Improved LSB Method using Standard Deviation Scale

Nadia M. Mohammed **Riyam J. Essa** **Amina M. Najim**
nadia.m.mohammed@uomosul.edu.iq
College of Computer Science and Mathematics
University of Mosul, Mosul, Iraq

Received on: 12/10/2012 **Accepted on: 30/01/2013**

ABSTRACT

The concealment of the most important means used by the security institutions with critical communications in all countries of the world, they provided the technology of high security, especially in the communication networks and the Internet. Summarized algorithm concealment improved transfer confidential letter to the formula of (Binary) and then encrypted using a key agreed upon by the two parties (sender and recipient), followed by the division of the image to be hide data where clips blocks size (8*8) and account values standard deviation (Standard Deviation (STD)) for each section are then finding less and the largest value of a standard deviation in addition to the median value, then isolate sections where the value of the standard deviation less Oomsawih of median value to be key concealment (ie be adopted as locations to hide) and that by including all bit of message into the (LSB) for each section of the selected sections. Used measures of efficiency (PSNR), (MSE), (BER) for the purpose of measuring the efficiency of the algorithm and the adoption of improved language (MATLAB).

Keywords: STD, LSB.

تحسين طريقة LSB باستخدام مقياس الانحراف المعياري

نادية معن محمد **ريام جاسم عيسى** **آمنة مؤيد نجم**
كلية علوم الحاسوب والرياضيات
جامعة الموصل، الموصل، العراق

تاريخ استلام البحث: 2012/10/12 **تاريخ قبول البحث: 2013/01/30**

الملخص

يعد الإخفاء من أهم الوسائل التي تلجأ أليها المؤسسات الأمنية ذات الاتصالات الحساسة في جميع دول العالم، لما توفره هذه التقنية من أمنية عالية وخصوصا في اتصالات الشبكات والانترنت. تتلخص خوارزمية الإخفاء المحسنة بتحويل الرسالة السرية إلى صيغة الـ(Binary) ومن ثم تشفيرها باستخدام مفتاح متفق عليه من قبل الطرفين(المرسل والمستلم) يليها تقسيم الصورة المراد إخفاء البيانات فيها إلى مقاطع blocks بحجم (8*8) وحساب قيم الانحراف المعياري ((Standard Deviation (STD) لكل مقطع بعدها يتم إيجاد اقل واكبر قيمة للانحراف معياري بالإضافة إلى القيمة الوسطية، ثم عزل المقاطع التي تكون فيها قيمة الانحراف المعياري اقل أو مساوية من القيمة الوسطية لتكون مفتاح الإخفاء (أي يتم اعتمادها بوصفها مواقع للإخفاء) وذلك من خلال تضمين كل bit من الرسالة داخل الـ (LSB) لكل مقطع من المقاطع المحددة. استخدمت مقاييس الكفاءة (PSNR) و (MSE) و(BER) لغرض قياس كفاءة الخوارزمية المحسنة وباعتماد لغة (MATLAB).

الكلمات المفتاحية: الانحراف المعياري، الخلية الاقل اهمية.

1- مقدمة

منذ العهود القديمة كانت هناك حاجة ملحة لإيجاد وسائل سرية للحفاظ على أمنية الرسائل المرسلة وخصوصاً في وقت الحروب وظهرت هناك طرائق مختلفة في هذا المجال ولكن مع تطور وسائل الاتصال وتطور علم الحاسوب أصبحت هناك حاجة ملحة لإيجاد وسائل أكثر تطوراً لخدمة هذا الغرض فكان ظهور علم جديد وهو علم التشفير وبالرغم من كونه طريقة جيدة لحفظ المعلومات إلا انه سهل الاكتشاف ويمكن لأي متطفل التلاعب به فكانت الحاجة إلى تقنية أكثر تطورا وأكثر سرية وحفاظا على المعلومات وخصوصا مع ظهور وتطور شبكة الانترنيت فتم اللجوء إلى علم آخر وهو علم الإخفاء الذي يعتمد على مبدأ أن الرسالة المرسلة تكون غير مرئية لأي شخص بواسطة إخفائها داخل إحدى وسائل الاتصال (الصوت، الصورة، النص والفيديو). ولإخفاء المعلومات أهمية كبيرة فهي تعد عاملا مساعدا على إضفاء حماية وأمناً على المعلومات وذلك بسبب عدم وضوح المعلومات للعيان عاملاً مساعداً على إضفاء حمايةً وأمناً على المعلومات.[1]

تم اقتراح العديد من تقنيات الإخفاء خلال السنوات الأخيرة، ومعظمها يمكن تحديدها كأنظمة الإبدال Transposition، فأنها تحاول إبدال أجزاء من الغطاء برسالة سرية. وتم حديثا تطوير تقنيات الكتابة المغطاة للحصول على أمنية عالية وقوية.[2]

ويمكن الدمج بين التشفير والإخفاء لزيادة مستوى السرية لأنه حتى وان تم اكتشاف الرسالة المخفية فأن رؤيتها بصيغتها المبعثرة تُزيل الشك بوجود الإخفاء وهذا ما تم استخدامه فعلا في بحثنا هذا. حيث تم في هذا البحث تحسين طريقة الإخفاء في الخلية الأقل أهمية Least Significant Bit(LSB) وذلك بدمجها مع مقياس الانحراف المعياري (وهو من المقاييس المهمة والمستخدمة بكثرة في قياس مدى تشتت الصورة) لإخفاء بيانات سرية بعد تشفيرها باستخدام عملية XOR في صور ملونة من نوع (JPG,GIF).

2- الانحراف المعياري

يمثل هذا المقياس مدى تشتت القيم وانحرافها عن وسطها الحسابي، وهو الجذر التربيعي للتباين ويرمز له بالحرف اليوناني "سيجما"(σ)، (وعلى العكس يساوي التباين مربع الانحراف المعياري ويرمز إليه غالبا بالرمز(σ^2)). توضح الصيغة التالية كيفية حساب الانحراف المعياري لمجموعة من القيم n (100=n) ومتوسط مجتمعها μ_p:-[3]

$$\sigma = [(1/100)({d_1}^2 + {d_2}^2 + \cdots + {d_{100}}^2)]^{1/2}$$
$$\sigma = \{(1/100)[(x_1 - \mu_p)^2) + (x_2 - \mu_p)^2) + \cdots + (x_{100} - \mu_p)^2]\}^{1/2}$$

وببساطة يمكن توضيح معنى التباين بأنه عبارة عن متوسط مربع درجات بعد كل قيمة عن الوسط الحسابي. ويمثل الانحراف المعياري الجذر التربيعي للتباين.[4]

3- طريقة الإخفاء في الخلية الثنائية الأقل أهمية (LSB)

تعد الصور الرقمية من أكثر الوسائط المتعددة استخداما في الكتابة المغطاة بوصفها حاملا للبيانات السرية وذلك بسبب انتشارها الواسع على الإنترنت. ومن أكثر طرائق الإخفاء شيوعا في تضمين البيانات ضمن الصور الرقمية هي طريقة إدخال الخلية الثنائية الأقل أهمية (LSB)، إذ تستبدل الخلية الثنائية الأقل أهمية من كل نقطة ضوئية في الصور الرقمية بخلية ثنائية من البيانات السرية. وفيها تبدو الصورة الناتجة بعد عملية الإخفاء مماثلة للصورة الغطاء بالنسبة للعين البشرية. [1][2]

يمكن أن تستبدل الخلية الثنائية الأقل أهمية الأولى والثانية من النقاط الضوئية مع بقاء العين البشرية غير قادرة على تمييز الفرق بين الصورتين وهناك أشكال متعددة لهذه الطريقة إذ يمكن نشر الرسالة السرية بصورة عشوائية على الغطاء باستخدام مفتاح سري يعد بذرة (Seed) لمولد أرقام عشوائية أو يمكن اختيار بعض المناطق الأقل تأثراً بالتشوهات لتضمين البيانات في الصورة بالاعتماد على خصائص الرؤية للإنسان.[5][6][7]

4- خوارزمية الإخفاء المحسنة

تتضمن خوارزمية الإخفاء إخفاء نص مشفر بطريقة(XOR) داخل صورة ذات امتداد (JPG,GIF) باستخدام طريقة LSB والانحراف المعياري وكما يأتي:-

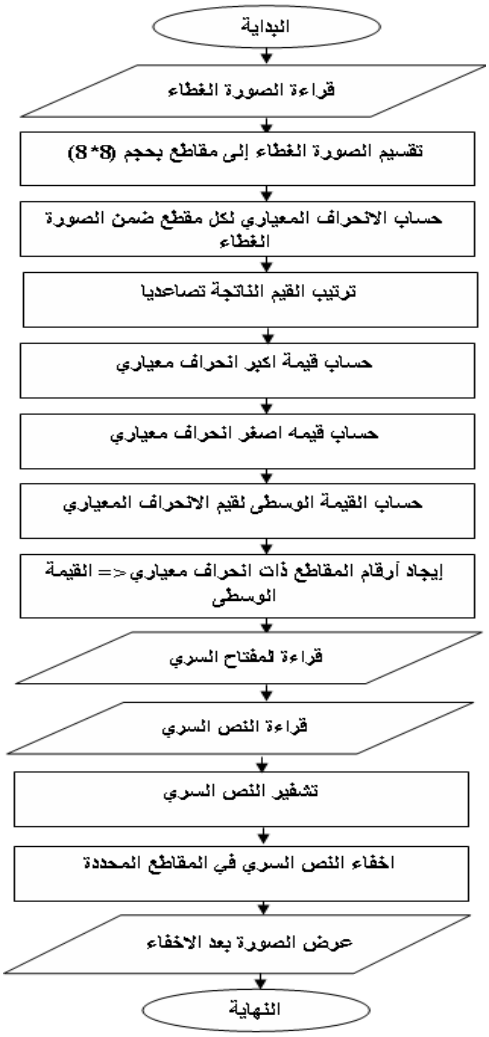
1. قراءة الصورة الغطاء.
2. تقسيم الصورة الغطاء إلى مجموعة من المقاطع (blocks)، حجم كل مقطع (8*8).
3. حساب الانحراف المعياري لكل مقطع (block).
4. ترتيب قيم الانحراف المعياري الناتجة تصاعديا.
5. حساب قيمة اكبر انحراف معياري (maxstd).
6. حساب قيمة اصغر انحراف معياري (minstd).
7. حساب القيمة الوسطى لقيم الانحراف المعياري (midstd).
8. تحديد المقاطع التي تملك قيم انحراف معياري اقل أو مساوية للقيمة الوسطى للانحراف المعياري لاعتمادها كمفتاح (key) للإخفاء.
9. قراءة المفتاح السري(الخاص بعملية التشفير).
10. تحويل المفتاح السري إلى صيغة ثنائية (binary).
11. قراءة النص السري.
12. تحويل النص السري إلى صيغة ثنائية (binary).
13. أجراء عملية XOR المنطقية بين النص السري والمفتاح السري للحصول على النص المشفر.
14. إخفاء النص المشفر ضمن المقاطع المحددة وذلك باستخدام طريقة إخفاء البت الأقل الأهمية (LSB)، حيث يتم إخفاء bit 2 من النص في الـ bit الثانية والثالثة من كل byte فردي ضمن المقطع المحدد.
15. عرض الصورة بعد الإخفاء (Stego).
16. خزن الصورة بعد الإخفاء (Stego). لاحظ الشكل (1).

5- خوارزمية الاسترجاع المحسنة

تتضمن خوارزمية الاسترجاع استرجاع النص السري من صورة ذات امتداد(JPG, GIF) باستخدام طريقة (LSB) والانحراف المعياري وكما يلي:-

1. قراءة الصورة بعد الإخفاء (stego).
2. قراءة الصورة الغطاء.
3. تقسيم الصورة بعد الإخفاء (stego) إلى مجموعة من المقاطع (blocks) بحجم (8*8).
4. تقسيم الصورة الغطاء إلى مجموعة من المقاطع (blocks) بحجم (8*8).

5. حساب الانحراف المعياري لكل مقطع (STD) للصورة الغطاء.
6. ترتيب قيم الانحراف المعياري الناتجة تصاعديا.
7. إيجاد قيمة اكبر انحراف معياري (maxstd).
8. إيجاد قيمة اصغر انحراف معياري (minstd) .
9. إيجاد القيمة الوسطى لقيم الانحراف المعياري (midstd).
10. تحديد أرقام الـ blocks التي تملك (midstd => STD) ضمن الصورة (stego).
11. استرجاع الـ bit الثاني والثالث من كل byte فردي ضمن الـblocks المحددة لاسترجاع النص المشفر.
12. أجراء عملية XOR المنطقية بين النص المشفر والمفتاح السري المتفق عليه للحصول على النص السري.
13. تحويل النص السري من صيغة ثنائية (binary) إلى صيغة حرفية (char).
14. عرض النص السري. لاحظ الشكل (2).



الشكل (1). المخطط الانسيابي لخوارزمية الإخفاء المحسنة

الشكل (2). المخطط الانسيابي لخوارزمية الاسترجاع المحسنة

6- مناقشة النتائج

يعتبر الانحراف المعياري (σ) من أدق وأفضل مقاييس التشتت (مقاييس عددية لقياس مقدار التفاوت بين البيانات) لسهولة التعامل معه في التحليل، وتم حسابه عن طريق القانون التالي[4]:

$$ó = \sqrt{(\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}} \quad \ldots(3)$$

علما أن : $i : 1 \rightarrow n$، x_i : نقاط الصورة، n : عدد نقاط الصورة، $\overline{x}$: الوسط الحسابي:

$$\overline{x} = (\sum_{i=1}^{n} x_i / n) \quad \ldots(4)$$

كما تم حساب قيمة (BER) لمعرفة فيما إذا كان النص قد استرجع بالكامل أم لا (معرفة عدد bits الخطأ التي تم استرجاعها)، عن طريق القانون الآتي [7][8][9]:-

BER=(no. of wrong bit/no.of original bit)*100 …(5)

وحساب قيمة (PSNR) للصورة الناتجة (الصورة بعد الإخفاء) حسب القانون التالي[8][9]:-

PSNR=10log10[Cmax^2/MSE] …(6)

أما حساب (MSE) فتم عن طريق القانون التالي[9]:-

MSE=1/N*M*(S-C)^2 …(7)

علما أن:

N وM: أبعاد الصورة. C وS : تمثلان الصورة الغطاء والصورة بعد الإخفاء على التوالي.

Cmax: أعلى قيمة لونية في الصورة .

طبقت الخوارزمية المحسنة على أكثر من صورة من نوع (JPG,.GIF.) وقِيسَ الانحراف المعياري لمقاطع الصورة لاحظ الجدول رقم (1)، وتم الاستنتاج إلى انه كلما قلت قيمة الانحراف المعياري كلما قلت قيمة التشتت لبيانات الصورة مما يشير إلى نتائج أفضل. كما إن قيمة BER الناتجة كانت معظمها مساوية للصفر مما يعني استرجاع النص بالكامل وبشكل صحيح وبدون أي أخطاء.

نلاحظ من الجدول رقم (2) إن قيم (PSNR) كانت عالية مما يشير إلى عدم وجود أي تغيير واضح للعيان في الصورة بعد الإخفاء مقارنة مع الصورة الغطاء لاحظ الشكلين (3)(4).

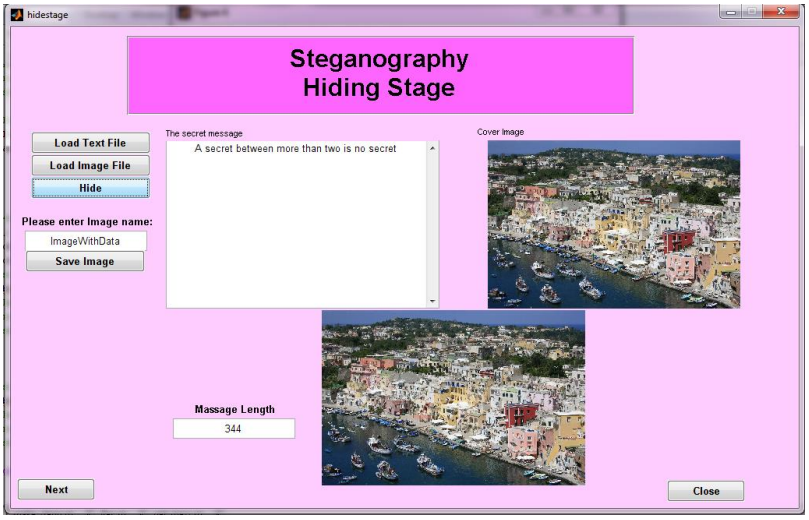
جدول رقم (1). قيم الانحراف المعياري لكل مقطع

أرقام المقاطع (Blocks)	قيم الانحراف المعياري (STD)
1	1.94
2	1.63
3	25.81
4	20.31
5	5.12
6	0.76
7	1.03
8	1.95
9	40.86
10	38.63
11	30.28
12	27.82
13	34.66
14	34.30
15	40.93
16	30.35
17	36.29
18	38.81
19	25.84
20	35.41
21	36.87
22	20.59
23	26.86
24	54.39
25	16.89
26	23.86
27	31.60
28	33.21
29	28.42
30	25.91
31	28.61
32	53.57
33	34.32
34	35.44
35	26.51
36	40.20

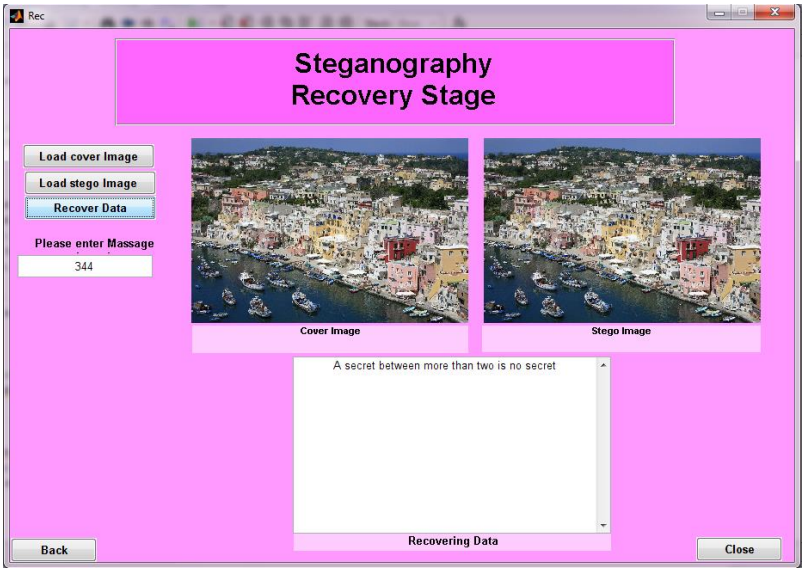
أرقام المقاطع (Blocks)	قيم الانحراف المعياري (STD)
37	26.10
38	31.68
39	47.77
40	38.44
41	23.71
42	27.18
43	31.37
44	37.59
45	30.50
46	36.96
47	54.00
48	25.88
49	9.33
50	35.35
51	33.51
52	10.78
53	44.84
54	38.57
55	21.98
56	32.34
57	29.07
58	28.57
59	24.30
60	43.85
61	7.63
62	10.54
63	41.73
64	36.06

الجدول (2). نتائج التنفيذ

BER	PSNR	MSE	اسم الصورة
0	59.198	0.021	NR1.JPG
0	60.782	0.018	NR2. JPG
0	55.088	0.029	NR3. JPG
0	52.009	0.036	NA1.GIF
0	52.156	0.035	NA2.GIF
0	48.442	0.048	NA3.GIF



الشكل (3). واجهة المرسل



الشكل (4). واجهة المستلم

7- مثال تطبيقي

7-1 مرحلة التشفير

في هذه المرحلة سيتم تشفير النص السري باستخدام عملية (XOR) لاحظ الأشكال (5)، (6)، (7):

A secret between more than two is no secret

01000001001000000111001101100101011000110111001001100101011101000001000000110001001100101011101000111011101100101011001010110111000100000011011010110111101110010011001010010000001110100011010000011000010110111000100000011101000111011101101111001000000110100101110011001000000110111001101111001000000111001101100101011000011011100100110010101110100

الشكل (5). النص السري والصيغة الثنائية له

Hi
1001000101001001

الشكل (6). المفتاح السري والصيغة الثنائية له

10010010011011010100011111011011000101110010001101011011000111001000100101100010010100000100000010100100000101000001010001101111010110110101100111011110000011111101010001111110011001110010111001000100011000101001010011100111010010110111111001100111111001010001101100000011111101010000110100111

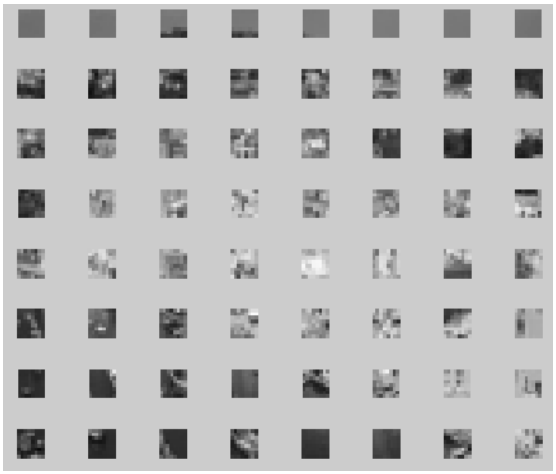
الشكل (7). النص المشفر بعملية XOR

7-2 مرحلة الإخفاء

في هذه المرحلة سيتم إخفاء النص السري المشفر وباستخدام طريقة LSB لمواقع تُحدد بالاعتماد على الانحراف المعياري، حيث يتم أولا قراءة الصورة الغطاء وتحويلها إلى صورة رمادية (Gray)، لاحظ الشكلين (8)،(9). ثم تقسيم الصورة الناتجة إلى مقاطع بحجم (8*8) لاحظ الشكل (10).

الشكل(8). الصورة الغطاء

الشكل (9). الصورة الغطاء Gray



الشكل (10). الصورة الغطاء بعد تقسيمها إلى مقاطع

- حساب قيم الانحراف المعياري لجميع المقاطع، لاحظ الشكلين (11)،(12).

1.94	1.63	25.81	20.31	5.12	0.76	1.03	1.95
40.86	38.63	30.28	27.82	34.66	34.30	40.93	30.35
36.29	38.81	25.84	35.41	36.87	20.59	26.86	54.39
16.89	23.86	31.60	33.21	28.42	25.91	28.61	53.57
34.32	35.44	26.51	40.20	26.10	31.68	47.77	38.44
23.71	27.18	31.37	37.59	30.50	36.96	54.00	25.88
9.33	35.35	33.51	10.78	44.84	38.57	21.98	32.34
29.07	28.57	24.30	43.85	7.63	10.54	41.73	36.06

الشكل (11). قيم الانحراف المعياري

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64

الشكل (12). أرقام المقاطع

- تحديد اقل وأعلى قيمة للانحراف المعياري بالإضافة إلى القيمة الوسطى لقيم الانحراف المعياري:

Min_STD=0.7664 Max_STD= 54.3901 Mid_STD=30.3533

- تحديد المقاطع التي تكون فيها قيمة الانحراف المعياري اقل أو مساوية للقيمة الوسطية للانحراف المعياري لتعتمد كمفتاح للإخفاء لاحظ الشكل (13).

1	2	3	4	5	6	7	8
11	12	16	19	22	23	25	26
29	30	31	35	37	41	42	48
49	52	55	57	58	59	61	62

الشكل (13). أرقام المقاطع المحددة

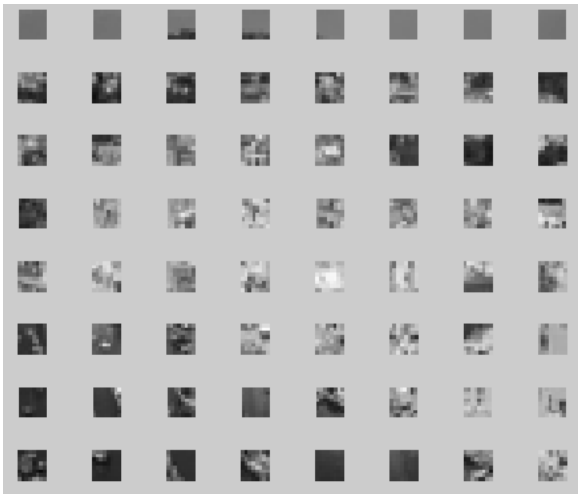
- إخفاء النص السري ضمن الصورة الغطاء والحصول على الصورة بعد الإخفاء لاحظ الشكل (14).

الشكل (14). الصورة بعد الإخفاء

3-7 مرحلة الاسترجاع

في هذه المرحلة سيتم استرجاع النص السري(المشفر) من المقاطع المحددة وكالاتي:

- قراءة الصورة بعد الإخفاء الموضحة بالشكل (14).
- تقسيم الصورة بعد الإخفاء إلى مجموعة من المقاطع بحجم(8*8)لاحظ الشكل (15).



الشكل (15). الصورة بعد الإخفاء بعد تقسيمها إلى مقاطع

- الصورة الغطاء وتحويلها إلى صورة رمادية (Gray)، كما في الشكلين (9, 8).
- تقسيم الصورة الغطاء إلى مجموعة من المقاطع بحجم (8*8) كما في الشكل (10).
- حساب قيم الانحراف المعياري لكل مقطع. كما في الشكلين (11)، (12).
- تحديد اقل وأعلى قيمة للانحراف المعياري بالإضافة إلى القيمة الوسطى لقيم الانحراف المعياري:

Min_STD=0.7664 Max_STD= 54.3901 Mid_STD=30.3533

- تحديد المقاطع التي تكون فيها قيمة الانحراف المعياري <= للقيمة الوسطية للانحراف المعياري (مفتاح للإخفاء) لاسترجاع النص السري لاحظ الشكل(16).

1001001001101101010001111011011000101110010001101011011000111001000101001011000100101000010000001010010000101000010100011011110101101101011001101111000011111101010001111100110011100101011100101100100010001100010100101001110011101001011011111100110010010001000111111100110010001100101101111111001100111111001010001101100000011111110101000110100111

الشكل (16). النص المشفر المسترجع

4-7 مرحلة فك الشفرة

في هذه المرحلة سيتم فك تشفير النص السري المسترجع لاحظ الشكلين (17)، (18):

01000001001000000111001101100101011000110111001001100101011101
00001000000110001001100101011101000111011101100101011001010110
11100010000001101101011011110111001001100101001000000111010001
10100001100001011011100010000001110100011101110110111100100000
01101001011100110010000001101110011011110010000001110011011001
01011000110111001001100101011101 00

الشكل (17). النص السري بعد فك شفرته باستخدام XOR

A secret between more than two is no secret

الشكل (18). النص السري الأصلي

8- الاستنتاجات

1- أثبتت النتائج كفاءة الخوارزمية، حيث إن المعلومات المخفية لم تحدث أي تشويه على ملفات الغطاء المستخدمة.

2- أدى تقسيم الصورة الغطاء إلى مجموعة من المقاطع إلى زيادة قوة الخوارزمية المحسنة.

3- أدى تشفير النص المخفي إلى زيادة سرية الخوارزمية المحسنة.

4- تم استرجاع النص بالكامل وهذا واضح من خلال قيم المقياس BER.

5- إن الإخفاء في المقاطع ذات قيمة <= MidSTD أفضل من الإخفاء في المقاطع ذات قيمة > MidSTD، حيث تمتاز المقاطع ذات قيمة <= MidSTD بتشتت قليل للبيانات.

6- كلما قلت قيمة الانحراف المعياري (STD) كلما زادت قيمPSNR.

المصادر

[1] الحمامي، علاء حسين والحمامي، محمد علاء، (2008)، "إخفاء المعلومات: الكتابة المخفية والعلامة المائية "، إثراء للنشر والتوزيع، الشارقة.

[2] Hsing, C., Jeng, S., (2010), "Transforming LSB Substitution for Image-based Steganography in Matching Algorithms", Journal of Information Science and Engineering.

[3] Hogg, R.V., McKean, J.W., Craig, A.T., (2005), "Introduction to Mathematical Statistics", Pearson Education International, Sixth Edition.

[4] جيبيليسكو، ستان، (2009)، "الإحصاء بوضوح"، ترجمة خالد العمري، دار الفاروق للاستثمارات الثقافية.

[5] Toumazis, Alex, (2009), "Steganography", Journal of Information Science and Engineering.

[6] Gutte, R.S., Chincholkar Y.D., (2012), "Comparison of Steganography at One LSB and Two LSB Positions", International Journal of Computer Applications, Vol. 49, No.11.

[7] محمد، همسة معن، نادية معن، شيماء شكيب، (2011)، "طريقة خوارزمية جينية مثلى للإخفاء"، المؤتمر الرابع لكلية علوم الحاسوب والرياضيات، جامعة الموصل.

[8] الجواهرجي، شيماء شكيب، (2004)، "الإخفاء في ملف صوت مكبوس"، بحث ماجستير، قسم علوم الحاسوب، كلية علوم الحاسوب والرياضيات، جامعة الموصل.

[9] سلو، أميرة بيبو، (2009)، "تقنيات إخفاء المعلومات باستخدام الشبكات العصبية وبروتوكولات الشبكة"، بحث ماجستير، قسم علوم الحاسوب، كلية علوم الحاسوب والرياضيات، جامعة الموصل.