or KB

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

Case five: Transparency

In this proposed technique is very high transparency, because not seen in human vision and not clear for attack. Especial when the text without kashida and single-double quotation. As depicted in figure (13) and figure (14).

cover	لدى السياب رموز شخصية متفردة خلقها ليلقيها على أذهننا ولنتفاعل معه من خلالها و نتأثر فله المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق المتلقي. من بين هذه الرموز الشخصية التي انفردا السياب ذكرنا رمز الموت و البعث الذي تردد في ديوانه كثيرا. و جيكورمسقط رأس الشاعر و بويب النهر الذي ذكره السياب كثيرا فى ديوانه و هو يجري قرب منازل أهله وذويه والمطر أو الماء و وفيقة ابنة عم الشاعر و حبيبته التي تزوجت من
Stego-cover Layer one kashida	لدى السياب رموز شخصية متفردة خلقها ليلقيها على أذهننا ولنتفاعل معه من خلالها و نتأثر فله المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق المتلقي. من بين هذه الرموز الشخصية التي انفردا السياب ذكرنا رمز الموت و البعث الذي تردد في ديوانه كثيرا. و جيكورمسقط رأس الشاعر و بويب النهر الذي ذكره السياب كثيرا فى ديوانه و هو يجري قرب منازل أهله وذويه والمطر أو الماء و وفيقة ابنة عم الشاعر و حبيبته التي تزوجت

Figure (13): Proposed technique example Transparency in layer one.

cover	في ثنايا قصائده التي قالها و هو يشهد مقاوما حركة المد الشيوعي في العراق فقد اعتمد فيها اللجوء إلى أسطورة أدونيس و عشتار وكان يستمد من الأسطورتين وأمثالها شعوره بأن الخصب لابد أن يخلف. لدى السياب رموز شخصية متفردة خلقها ليلقيها على أذهننا ولنتفاعل معه من خلالها و نتأثر فله المقدرة الكافية لكي يجعل من هذه الرموز الشخصية رموزا مشتركة ذات وقع و تأثير في أعماق
Stego-cover Layer one Single-double quotation	ثنايا قصائده التي قالها و هو يشهد مقاوما حركة المد الشيوعي في العراق فقد اعتمد فيها اللجوء إلى أسطورة أدونيس و عشتار وكان يستمد من الأسطورتين وأمثالها شعوره بأن الخصب لابد أن يخلف.العربي والحضارة العربية إلا أن هذا "الموت" "ليستقيق" و لابد لأنه لا يمكن أن يحيا دون "الماضي" فهما يبيان معا من "القبر وهذا" البعث يتطلق أيضا في ثنايا قصائده التي قالها" و هو يشهد "مقاوما" حركة المد الشيوعي في "العراق" فقد اعتمد فيها اللجوء إلى أسطورة" أدونيس و "عشتار" وكان "يستمد" من "الأسطورتين" وأمثالها" شعوره بأن "الخصب" لدى السياب رموز "شخصية" متفردة خلقها ليلقيها على "أذهنا" ولنتفاعل معه من "خلالها" و نتأثر فله المقدرة الكافية لكي يجعل من هذه "الرموز" "الشخصية" رموزا مشتركة ذات

Figure (14): Proposed technique example Transparency in layer one.

Case six: Capacity

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

Hidden Ratio = amount of hidden data / carrier file size
For example:
Hide ratio = 22KB/53 KB = 0.4150
Hide ratio = 22 KB/ 80 KB = 0.257

VI. CONCLUSION:

In this paper a new layers Arabic language steganography is implemented using the FFT implementation and Kashida as an embed process, and RSVD as random location generator to embed the Arabic secret message in the Arabic script. We present some conclusions bellow:

1. Applying Steganography methods to document (text) files as a cover which is written by Arabic language is difficult, due to the visually sensitivity of Arabic letters to any miner change as in case one.
2. The RSVD is fast search algorithm, which is improved to use as means to allocate randomly positions in the cover media (Arabic scripts) to perform the embedding operation.
3. As embedding methods, usually frequency method is harder against attack than time domain method, so using FFT and Kashida as embedding method, which improve its security agents attack.
4. Algorithm robustness: The proposed algorithm prohibits any change to 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. Algorithm transparency: the proposed algorithm improvers the transparency property by hiding secret message inside the Arabic script using FFT. In addition another layer of hiding is applied using Kashida.
6. Algorithm security: the proposed algorithm improvers the security property by hiding secret message inside the Arabic script using FFT and apply kashida as first layer then apply single or double quotation as second layer to the rest Arabic script.

7. Algorithm Capacity: This algorithm is more capacity after hide a secret message in cover Arabic text as the equation is:

Hidden Ratio = amount of hidden data / carrier file size

References

[1] Hana'a M. Salman, " A Natural Language Steganography Technique for Text Hiding Using LSB's", Eng.&Tech. Vol.26,No3,2008.

[2] Xiaoxi Hu, Gang Luo, Yongjing Lu, and Lingyun 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] Eric Cole , Ronald D. Krutz," Hiding in Plain Sight: Steganography and the Art of Covert Communication", Wiley publishing, In.2003.

[5] 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.

[6] 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.

[7] Adnan Abdul-Aziz Gutub, Wael Al-Alwani, and Abdulelah 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.

[8] A. Ali and F. Moayad, "Arabic text steganography using kashida extensions with Huffman code," Journal of Applied Sciences, vol. 10, pp. 436-439, 2010.

[9] A. Odeh, K. Elleithy, and M. Faezipour, "Steganography in Arabic text using Kashida variation algorithm (KVA)," in Systems, Applications and Technology Conference (LISAT), 2013 IEEE Long Island, 2013, pp. 1-6.

[10] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, Michael Metcalf," Numerical-Recipes-in-C-Second-Edition.", Cambridge University Press; 2 edition (October 30, 1992).

[11] K. Mounika, D. Sri Navya Lakshmi, K. Alekya," SVD Based Image Compression", International Journal of Engineering Research and General Science Volume 3, Issue 2, March-April, 2015, ISSN 2091-2730.