Journal of Theoretical and Applied Information Technology
31st March 2015. Vol.73 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

NEW TECHNIQUE FOR DUPLICATION THE REMOTE DATABASES BASED ON MULTI-AGENTS

[1] **NOREDDINE GHERABI**

Department of Electrical and Computer Engineering
National School of Applied Sciences, Fes, Morocco
E-mail : [1] gherabi@gmail.com

ABSTRACT

Most enterprises use multiple sites to store remote data. Therefore, it is important to centralize the data in real time with efficient systems that duplicate the data to a central site. In this paper, I present a new approach to duplicate the data stored in multiple remote sites using multi-agents. These facilitate the transfer of data and schema of a database automatically and quickly, we exploit the communication between multiple agents to duplicate data and schema. A prototype has been developed using multiple mobile agents to extract the Relational-Object schema and corresponding data from a remote site, and then make a comparison with the schema from the central site to duplicate the data. The experimental results were highly significant, demonstrating that the proposed method is feasible and effective.

.**Keywords:** *Component; Web Semantic, Web Services.*

1. INTRODUCTION AND RELATED WORK

The evolution of the information technology industry requires huge storage of data in multiple locations. The main problem lies in the backup and restoration of data in real time; and most existing systems do not meet the requirement of enterprises to centralize data in a quick and efficient manner. Our work focuses on replication data from multiple sites based on mobile agents.
At this time, the data on the web continues to grow at an exponential rate. However, most of the databases are still managed by the relational systems.
Several techniques have been developed in databases and object-relational databases, namely migration techniques and conversion [1],[2] optimization [3],[4] or may be replication.

Currently, many studies are focused on the discovery of migration and extracting conceptual schema from a logical RDB schema [5], [6][7]. For example our work [8] proposes a method to convert a relational database to an OWL description. Cristian Perez de Laborda et al [9] introduces Web Ontology Language (OWL) for a logical representation of relational data and schema components, which is particularly suitable for exchanging elements among systems database remote. Fonkam et al [10] proposes an algorithm for converting RDB schemas into conceptual models and Cullot, et al [11], have created a tool for automatic migration of databases to ontology.

However, the extraction of a conceptual schema diagram logical RDB has been extensively studied [12], [13]. These conversions are usually specified by rules that describe how to derive constructions RDB (eg, relationships, keys), classify and identify relationships.

The purpose of this article is to demonstrate an effective prototype based on the extraction of database schemas exist in multiple sites. In this technique, I used several agents to distribute the tasks that will be the communication with sites, extracting and duplication schemas, or may be data, in a central location. The techniques for mobile agents are used in several works [14][15][16] and resolved several problems

I developed an agent called super agent to manage access to the central site, and another agent communication to facilitate communication with the sites and the central site. To complete the process, another agent created in each site to extract the schema and the local database, by focusing on the conceptual model of this database.

The remainder of this paper is organized as follows. Section two gives an overview of our methodology in details; this section explains the proposed method for extracting the schema and data of database and describes how to duplicate it.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

After in section three, I present the application of the approach on several cases from real world domains. Finally, Section four concludes the paper.

2. PROPOSED APPROACH

In this section, I present the proposed method for duplicating schemes of multiple sites in a central site (See Fig.1).

The first step is to start the necessary agents to execute administrative tasks, communication tasks and extraction schemes. The super agent begins to analyze the bases found in sites and then creates an agent that facilitates communication with other agent's sites, each site contains a schema of database and each schema is extracted by the corresponding agent then sending the schema into super agent to compare it with the central site

2.1 The process for extracting and duplicating RDB.

2.1.1 Mechanism for extracting RDB

In this section, I present the mechanism for extracting Metadata for multiple sites based on communication between multi-agents.

This mechanism is developed as follows:

Input: The RDB Sites

Output: The schema of each RDB and Dataset

Steps:

Step 1: Create Super Agent and Communication Agent.

Step 2: Start the agents

Step 3: Communication agent accesses the site database and starts a new agent site for extracting the relation database schema.

Step 4: Site Agent sending the result of extraction to another agent of verification, to verify the existence of the schema found in the central site.

Step 5: Super agent receives the relational schema of the site and duplicates it into the central site

End

2.1.2 Agent of communication in JADE

A basic thing of multi-agent systems is that agents communicate and interact with easy and efficient way. The agents are used to exchange messages and understand the message structure, because the

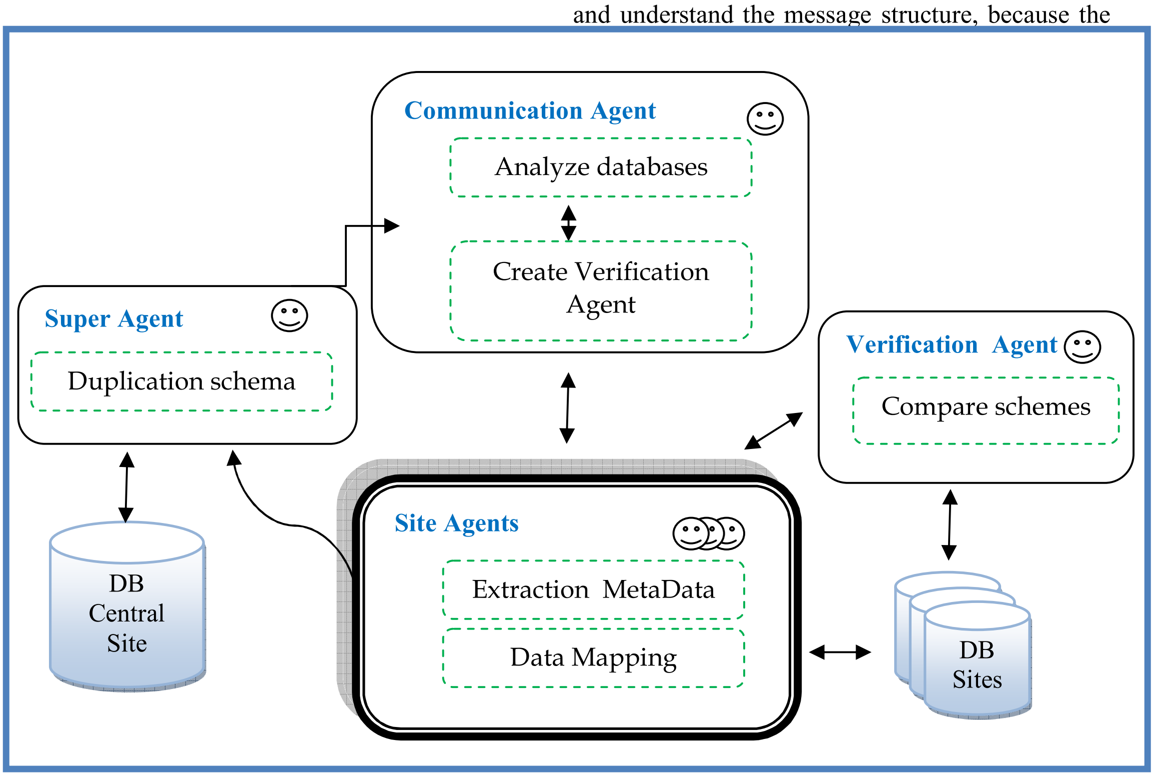

Fig.1: *Architecture To Extract And Duplicate The Schema And Data*

Journal of Theoretical and Applied Information Technology
31st March 2015. Vol.73 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Using JADE [17],[18],[19] in Java to create and communicate mobile agents. JADE complies with FIPA standards and, ideally, Jade agents may interact with agents in other languages and running on other platforms.

A message JADE contains several pieces in addition to the content, for example: receivers, the sender and the message type. In JADE, messages strictly comply with the ACL (Agent Communication Language) that provides several options for coding the actual content. Specifically, Jade supports the FIPA SL (Semantic Language), actions and predicates. There is a possibility to serialize Java objects.

If the sender or the recipient does not belong to the same platform, the ACL message is automatically converted to FIPA compliant string format. Therefore, this conversion is not visible to the performers agents who need to deal with the same class of Java object;

In this work, the message content is in the form of serialized objects, because the content of the message can be retrieved from objects of an object-oriented database.

2.1.3 Extracting MetaData of RDB (MTR).

The process started by extracting the basic Metadata information about the RDB in each site, including relations and fields properties. The technique of extraction schemas is developed in our work [1]

In my method, each RDB schema is represented as a set of elements (Table name (T_N), set of fields (T_F), Primary Keys (T_P), Foreign Keys (T_F) and Unique Keys (T_U))

$$MTR=\{T/\,T:=T_N, T_{Fl}, T_P, T_F, T_U\}$$

- T_N is the name of the table and T_F describes the set of fields of each table:

$$T_F=\{F|\,F:=F_N, F_T, F_L, F_{Nl}, F_D\}$$

Where **F** is the field of the table T, $\mathbf{F_N}$ is the name of F, $\mathbf{F_T}$ its type, $\mathbf{F_L}$ is the data length of the field F, $\mathbf{F_{Nl}}$ is nullable or not and $\mathbf{F_D}$ denotes the default value.

- T_P is a primary key of the table (single key or composite keys)

- T_F denotes the set of foreign key(s) of each table T.

$T_F(T)$ = {FKn,T_P(T')}, where FKn represents foreign key field name and T_{PK}(T') name of an exporting (i.e., referenced) the second relation T' that contains the referenced T_K.

- Relationships (RS): A table T has a set of relationships RS.

  Each relationship (rel ϵ RS) between a table T and another table T' is defined as:

  RS(T,T') := {rel | rel := (T_P(T), T, T_F (T'),T',Ca)}

  Where T_P(T) is the primary key of T, T_F (T') is the foreign key representing the relationship in T' and Ca the cardinality of the source table T

In Database, the DatabaseMetaData interface is used for retrieving the structure of the database. Figure 2 shows an overview of some instructions for extraction Metadata.

MTRDB: getMetaData		
F_N: COLUMN_NAM	F_T	TYPE_NAME
	F_L	COLUMN_SIZE
	F_{Nl}	IS_NULLABLE
	F_D	COLUMN_DEF
T_P	getPrimaryKeys	
T_F	getImportedKeys	
RS	T_P(T)	PKCOLUMN_NAME
	T_F(T)	FKCOLUMN_NAME

Fig.2: Instructions For Extraction The MTR Schema

2.1.4 Algorithm for extraction of MTR using Multi-Agents

This section presents the algorithm for extracting Metadata of Database; this algorithm is used to extract all information about Metadata of RDB, which contains the names of the tables, columns and integrity constraints of all the tables extracted from an RDB.

The algorithm for extraction the MTR from RDB is as follows:

```
Algorithm MTR (BD: RDB) return MTR
Start Agent Site: AgentS;
MTR: = null; // a set to store RDB
For each table T ϵ RDB do
  TN := Extract name of ( T)
For each table TF ϵ T do
  FN := AgentS→ExtractClumn(T)
```

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

F_T := AgentS→ExtractColumnType(T)

F_L := AgentS→Extractlengthofthecolumn (T)

End For

T_P := AgentS→ExtractPrimaryKeys (T)

T_F := AgentS→ExtractForeignKeys(T)

T_U := AgentS→ExtractUniqueKeys (T)

End For

For each set of tables (T, T'); create element for storing the prosperities of the relationships between T and T'.

T_F (T'):= AgentS→ExtractForeignKey(T')

T':= AgentS→ExtractTable (T')

C:= AgentS→ExtractCardinalitySource(T)

RS:= { T_P ,T, T_F (T'),T',C}

End For

MTR: = {T, F_N,F_T,F_L,T_P,T_F,T_U, RS}

AgentS send the MTR schema

End algorithm

2.2 Duplicating schema and data

2.2.1 Process for duplicating.

The duplication process begins by reading the MTR scheme of each site and transforms it into serialized objects, these objects will send to an agent for the verification, the verification agent in its turn examines the MTR scheme and verify its existence in the central site. If the schema exists in the central site, data will be duplicated by replacing the old schema of the central site; else, the central agent creates the new schema.

The following diagram, in figure 3, summarizes the verification steps and duplication:

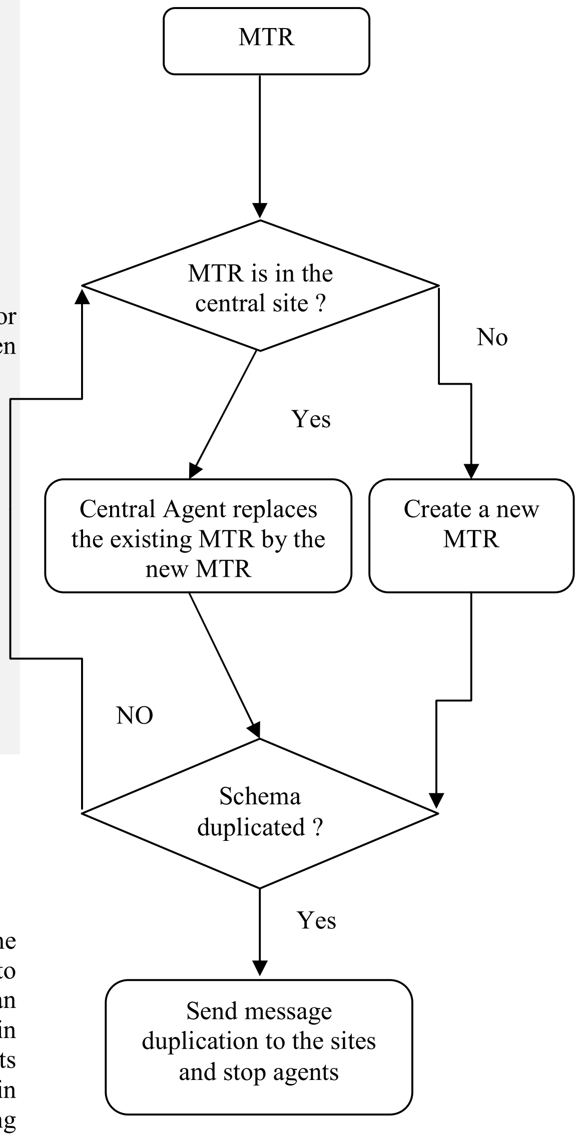

Fig.3: Verification Steps And Duplicating Of MTR

2.2.2 Techniques for duplicating

The proposed process is based on communication between agents such that each agent executes a specified task and sending the results to another agent to complete the process. For a good duplication schemes and data, several techniques of distributed databases can be used to copy schema and data from a database; namely materialized views, synonyms, stored procedures, or may be snapshots.

In this paper, I use the concept of service of a database (bd@) to communicate the remote

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

databases, based on the instructions of DBMS oracle.

The pseudo-code to duplicate schemas and data is as follows:

```
S← Current Site
For each S
 M← Get Metadata of RDB ( MTR)
If ( M exists in a central site )

// Replace the existing MTR in the central site by the MTR of the remote site

  For each T in MTR
    COPY FROM
UserSite/PasswordUserSite@BD-Site
    TO
AdminCentral_Site/PasswordAdminCentral_Site
    @BD-Central-Site
    REPLACE T
        USING Select * from T@BD_Service_Site
…

  End For each

Else If ( M not exists in a central site )

// Create the MTR in the central site

   For each T in MTR
   COPY FROM
UserSite/PasswordUserSite@BD-Site
   TO
AdminCentral_Site/PasswordAdminCentral_Site
   @BD-Central-Site
   CREATE T
       USING Select * from T@BD_Service_Site
…

  End For each

End For each
```

In this algorithm, the COPY command is regularly duplicate data on the central server.
The database is identified by the service named BD_Service_Site to communicate with the remote site. During the connection, a session should be initiated by an account and password.
To run the duplication correctly, the user must have the necessary privileges to open a session and duplicate data

- **CREATE:** If the destination table exists, COPY generates an error. Otherwise, the table is created and the data will be copied (**CREATE + INSERT)**
- **REPLACE**: If the destination table exists, copy deletes and recreates the table with the new data. Otherwise, it creates the table and insert the data **(DROP] + CREATE + INSERT)**

3. TEST AND EVALUATION

The proposed method in this paper is validated by a prototype created using java tools in communication with remote databases stored in Oracle.
For a good test of the prototype, we take a simple example of database shown in Figure 4
The prototype starts sending messages to the sites to start the agents and extract the existing MTR schema at each site by displaying the list of tables and views of each site.

Journal of Theoretical and Applied Information Technology
31st March 2015. Vol.73 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

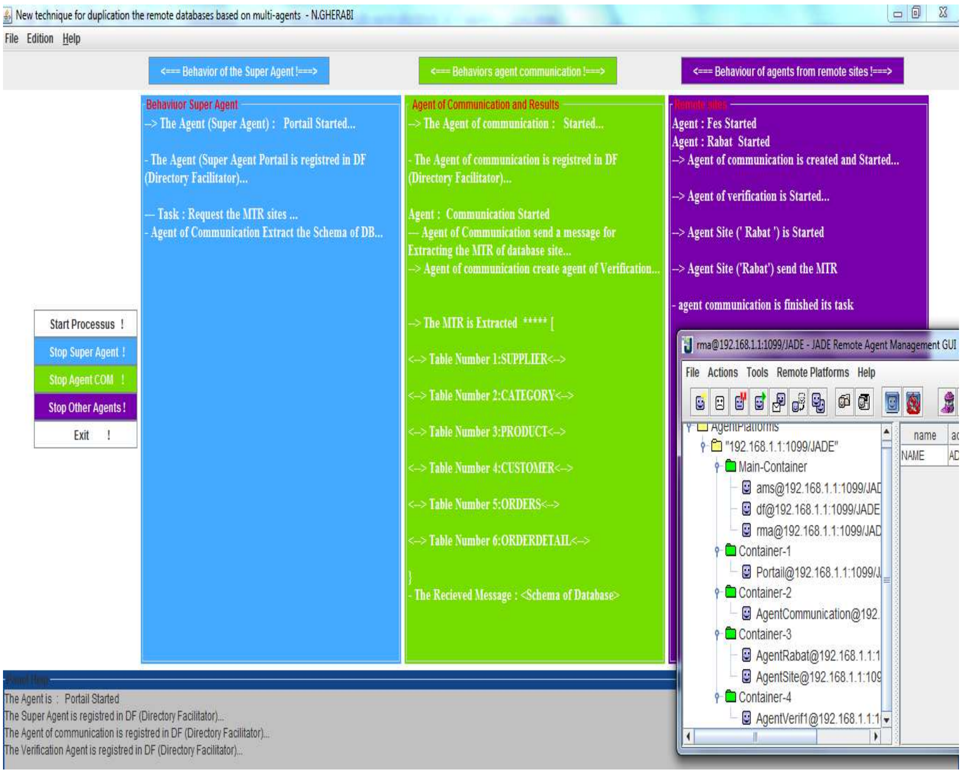

Fig.5: The Result Of Extracting And Duplicating Schema

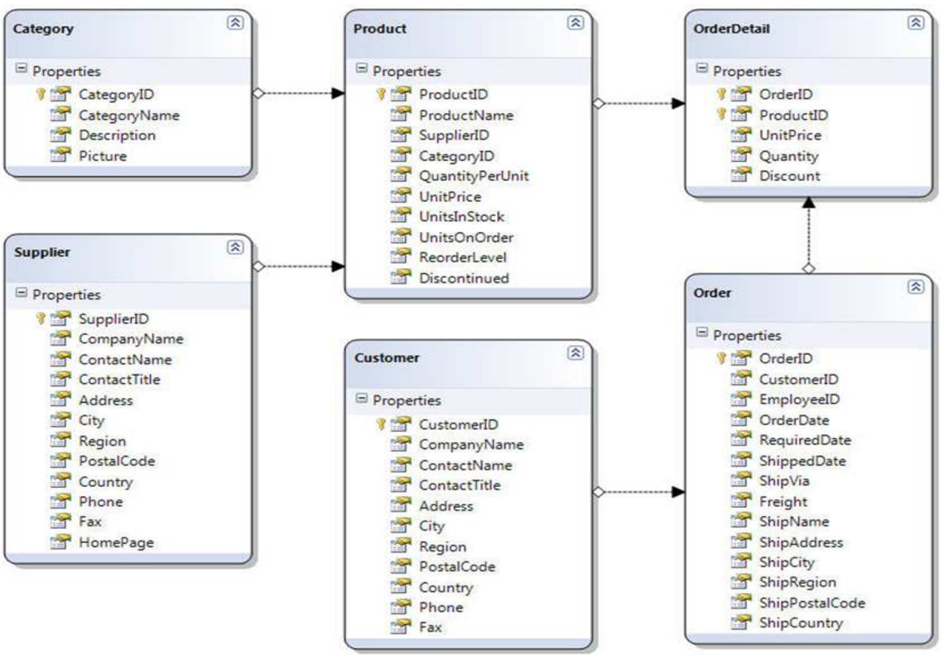

Fig.4: Example Of RDB

Journal of Theoretical and Applied Information Technology
31st March 2015. Vol.73 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The agents of sites send the MTR schema to the communication agent for verifying the existence of the schema in the central site. After having verified the MTR schema, the super agent duplicates the schema and shows the result of duplication

Figure 5 shows an overview of MTR generated by the system. The system displays the list of exported tables and also displays the behavior of agents during treatment.

For better appreciation of my approach, we calculate the execution time with the use of agents and without agents. The table in figure 6 shows the total time calculated to run the tasks of the four agents and the table in figure 7 shows the execution time without agents.

We note that the time of execution using the multi-agent system is reduced compared to a simple system of duplication.

		with Agents (Execution Time in Milliseconds)			
Duplication	Super Agent	Agent Site	Comm Agent	Verif Agent	Total Time
Dupl N°1	93	15	891	31	**1030**
Dupl N°2	63	17	760	16	**856**
Dupl N°3	165	17	781	16	**979**
Dupl N°4	78	16	749	2	**845**

Fig.6: The Total Time (In Milliseconds) To Execute The Tasks Of The Four Agents

	without Agents (Execution Time in Milliseconds)
Duplication	Total Time
Dupl N°1	**3200**
Dupl N°2	**4342**
Dupl N°3	**7540**
Dupl N°4	**5137**

Fig.7: The Time (In Milliseconds) To Execute The Duplication Without Agents

The use of agents will reduce the execution time and the tasks will be distributed, in same time, on agents for a good performance, unlike a single-system tasks that will be in charge to execute all tasks in series. The figure 8 illustrates the performance of the agents in the duplication of data and schema.

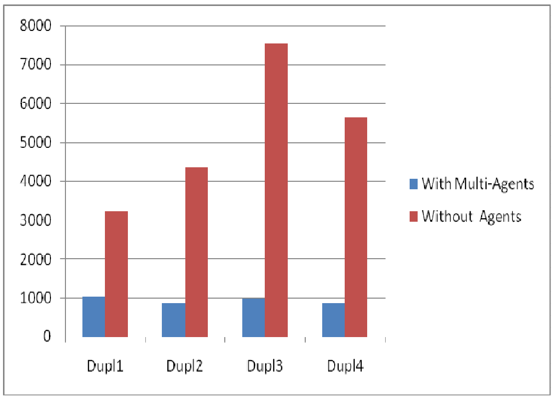

Fig.8: The Difference Between The Duplication Using Multi-Agent And Without Agents.

In my system, the agent communication plays a very important role in the execution of tasks; it is he who facilitates communication between agents. Figure 9 shows the charge distribution for each agent.

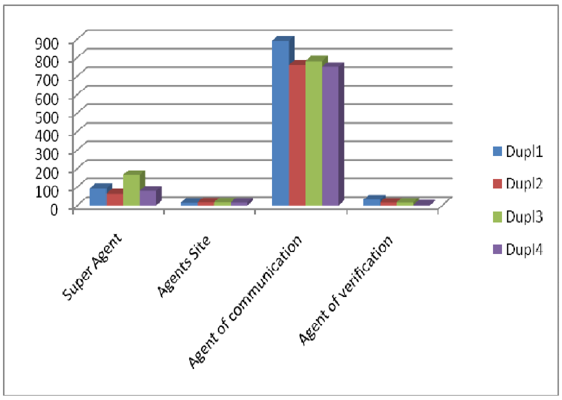

Fig.9: The Distribution Charge Of Each Agent

I believe that this article will be among the best papers to present algorithms that automatically duplicate the schema directly from several sites into a central site, retaining semantic information.

4. CONCLUSION

In this article, I presented an automated method based on a prototype for the extraction and duplication of several remote databases using the techniques of intelligent agents. Several agents have been created to good communication and to distribute the different tasks; namely extraction, verification and duplication.
This method is approved experimentally on real databases existing in remote sites. The results demonstrate that the proposed approach is practical

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and solves problems of duplication and extraction schemas and data in the distributed database.

REFERENCES:

[1] N. Gherabi, M. Bahaj "Discovering new technique for mapping relational database based on semantic web technology", 978-1-4673-1949-2, IEEE (2012)

[2] N. Gherabi, M. Bahaj "Conversion database of the shapes into XML data for shape matching" World of Computer Science and Information Technology Journal **(**WCSIT**)** Vol. 2, No. 1,22-28, 2012

[3] Matthias Jarke, Jurgen Koch "Query Optimization in Database Systems" ACM Computing Surveys (CSUR), Volume 16 Issue 2 (1984)

[4] Philip A. Bernstein, N.Goodman, E.Wong" Query processing in a system for distributed databases" ACM Transactions on Database Systems, Volume 6 Issue 4.

[5] Fahrner, C. and Vossen, G.: Transforming Relational Database Schemas into Object-Oriented Schemas According to ODMG-93.In 4th Int. Conf. on Deductive and Object-Oriented Databases,pp. 429–446, 1995.

[6] Dragut, E., Lawrence, R.: Composing mappings between schemas using a reference ontology. In Proceedings of International Conference on Ontologies, Databases and Applications of Semantics (ODBASE'04). (2004) 783-800

[7] Noreddine Gherabi and Mohamed Bahaj. Robust Representation for Conversion UML Class into XML Document using DOM. *International Journal of Computer Applications* 33(9):22-29, November 2011

[8] N. Gherabi, M. Bahaj**,** K. Addakiri "Mapping relational database into OWL Structure with data semantic preservation" **(IJCSIS)** 2012

[9] Cristian Perez de Laborda and Stefan Conrad. Relational.OWL - A Data and Schema Representation Format Based on OWL. In Second Asia-Pacific Conference on Conceptual Modelling (APCCM2005), volume 43 of CRPIT, pages 89–96, Newcastle, Australia, 2005.

[10] Fonkam, M. M. and Gray, W. A.: An Approach to Eliciting the Semantics of Relational Databases. In 4th Int. Conf. on Advanced Info. Syst. Eng., vol. 593, pp. 463–480, 1992.

[11] Cullot, N., Ghawi, R., Yetongnon, K..,"DB2OWL: A Tool for Automatic Database to Ontology Mapping", In Proc. of 15th Italian Symposium on Advanced Database Systems (SEBD 2007), pages 491-494, Torre Canne, Italy, 17-20 June 2007.

[12] Wu, Z., Chen, H., Wang, H., Wang, Y., Mao, Y., Tang, J., Zhou, C., "Dartgrid: a Semantic Web Toolkit for Integrating Heterogeneous Relational Databases", Semantic Web Challenge at 4th International Semantic Web Conference (ISWC 2006), Athens, USA, 5-9 November 2006.

[13] Alhajj, R.: Extracting the Extended Entity-Relationship Model from a Legacy Relational Database. Inf. Syst, vol. 28, pp. 597–618, 2003.

[14] R. Gupta and G. Kansal, "A Survey on Comparative Study of Mobile Agent Platforms" in International Journal of EngineeringScience and Technology (IJEST). ISSN: 0975-5462 Vol. 3 No. 3 March 2011.

[15] T. Kuster, M. Lutzenberger, A. Hebler, and B. Hirsch, "Integrating Process Modelling into Multi-Agent System Engineering" in Multiagent and Grid Systems, International Journal, pp. 105-124, January 2012

[16] Y. Labrou, T. Finin and Y. Peng,"The Interoperability Problem: Bringing together Mobile Agents and Agent Communication Languages" (IEEE)

[17] F. Bellifemine, A. Poggi and G. Rimassa, "JADE: A FIPA compliant agent framework". Practical Applications of Intelligent Agents and Multi-Agents

[18] Bellifemine, F. & Caire, G. & Greenwood D.: Developing Multi-Agent Systems with JADE, John Wiley & Sons Ltd, England, 145-161 (2007)

[19] Parul Ahuja1, Vivek Sharma "A JADE Implemented Mobile Agent Based Host Platform Security" Computer Engineering and Intelligent System Vol 3, No.7, 2012