

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

(( يَا أَيُّهَا الَّذِينَ آمَنُوا إِذَا قِيلَ لَكُمْ تَفَسَّحُوا فِي الْمَجَالِسِ فَافْسَحُوا  
يَفْسَحِ اللَّهُ لَكُمْ وَإِذَا قِيلَ انشُرُوا فَانشُرُوا يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا  
مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ )) .

صدق الله العظيم ،،

المجادلة - الآية (11).

# *Dedication*

❖ *To my family:*

*My parents,  
Who taught me how  
To find my way in the life*

'''

*My aunt, brother, and sister*

❖ *To all those who I love and respect them very much*

## **ACKNOWLEDGMENTS**

I would like to express my sincere gratitude to my supervising Dr. Awad Mohammed Awadelkarim, for his constant encouragement and valuable guidance throughout my research, where he gave me the advice and assistance without boredom or shows any kind of dissatisfaction. Thank him very much.

I am also grateful to all members of the Department of Management Information System, MIS in the Nile Valley University, especially, Yassir Mousa, and Osama, for their unlimited assistance throughout my study.

I am immensely thankful to Mujahed Hashim, Greeballa Abdulmajeed, and, Gamal Abdelgader; for their great help, and this work could not have been accomplished without their support

My heartfelt and especially thanks go to my family, father, mother, aunt, brother and sister and to my friends, especially, Rasha, Suad, Amira and Mervat.

## **ABSTRACT**

In Multidatabase Systems (MDBS), pre-existing autonomous local database systems that manage by heterogenous Database Management System (DBMS) in distributed environment are integrated together. Consequently, in such surroundings, global query optimization should be considered to achieve the overall system performance. Concurrently, security consideration is enlarged greater than before. Therefore, autonomous nature of such MDBS causes additional challenges in optimizing query process as well as offering adequate security features. Although, the majority of algorithms that have been suggested and used for optimizing query in such context are lacking of security aspects. Thus, this study proposes a security enhancement to query optimization algorithms.

The proposed security enhancement considered access control mechanisms, which maintain the database confidentiality and integrity, as a crucial security aspect that ought to associate with query optimization process. Moreover, an appraisal experiment is conducted using ORACLE-9i in order to confirm the feasibility and robustness of such enhancement. The experimental results shown that the enchantment approach is entirely adequate and promising, particularly in term of estimating local cost parameters. Furthermore, statistical measures confirm that the derived cost formula is significant

## المستخلص

نظم قواعد البيانات المتعددة (Multi Database Systems) هي عبارة عن مجموعة أنظمة قواعد بيانات تتميز بالاستقلالية (Autonomous) عن بعضها البعض حيث تدار هذه الأنظمة بواسطة نظم إدارة قواعد بيانات غير متجانسة (Heterogeneous)، يتم تجميع تلك الأنظمة مع بعضها البعض في البيئات الموزعة على هيئة تجعلها متكاملة معاً . وعليه لابد من اعتبار أمثلية الاستفسارات في مثل ذلك النوع من الانظمة وذلك لتحقيق الاستفادة القصوى منها فيما يتعلق بالأداء العام لها ، وفي ذات الوقت اعتبار بعض الجوانب الأمنية والتي من الممكن أن تتعلق بعملية تحسين الاستفسار، لكن الطبيعة المستقلة لمثل تلك الأنظمة تتسبب في خلق وإيجاد العديد من الصعوبات والتحديات والتي من المحتمل أن تجعل الوصول إلى مثل ذلك الهدف من الصعوبة بمكان وذلك على الرغم من وجود العديد من الخوارزميات المقترحة والمستخدمه في عملية تحسين الاستفسارات، لكننا نجد أن غالبية هذه الخوارزميات لم تأخذ في اعتبارها محاولة توفير بعض الجوانب الأمنية ذات الصلة بتلك العملية.

ولهذا كان السبب الأساسي لهذه الدراسة محاولة تلافى ذلك بعمل التعزيز المطلوب والذي يتعلق باعتبار بعض تلك الجوانب الأمنية المشار إليها ، مثل آلية التحكم بالوصول كأحد جوانب السرية والتي تعمل على الحفاظ على صحة وأمن البيانات. من ناحية أخرى، تم تقييم هذه الدراسة باستخدام أحد نظم إدارة قواعد البيانات وهو (Oracle9i) للتأكد من جدوى وفعالية الجزء المضاف. وقد أظهرت نتائج تلك التجارب أهمية هذا التحسين خاصة فيما يتعلق بعملية تقدير التكلفة المحلية للمعاملات ،كذلك أكدت النتائج الإحصائية أهمية صيغ التكلفة المشتقة للاستفسارات.

## TABLE OF CONTENTS

الآية.....	I
DEDICATION.....	II
ACKNOWLEDGMENTS.....	III
ABSTRACT .....	IV
المستخلص .....	V
TABLE OF CONTENTS .....	VI
LIST OF FIGURES .....	IX
LIST OF TABLES .....	X
LIST OF ABBREVIATIONS .....	XI
CHAPTER 1 – INTRODUCTION .....	
1.1 Preface .....	1
1.2 Problem Definition .....	1
1.3 Objectives .....	2
1.3.1 General Objective .....	2
1.3.2 Specific Objective .....	2
1.4 Importance of the Study .....	2
1.5 Scope of the Study .....	2
1.6 Thesis Organization .....	3
CHAPTER 2 – BACKGROUND AND RELATED WORKS .....	
2.1 Background .....	4
2.1.1 Integration and Multidatabase System .....	4
2.1.2 Query Processing .....	6
2.1.3 Optimization Types .....	6
2.1.4 Global Optimization .....	7
2.1.5 Global Optimization Issues .....	7
2.1.6 Information Security .....	8
	10

2.1.7	Database Security .....	11
2.2	Related Works .....	14
2.3	Research Methodology .....	14
2.4	Research Development Phases .....	
CHAPTER 3 – THE ADOPTED ALGORITHM.....		
3.1	Preface .....	15
3.2	Justifications for Choosing Such Algorithm .....	15
3.3	An Overview of the Algorithm .....	16
3.3.1	Algorithm Description .....	16
3.3.2	Classification of Queries .....	17
3.3.3	Sampling Queries .....	20
3.3.4	Derivation of Cost Estimation Formula .....	21
3.3.5	The Experimental Results .....	31
3.4	Summary .....	34
CHAPTER 4 – SECURITY ENHANCEMENT TO QUERY OPTIMIZATION LGORITHMS		
4.1	Preface .....	35
4.2	Handling and Managing Access Request .....	35
4.2.1	Handling Access Request .....	35
4.2.2	Managing Access Request .....	36
4.3	The Proposed Security Enhancement Approach .....	37
4.3.1	Query Optimizing .....	37
4.3.2	The Specification of Handling Access Request and Optimizing Query .....	43
CHAPTER 5– THE IMPLEMENTATION AND EVALUATION OF THE PROPOSED SECURITY ENHANCEMENT.....		
5.1	Preface .....	47
5.2	Experimental Environment Details .....	7
5.3	Experimental Results .....	49
5.3.1	The Enhancement Algorithm Evaluation .....	49
5.3.2	Feasibility of Handling Access Request Strategy .....	52

CHAPTER 6– CONCLUSION AND RECOMMENDATIONS .....	
6.1 Conclusion .....	54
6.2 Suggestions Future Works .....	55
REFERENCES .....	56
APPENDIX	
APPENDIX A: Drawn Sample Queries of Class (G <sub>11</sub> ) .....	i
APPENDIX B: Drawn Sample Queries of Class (G <sub>12</sub> ) .....	iv
APPENDIX C: Drawn Sample Queries of Class (G <sub>13</sub> ) .....	vii
APPENDIX D: Drawn Sample Queries of Class (G <sub>14</sub> ) .....	x
APPENDIX E: Drawn Sample Queries of Class (G <sub>15</sub> ) .....	xiii
APPENDIX F: Drawn Sample Queries of Class (G <sub>21</sub> ) .....	xvi
APPENDIX G: Drawn Sample Queries of Class (G <sub>22</sub> ) .....	xx
APPENDIX H: Drawn Sample Queries of Class (G <sub>23</sub> ) .....	xxiii
APPENDIX I: Drawn Sample Queries of Class (G <sub>24</sub> ) .....	xxvi
APPENDIX J: Various Explanatory Variables and Queries Observed Cost of class (G <sub>11</sub> ) .....	xxix
APPENDIX K: Various Explanatory Variables and Queries Observed Cost of class (G <sub>12</sub> ) .....	xxx
APPENDIX L: Various Explanatory Variables and Queries Observed Cost of class (G <sub>13</sub> ) .....	xxxii
APPENDIX M: Various Explanatory Variables and Queries Observed Cost of class (G <sub>14</sub> ) .....	xxxiii
APPENDIX N: Various Explanatory Variables and Queries Observed Cost of class (G <sub>15</sub> ) .....	xxxiv
APPENDIX O: Various Explanatory Variables and Queries Observed Cost of class (G <sub>21</sub> ) .....	xxxvi
APPENDIX P: Various Explanatory Variables and Queries Observed Cost of class (G <sub>22</sub> ) .....	xxxviii
APPENDIX Q: Various Explanatory Variables and Queries Observed Cost of class (G <sub>23</sub> ) .....	xxxviii
APPENDIX R: Various Explanatory Variables and Queries Observed Cost of class (G <sub>24</sub> ) .....	xl
APPENDIX 1: Setting up Data Label- Label Component-Policy1	xlii
APPENDIX 2: Setting up Data Label of Policy1	xlii
APPENDIX 3: Setting up Use's Label of Policy1	Xliii
APPENDIX 4: Detail of Use's Label of Policy 1	xliii
APPENDIX 5: Data Retrieving of Policy 1	xliv
APPENDIX 6: Setting up Access Control Options of Policy1	xliv
APPENDIX 7: Setting Up Row Label Security Privileges of Policy1	xlvi
APPENDIX 8: Message Error Shows Violation of Restrictions on Policy1	xlvi
APPENDIX 9: Setting up Data Label- Label Component on Policy2	xlvi
APPENDIX 10: Setting up Data Label of Policy2	xlvi
APPENDIX 11: Detail of Use's Label of Policy 2	xlvi
APPENDIX 12: Setting up Access Control Options of Policy2	xlvi
APPENDIX 13: Result of Read Request of Policy 2	xlvi
APPENDIX 14:Detail of Use's Label of Policy 3	xlvi
APPENDIX 15: Result of Read Request of Policy 3	xlvi



## LIST OF FIGURES

<b>Figure</b>	<b>Page</b>
Figure 2.1: The CIA Triad .....	8
Figure 3.1: The Idea of Query Sampling Method .....	17
Figure 3.2: Selection of Variables for Regression Model .....	29
Figure 3.3: The Structure of Query Sampling Method Algorithm .....	33
Figure 4.1: The Sequence of the Proposed Algorithm .....	38
Figure 4.2-a: The Sequence of Handling and Managing Access Requests .....	46
Figure 4.2-b: The Sequence of Handling and Managing Access Requests.....	47

## LIST OF TABLES

<b>Table</b>	<b>Page</b>
Table 3.1 Some Common Rules for Unary Queries.....	18
Table 3.2: Some Common Rules for Join Queries.....	19
Table 3.3: Representative Query Forms for Common Unary Query Classes.....	23
Table 3.4: Representative Query Forms for Common Join Query Classes.....	23
Table 5.1: Proposed Tables in Experimental Databases .....	49
Table 5.2: Cardinalities of the Tables in Experimental Database.....	49
Table 5.3: Derived Cost Formula for Query Classes on ORACLE 9i for the Original Algorithm.....	51
Table 5.4: Statistical Measures for Local Cost Formula on ORACLE 9i for the Original Algorithm .....	51
Table 5.5: Derived Cost Formula for Query Classes on ORACLE 9i for the Enhancement Algorithm .....	52
Table 5.6: Statistical Measures for Local Cost Formula on ORACLE 9i for the Enhancement Algorithm .....	52

## LIST OF ABBREVIATIONS

MDBS	MultiDataBase System
DB	Data Base
$R_i$	A table in Data Base
G	Set of all component (unary and join) query
$R_{i.a_n}$	Denote a column of $R_i$
C	A constant in the domain of column $R_{i.a_n}$
$G_{1.1}$	Unary queries whose qualifications have at least one conjunct $R_{i.a_n} = C$ where $R_{i.a_n}$ is clustered- indexed
$G_{1.2}$	Unary queries whose qualifications have at least one conjunct $R_{i.a_n} = C$ where $R_{i.a_n}$ is indexed
$G_{1.3}$	Unary queries whose qualifications have at least one conjunct $R_{i.a_n} \theta C$ where $R_{i.a_n}$ is clustered – indexed
$G_{1.4}$	Unary queries whose qualifications have at least one conjunct $R_{i.a_n} \theta C$ where $R_{i.a_n}$ is indexed
$G_{1.5}$	Unary queries whose qualifications have at least one conjunct $R_{i.a_n} \theta C$ where $R_{i.a_n}$ is indexed
$G_{2.1}$	Join queries whose qualifications have at least one conjunct $R_{i.a_n} = R_{j.a_m}$ where either $R_{i.a_n}$ or $R_{j.a_m}$ is clustered-indexed
$G_{2.2}$	Join queries whose qualifications have at least one conjunct $R_{i.a_n} = R_{j.a_m}$ where either $R_{i.a_n}$ or $R_{j.a_m}$ is indexed
$G_{2.3}$	Join queries whose qualifications have at least one index-usable for at least one operand table
DAC	Dictionary Access Control
MAC	Mandatory Access Control
LBAC	Label Base Access Control
BLP	Bell-Lapadula model
TS	Top-Secret
S	Secret

C	Confidential
U	Unclassified
$V_{UB}$	Basic Explanatory Variables
$V_{JB}$	Secondary Explanatory Variables
$R_u$	The operand table for a unary query;
$R_{j1}$	The first operand table for a join query
$R_{j2}$	The second operand table for a join query
$N_u$	Cardinality of unary operand table
$N_{J1}$	Cardinality of the first join operand table
$N_{J2}$	Cardinality of the second join operand table
$L_u$	The tuple length of $R_u$
$L_{J1}$	The tuple length of $R_{j1}$
$L_{J2}$	The tuple length of $R_{j2}$
$RL_u$	The tuple length of the result table for a unary query
$RL_J$	The tuple length of the result table for a join query
$TN_U$	The cardinality of the intermediate table for a unary query
$RN_U$	The cardinality of the result table for the unary query
$RN_J$	The cardinality of the result table for a join query
$TN_{J1}$	The size of the intermediate table obtained by performing the conjunction of all separable conjunctive terms on $R_{J1}$
$TN_{J2}$	The size of the intermediate table obtained by performing the conjunction of all separable conjunctive terms on $R_{J2}$
$TN_{J12}$	The size of the cartesian product of the intermediate table
$Z_U$	Real physical size of the operand table of a unary table.
$RZ_U$	Real physical size of the result table of a unary query
$Z_{J1}$	The physical size of the first operand table of join query
$Z_{J2}$	The physical size of the second operand of join query
$RZ_J$	The physical size of the result table of join query
Y	Query cost