Authentication of Deoxyribo Nucleic Acid (DNA)

Naglaa Badi Ibrahim

Yassin Hikmat Ismail

College of Computer Science and Mathematics
University of Mosul, Mosul, Iraq

Received on: 17/09/2012

Accepted on: 30/01/2013

ABSTRACT

Deoxyribo Nucleic Acid (DNA) contains many of the nitrogenous bases that are random and have the ability to store a large amount of information. DNA sent either in natural form in the test tube (after being processed) or in digital form (a series of nitrogenous bases), in both cases, the DNA bar can be vulnerable to intrusion by unauthorized destination. This research aims to authenticate the acid and thus the receiving side can ensure the credibility of the acid and that it had not been altered.

Keywords: Deoxyribo Nucleic Acid, Authentication.

توثيق الحامض النووي الرايبي منقوص الأوكسجين (DNA)

ياسين حكمت إسماعيل

نجلاء بديع إبراهيم

كلية علوم الحاسوب والرياضيات

جامعة الموصل، الموصل، العراق

تاريخ قبول البحث: 2013/01/30

تاريخ استلام البحث: 2012/09/17

الملخص

يحتوي الحامض النووي الرايبي منقوص الأوكسجين (DNA) على العديد من القواعد النتروجينية ذات الطبيعة العشوائية والتي لها القابلية على تخزين كمية كبيرة من المعلومات. يتم إرسال شريط الـ DNA إما بصيغة طبيعية في أنبوبة اختبار (بعد تصنيعه) أو بصورة رقمية (عبارة عن سلسلة من القواعد النتروجينية)، في كلتا الحالتين فإن شريط الـ DNA يكون عرضة للتطفل من قبل جهة غير مخولة. يهدف البحث إلى تحقيق التوثيق (Authentication) للحامض وبالتالي بإمكان الجهة المستلمة التأكد من مصداقية الحامض و أنه لم يحدث له تغيير خلال نقله.

الكلمات المفتاحية: الحامض النووي الرايبي منقوص الأوكسجين، التوثيق.

1. المقدمة:

الحامض النووي الرايبي منقوص الأوكسجين (DNA) حظي باهتمام كبير نظراً لقابلية التخزين الكبيرة للحامض إذ أن غراماً واحداً منه له القابلية على تخزين بيانات تقدر بـ 10^8 (Tera Bytes) هذه القابلية لتخزين المعلومات تفوق وسائط التخزين المعروفة كلها (الكهربائية، المغناطيسية، الضوئية) [8] [11]. الأحماض النووية هي مركبات كيميائية معقدة التركيب توجد في جميع الأحياء وهي ذات أهمية كبيرة. 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] [5]

[4]:

1. قاعدة نيتروجينية: وهي مركب حلقي يحتوي على النيتروجين بالإضافة إلى الكربون والهيدروجين والأكسجين (عدا الأدينين حيث لا يحتوي على الأكسجين) ويوجد منها نوعان هما:
أولاً: البيريميديات: وتتكون من حلقة واحدة وتشمل القواعد الآتية:
أ. الثايمين T ، ب. السيتوسين C
ثانياً: البيورينات: وتتكون من حلقتين وتشمل القواعد الآتية:
أ. الأدينين A ، ب. الكوانين G
2. سكر خماسي الكربون: وهو سكر الرايبوز منقوص الأكسجين الذي يختلف عن الرايبوز بفقدانه ذرة أكسجين واحدة، وصيغته الجزيئية هي (C₅H₁₀O₄).
3. حامض الفسفوريك.

قدم واطسن وكريك [1] [5] عام 1953 نموذجاً للحامض النووي DNA مؤلفاً من سلسلتين أو شريطين ملتفين على هيئة سلم حلزوني ترتبط فيه إحدى القواعد النيتروجينية في أحد شقي الحلزون مع القاعدة النيتروجينية للشق الآخر بواسطة أواصر هيدروجينية.

إن ارتباط القواعد النيتروجينية بين الشقين لا يكون عشوائياً بل مقيداً، فالأدينين في أحد الشريطين يرتبط دائماً مع الثايمين في الشريط الآخر بأصرتين هيدروجينيتين ويرتبط السيتوسين في أحد الشريطين مع الكوانين في الشريط الآخر بثلاث أواصر هيدروجينية وكما موضح بالشكل (1).

الشكل (1). نموذج العالمان واطسن وكريك

2. الدراسات السابقة

أول من استخدم الحامض النووي الرايبوزي منقوص الأكسجين (DNA) في مجال الحساب هو العالم أديلمان [11] عام 1994 لحل مشكلة إيجاد أفضل مسار، إذ وجد بأن الحامض يحتوي على خاصية المعالجة المتوازية والتي توفر سرعة عالية جداً إذا ما تم استغلالها في المجالات الحسابية. في عام 1995 قام بونيه وآخرون [7] باستخدام حامض الـ DNA لكسر نظام تشفير البيانات القياسي (DES). بعد ذلك ظهرت العديد من المحاولات [10] [13] [12] [6] [9] في بناء أنظمة تشفير للبيانات تعتمد على حامض الـ DNA إذ استخدمت خاصية عشوائية تسلسل القواعد النيتروجينية للحامض وعددها مفتاحاً لنظام تشفير المرة الواحدة (One-time Pad)

(System). نظرياً نظام التشفير الذي يعتمد خاصيتي عشوائية مفتاح التشفير واستخدامه لمرة واحدة لا يمكن أن يكسر [2].

3. الوثوقية (Authentication)

يمكن تعريف الوثوقية على أنها إحدى خدمات الأمانة والتي تتضمن أن الاتصال موثوق به، وهذا يعني التأكد من أن المعلومات قادمة من جهة مخولة فضلاً على أنها لم تتعرض أثناء مرورها في قناة النقل لأي تأثير من قبل جهة غير مخولة [3] [2].

نفرض أن المرسل يقوم بإرسال حامض الـ DNA إلى المستلم من خلال قناة النقل (Transmission Channel) والواقعة تحت سيطرة الخصم كما هو موضح بالشكل (2).

الشكل (2). سيطرة الخصم على قناة النقل

حيث يستطيع الخصم القيام بنشاطات مختلفة على قناة النقل مثل:

1. التصنت والاستماع للمعلومات المرسلة عبر قناة النقل.
 2. منع مرور البيانات.
 3. تخزين البيانات وإعادة إرسالها في وقت آخر خلال فترة إرسال كاذبة.
 4. احتمالية انتحال الخصم لشخصية أحد أطراف الاتصال (المرسل أو المستلم).
 5. تغيير محتوى الرسالة بحذف، أو حشر أو إعادة ترتيب أجزاء مختلفة من الرسالة.
- لذلك يجب على المستلم أن يتحقق مما يأتي [3] [2]:

1. الرسالة مرسله من قبل المرسل الحقيقي.
2. محتوى الرسالة لم يتغير خلال الإرسال في قناة النقل.

4. هدف البحث:

نظراً لقابلية التخزين الكبيرة والطبيعة العشوائية لتسلسل القواعد النروجينية المكونة لحامض الـ DNA فقد تم اعتماده في مجال تشفير البيانات حيث ظهر فرع جديد هو التشفير المعتمد على الـ DNA (DNA-based cryptography) [6] [10] [13]، كذلك فقد تم اعتماده في مجال إخفاء المعلومات (Steganography) [8] [12]. مما سبق نجد بأن حامض الـ DNA المرسل قد يحتوي على بيانات سرية مهمة فإذا حدث للقواعد النروجينية المكونة للحامض تغيير (إضافة، حذف، إعادة ترتيب) أثناء إرساله في قناة النقل من قبل جهة غير مخولة فسوف يؤدي إلى فقدان المصادقية للمعلومات المهمة التي يحتويها الحامض وعدم إمكانية الجهة المستلمة من التأكد من أن المعلومات المستلمة صحيحة ولم تتعرض إلى تغيير. يهدف البحث إلى تحقيق الوثوقية

للحامض النووي الرايبى منقوص الأوكسجين (DNA) وبالتالي بإمكان الجهة المستلمة التحقق من صحة الحامض المستلم وأنه لم يحدث له تغيير من قبل المتطفل.

5. الطريقة المقترحة

تتضمن الطريقة استخدام فكرة دوال الترميز (Hash Function) لإيجاد ملخص الرسالة (Message Digest) للحامض [3] وبالتالي استخدام ملخص الرسالة لتحقيق الوثوقية لحامض الـ DNA. تتضمن الطريقة مجموعة من الخطوات تنفذ عند الجهة المرسله والمستلمة وكما موضح:

الخطوات المنفذة لدى الجهة المرسله هي:

1. قراءة سلسلة قواعد حامض الـ DNA المراد تحقيق الوثوقية له.
2. حساب أعداد القواعد النتروجينية في سلسلة الحامض المراد تحقيق الوثوقية له وتمثيلها بثمان وحدات (8-digits) وعلى النحو الآتي:

- أ. حساب عدد القاعدة T والواقعة في المواقع الفردية ضمن سلسلة الحامض.
- ب. حساب عدد القاعدة T والواقعة في المواقع الزوجية ضمن سلسلة الحامض.
- ت. حساب عدد القاعدة A والواقعة في المواقع الفردية ضمن سلسلة الحامض.
- ث. حساب عدد القاعدة A والواقعة في المواقع الزوجية ضمن سلسلة الحامض.
- ج. حساب عدد القاعدة C والواقعة في المواقع الفردية ضمن سلسلة الحامض.
- د. حساب عدد القاعدة C والواقعة في المواقع الزوجية ضمن سلسلة الحامض.
- ذ. حساب عدد القاعدة G والواقعة في المواقع الفردية ضمن سلسلة الحامض.
- ر. حساب عدد القاعدة G والواقعة في المواقع الزوجية ضمن سلسلة الحامض.
- ز. حساب عدد المقاطع الثنائية لكل القواعد ضمن سلسلة الحامض.
- س. حساب عدد المقاطع الثلاثية لكل القواعد ضمن سلسلة الحامض.
- ش. حساب عدد المقاطع الرباعية لكل القواعد ضمن سلسلة الحامض.
- ص. حساب عدد المقاطع الخماسية لكل القواعد ضمن سلسلة الحامض.
- ع. حساب عدد المقاطع السداسية لكل القواعد ضمن سلسلة الحامض.

3. إيجاد قيم المعادلات التالية:

$$أ \oplus ت \oplus ج \oplus ذ = \text{فردية}$$

$$ب \oplus ث \oplus د \oplus ر = \text{زوجية}$$

$$ز \oplus س \oplus ش \oplus ص \oplus ع = \text{مقاطع}$$

$$\text{فردية} \oplus \text{زوجية} \oplus \text{مقاطع} = \text{ملخص الرسالة}^*$$

4. ترميز ملخص الرسالة (النظام الثنائي) الى صيغة القواعد النتروجينية حيث:

$$G = 11, C = 10, T = 01, A = 00$$

وبالتالي يكون ملخص الرسالة عبارة عن 4 قواعد نتروجينية.

5. حشر قيمة ملخص الرسالة في بداية ونهاية السلسلة وثلاثة مواقع أخرى متتق عليها بين الطرفين.

6. إرسال سلسلة القواعد النتروجينية المكونة للحامض بصيغتها الجديدة.

الخطوات المنفذة لدى الجهة المستلمة هي:

1. قراءة سلسلة القواعد النتروجينية للحامض المراد التأكد من وثوقيته.
 2. حذف قيم ملخص الرسالة (أربعة قواعد نتروجينية) من بداية ونهاية السلسلة والمواضع الثلاثة المنق علىها، ومقارنة تلك القيم الخمسة إذ يجب أن تكون متطابقة جميعها وبخلافه يعد الحامض غير موثوق به.
 3. إجراء عملية الحساب لأعداد القواعد النتروجينية للسلسلة الحامض الناتجة من الخطوة 2 أعلاه وبنفس الأسلوب المتبع في الخطوة 2 عند الجهة المرسل.
 - * رقم ممثل بالنظام الثنائي بثمان وحدات (8-digits).
 4. حساب قيم المعادلات وبنفس الأسلوب المتبع في الخطوة 3 عند الجهة المرسل.
 5. ترميز ملخص الرسالة (النظام الثنائي) الى صيغة القواعد النتروجينية وبالتالي الحصول على أربع قواعد نتروجينية تمثل صيغة ملخص الرسالة.
 6. مقارنة قيمة ملخص الرسالة الناتجة من الخطوة 5 مع قيمة ملخص الرسالة المحذوفة في الخطوة 2 أعلاه، إذا تساوت القيمتان دل ذلك على أن الحامض المستلم موثوق به وإلا فمعنى ذلك أن الحامض المستلم غير موثوق به (أي حصل تغيير في القواعد النتروجينية المكونة للحامض).
- يمكن توضيح الخطوات المنفذة لدى الجهة المرسل والجهة المستلمة بالمخططين الانسيابيين في الشكلين (3, 4).

الشكل (3). الخطوات المنفذة لدى الجهة المرسل

الشكل (4). الخطوات المنفذة لدى الجهة المستلمة

7. الاستنتاجات

في هذا البحث تم تقديم طريقة مقترحة لتحقيق الوثوقية للحامض النووي الرايبي منقوص الأوكسجين (DNA). أن حامض الـ DNA يستخدم لحزن بيانات مهمة قد تكون مشفرة أو مخفية بصيغة قواعد نتروجينية. إذا حدث تغيير لشريط الـ DNA من قبل جهة غير مخولة فلا يمكن للجهة المستلمة اكتشاف ذلك التغيير، لهذا فقد وفرت الطريقة المقترحة الوثوقية للحامض وبالتالي بالإمكان اكتشاف أي تغيير يحدث في القواعد النتروجينية المكونة للحامض.

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 - الأردن، (2007).
- [3] العبيدي ياسين حكمت، "دراسة الوثوقية وتحقيقها باستخدام التوقييع الرقمية"، رسالة ماجستير، جامعة الموصل/كلية علوم الحاسوب والرياضيات، 2004.
- [4] تاج الدين سعد جابر، "علم الوراثة"، الطبعة الثانية، جامعة البصرة، البصرة - العراق، (2000).
- [5] شكاره مكرم ضياء، "علم الوراثة"، الطبعة الأولى، دار المسيرة، عمان - الأردن، (2000).
- [6] Beenish Anam, and et al, "Review on the Advancements of DNA Cryptography", SKIMA Paro, Bhutan, 2010.
- [7] Dan Boneh, Christopher Dunworth, and Richard J. Lipton, "Breaking DES using a molecular Computer", Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, 1996.
- [8] 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.
- [9] Guangzhao Cui, and et al., "An Encryption Scheme Using DNA Technology", IEEE 37 BIC-TA, 2008.
- [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", IEEE, 2009.