

هجرة البيانات بين قواعد البيانات غير المتجانسة  
Data Migration Between Heterogeneous Databases

محمد قاسم ياسين  
شركة نفط الجنوب

د. طالب عبد الصمد عبيد  
كلية تقنية المعلومات / جامعة البصرة

الملخص

التطور الكبير في تقنيات المعلومات والاتصال وظهور نظم قواعد البيانات الموزعة وتوزيع البيانات على أماكن متعددة أدى إلى الحاجة لنقل البيانات من نظم قواعد بيانات التقليدية (القديمة) إلى نظم قواعد بيانات حديثة وفعالة فضلاً عن الحاجة الملحة لنقل البيانات من قاعدة بيانات إلى أخرى قد تختلف بنوع نظام إدارة قواعد البيانات أو تختلف في النوع البياني لحقول جداولها وذلك لتوفير سرعة عالية لنقل البيانات و التخلص من التعقيدات الناتجة من عدم التماثل في بيئة قواعد البيانات الموزعة.

في عملنا هذا جرى بناء برنامج تطبيقي باستخدام لغة فصول بيسك دوت نت VisualBasic.net يوفر إمكانية سهلة وفعالة ودقيقة جداً لنقل البيانات بين قواعد البيانات المتماثلة و غير المتماثلة على حدٍ سواء. يحتوي البرنامج التطبيقي على مجموعة من النهج المعالجة تستخدم لنقل البيانات من قاعدة بيانات إلى أخرى تختلف في نظام ادارة قاعدة البيانات او هيكل الجداول والحقول او النوع البياني.

الكلمات المفتاحية : هجرة البيانات، قواعد البيانات الموزعة، قواعد البيانات المتجانسة وغير المتجانسة .

Abstract

The great development in information and communication technologies and the emergence of distributed database system that allocated the data on different locations. However, there is a need to transfer data from the one database systems to another database system effectively. Actually, the database may differ in the managements database system or the structure of fields' and the data type. Consequently, it is provided high availability of data and get rid of the complexities arising from the asymmetry in distributed databases environment.

In this work we have focused on building a software application using VisualBasic.net to provide an easy, efficient and very accurate for the transfer data between the databases whether it is symmetric or asymmetric. The application program is a set of procedures that perform the data transfer between distinct database whether that are differ in the database management system or the structure of tables besides field's data types

**Keywords:** - Data Migration, Distributed database, Homogenous and Heterogeneous database

1. المقدمة

إن تطور أنظمة المعلومات الحاسوبية وانتشارها على مواقع متعددة وظهور الحاجة للاستفادة من بيانات تلك الأنظمة في أماكن مختلفة أدى إلى ظهور نظم قواعد البيانات الموزعة والتي تعرف بأنها مجموعة من الأنظمة المستقلة في مواقع متباعدة ومرتبطة بشبكة حاسوبية مع بعضها حيث تبدو بالنسبة إلى المستخدمين كأنها نظام واحد متماسك [1]. وتمتاز قواعد البيانات الموزعة بالاستقلالية المحلية لكل قاعدة بيانات مشاركة في البيئة الموزعة، واستقلالية الموقع حيث يستطيع المستخدم الوصول إلى قاعدة البيانات من أي موقع ومعالجة البيانات والاستعلامات واسترجاع البيانات من المواقع المختلفة [2].

التطور السريع في نظم إدارة قواعد البيانات وتعدد أنواع نظم الإدارة والأنواع البيانية لحقول الجداول أدى إلى تصنيف نظم قواعد البيانات الموزعة إلى صنفين حسب التغيرات وعدم التغيرات في نوع نظام إدارة البيانات أو النموذج البياني أو تنظيمات هياكل البيانات وأنواع مفاتيح الربط. يطلق على الصنف الأول قواعد البيانات المتجانسة Homogeneous Database حيث يمتاز هذا النوع من قواعد البيانات في تماثله في كل المواقع من ناحية نظم إدارة قواعد البيانات والنوع البياني لحقولها و نظم التشغيل. إن بناء نظام قواعد بيانات موزعة لهذا النوع يكون أكثر سهولة من حيث التصميم والإدارة وتكامل البيانات بين المواقع المختلفة في بيئة العمل. إما الصنف الآخر هو قواعد البيانات غير المتجانسة Heterogeneous Database و تطلق على قواعد البيانات الموزعة بغير المتجانسة عندما تكون قواعد البيانات في البيئة الموزعة المتواجدة على الشبكة مختلفة و يكون الاختلاف إما في أنواع نظم إدارة قواعد البيانات أو النموذج البياني والنوع البياني للحقول ويمتاز هذا النوع بصعوبة الإدارة وتكامل البيانات بين المواقع المختلفة [3] ، [1].

إن الاستفادة من التطور في نظم قواعد البيانات المختلفة على صعيد العديد من المؤسسات أظهرت الحاجة إلى البحث عن وسائل لنقل البيانات المتواجدة في قواعد بيانات قديمة إلى قواعد بيانات جديدة. إن عملية إعادة إدخال البيانات من جديد تكون غير فعالة ومصاحبة للكثير من الأخطاء لذلك يجب البحث عن طرق أكثر ملائمة إلى نقل البيانات بين قواعد البيانات المختلفة بطريقة فعالة، آمنة، وأسرع مع الابتعاد عن احتمالية وقوع الأخطاء في نقل تلك البيانات. حيث ظهرت العديد من البحوث في هذا المجال منها:

قدم مجموعة من الباحثين Patel S., et al [4] بحثاً استند على بناء أداة لعملية نقل الهيكل البياني والبيانات من قاعدة بيانات إلى أخرى تختلف في نظام الإدارة، حيث يقوم المستخدم بتحديد نوع نظام إدارة قاعدة بيانات المصدر مثل (Access, SQL Server, Oracle) حيث يقوم البرنامج بفتح الاتصال معها ثم يقوم المستخدم بتحديد قاعدة بيانات الهدف وفتح الاتصال معها ويقوم البرنامج بتحويل البيانات.

عام 2013 تطرق بحث [5] Vitthal S. A., et al لبناء أداة تعمل على إنشاء ونقل بيانات الجداول من قاعدة بيانات إلى أخرى غير متماثلة في نظام الإدارة قواعد البيانات حيث تم بناء النظام على جزئين باستخدام لغة C#. الجزء الأول هو الواجهة الأمامية Front End و تم بناء الواجهات التفاعلية لتكون أكثر جاذبية وتفاعلية، أما الجزء الآخر الجهة الخلفية Back End حيث احتوى على شفرات الاتصال بقواعد البيانات وشفرات إنشاء الجداول. تدعم تلك الأداة الاتصال بأنواع نظم قواعد البيانات (Access, SQL Server, Oracle).

عام 2006 تطرق الباحث عقيل فرحان [6] لعملية تحويل البيانات بين الأنواع المختلفة من قواعد البيانات غير المتجانسة أي بناء قواعد بيانات جديدة دون الحاجة إلى إنشائها يدوياً وإعادة إدخال البيانات لها.

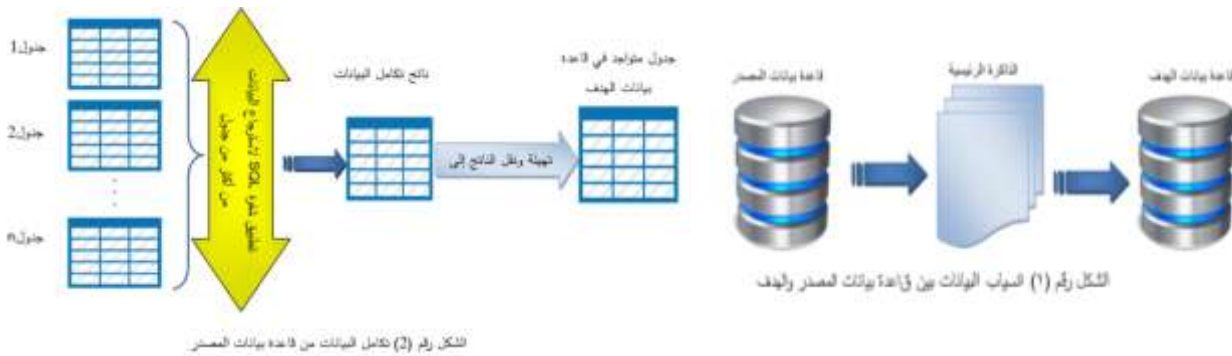
عام 2003 قدم بحث [7] من قبل Qinghong Line و تضمن تصميم وتطبيق نظام لنقل بيانات بين قواعد البيانات غير المتماثلة. جرى تطبيق البرنامج على قاعدتي بيانات تحتوي معلومات إدارة مدرسة، وكانت كلا القاعدتين تختلف في الهيكل البياني للجداول.

## 2. النظام المقترح لنقل البيانات بين قواعد البيانات غير المتماثلة

إن عملية نقل البيانات المتواجدة في قاعدة بيانات المصدر إلى قاعدة البيانات الهدف من الأمور الأكثر تعقيداً وذلك بسبب الاختلاف في نظام إدارة البيانات أو الهيكل البياني والنوع البياني للحقول لكلا قاعدتي البيانات. جرى التركيز في هذا البحث على بناء أداة تقوم بنقل البيانات بين قواعد البيانات المتجانسة و غير المتجانسة في نظام إدارة قاعدة البيانات وكذلك هيكل الجداول و النوع البياني للحقول.

الشكل رقم (1) يوضح انسياب البيانات من قاعدة بيانات المصدر إلى قاعدة بيانات الهدف. في بادئ الأمر يتم إجراء عملية تكامل بين جداول قاعدة بيانات المصدر (عملية تجميع حقول البيانات من الجداول المختلفة) في حالة كون البيانات موزعة على جداول مختلفة وبعد ذلك يتم استنساخ البيانات من قاعدة بيانات المصدر إلى الذاكرة الرئيسية ومن ثم يتم إجراء عملية تهيئة لتلك البيانات لتكون متماثلة مع النوع البياني لحقول قاعدة بيانات الهدف وأخيراً يتم نقلها إلى قاعدة بيانات الهدف بعد إتمام تهيئة البيانات الموجودة في الذاكرة بشكل بيانات المصدر.

الشكل رقم (2) يوضح تكامل البيانات من جداول قاعدة بيانات المصدر في هيكل موحد ليتم إجراء عملية التهيئة عليها ثم النقل.



### 2.1- مراحل نقل البيانات بين قواعد البيانات غير المتماثلة في البرنامج المقترح

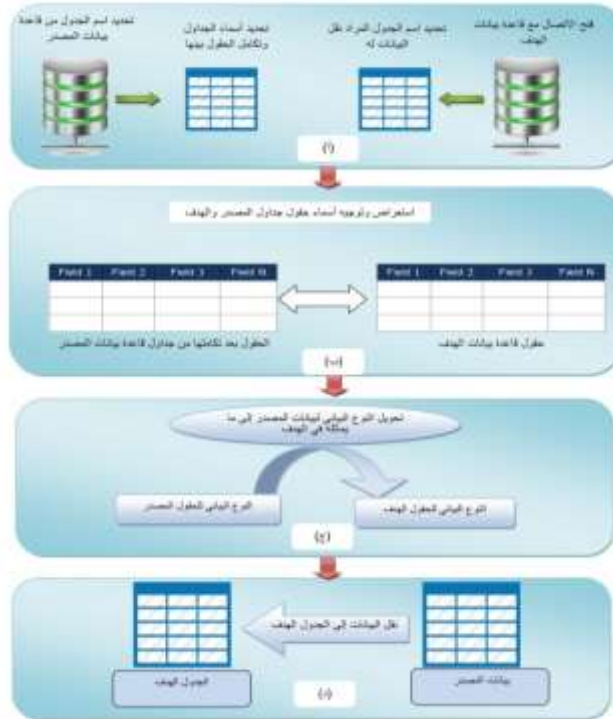
إن عملية نقل البيانات بين قاعدتي بيانات تمر بعدة مراحل متسلسلة:

أ- مرحلة فتح الاتصال وجلب المعلومات التي تصف الجداول والحقول في قاعدة البيانات وتتضمن هذه المرحلة العمليات التالية الشكل رقم (3- أ) يوضح هذه الخطوة:

1. تحديد مزود قاعدة البيانات.
2. تحديد اسم قاعدة البيانات.
3. فتح الاتصال بقاعدة البيانات.
4. جلب أسماء الجداول المتواجدة داخل قاعدة البيانات.
5. جلب البيانات الوصفية الخاصة بالجدول متضمنة معلومات أسماء الحقول والأنواع البيانية لتلك الحقول بعد إن يتم اختيار الجدول.

ب- مرحلة إجراء مطابقة بين أسماء حقول قاعدة بيانات المصدر والهدف حيث يتم التخلص من الاختلاف في هيكل قاعدتي بيانات المصدر والهدف وذلك من خلال عمل مطابقة بين حقول جداول قاعدة بيانات المصدر وحقول جدول قاعدة بيانات الهدف والشكل رقم (3- ب) يوضح هذه الخطوة.

ج-مرحلة تهيئة البيانات قبل النقل  
 قبل عملية النقل يتم التخلص من الاختلاف في النوع البياني بين البيانات وذلك بإدخال البيانات قبل تمريرها إلى قاعدة البيانات الهدف إلى نهج معالجة يعمل على تعديل الأنواع البيانية لبيانات المصدر لتكون مطابقة إلى النوع البياني لبيانات الهدف والشكل رقم (3- ج) يوضح هذه الخطوة،



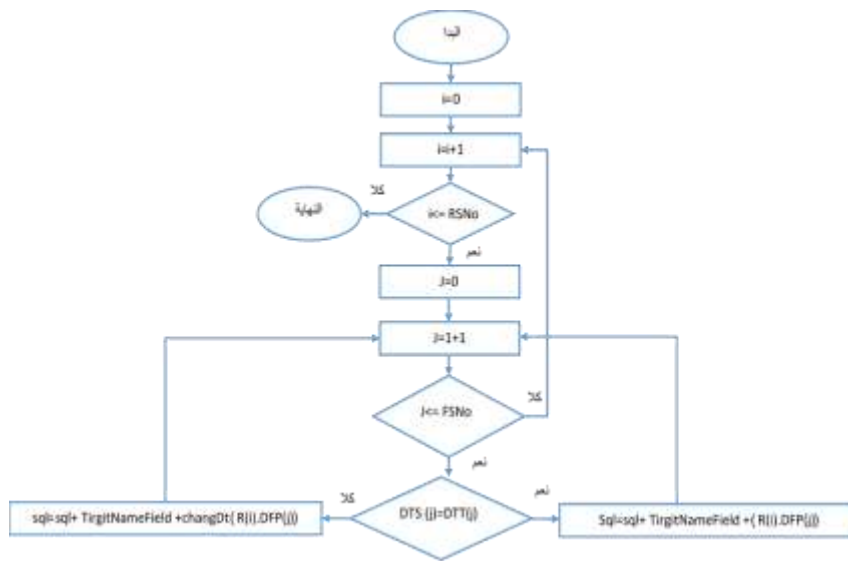
شكل (3) مراحل نقل البيانات من قاعدة بيانات المصدر

مرحلة تهيئة البيانات ممكن ان نجريها اما بالتهيئة الافقية او التهيئة العامودية.

- التهيئة الافقية يجري الفحص والتحويل للنوع البياني للبيانات وخزنة سطر بسطر. والمخطط رقم (1) يبين خوارزمية التهيئة الافقية.

عدد مرات فحص النوع البياني=عدد الاعمدة\*عدد الاسطر.

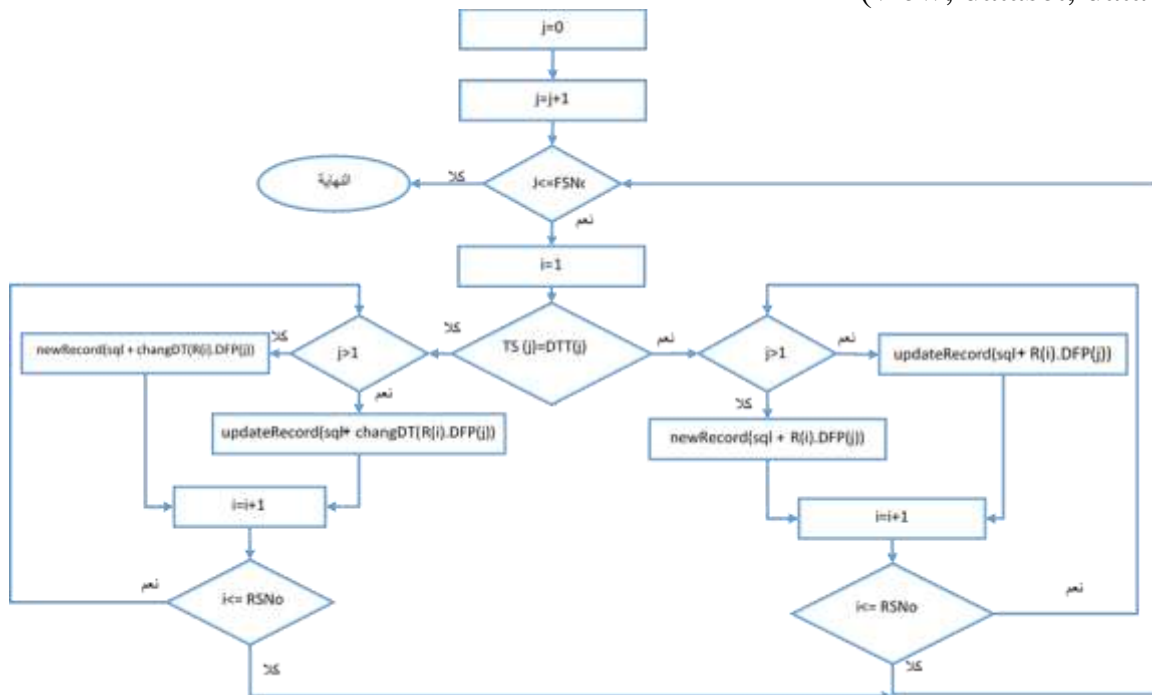
عدد مرات الاتصال بالملقم لحفظ البيانات=عدد الاسطر



مخطط رقم (1) خوارزمية التهيئة الافقية

- التهيئة العامودية يجري فحص وتحويل البيانات و تخزينها عمود بعمود، تكون اكثر سرعة عندما يكون هناك عدد قليل من الاعمدة. والشفرة البرمجية اكثر تعقيداً من التهيئة الافقية. المخطط رقم (2) يمثل خوارزمية التهيئة العامودية.

عدد مرات فحص النوع البياني = عدد الاعمدة  
 عدد مرات الاتصال بالملقم لحفظ البيانات = عدد الاعمدة \* عدد الاسطر  
 (ومن الممكن تقليل الاتصال بالملقم من خلال خزن البيانات على الذاكرة باستخدام الادوات مثل data grid (view, dataset, data table



مخطط رقم (2) خوارزمية التهيئة العامودية

R(i): موقع السجل في قاعدة بيانات المصدر  
 DTS(j): مصفوف تحتوي على النوع البياني لقاعدة بيانات المصدر (المنقول منها)  
 DTT(j): مصفوفة تحتوي على النوع البياني لقاعدة بيانات الهدف (المنقول لها)  
 DFP(j): موقع الحقل المراد نقل بياناته من قاعدة بيانات المصدر  
 FSNo: عدد الحقول المراد تحويلها  
 RDNO: عدد السجلات المراد نقلها  
 j: عداد للتنقل بين الاعمدة  
 i: عداد للتنقل بين الصفوف  
 R(i).DFP(j): قيمة الحقل المراد نقلها  
 R(i): تمثل رقم السجل، DFP(j): تمثل رقم العمود

د-مرحلة نقل البيانات إلى قاعدة بيانات الهدف  
 بعد التخلص من الاختلاف في النوع البياني للبيانات يجري نقلها إلى قاعدة بيانات الهدف او عرضها في احدى ادوات عرض البيانات ومن ثم نقلها والشكل رقم (3- د) يوضح هذه الخطوة.

## 2.2 - تنفيذ البرنامج

تم استخدام لغة فصول بيسك دوت نت VB.Net في بناء البرنامج التطبيقي ولذلك لما يقدمه من إمكانيات في تصميم الواجهات التفاعلية و الدور الكبير في تسهيل عملية تعامل المستخدم مع البرنامج، بالإضافة إلى إمكانيته في القدرة على الاتصال مع العديد من أنظمة قواعد البيانات المختلفة.  
 عند تنفيذ النظام التطبيقي تظهر لنا النافذة كما في الشكل رقم (4) والتي تحتوي على أربعة أقسام:



شكل رقم (4) النافذة الرئيسية للبرنامج التطبيقي

أ. يختص القسم الأول بمعلومات قاعدة بيانات المصدر التي سوف يتم نقل البيانات منها بالإضافة إلى جزء يتم فيه عرض البيانات التي سوف يتم نقلها إلى قاعدة بيانات الهدف.

ب. يحتوي القسم الثاني على معلومات قاعدة بيانات الهدف التي سوف يتم تحويل البيانات إليها بالإضافة إلى جزء يتم فيه عرض بيانات الجدول الذي سوف يتم إضافة بيانات له.

ج. يتضمن القسم الثالث أسماء الحقول التي تنتمي إلى جداول قاعدة بيانات المصدر وحقول جدول قاعدة البيانات الهدف والأنواع البيانات وكذلك يتضمن جدول يتم وضع فيه كل حقل من قاعدة بيانات المصدر وما يقابله في قاعدة بيانات الهدف.

د. القسم الرابع خاص بالبيانات التي تم تنسيقها لتكون جاهزة للنقل.

### 2.3 - إجراء تحويل للبيانات بين قاعدتين بيانات غير متماثلتين

إن عملية إجراء نقل للبيانات بين قاعدتي بيانات تكون بإجراء مجموعة من الخطوات المتسلسلة وهي كما يلي:

1. الخطوة الأولى هي مرحلة جلب البيانات من قاعدة بيانات المصدر إلى ذاكرة الحاسوب مؤقتا وهي كالتالي:

أ. يتم اختيار نوع قاعدة البيانات من القائمة المنسدلة (حيث يسمح النظام بالاتصال بقواعد البيانات (SQL Server, Oracle, SQL compact, Access).

ب. يتم الضغط على زر فتح الاتصال بقاعدة البيانات (التي تم تحديدها من القائمة المنسدلة) حيث تظهر نافذة تطلب منا تحديد موقع خزن قاعدة البيانات في حالة قواعد البيانات (Access, SQL compact) أو اسم

الخادم Server في حالة قواعد البيانات (Oracle, SQL Server)

ت. تظهر نافذة تطلب منا إدخال اسم قاعدة البيانات وعند إدخال اسم قاعدة البيانات تظهر بعدها نافذة إدخال اسم المستخدم ومن ثم نافذة إدخال رمز المرور للتأكد من صلاحية المستخدم بالوصول إلى قاعدة البيانات.

ث. يقوم البرنامج بفتح الاتصال مع قاعدة البيانات بعد التحقق من اسم المستخدم ورمز المرور وجلب معلومات الجداول التي تحتويها قاعدة البيانات. والشكل (5) يوضح مراحل الاتصال بقاعدة البيانات.



شكل (5) مراحل الاتصال بقاعدة البيانات

2. بعد فتح الاتصال مع قاعدة البيانات المصدر يتم اختيار اسم الجدول الذي سوف تنقل بياناته من خلال قائمة أسماء الجداول ليتم بعدها عرض بيانات ذلك الجدول، بالإضافة إلى ذلك يوفر النظام إمكانية جلب البيانات من أكثر من جدول وذلك من خلال تمرير شفرة SQL التي يجري بنائها او كتابتها تحت القسم الأول والضغط على زر نفذ شفرة sql لتعرض البيانات في الإدارة (Data Grid View).

إن عملية جلب البيانات يتم من خلال الكيانات التي يوفرها فجول بيسك دوت نت VB.net ويعتبر الكيان (Data Adapter) المسئول عن جلب البيانات و الكيانات (Dataset, Datatable) هما عبارة عن مخازن وسيطة لتخزين البيانات التي تم جلبها من قاعدة البيانات. الشكل (6) يبين القسم الأول بعد عملية فتح الاتصال بقاعدة بيانات المصدر.

ID_I	ID_S	DATE_S	DATE_D	BINARY_I	BINARY_S	NUMBER_F	INVARCHAR2_F
1	1	1999/10/28	07/05/1993	1	1	321	321
2	2	1999/10/28	07/05/1993	0	0	123	123
3	3	1999/10/28	07/05/1993	1	1	123.20	123.2
4	4	1999/10/28	07/05/1993	0	0	247	247
5	5	1999/10/28	07/05/1993	0	0	444	444
7	7	1999/10/28	07/05/1993	1	1	213	213
8	8	1999/10/28	07/05/1993	0	0	55.70	55.77
9	9	1999/10/28	07/05/1993	0	0	8888E.20	8888.2
10	10	1999/10/28	07/05/1993	1	1	77.66	77.66

الشكل (6) القسم الأول بعد فتح الاتصال بقاعدة بيانات المصدر



3. يتم إجراء نفس الخطوات السابقة لفتح الاتصال مع قاعدة بيانات الهدف وجلب معلومات الجداول وعرض البيانات في الإدارة (DataGridView). والشكل (7) يبين القسم الأول والثاني بعد فتح الاتصال بكلاً قاعدتي بيانات المصدر والهدف.

4. بعد إن يتم اختيار البيانات المراد نقلها وتحديد الجدول المراد نقل البيانات له في قاعدة بيانات الهدف يتم إجراء عملية مقابلة الحقول إي بيانات الحقل ذات الاسم (firstn) في جدول المصدر يجب إن ينقل بياناته إلى الحقل (fname) في جدول الهدف وذلك للتخلص من التغيرات في أسماء الحقول، وكما موضح في الشكل رقم (8)

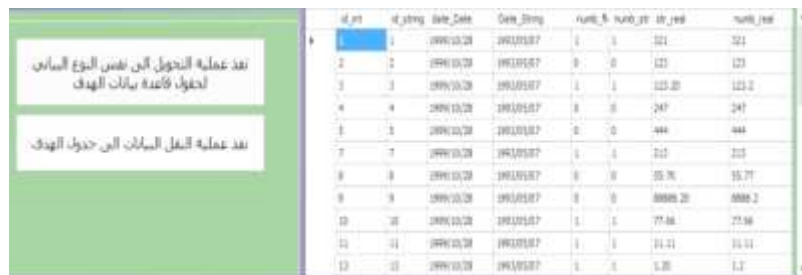


الشكل (7) البيانات بعد فتح الاتصال بقاعدتي بيانات المصدر والهدف



الشكل (8)مطابقة أسماء الحقول

5. بعد اختيار الحقول المراد نقل بياناتها بين قاعدتي البيانات يتم الضغط على زر تهيئة البيانات الذي يظهر في القسم الرابع في الشكل رقم (4) ليتم تنفيذ نهج يقوم بتغيير النوع البياني للبيانات وتهيئتها للنقل إلى قاعدة بيانات الهدف، ليظهر القسم الرابع كما في الشكل (9).



الشكل (9) البيانات المتماثلة التي سوف تنقل الى الجدول الهدف

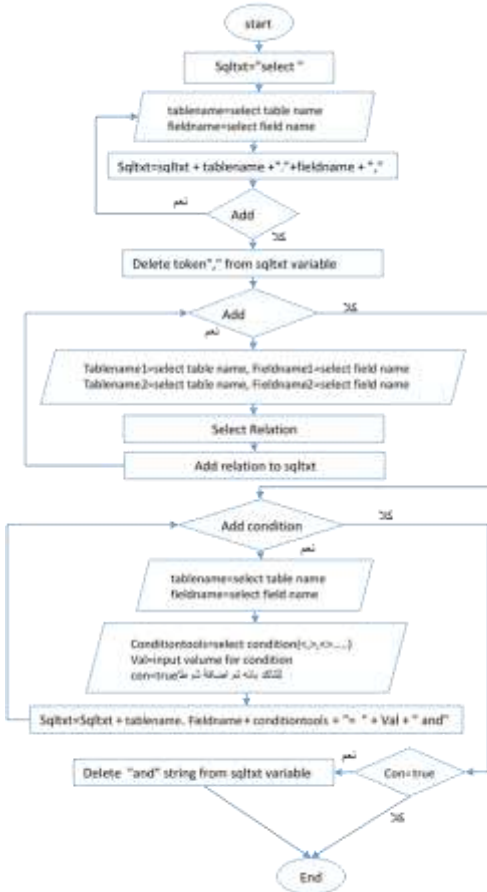
6. بعد عملية التحويل بيانات المصدر الى نفس النوع البياني لحقول جدول قاعدة بيانات الهدف وظهرها في (DataGridView) في القسم الرابع يتم الضغط على زر (نفيذ عملية نقل البيانات الى جدول الهدف) ليتم عملية نقل البيانات وتظهر لنا نافذة تخبرنا بعدد السجلات التي جرى نقلها وعدد السجلات التي لم يجري نقلها بصورة صحيحة.

## 2.4- بناء شفرة SQL

ان كتابة شفرة Sql تحتاج الى مبرمج ذات خبرة كافية في تلك اللغة، وفر البرنامج آلية سهلة لبناء شفرة الاستعلام التي تعمل على إرجاع البيانات وعرضها حسب شرط معين والمخطط التالي يوضح آلية بناء الشفرة.

### 2.4.1- مراحل تنفيذ بناء الشفرة في النظام

عند الضغط على زر بناء عبارة sql تظهر نافذة من خلالها يجري اضافة اضافة الحقول الى الشفرة بعدها الضغط على زر اصف الحقول الى الشفرة. وعند الانتهاء من اضافة الحقول يتم الضغط على زر الانتهاء من اضافة الحقول كما في الشكل رقم (10) وبعدها تظهر مجموعة من الادوات لاضافة شرط على استرجاع البيانات من الجداول كما في الشكل رقم (11)



مخطط رقم (3) آلية بناء شفرة SQL



شكل رقم (10) تحديد الحقول في نافذة بناء شفرة sql شكل رقم (11) تحديد شرط على الحقول في نافذة بناء شفرة sql

### 3- الاستنتاجات

التطور في تقنيات المعلومات والاتصالات وظهور العديد من أنظمة إدارة قواعد البيانات المتجانسة وغير المتجانسة الحديثة التي تدعم الكثير من الوظائف، إن العديد من المؤسسات اتجهت نحو استخدام تلك الأنظمة الحديثة لكن واجهت مجموعة من المشاكل منها كيفية الاستفادة من البيانات المخزونة في أنظمة قواعد البيانات القديمة في قواعد البيانات الحديثة بدون محاولة إعادة إدخال البيانات يدوياً أو نقل البيانات من قاعدة بيانات إلى أخرى تختلف في الهيكل البنائي لجداولها، ولإنجاز عملية نقل البيانات بين أنظمة قواعد البيانات المختلفة يجب توفير برنامج سهل التعامل يقوم بعملية نقل البيانات بين قواعد البيانات المختلفة ويحل مشكلة غير التجانس في البيانات المتواجدة في قواعد البيانات المختلفة، حيث يتم استخدام ذلك البرنامج وذلك للحفاظ على الوقت والجهد المبذول من عملية إعادة الإدخال للبيانات والتخلص من الأخطاء الناتجة من الإدخال اليدوي، حيث يعتبر البرنامج التطبيقي كمرر لنقل البيانات بين قواعد البيانات المختلفة.

تم تطبيق البرنامج التطبيقي على قواعد بيانات افتراضية وتم بنائها باستخدام أنظمة إدارة قواعد البيانات (SQL Server, Oracle, SQL compact, Access)، ويحل النظام مشكلة نقل البيانات بين قواعد البيانات التي تختلف في نظام الإدارة والنوع البياني للحقول فضلاً عن امتياز هذا البحث عن البحوث الأخرى بأنه يعالج مشكلة غير التجانس في تنظيمات هيكل الجداول (اختلاف أسماء الحقول، عدد الحقول في جدول المصدر والهدف)، وتوفير نافذة لربط الجداول واسترجاع البيانات من قاعدة بيانات المصدر ونقلها لقاعدة بيانات الهدف، عرض البيانات مع إمكانية إجراء عملية انتقاء للبيانات (filter) قبل نقلها، وإمكانية التعديل على السجلات عند عرضها قبل نقلها إلى قاعدة بيانات الهدف، فضلاً عن التطرق إلى طريقتين لحل مشكلة غير التجانس في النوع البياني.

- [1]-Database System a practical approach to design, Implementation, and management 4th ed.-Thomas M. Connolly and Carolyn E. Begg - Pearson Education -2005
- [2]- Modern database management - 10th ed-Jeffrey Hoffer and Heikki Topi-Pearson Education -2011
- [3]- Distributed Database Management Systems -Saeed K. Rahimi and Frank S. Haug, Wiley- 2010
- [4]- Data Migration System in Heterogeneous Database - Shashikant Patel, Sagar Wakchaure, Mahendra Pingale and Saba Siraj, - International Journal of Research in Engineering and Technology -2014- Vol.(3) - P 305,307
- [5]- Data Migration System in Heterogeneous Database - Shinde Anita Vitthal, Thite Vaishali Baban, Roshni Warade and Krupali Chaudhari -International Journal of Research in Engineering and Technology - 2013 - Vol.(2),P 88-92
- [6]- بناء نظام لتحويل جداول قاعدة البيانات - عقيل فرحان - أطروحة ماجستير جامعة التكنولوجيا قسم علوم الحاسبات- 2006
- [7]-XML Data Transfer Between Heterogeneous Database System - Qinghong Line - Master's thesis- Concoradia University-Computer Science– 2003