P- ISSN 1991-8941 E-ISSN 2706-6703
2013,(7), (1) :233-239

(JUAPS) مجلة جامعة الانبار للعلوم الصرفة

Open Access

تحسين طرق التشفير الإبدالية باعتماد تشفير الـDNA

ياسين حكمت إسماعيل نجلاء بديع ابراهيم

جامعة الموصل- كلية علوم الحاسوب والرياضيات

معلومات البحث:

تاريخ التسليم: 2012/9/11
تاريخ القبول: 2013/1/23
تاريخ النشر: 30/ 11 / 2013
DOI: 10.37652/juaps.2013.83081

الكلمات المفتاحية:

تشفير،
تحسين،
زيادة الامنية،
.DNA

الخلاصة:

أظهرت الدراسات الحديثة على الحامض النووي الرايبي منقوص الأوكسجين (DNA) العديد من الميزات المهمة منها الطبيعة العشوائية لتسلسل القواعد النتروجينية المكونة للحامض و قابلية الخزن الكبيرة للمعلومات والتي أدت الى اعتماده في مجال التشفير حيث ظهر فرع جديد وهو تشفير الـ DNA. يهدف البحث الى استخدام مفهوم تشفير الـ DNA في تحسين و زيادة أمنية طرق التشفير الأبدالية.

المقدمة :

تشفير الحامض النووي الرايبي منقوص الأوكسجين (DNA) حظي بأهتمام كبير نظرا" لقابلية الخزن الكبيرة للحامض حيث أن غرام واحد منه له القابلية على خزن بيانات تقدر بـ 10^8 (Tera Bytes) هذه القابلية لخزن المعلومات تفوق كل وسائط الخزن المعروفة (الكهربائية، المغناطيسية، الضوئية) [7] [10].

الأحماض النووية هي مركبات كيميائية معقدة التركيب توجد في جميع الأحياء وهي ذات أهمية كبيرة لها، إن جزيئات الحامض النووي الرايبي منقوص الأوكسجين مؤلفة من عدد كبير من الوحدات الأصغر تعرف بالنيوكليوتيدات (Nucleotides) [1].

يتألف كل جزيء من النيوكليوتيد من ثلاث جزيئات أبسط مرتبط بعضها ببعض مباشرة وهي [1] [3] [4] :

1. قاعدة نتروجينية : وهي مركب حلقي يحتوي على النتروجين بالإضافة إلى الكاربون والهيدروجين والأوكسجين (عدا الأدنين حيث لا يحتوي على الأوكسجين) و يوجد منها نوعين هما :

أولا : البريميدينات : وتتكون من حلقة واحدة وتشمل القواعد التالية :

أ.الثايمين T، ب.السايتوسين C

ثانيا : البيورينات : وتتكون من حلقتين وتشمل القواعد التالية:

أ.الأدنين A، ب.الكوانين G

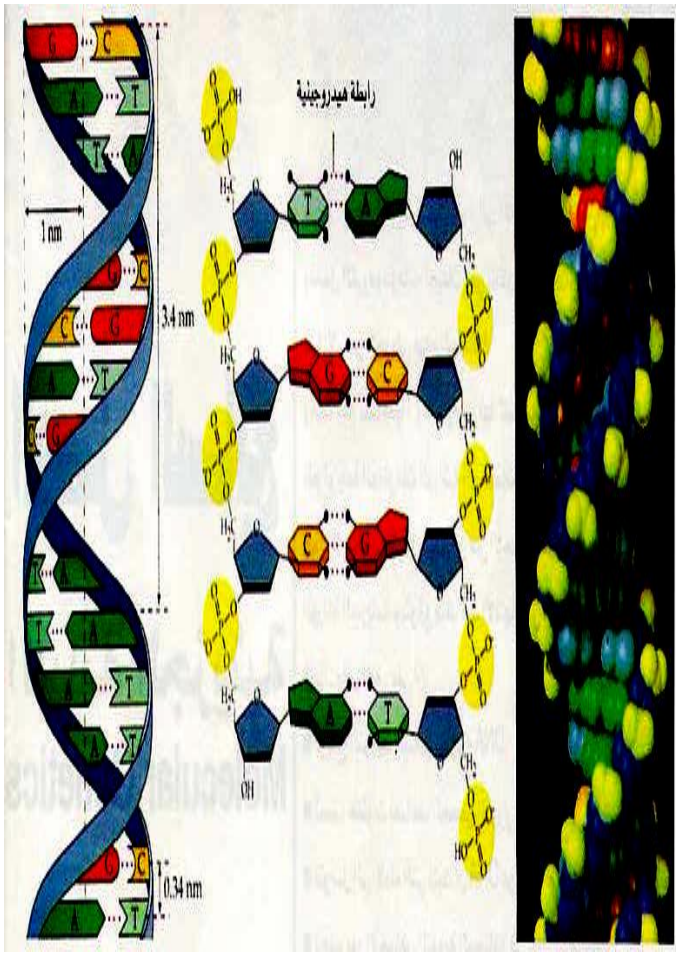
2. سكر خماسي الكاربون : وهو سكر الرايبوز منقوص الأوكسجين الذي يختلف عن الرايبوز بفقدانه ذرة أوكسجين واحدة، وصيغته الجزيئية هي (C5H10O4).

3. حامض الفوسفوريك.

قدم واطسن وكريك عام 1953 نموذجا" للحامض النووي DNA مؤلف من سلسلتين أو شريطين ملتفين على هيئة سلم حلزوني ترتبط فيه إحدى القواعد النتروجينية في أحد شقي الحلزون مع القاعدة النتروجينية للشق الاخر بواسطة أواصر هيدروجينية. إن إرتباط القواعد النتروجينية بين الشقين لا يكون عشوائيا" بل مقيدا"، فالأدنين في أحد الشريطين يرتبط دائما مع الثايمين في الشريط الاخر باصرتين هيدروجينيتين ويرتبط السايتوسين في أحد الشريطين مع الكوانين في الشريط الاخر بثلاث أواصر هيدروجينية وكما موضح بالشكل رقم (1) [1] [4].

* Corresponding author at: Mosul University - College of Computer Science and Mathematics;
ORCID: https://orcid.org/0000-0001-5859-6212.Mobil:777777
E-mail address:

P- ISSN 1991-8941 E-ISSN 2706-6703 مجلة جامعة الانبار للعلوم الصرفة (JUAPS) Open Access
2013,(7), (1) :233-239



الشكل رقم (1) يوضح نموذج العالمان واطسن و كريك

الدراسات السابقة :

أول من أستخدم الحامض النووي الرايبي منقوص الأوكسجين (DNA) في مجال الحساب هو العالم أديلمان عام 1994 لحل مشكلة إيجاد أفضل مسار، حيث وجد بأن الحامض يحتوي على خاصية المعالجة المتوازية والتي توفر سرعة عالية جدا إذا ما تم أستغلالها في المجالات الحسابية [11]. في عام 1995 قام بونييه و أخرون بأستخدام حامض الـ DNA لكسر نظام تشفير البيانات القياسي (DES) [6]. بعد ذلك ظهرت العديد من المحاولات في بناء أنظمة تشفير للبيانات تعتمد على حامض الـ DNA حيث استخدمت خاصية عشوائية تسلسل القواعد النتروجينية للحامض وأعتبارها مفتاح لنظام تشفير المرة الواحدة (One-time Pad System) [10] [13] [12] [5] [8] . نظريا" نظام التشفير الذي يعتمد خاصيتي عشوائية مفتاح التشفير واستخدامه لمرة واحدة لا يمكن أن يكسر [2].

إن استخدام الحامض النووي الرايبي منقوص الأوكسجين في مجال التشفير يوفر مستويين من الحماية للبيانات المشفرة، الأول يكمن في صعوبة إجراء وأستخدام العمليات والتقنيات البايولوجية والمستوى الثاني يتعلق بصعوبة حل و تحليل العمليات الرياضية المستخدمة في تشفير البيانات [5].

أنواع أنظمة التشفير :

يمكن تقسيم أنظمة التشفير من حيث آلية تعاملها مع أحرف النص الواضح لغرض الحصول على النص المشفر الى نوعيين [2] :

1. أنظمة التشفير الأبدالية (Transposition Cipher System) : يتم في هذا التشفير أعادة ترتيب حروف الرسالة الواضحة بحيث تبقى بدون تغيير فقط يتم تغيير مواقعها ضمن النص، حيث يتم تشفير " SEND HELP SOON" الى "SLEPNSDOHOEN". أي أن أحصائية أحرف النص الواضح تكون مساوية لأحصائية النص المشفر فمثلا يحتوي النص الواضح في المثال أعلاه على حرفين "N" وكذلك النص المشفر الناتج .
2. أنظمة التشفير التعويضية (Substitution Cipher System) : هنا يتم استبدال حروف النص الواضح بحروف أخرى أو أعداد أو رموز، فمثلا" يتم تشفير كلمة "COMPUTER" الى "XRSYMHZK".

يهدف البحث الى تقديم طريقة مقترحة لتشفير البيانات تعتمد على أستخدام فكرة الحامض النووي الرايبي منقوص الأوكسجين (DNA) في تطوير و زيادة كفاءة أنظمة التشفير الأبدالية، حيث أن النص الناتج بأستخدام طرق التشفير الأبدالية يحتوي على نفس أحصائية أحرف النص الواضح و بالتالي فمن السهولة تحليل وكسر خوارزمية التشفير. الطريقة المقترحة أدت الى القضاء على مشكلة أحصائية الأحرف و وفرت مستوى عالي من السرية للنص المشفر الناتج.

الطريقة المقترحة :

تتضمن الطريقة المقترحة استخدام أحدى طرق الكبس بدون فقدان (Run Length) [9] لكبس أحرف النص الواضح (بعد تحويلها

الى النظام الثنائي) حيث أن استخدام عملية الكبس بدون فقدان تؤدي الى تقليل في حجم النص المشفر الناتج وهذه صفة مميزة ومهمة وذلك لأن معظم طرق التشفير باعتماد الحامض النووي الرايبي منقوص الأوكسجين (DNA) يكون فيها النص المشفر الناتج ذات حجم أكبر بكثير من النص الواضح المراد تشفيره. يمكن توضيح خطوات والية عملية التشفير بالنقاط التالية :

- قراءة النص الواضح وتحويله الى نظام الاسكي (ASCII Code) ومن ثم الى النظام الثنائي.
- قراءة سلسلة حامض الـ DNA و المتفق عليها بين الطرفين.
- إجراء عملية الإبدال العمودي لسلسلة النظام الثنائي الناتجة من الخطوة 1، وحسب مفتاح متفق عليه بين المرسل والمستلم وهو جزء من تسلسل القواعد النتروجينية في الحامض الناتج من الخطوة 2 (مثلا" أول 6 قواعد من الحامض).
- كبس سلسلة النظام الثنائي الناتجة من الخطوة 3 باستخدام خوارزمية الـ Run Length وهي من خوارزميات الكبس بدون فقدان.
- ترميز السلسلة الرقمية الناتجة من الخطوة 4 باستخدام الحامض في الخطوة 2، تتم عملية الترميز بقراءة عنصر من السلسلة الثنائية مع قاعدة نتروجينية من الحامض، فإذا كانت قيمة عنصر السلسة هو 0 يتم أتخذ المتمم للقاعدة النتروجينية وإذا 1 تنزل القاعدة النتروجينية كما هي، أما اذا كانت رقم (ليس 0 أو 1) وهو مانحصل عليه نتيجة عملية الكبس فيتم انزال القاعدة النتروجينية كما هي متبوعة بنفس الرقم.
- إجراء عملية الابدال العمودي للسلسلة المرمزة الناتجة وذلك لغرض زيادة الأمنية.

إن عملية فك التشفير تتم بصورة معاكسة لخطوات عملية التشفير الموضحة في النقاط أعلاه. يمكن توضيح آلية عمل الطريقة المقترحة بالمخطط الانسيابي في الشكل (2).

مناقشة النتائج :

لغرض مناقشة النتائج وتحليلها وبيان كفاءة الطريقة المقترحة في القضاء على مشكلة أحصائية الأحرف سوف يتم توضيح نتائج التشفير باستخدام طريقة التشفير الإبدالي ومقارنة تلك النتائج مع نتائج التشفير بأستخدام الطريقة المقترحة ومن خلال المثال التالي :

نفرض أن لدينا النص الواضح التالي :

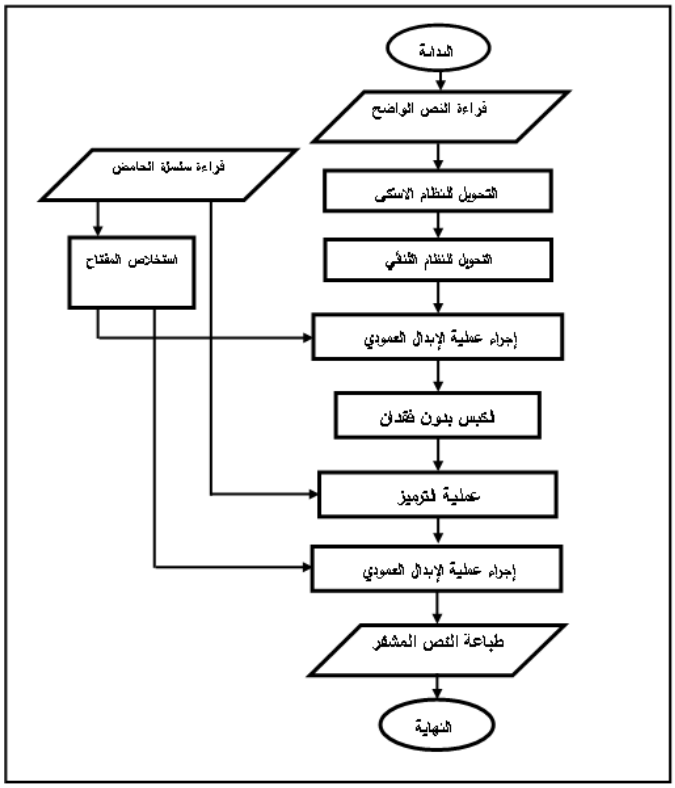
“The biological research in the field of information technology paves the exploitation of storing capabilities parallelism and also in conservative cryptography which enhances the security features for data transmission DNA is the gene information which encodes information of all living beings Though the DNA computing has its application in the field of huge information storage massive parallel processing low energy consumption which have been proposed and proved by the researchers and soon the molecular computer can replace the existing silicon computer and it exploits the world smallest computer The combination of DNA molecules can be interpreted as a result to give a solution to a specific problem “

أولا" : عند إستخدام طريقة من طرق التشفير الإبدالي مثلا" طريقة الإبدال العمودي (Column Transposition) وحسب مفتاح معين كانت نتيجة التشفير كالتالي :

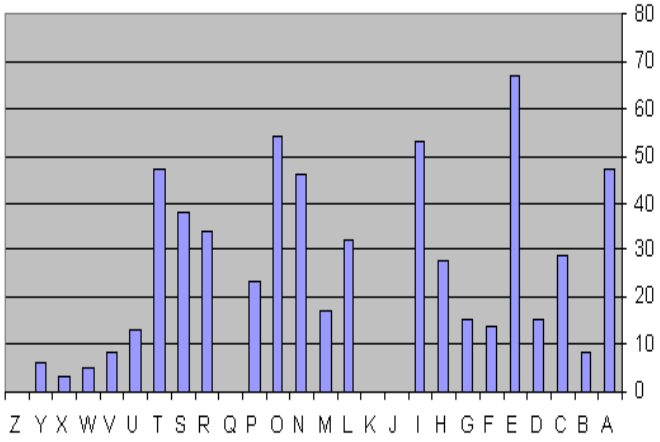
“Tebooia eerhi h il fifraintcnlgpvsteepotto fsoigcpblte aalls nas ncnevtv rporpywihehne h euiyetrsfrdt rnmsinDAi h eeifrainhc noe nomto fallvn ensTog hDAcmuighsisapiaini h il h ilgclrsac ntefedo nomto ehooyae h xliaino trn aaiiisprleimadloi osraiecytgah hc nacstescrtfaue o aatasiso N stegn nomtowihecdsifraino l iigbig huhteN optn a t plcto ntefedo uenomto trg asv aallpoesn o nryosmto hc aebe rpsdadpoe yteeerhr n ontemlclrcmue a elc heitn iio optradi xlistewrdsalscmue h obnto fDAmlclscnbitrrtda eutt ieaslto oaseii rbefhgifrainsoaemsieprle rcsiglweegcnupinwihhv enpooe n rvdb hrsacesadso h oeua optrcnrpaetexsigslcncmue n tepot h ol mletoptrTecmiaino N oeue a enepee sarsl ogv ouint pcfcpolm”

وعند حساب تكرارات الأحرف للنص الناتج و إيجاد تحليل التكرارات **(frequency analysis)** كانت النتيجة كما موضحة بالشكل رقم (3).

P- ISSN 1991-8941 E-ISSN 2706-6703 (JUAPS) مجلة جامعة الانبار للعلوم الصرفة Open Access
2013,(7), (1) :233-239



الشكل رقم (2) المخطط الانسيابي للطريقة المقترحة



شكل رقم (3) يوضح تحليل تكرارات الأحرف في الطرق الإبدالية

حيث نلاحظ أن هنالك تفاوت في أعداد تكرارات الأحرف في النص المشفر الناتج فالأحرف (A,E,I,O,T) هي الأكثر تكرارا" بينما الأحرف (J,K,Q,X,Z) هي الأقل تكرارا" وهي نفسها الميزات والصفات لتكرارات الأحرف لأي نص واضح مكتوب باللغة الإنكليزية مما يسهل على المتطفل تحليل النص المشفر وبالتالي كسر التشفير.

ثانيا" : لغرض تطبيق الخوارزمية المقترحة والمعتمدة على الحامض النووي الرايبي منقوص الأوكسجين يتم إعتماد نفس النص الواضح المستخدم في التشفير بالطرق الإبدالية. تتضمن الخطوة الأولى في الخوارزمية المقترحة قراءة النص الواضح وتحويله الى النظام الاسكي ومن ثم إلى النظام الثنائي، حيث كانت النتيجة الحصول على السلسلة الثنائية التالية :

"10101001101000110010110000011000101101001110111111011
01101111110011111010011100011110000111011001000001110
01100101111001111001011100001111001011000111101000100
01101001110111010000011101001101000110010110000011001
11010011100101110110011001001000001101111110011010000
10100111011101100110110111111100101101101110000111010
11010011101111110111010000011101001100101110001111010
11011101101111110110011011111100111111100111100001100
11110110110010111100111000001110100110100011001011000
11001011111000111000011011001101111101001111010011000
11110100110100111011111101110100000110111111001101000
11100111110100110111111100101101001110111011001111000
11000111100001111000011000011100010110100111011001101
11110100110100111001011110011100000111000011000011110
01100001110110011011001100101111011001101001111001111
01100000110000111011101100100110000111011001110011111
11100000110100111011101000001100011110111111011101110
11100101111001011101101100001111010011010011110110011
01100000110001111100101111001111000011101001101111111
11111001011000011110000110100011110011000001110111110
00110100111000111101000100000110010111011101101000110
01110111011000111100101111001110000011101001101000110
01100000111001111001011100011111010111100101101001111
00111100111001101100101110000111101001110101111001011
10111100111000001100110111011111100101000001100100110
01111010011000011000001110100111001011000011101111011
11110110111010011100111111001111010011101111111011010
00100010010011101000001100000110100111001110000011010
01101000110010110000011001111100101110111011000101100
11010011101110110011011011111110010110110111000011110
01101001110111111011101110111110100011010011100011110
00100000110010111011101100011110111111001001100100111
11100000110100111011101100110110111111001011011011000
01111010011010011101111110111010000011011111100110100
01100001110011110100110100111011111101110100000110100
10111010000011101001101000110010110000011001101101001
00101110110011001001000001101111110011010000011010001
01011100111110010111010011101110110011011011111110010
01101110000011110100110100111011111101110100000111001
10100110111111100101100001110011111001011000001101101
00001111001111100111101001111011011001011000001110000
00001111001011000011101100110110011001011101100100000
10000111001011011111100011110010111001111100111101001
10111011001111000001101100110111111101111000001100101
01110110010111100101100111111100111000111101111110111
11001111101011101101111000011101001101001110111111101
10000011101111101000110100111000111010001000001101000
10000111101101100101100000110001011001011100101110111
00000111000011100101101111111000011011111110011110011
10010010000011000011101110110010010000011100001110011
01111111011011001011100100100000110001011110011000001
01001101000110010111100101100101111001111001011100001"

1001011000111101000110010111100101110011100000110000111
0111011001001000001110011110111111011111101110100000111
0100110100011001011000001101101110111111011001100101110
0011111010111011001100001111001010000011000111101111110
1101111000011101011110100110010111100101000001100011110
0001110111010000011100101100101111000011011001100001110
0011110010110000011101001101000110010111001011111000110
1001111001111101001101001110111011001111000001110011110
1001110110011010011100011110111111011101000001100011110
1111110110111100001110101111010011001011110010100000110
0001110111011001001000001101001111010010000011001011111
0001110000110110011011111101001111010011100111000001110
10011010001100101100000111011111101111111001011011001100
1001000001110011110110111000011101100110110011001011110
01111101001100011110111111101101111000011101011110100110
0101111001010000010101001101000110010110000011000111101
1111101101110001011010011101110110000111101001101001110
111111011101000001101111111001101000000100010010011101000
0011000001101101110111111011001100101110001111101011101
1001100101111001110000011000111100001110111010000011000
1011001011101001110111011101001100101111001011100001110
0101100101111010011001011100100100000110000111100111000
0011000011000001110010110010111100111110101110110011101
0010000011101001101111100000110011111010011110110110010
11000001100001100000111001111011111110110011101011110100
110100111011111110111010000011101001101111100000011000011
0000011100111110000110010111000111101001110011011010011
10001110000011100001110010110111111100010110110011001011
101101”

بعدها يتم قراءة سلسلة الحامض والمتفق عليها بين الطرفين من إحدى مواقع الانترنت الخاصة بقواعد بيانات الحوامض النووية مثلا" NCBI ونفرض أن الحامض المتفق عليه هو :

“*Homo sapiens FOSMID clone ABC24-1954N7 from chromosome 1*”

الخطوة المهمة ضمن عملية التشفير في الطريقة المقترحة هي عملية ترميز سلسلة الأرقام الثنائية وتحويلها إلى سلسلة من القواعد النتروجينية بالأعتماد على الحامض المتفق عليه بين الطرفين ولتوضيح عملية الترميز نفرض المثال التالي :

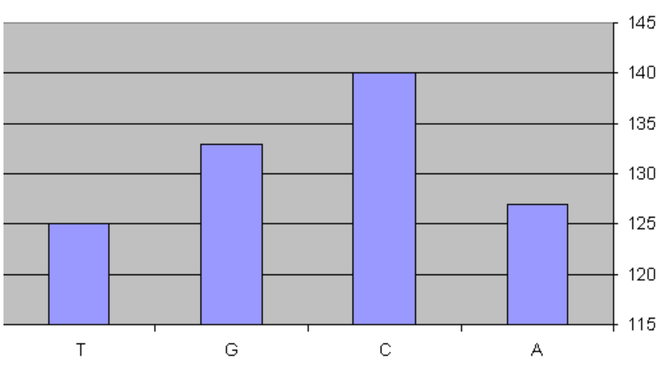
سلسلة الأرقام الثنائية	سلسلة الحامض المتفق عليه	السلسلة المرمزة الناتجة
1	A	A
0	T	A
1	C	C
1	C	C
0	A	T
1	G	G
0	A	T
0	C	G

نلاحظ في عملية الترميز أنه إذا كانت قيمة الرقم الثنائي صفر يتم أخذ المتمم للقاعدة النتروجينية في سلسلة الحامض أما إذا كانت القيمة في سلسلة الأرقام الثنائية هي واحد فتنزل القاعدة النتروجينية كما هي وبدون تغيير.

عند الإنتهاء من تنفيذ الخوارزمية المقترحة يكون النص المشفر عبارة عن سلسلة من القواعد النتروجينية و كما يلي :

“CCCGCCTATCATTGCTCAGTACAAAGCGCAAA
TAACCGATGACTGAGCGTTGATCATCAATAATC
ACTGACGGCTAAGTGTTGCCGAGAGTTGAGTAG
TGCGGCTTATAGCCGATCACATATCGTTGCAAA
TAGTTGCCGAGAGCGCAGAAATCAGCACTGACC
GCCGAGAACACCTTACCGCGGATGACCTAGAGC
AAAGCGCAAATAAACGACTGCTGACAGCGGAC
GATTGAGTAGTGCGGCTTATCAACTACTGCATA
GAACCGACCGCTGACCGCCGAGAACACCTTAG
AACGCCTTACGAAGTAGATCAATCACTGATCAT
CAAGTATCACGTCTTATTAATACTTATTATTGAT
CATCAATAACCACATACCGTTGAGTAGAATTGC
GAAGTAAGCATCATTGCTGCTTACGACAAAGCG
AGGACCGCCGAGAACACCTTACCGCAACTTATT
ATTGATAACTGCGGATGACCGATAGCCGATGAT
TGATTACGGCAAATGGCATATCACAG”

وعند حساب تكرارات الأحرف للنص الناتج و إيجاد تحليل التكرارات **(frequency analysis)** كانت النتيجة كما موضحة بالشكل رقم (4)



شكل رقم (4) يوضح تحليل تكرارات الأحرف في الطريقة المقترحة

من خلال مقارنة نتائج التحليل الإحصائي لتكرارات الأحرف للنص المشفر في الطرق الإبدالية و الطريقة المقترحة، نلاحظ بأن الطريقة

المقترحة أعطت نتائج متقاربة بالنسبة لتكرارت الأحرف وكذلك فأن الأحرف المكونة للنص المشفر هي ليست أحرف الهجائية الإنكليزية (26حرف) بل فقط أربعة أحرف (A,C,G,T) تمثل سلسلة من القواعد النتروجينية.

بالتالي يمكن القول بأن الطريقة المقترحة قد حلت مشكلة إحصائية الأحرف و وفرت مستوى عالي من السرية للنص المشفر الناتج.

الاستنتاجات :

في هذا البحث تم استخدام الحامض النووي الرايبي منقوص الأوكسجين (DNA) في تحسين وزيادة كفاءة طرق التشفير الأبدالية (Transposition Cipher System) والتي أصبح استخدامها في الوقت الحالي قليل أو منعدم نظرا" لزيادة خبرة وتطور أجهزة وتقنيات محللي الشفرات. هنالك العديد من الميزات لطرق التشفير الأبدالية منها سهولة استخدامها و سرعتها العالية في إجراء عملية التشفير حيث أن الطريقة المقترحة فتحت المجال في إمكانية استخدام طرق التشفير الأبدالية والاستفادة من ميزاتها و الحصول على درجة عالية من الأمنية للنص المشفر الناتج.

المصادر :

1. الجلبي قصي عبد القادر، " الأحماض النووية "، دار الحكمة للطباعة والنشر، الموصل – العراق، (1991).

2. الحمامي علاء حسين، " تكنولوجيا أمنية المعلومات وأنظمة حمايتها "، الطبعة الأولى، دار وائل للنشر، عمان – الاردن، (2007).

3. تاج الدين سعد جابر " " علم الوراثة "، الطبعة الثانية، جامعة البصرة، البصرة – العراق، (2000).

4. شكارة مكرم ضياء، " علم الوراثة "، الطبعة الاولى، دار المسيرة، عمان – الاردن، (2000).

5. Beenish Anam, and et al, “Review on the Advancements of DNA Cryptography“, SKIMA Paro, Bhutan , 2010.

6. Dan Boneh, Christopher Dunworth, and Richard J. Lipton, “Breaking DES using a molecular Computer", Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, 1996.

7. G. Cui, Y. Liu, and X. Zhang , “New direction of data storage: DNA molecular storage technology,” *Computer Engineering and Application*, vol. 42, no. 26, pp. 29–32, 2006.

8. Guangzhao Cui , and et al., “An Encryption Scheme Using DNA Technology“ , IEEE 37 BIC-TA , 2008

9. Guy E. Blelloch, “ Introduction to Data Compression “ , Carnegie Mellon University, Guy Blelloch, 2010.

10. J. Chen, “A DNA-based, biomolecular cryptography design”, IEEE International Symposium on Circuits and Systems (ISCAS), 2003.

11. L. M. Adleman. Molecular computation of solutions to combinatorial problems. Science, 1994.

12. Monica BRDOA, Olga TORNEA, “ DNA Secret Writing Techniques “ , IEEE , 2010.

13. Xing Wang, Qiang Zhang, “ DNA computing-based cryptography “, 1EEE, 2009.

P- ISSN 1991-8941 E-ISSN 2706-6703 (*JUAPS*) مجلة جامعة الانبار للعلوم الصرفة Open Access
2013,(7), (1) :233-239

ENHANCING TRANSPOSITION CIPHER METHODS USING DNA ENCRYPTION

YASEEN H. ASNAEEL **NAJLA B. IBRAHEAM**

ABSTRACT

Recent studies have shown the Deoxyribo nucleic acid (DNA) have several important features including the indiscriminate nature of the sequence of nitrogenous bases consisting the acid and large storage capability of the information that led to its adoption in the field of encryption where the appearance of a new branch which is encryption of DNA. The research aims to use the concept of DNA encryption to improve and increase the security of transposition cipher methods.