Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Computer Science

Protection of Relational Databases by Means of Watermarking

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Doctor of Philosophy in
Computer Science**

By

Asmaa Mowafaq Mohammed AlQassab

Supervised by

Prof. Dr. Mafaz Mohsin Khalil Alanezi

2022 A.D. **1444 A.H.**

Abstract

For a long time, the watermarking scheme to meet image, audio and video data type security was considered ever-popular, but during the last several years, "Relational Database" watermarking was really in mind and understudy. Generally, relational databases differ from multimedia objects in data characteristics; this is why techniques developed for watermarking multimedia objects are not applicable for watermarking relational databases. Multimedia data are highly correlated unlike the relational databases which comprised of isolated objects (tuples) that may be usually modified, deleted, or inserted with malign or even benign intentions. Such operations on relational databases' tuples cannot be supported by any of the existing techniques designed for multimedia watermarking; therefore, an alternative class of watermarking approaches is needed.

Watermarking implies the embedding of some data in a manner that it is easily accessible by a user who is authenticated rather than other users, bearing in mind that the underlying data may experience some changes as a result of that data embedding. For watermarking relational database, there are some limitations regarding watermark embedding; with such a well-designed tabular data structure, the entire database renders to be useless in case of attributes values are changed, even minimal alteration can affect data usage. Besides, existing techniques are so fragile to be utilized for the database data which keeps updating, and finally the watermark detection are non-blind.

Even though various relational database watermarking approaches had been suggested, nonetheless, these techniques do not sturdy enough against malignant attacks that can result in false insertion, modification or elimination leading to data performance and quality deterioration. Altering the attributes values is the most significant point in watermarking relational databases and has to be taken into consideration. If it is not possible to watermark an attribute without affecting its value, then a watermark must not be inserted. So, with the aim of advancing and improving existing relational database watermarking techniques, a new robust and reversible relational database watermarking method based on DE is suggested by which the relational database data type is treated as

any other data type and embedded as meta data within a multimedia object (colored image).

By applying the suggested method, the quality of both embedded and underlying data is guaranteed and better measurements outcomes are achieved not only after the embedding process of the watermark but also after experimenting various known types of image watermarking attacks. Metrics utilized for evaluating the results include: SSIM, PSNR, MSE and IF. The suggested method shows that these metrices have average values of: 1, 78, 0 and 1 respectively. Besides, AR and BER metrics utilized for analyzing to which extent the embedded watermark has been affected, where the results are 1 and 0 respectively which show that the embedded data are completely retrieved. Also, the suggested method succeeds in resisting 7 attacks including: Bilateral Filtering, Histogram Equalization, Intensity Adjustment, Invert Image, Salt and Pepper Noise, Gaussian Noise and Sharpening. Metrics utilized for evaluating the results include: SSIM, PSNR, MSE and IF with still accepted average values. AR and BER metrics utilized for analyzing to which extent the embedded watermark has been affected by attacks, where the results show that, with the first five attacks the embedded data are completely retrieved with AR value of 1 and BER value of 0, and with the last two attacks about 25% can be retrieved with AR and BER values close to 1 and 0 respectively.

وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

حماية قواعد البيانات العلائقية

بواسطة العلامة المائية

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة دكتوراه فلسفة في
علوم الحاسوب

من قبل

أسماء موفق محمد القصاب

بإشراف

أ.د.مفاز محسن خليل العنزي

١٤٤٤ هـ ٢٠٢٢ م

المستخلص

لطالما كانت تقنية "العلامة المائية" الأوسع شهرة فيما يتعلق بأمنية البيانات سواء أكان نوعها صورة ، صوت ، فيديو أو نص والتي حظيت بشعبية كبيرة ; ولكن في السنوات الأخيرة لم تلبث "قواعد البيانات العلائقية" الا وأن أصبحت جزء من تلك الأنواع وحظيت بذات الشعبية. بصورة عامة فان قواعد البيانات العلائقية تختلف في خصائص بياناتها عن الوسائط المتعددة ولذا فان تقنيات العلامة المائية المطورة للوسائط المتعددة لا يمكن تطبيقها على قواعد البيانات العلائقية، حيث تكون البيانات في الوسائط المتعددة على درجة عالية من الترابط على خلاف البيانات في قواعد البيانات العلائقية التي تكون معزولة عن بعضها وقابلة للتعديل، الحذف او الاضافة بشكل طبيعي او متعمد. هذا النوع من العمليات التي تجري بصورة دورية على البيانات في قواعد البيانات العلائقية غير قابل للدعم من قبل أي من تقنيات العلامة المائية المصممة للوسائط المتعددة، وبالتالي ظهرت الحاجة الى تقنيات بديلة.

تشير العلامة المائية الى تلك التقنية التي تتضمن عملية تضمين بعض البيانات بشكل يمكن فيه للأشخاص المخولين فقط الوصول اليها، مع الأخذ بنظر الإعتبار أن عملية التضمين هذه من الممكن أن تتسبب بإحداث تغييرات للبيانات الأصلية (التي يتم التضمين خلالها); وفيما يخص قواعد البيانات العلائقية فان طمر العلامة المائية فيها يخضع لقيود تتعلق بهيكلية البيانات من جداول وسمات والتي من الممكن ان تجعل قواعد البيانات تفقد جدواها في حال حصول تغيير لاي من السمات حتى وان كان اقل مايمكن. الى جانب انه التقنيات المتاحة تعتبر هشة جدا اذا ما استخدمت لقواعد البيانات التي تكون في حالة تحديث مستمر.

تم اقتراح العديد من الطرق لتأمين الحماية لقواعد البيانات العلائقية بإستخدام العلامة المائية، ولكن لا تزال مثل هذه الطرق ليست بالقوة الكافية أمام الهجمات المتعمدة التي قد تتسبب في تغيير أو حذف أو إضافة غير مصرح لها للبيانات والتي من شأنها أن تؤدي إلى تدهور جودة وكفاءة البيانات. ان تغيير قيم السمات هو الامر الأكثر أهمية عند تطبيق العلامة المائية على قواعد البيانات العلائقية والذي يجب اخذه بنظر الاعتبار. وفي حال لا يكون بالإمكان ادخال العلامة المائية من دون التأثير على قيم تلك السمات فعندها يتوجب عدم ادخال العلامة المائية. لذا وبهدف تطوير وتحسين طرق العلامة المائية المستخدمة لحماية قواعد البيانات

العلائقية، تم اقتراح تقنية علامة مائية جديدة قوية قابلة للعكس باستخدام تقنية DE كتقنية يتم فيها التعامل مع بيانات قواعد البيانات العلائقية كأي نوع اخر من البيانات ليتم طمرها في غطاء(صور ملونة).

من خلال تطبيق التقنية المقترحة تم ضمان جودة كلا من البيانات المطمورة والبيانات الاصلية (بيانات الغطاء)، كما تم الحصول على نتائج أفضل لم تقتصر على نتائج مابعد عملية تضمين العلامة المائية وانما شملت اختبار التقنية المقترحة لسبعة أنواع من الهجمات المتعارف عليها والتي تضمنت: (Bilateral Filtering, Histogram Equalization, Intensity Adjustment, Invert Image, Salt and Pepper Noise, Gaussian Noise, Sharpening). وقد تم اعتماد مقاييس للمقارنة بين الصور قبل وبعد عملية التضمين تضمنت (SSIM, PSNR, MSE, IF) بمعدل قيم (1, 78, 0, 1) على التوالي، بالإضافة الى مقياسي (AR , BER) لاختبار مدى تاثر العلامة المائية وبمعدل قيم (1,0) على التوالي والتي تشير الى إمكانية استعادة العلامة المائية بالكامل. كذلك تم اعتماد المقاييس ذاتها للمقارنة بين الصور قبل وبعد التعرض للهجمات، والتي بينت أيضا مدى فاعلية الطريقة المقترحة وقدرتها على استعادة العلامة المائية بالكامل في الانواع الخمس الأولى من الهجمات، وما يقارب 25% في النوعين الأخيرين.