Copyright Protection for Text Documents

Dujan B. Taha **Farqad H. AbdulRaheem**

dujan_taha@uomosul.edu.iq

College of Computer Sciences and Mathematics *Technical College*

University of Mosul, Iraq *Foundation of Technical Education*

Received on: 08/03/2007 **Accepted on: 30/05/2007**

ABSTRACT

Due to the rapid development in communication and multimedia technology, different techniques are used for copyright protection and monitoring illegal copying of these media. One of the most important techniques is digital watermarking.

In this research, an algorithm for copyright protection of text documents was designed and implemented.

The algorithm includes three steps: Watermark generation using the original text, watermark embedding using space embedding approach and finally watermark extracting and comparing it with the original one using normalized correlation.

Experimental results are presented to demonstrate the robustness of the algorithm against common attacks which alter the embedded text like changing some letters of the text, changing structure of text sentences, or changing the number of spaces between the words …etc.

Matlab Ver. 6.5 was used to implement the algorithm under Windows.

Keywords: Copyright Protection, Text Documents, Digital Watermarking.

حماية حقوق الملكية للوثائق النصية

دجان بشير طه **فرقد حامد عبد الرحيم**

كلية علوم الحاسبات والرياضيات، جامعة الموصل *الكلية التقنية، هيئة التعليم التقني*

تاريخ استلام البحث: 2007/03/08 **تاريخ قبول البحث: 2007/05/30**

الملخص

إن التطور السريع الحاصل في الاتصالات وتقنيات الوسائط المتعددة أدى إلى الحاجة إلى استخدام تقنيات لحماية حقوق الملكية ومراقبة النسخ غير الشرعي لتلك الوسائط ومن أهم هذه التقنيات هي العلامة المائية الرقمية .

تم في هذا البحث تصميم وتنفيذ خوارزمية لغرض توثيق النصوص باستخدام العلامة المائية. تشمل الخوارزمية ثلاث مراحل : توليد العلامة المائية بالاعتماد على النص الأصلي ,

تضمين العلامة المائية داخل النص باستخدام أسلوب التضمين بالفراغات، وأخيرا استرجاع العلامة المائية ومقارنتها مع العلامة المائية الأصلية باستخدام مقياس الترابط.

أثبتت النتائج العملية أن الطريقة المستخدمة تقاوم مختلف أنواع الهجوم الذي قد يتعرض له النص بعد تضمين العلامة المائية فيه كالتغيير في أحرف النص وكلماته ، التغيير في تركيب الجمل أو تغيير عدد الفراغات بين الكلمات ... الخ .

تم استخدام MatlabVer6.5 في تمثيل برامج الخوارزمية تحت نظام تشغيل Windows بإصداراته المختلفة.

الكلمات المفتاحية: حماية حقوق الملكية، الوثائق النصية، العلامة المائية الرقمية.

1. مقدمة في اخفاء المعلومات:

إن تقنية إخفاء المعلومات Information Hiding البسيطة استخدمت منذ مئات السنين , ولكن مع ازدياد استخدام الملفات بالصيغة الالكترونية , أصبح من الضروري تطوير تقنيات جديدة لإخفاء المعلومات [1] .

إن إخفاء المعلومات عنوان عام لنوعين من التقنيات , النوع الأول يستخدم لحماية المعلومات من المراقبين والمهاجمين ويدعى بالكتابة المغطاة (covered writing) أو ما يطلق عليه مصطلح نظام التغطية (steganography) ، أما النوع الثاني فيستخدم لإثبات حقوق الملكية الفكرية (intellectual property) أو لضمان الوثوقية (authentication) ، وهو ما يدعى بالعلامات المائية الرقمية (digital watermarking) وهي موضوع البحث [3,2].

ومن الاستخدامات الرئيسة لإخفاء المعلومات علامات حقوق النسخ للرسالة التي تدرج وتستخدم للتأكيد على حقوق الاستنساخ والنشر للوثيقة. ويمكن أن يقسم هذا إلى العلامات المائية وبصمة الإصبع [8]. تعطي العلامة المائية إعاقة لعملية الاستنساخ ولكنها لا تمنع الاستنساخ غير الشرعي ومن الممكن أن توضع إشارة مضمنة مع العلامة المائية لتعطيل عملية الاستنساخ لهذا المنتج في حالة طلب استنساخه من جهة خارجية.

إن علم إخفاء المعلومات Information Hiding والتشفير Cryptography يستخدمان مع اختلافهما في تأكيد وثوقية البيانات ففي التشفير، يمكن لأي شخص أن يكتشف أن هناك شخصين يتصلان بطريقة مشفرة أما إخفاء المعلومات فيخفي أصلا وجود الاتصال بحيث لا يمكن لأحد أن يلحظ وجود شخصين يتبادلان الرسائل عبر قنوات الاتصال وهذا ما جعل علم إخفاء المعلومات مناسبا في بعض المهام مثل إخفاء علامات حقوق النسخ إذ من السهولة إزالة

معلومات حقوق الطبع المشفرة ولكن عند تضمين تلك البيانات مع محتويات الملف نفسه فمن الصعب اكتشافها أو إزالتها [6] .

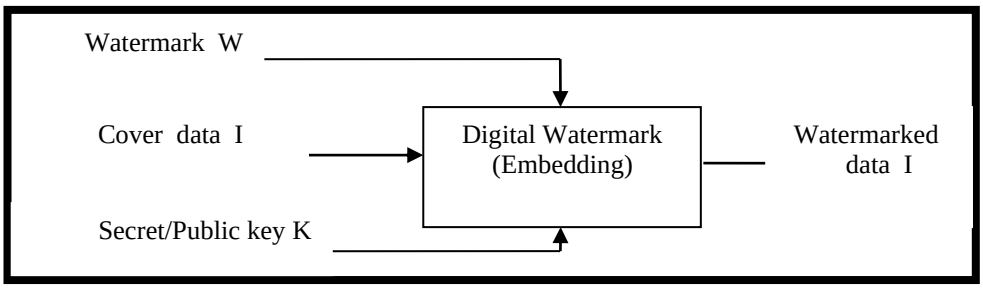
في يومنا هذا نرى أن الانتشار السريع في أنظمة الوسائط المتعددة يطرح سؤالا حول حماية حقوق الطبع , مراقبة إشارات الإزالة , وتطوير خوارزميات وتطبيقات الزمن الحقيقي التي قد تمنع النسخ غير الشرعي وتمنع إعادة توزيع النسخ المسروقة . ويكمن احد الحلول لهذه المسالة في علامات مائية للبيانات data watermarking تضاف إلى محتويات الوسائط المتعددة بطريقة ثابتة وغير مدركة , وبعد ذلك تكون الجهات المستخدمة لها قادرة على برهنة ملكية الملفات. إن تقنيات العلامات المائية قد طورت في البداية للصور الرقمية وللأفلام الفيديوية ووصلت مؤخرا إلى الصوت [5] .

2. المبادئ الأساسية للعلامة المائية :

تشترك كل نظريات العلامات المائية بنفس عمليتي التضمين embedding والاسترجاع extraction وتسمى أيضاً الكشف detection.

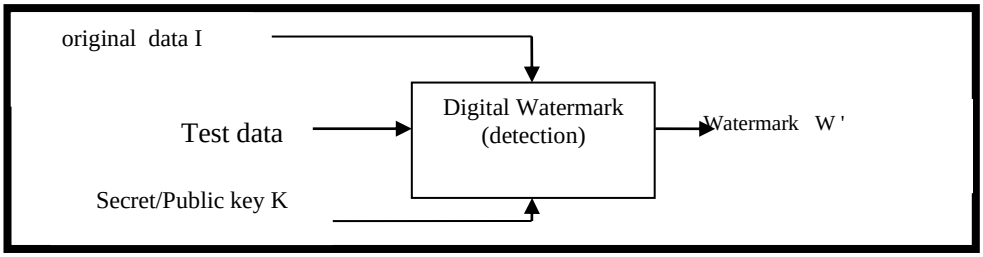
يوضح الشكل (1) عملية تضمين العلامة المائية . ويكون إدخال هذا النظام العلامة المائية ، عن طريق الغطاء Cover data والمفتاح key الذي يكون وجوده اختيارياً وهو إما أن يكون سريا Secret أو عاما Public وتكون العلامة المائية بعدة أشكال كأن تكون رقماً ، نصاً أو صورة. وفي معظم الأحيان يتم في عملية تضمين العلامة المائية استخدام مفتاح, الهدف من استخدامه جعل العلامة المائية صعبة الاختراق وتقاوم عمليات الحذف والإحلال بحيث لا يمكن قراءتها أو تغييرها من دون استخدام المفتاح نفسه .

ويكون الغطاء نصا ، صورة ، صورا متحركة أو ملفا صوتيا ، أما الإخراج فيكون بيانات الغطاء بعد تضمين العلامة المائية فيه .



الشكل (1) عملية تضمين العلامة المائية.

يوضح الشكل (2) عملية استرجاع العلامة المائية بإدخال بيانات الغطاء بعد تضمين العلامة المائية فيها ، والمفتاح اما ان يكون سريا أو عاما والبيانات الأصلية للغطاء قد تكون موجودة أو غير موجودة اعتمادا على الطريقة المستخدمة فعند استخدام بيانات الغطاء الأصل في عملية الاسترجاع تسمى العملية non-blind watermarking وعندما لا تستخدم البيانات الأصل في عملية الاسترجاع تسمى هذه العملية blind watermarking [8,7,5,4] .



الشكل (2) عملية كشف العلامة المائية واسترجاعها

3. تقنيات تضمين العلامة المائية في النص :

إن تضمين معلومات سرية في نص يعتبر من التحديات الكبيرة لان الملفات النصية تمتلك كمية قليلة من البيانات الفائضة التي يمكن إخفاء هذه المعلومات فيها وتعتبر هذه الحالة مشكلة في الملفات النصية. ومن المشاكل الأخرى في الملفات النصية هي أن تغيير النص واعادة صياغته من قبل شخص خارجي عملية سهلة جدا.

هناك عدة تقنيات لتضمين العلامة المائية في النص منها [6,1] :

1_ اتفاقية الترميز بتزحيف السطر Line shift coding protocol

الترميز بهذه الطريقة ببساطة هو تزحيف سطور مختلفة داخل الوثيقة إلى الأعلى أو الأسفل بنسبة كسرية صغيرة مثل 300/1 من الانج , والسطور المزحفة داخل الوثيقة تكون غير مكتشفة من قبل البشر, لأنها نسبة بسيطة ولكن من الممكن الكشف عنها عندما تقاس المسافة بين السطور بواسطة الحاسبة وفي العادة تستخدم في هذا السياق تقنيات ترميز مختلفة .

وعندما يجد الحاسوب أن السطر قد زحف إلى أعلى أو أسفل فمن المستطاع أن نمثل ذلك بمرتبة ثنائية 0 أو1 وإذا وضعنا كامل الوثيقة فمن الممكن أن ندرج عدداً من المراتب الثنائية وبعد ذلك سوف نمتلك القدرة على إخفاء معلومات كثيرة.

2_ اتفاقية الترميز بتزحيف الكلمات Word shift coding protocol

هذه الاتفاقية تستند إلى نفس قواعد تزحيف السطر والاختلاف الرئيسي هو زحف الكلمات يمينا أو يسارا , وليس السطور إلى أعلى أو أسفل .

وقواعد الاتفاقية تخبر المضمن عن أي من الكلمات سوف تزحف وبأي اتجاه يمينا أو يسارا , وبالمقابل فان تقنية فك الرموز تقيس المسافة بين الكلمات والتزحيف إلى اليسار يمكن أن يمثل بـ (0) والتزحيف إلى اليمين يمثل بـ (1) وبذلك يمكن تمثيل مجموعة من المراتب الثنائية .

3_ اتفاقية الترميز بالمظهر أو الصفات Feature coding protocol

وترميز الصفات بهذه الطريقة لا يختلف كثيراً مع الاتفاقيات السابقة والاختلاف هو أن الوثيقة تمر من خلال محلل لغوي (Parser) يفحص الوثيقة ويبني قواعد معينة خاصة بها استنادا إلى تحليله , وهذا المحلل اللغوي سيختار جميع الصفات التي يمكن أن تستخدم لإخفاء المعلومات وكل واحدة من هذه الصفات سوف تميز داخل الوثيقة , ومن الصفات التي يمكن أن يكشفها المحلل اللغوي ارتفاع حروف معينة ، أطوال الخطوط الأفقية في الحروف مثل الـ t , f ونلحظ أن تقنية تزحيف السطور والحروف تستخدم فيما لو كانت البيانات المراد إخفاؤها كبيرة .

4_ التعامل مع الفراغات في النص Spaces in Text

إن إحدى طرائق إخفاء البيانات في النص هي استخدام الفراغات ، وهذا يكون عن طريق إضافة فراغات بين الكلمات أو عند بدايات السطور أو نهاياتها ونتيجة هذه المعاملة يمكن أن تخزن عددا من البيانات , وتتم هذه العملية بإضافة كمية معينة من الفراغات في نهاية السطور وكمية الفراغات تطابق قيمة مرتبة ثنائية معينة . الحقيقة أن محررات النصوص تهمل الفراغات في نهاية السطور ولهذا السبب لن تلحظ من قبل المتطلع صدفة إلى الوثيقة.

4. وصف الخوارزمية وتصميم البرامج:

تتناول الفقرات التالية شرحا مفصلا للطريقة المستخدمة في عملية توليد العلامة المائية ، وتضمين العلامة المائية , واسترجاعها من النص عن طريق تحليل وتصميم البرامج المعدة لهذا الغرض .

1.4 توليد العلامة المائية Watermark Generation

إن الطريقة المستخدمة في توليد العلامة المائية للنص تتلخص بثلاث مراحل هي:-

1. تكوين مصفوفة الصفوف للنص .
2. تكوين مصفوفة الأعمدة للنص .
3. تكوين مصفوفة تكرار الحروف للنص .

- **تكوين مصفوفة الصفوف للنص:**

تبدأ العملية بقراءة النص الأصلي من الملف ثم تحويل الملف النصي إلى مصفوفة من الرموز حيث يمثل كل حرف في النص بما يقابله من ASCII code وتكون المصفوفة الناتجة من هذه العملية ذات بعد واحد , يتم بعدها تحويل هذه المصفوفة إلى مصفوفة ذات بعدين ليتسنى لنا التعامل معها كصفحة من الأرقام الصحيحة التي حولت بما يقابلها بالنظام الثنائي لينتج ASCII code لكل حرف في كل صف من المصفوفة ويتم بعدئذ حساب عدد القيم التي تساوي(1) في كل رمز فإذا كان العدد زوجياً فيعبر عن الرمز الحالي من النص الأصلي بـ (1) وبخلافه يعبر عنه بـ (0) .

بعد أن يمثل كل رمز في النص الأصلي بـ (1) أو(0) أي بمرتبة ثنائية واحدة يتم حساب عدد قيم الـ (1) في كل صف , فإذا كان العدد زوجياً فيعبر عن الصف الحالي بـ (1) وبخلافه يعبر عن الصف بـ (0) .

وبعد ذلك سوف ينتج عن كل صف (1) أو(0) أي يعبر عن كل صف بمرتبة ثنائية واحدة ويوضع هذا الناتج في مصفوفة ذات بعد واحد وتكون هذه المصفوفة هي المرحلة الأولى من توليد العلامة المائية.

- **تكوين مصفوفة الأعمدة:**

يتم توليد مصفوفة الأعمدة بنفس الأسلوب المستخدم في توليد مصفوفة الصفوف أي يتم حساب عدد قيم (1) في كل عمود في المصفوفة ويوضع هذا الناتج في مصفوفة ذات بعد واحد .

- **تكوين مصفوفة تكرار الحروف :**

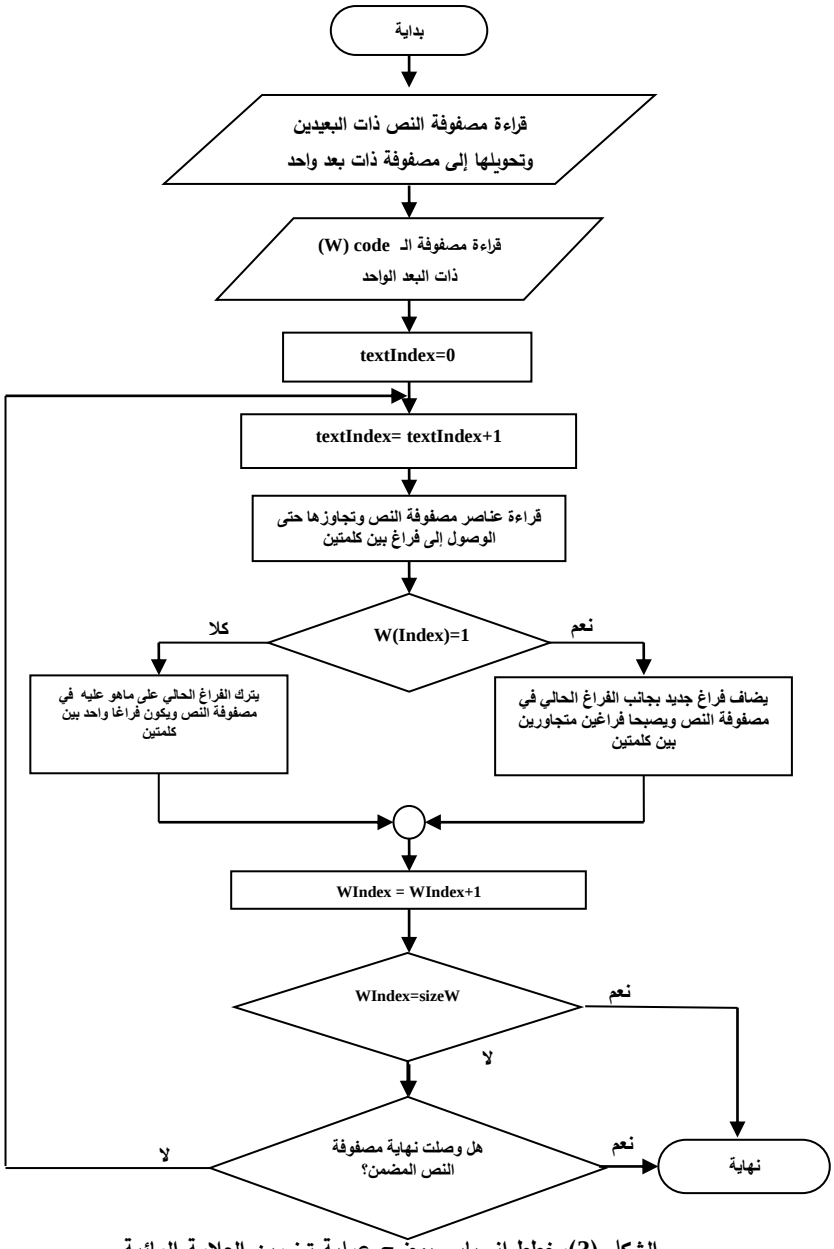
يتم في هذه المرحلة حساب تكرار كل حرف من حروف اللغة الإنكليزية في النص فإذا كان عدد تكرار الحرف زوجياً يعبر عنه بـ (1) وبخلافه يعبر عنه بـ (0) لينتج مرتبة ثنائية واحدة (1) أو (0) عن كل حرف من حروف اللغة الإنكليزية ويكون الناتج (26) مرتبة ثنائية توضع في مصفوفة ذات بعد واحد وهذه المصفوفة هي المرحلة الثالثة في عملية توليد العلامة المائية.

وأخيرا نربط المصفوفات الثلاث المذكورة انفا في مصفوفة ذات بعد واحد (W) تمثل العلامة المائية المولدة باستخدام النص الأصلي.

2.4 تضمين العلامة المائية Watermark Embedding

توجد عدة طرائق لتضمين العلامة المائية في النص كما تقدم في الفقرة 3 وفي هذا البحث تم اختيار طريقة التضمين بالفراغات تتابعيا في النص الأصلي وذلك باستخدام النص في هذه العملية وكذلك مصفوفة العلامة المائية الناتجة من التوليد (W) . ويوضح الشكل (3) عملية

تضمين العلامة المائية وذلك بفحص كل عنصر من مصفوفة العلامة المائية الناتجة من عملية التوليد و تفحص الفراغات بين الكلمات بالترتيب فإذا كان العنصر الحالي في المصفوفة (W) هو(0) فيبقى الفراغ المقابل له في النص على ما هو عليه ولا يتم تضمين أي شيء ثم يؤخذ الفراغ الذي يليه بين كلمتين ويؤخذ العنصر المقابل له من المصفوفة (W) فإذا كان هذا العنصر يساوي (1) فيتم في هذه الحالة إضافة (تضمين) فراغ آخر بجانب الفراغ الحالي في النص الأصلي ليصبحا فراغين متجاورين ويتم بعد هذه العملية تزحيف النص بأجمعه بقدر الفراغ المضاف وتستمر هذه العملية حتى نهاية المصفوفة (W) .



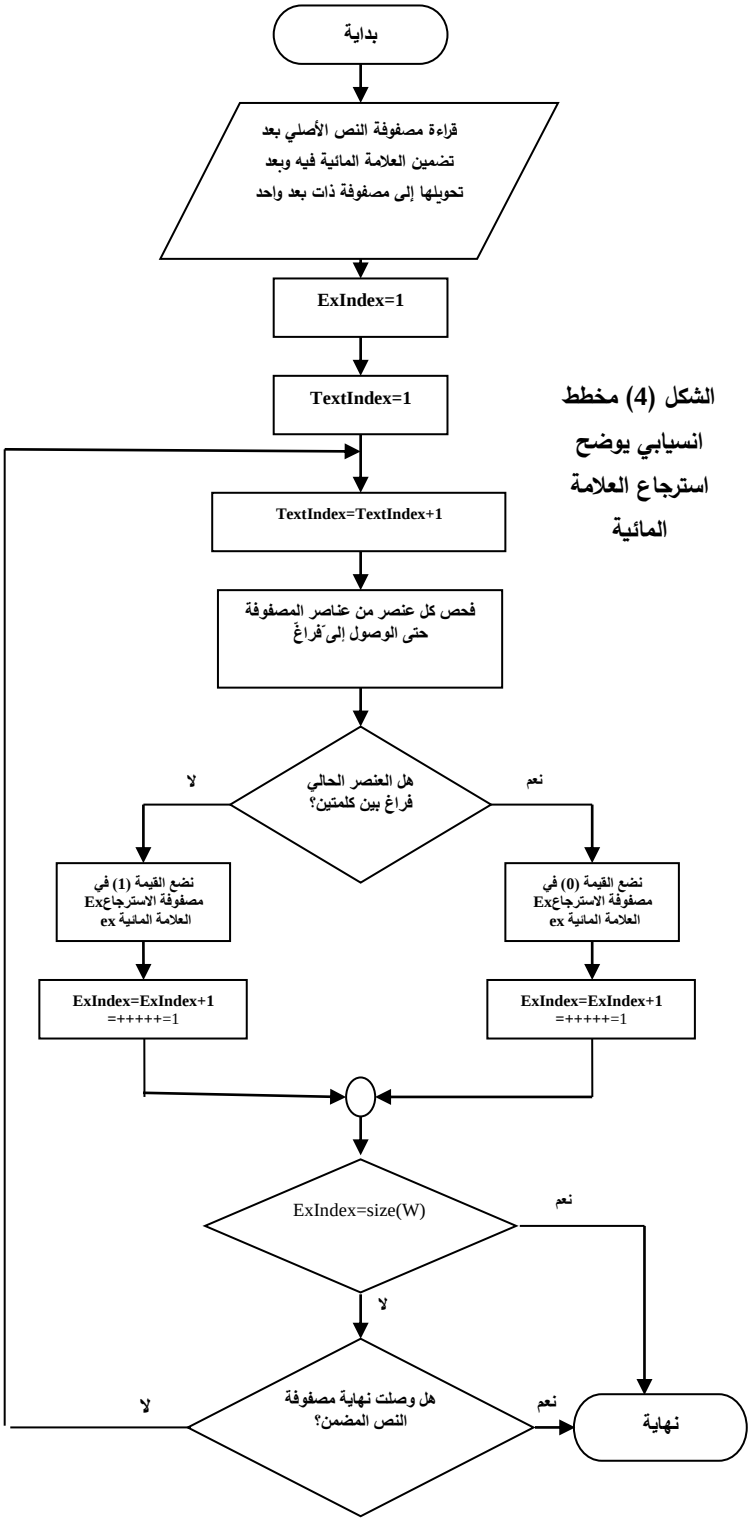
الشكل (3)مخطط انسيابي يوضح عملية تضمين العلامة المائية

3.4 استرجاع العلامة المائية Watermark Extracting

إن عملية استرجاع العلامة المائية هي العملية المعاكسة للتضمين , إذ يكون الإدخال في هذه العملية هو النص بعد تضمين العلامة المائية فيه والإخراج هو العلامة المائية المسترجعة التي تتمثل في البرنامج بالمصفوفة (Ex) بحيث تتم مطابقتها مع العلامة المائية الأصلية الناتجة من عملية التوليد (W) ويجب أن يكون التطابق تاما عندما لا يكون النص المضمن قد تعرض لأي تغيير بسبب التهديدات التي يمكن أن يتعرض إليها ، والتي سيرد ذكرها لاحقاً.

ويمثل الشكل (4) المخطط الانسيابي لعملية استرجاع العلامة المائية و يتلخص عمل خوارزمية استرجاع العلامة المائية بما يأتي :

تتم قراءة مصفوفة النص المضمن وكذلك يتم فحص الفراغات الموجودة في النص بين الكلمات ويعبر عن كل فراغ واحد بين كلمتين بالقيمة (0) ومن ثم وضع هذه القيمة في مصفوفة (Ex) ، ويتم التعبير عن كل فراغين متتابعين بين كلمتين بالقيمة (1) وتوضع هذه القيمة ايضا في المصفوفة (Ex)، وهي مصفوفة العلامة المائية الناتجة من عملية الاسترجاع .



الشكل (4) مخطط انسيابي يوضح استرجاع العلامة المائية

4.4 فحص العلامة المائية المسترجعة watermark testing

تتم مطابقة العلامة المائية الناتجة من خوارزمية الاسترجاع مع العلامة المائية الأصلية باستخدام مقاييس معينة أحدها هو مقياس الترابط (Normalized Correlation):

$$NC = \sum W_i \times W_i' / \sum W_i \times W_i$$

إذ أن :

W: تمثل العلامة المائية الاصلية .

W': تمثل العلامة المائية المسترجعة .

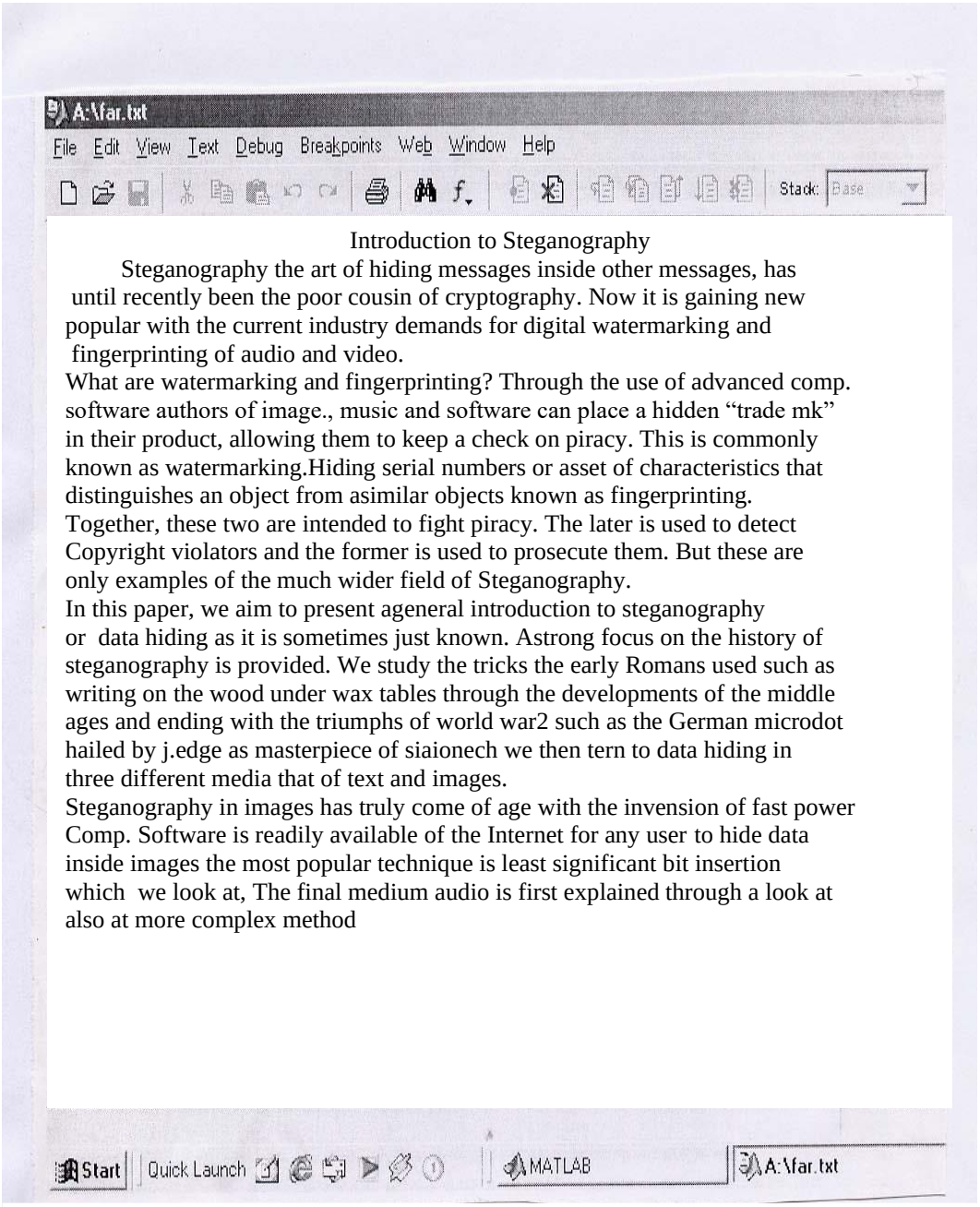
تكون قيمة NC بين [1,1-] وتعتبر هذه القيمة جيدة إذا كانت بين (0,1) وعندما لا يكون هناك أي تغيير في النص أو يكون الاسترجاع كاملاً فان هذه القيمة تساوي (1) [9].

4. النتائج:

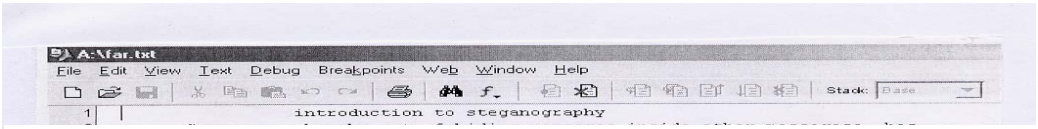
يمثل الشكل (5) النص الأصلي الذي قمنا بمعالجته بحسب طريقة البحث بينما يمثل الشكل (6) النص بعد تضمين العلامة المائية فيه.

5. التحقق من العلامة المائية Watermark verification

تم بعد ذلك إدخال كل من المصفوفة W التي تمثل العلامة المائية المولدة من النص الأصلي والمصفوفة Ex التي تمثل العلامة المائية المسترجعة من النص الأصلي بعد تضمين العلامة المائية فيه على دالة مقياس الترابط والتي تكون قيمة الناتج منها بين [1,1-] ، فتم الحصول على القيمة (1) ويعني ذلك أن العلامة المائية المسترجعة والمتمثلة بالمصفوفة Ex مطابقة تماما للعلامة المائية الأصلية والمتمثلة بالمصفوفة W.



الشكل (5) النص الأصلي



Introduction to Steganography

Steganography the art of hiding message inside other messages, has until recently been the poor cousin of cryptography. Now it is gaining new popular with the current industry demands for digital watermarking and fingerprinting of audio and video.

What are watermarking and fingerprinting? Through the use of advanced comp. software authors of image, music and software can place a hidden "trade mk" in their product, allowing them to keep a check on piracy. This is commonly known as watermarking.Hiding serial numbers or asset of characteristics that distinguishes an object from asimilar objects known as fingerprinting.

Together, these two are intended to fight piracy. The later is used to detect Copyright violators and the former is used to prosecute them. But these are only examples of the much wider field of Steganography.

الشكل (6) النص بعد تضمين العلامة المائية

6. التهديدات على النص المضمن Attacks on Embedding Text

قد يتعرض النص بوصفه وسطا من الأوساط المتعددة بعد تضمين العلامة المائية فيه إلى هجوم مقصود بقصد تغيير أو محو العلامة المائية وتتلخص التهديدات المستخدمة على النص بما يأتي :

1. تغيير في أحرف النص: وتشمل

- **تغيير بعض كلمات النص بكلمات أخرى بنفس عدد الحروف:**

قمنا بتغيير الكلمة(music) الواقعة في السطر(7) إلى الكلمة (sound) ثم قمنا باسترجاع العلامة المائية وإدخال النتائج على دالة مقياس الترابط فتم الحصول على القيمة (1).

ان تغيير كلمة واحدة أو مجموعة كلمات بكلمات أخرى لا تؤثر في العلامة المائية المسترجعة إذا كان عدد حروف الكلمات الجديدة بعدد حروف الكلمات الأصلية في النص المضمن وذلك لأن هذا النوع من التغيير لا يؤثر في الفراغات بين الكلمات .

- **تغيير بعض كلمات النص بكلمات أخرى بعدد متغير من الحروف :**

قمنا بتغيير الكلمة (.comp) الواقعة في السطر (6) من النص المضمن بالكلمة (computer) التي تحوي عدداً أكبر من الحروف والمختصر (mk) الواقع في السطر (7) من

النص المضمن بالكلمة (mark) والذي يحوي عدداً أكبر من الحروف **كما** قمنا بتغيير الكلمة (messages) الواقعة في السطر (2) من النص المضمن والتي وردت مرتين في نفس السطر بالكلمة (data) في كلا المرتين والتي تحوي عدداً أقل من الحروف ثم قمنا باسترجاع العلامة المائية وإدخال النتائج على دالة مقياس الترابط فتم الحصول على القيمة (1) .

نلاحظ مما سبق أن تغيير كلمة واحدة أو مجموعة كلمات لا يؤثر في العلامة المائية المسترجعة إذا كان عدد حروف الكلمات الجديدة أكبر أو أصغر من حروف الكلمات الأصلية في النص المضمن لأن هذا النوع من التغيير لا يؤثر في الفراغات بين الكلمات أيضاً .

- **تغيير في تركيب الجمل :**

قمنا بتغيير العبارة (of audio and video) الواقعة في السطر (5) من النص المضمن بالعبارة (of multimedia like text) كما قمنا بتغيير العبارة (through the use of) الواقعة في السطر (6) من النص المضمن بالعبارة (by using) ثم قمنا باسترجاع العلامة المائية وإدخال النتائج على دالة مقياس الترابط فتم الحصول على القيمة (0.75). ان التغيير في تركيب الجمل يؤثر في العلامة المائية المسترجعة بمقدار يعتمد على موقع الجملة وطول الجملة الجديدة قياساً بالجملة الأصلية وذلك لأن التغيير في تركيب الجمل يؤثر في الفراغات بين الكلمات أيضاً.

2. تزحيف أسطر في النص المضمن :

قمنا بتزحيف كل من السطرين (1) و (5) من النص المضمن ثم قمنا باسترجاع العلامة المائية و إدخال النتائج على دالة مقياس الترابط فتم الحصول على القيمة (1).

ان عدم تأثر العلامة المائية عند تزحيف الأسطر أمر طبيعي لأن الخوارزمية المقترحة في هذا البحث تقوم بقراءة النص المضمن في عملية استرجاع العلامة المائية وعندما يصادف وجود مجموعة فراغات متتابعة في النص المضمن فإن الخوارزمية تقوم بتجاوزها .

7. الاستنتاجات :

لقد تم في هذا البحث تصميم وتنفيذ طريقة لتوليد علامة مائية من النص الأصلي ومرت عملية التوليد بمراحل ثلاث كما تم تضمين هذه العلامة في النص بطريقة الفراغات ولوحظ في عملية التضمين ان عدد الفراغات المضافة إلى النص الأصلي في عملية تضمين العلامة المائية يكون بقدر عدد قيم الـ (1) الموجودة في مصفوفة العلامة المائية كما لوحظ ان حجم النص الأصلي بعد تضمين العلامة المائية فيه أكبر من حجم النص الأصلي بقدر عدد قيم الـ (1) في العلامة المائية. تمت عملية تضمين فراغ مع الفراغات الموجودة بين الكلمات فقط أي في حالة وجود فراغ

واحد بين كلمتين أما في حالة وجود مجموعة فراغات متتابعة فيتم تجاوزها حتى الوصول إلى فراغ واحد بين كلمتين . ان طريقة التضمين المستخدمة في هذا البحث لاتصلح مع النصوص الصغيرة جدا وذلك لقلة عدد الفراغات الموجودة بين الكلمات في هذه النصوص ولحل هذه المشكلة يعتمد أحد مستويات الحماية في توليد العلامة المائية ويهمل المستويان الآخران .

اما في عملية استرجاع العلامة المائية فقد لوحظ مقاومة الطريقة المستخدمة لمختلف انواع التهديدات التي قد يتعرض لها النص المضمن إذ أن التهديد بإضافة أو حذف كلمات في النص المضمن لا يؤثر في العلامة المائية المسترجعة وذلك لأن هذا النوع من التهديد لا يؤثر في عدد الفراغات بين الكلمات في النص. أما التهديد بتغيير بعض عبارات النص المضمن بعبارات أخرى فانه يؤثر تأثيرا واضحا في العلامة المائية المسترجعة وذلك لأن التغيير في عبارات النص يؤثر في عدد الفراغات بين الكلمات، ومدى تأثير هذا التهديد في العلامة المائية يعتمد على موقع العبارات التي يتم تغييرها وطول العبارات الجديدة قياساً بالعبارات الأصلية .

المصادر

[1] Bender, W.; Gruhl, D. ; N. Morimoto and A. Lu (1996) **Techniques for Data Hiding** ,IBM System Journal 35(3&4).

[2] Jian Zhao and Eckhard Kock (1996) **A Digital Watermarking System for Multimedia Copyright Protection**, ACM Multimedia, Boston, USA.

[3] Johnson, N.; Duric, Z., Jajoda (2000) **Information Hiding: Steganography and Watermarking Attacks and Countermeasure**, Kluwer Academic Publisher.

[4] Mark Noto (2001) **Hiding Text in MP3 Files** , SANS Institute.

[5] Nedeljko Cvegic,Tapio Seppanen, " Improving Audio Watermarking Scheme Using Psychoacoustic Watermark Filtering ".

[6] Neil, F. Johnson and Zoran Duric (2000) " Steganography and Digital Watermarking ", **http://www.jjtc.com/ Steganography/**.

[7] Sabrina, Sowers and Abdou Youssef (1998) **Testing Digital Watermarking Resistance to Destruction** , Information Hiding LNCS.

[8] Stefan Katzenbeisser ; Fabien A. P. Petitcolas (1999) "Information Hiding Techniques for Steganography and Digital Watermarking", **Arteech House**, Boston, London .

[9] Wah, W. (2003) "Image Watermaking and Data Hiding Techniques", Ph.D. Thesis, **Hong Kong University**.