

**MODELLING NETWORK MARKETING SEMI-STRUCTURED DATA USING
GRAPH DATABASE MANAGEMENT SYSTEM**

By

Teddy Aryono
2-1551-005

MASTER'S DEGREE
in

INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING & INFORMATION TECHNOLOGY



SWISS GERMAN UNIVERSITY

SWISS GERMAN UNIVERSITY
EduTown BSD City
Tangerang 15339
Indonesia

March 2017

Revision after the Thesis Defense on February 16, 2017

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in thesis.

Teddy Aryono

Student

Date

Approved By:

Dr. Maulahikmah Galinium, S. Kom, M. Sc.

Thesis Advisor

Date

Asep Sugih Suntana, Ph.D.

Thesis Co-Advisor

Date

Dr. Ir. Gembong Baskoro, M. Sc.

Dean

Date

Teddy Aryono

ABSTRACT

**MODELLING NETWORK MARKETING SEMI STRUCTURED DATA USING
GRAPH DATABASE MANAGEMENT SYSTEM**

By

Teddy Aryono, S. Kom
Dr. Maulahikmah Galinium, S. Kom, M. Sc., Advisor
Asep Sugih Suntana, Ph.D, Co-Advisor

SWISS GERMAN UNIVERSITY

Network Marketing Organizations (NMOs) are retailing selling channels that use independent distributors, to buy and resell product at retail and to recruit new distributors. When they recruit other distributors, it becomes their downlines and they earn referral fees. Compensation calculation is done by a system to serve real time information. Due to the network marketing model that require a heavily related data and shape a giant tree, the system needs to provide an efficient method to store the data and do calculations among them.

This research simplifies a storing method for semi-structured data in network marketing by modelling a data structure of network marketing compensation schema into graph to be implemented and stored in Graph Database Management System. Graph database is a technology to store data with graph structure where each data stored in it is formed a node. Graph database is excellent at managing highly connected data and complex query.

Performance measurement is measured from speed to run a query to retrieve data, how much storage occupied for particular amount of data, and how CPU and memory are utilized. Speed measurement is accomplished by running a query that is similar to the RDBMS machine experiment and experiment GDBMS machine. Storage occupation measurement is done by importing data to both engine and measure the physical data using Operating System (OS) utility tools. CPU and memory utilization is done by utilizing top command of UNIX while running the query.

To run 61 levels of graph traversal, Neo4j is 1425.93 times faster than MySQL. However, the storage occupied by MySQL is 47.82% more efficient than Neo4j at storing 100,000 records/nodes. In term of CPU and memory usage, to run 61 levels of traversal, Neo4j needs 56.72% less CPU and 55.86% less memory to be consumed than MySQL.

Keywords: network marketing, graph database, relational database, Neo4j, Cypher Query Language, performance measurement.



DEDICATION

I dedicate this work for my family, especially for my parents. I also dedicate this work to all my colleagues, my business partners and all staff in PT. KOTAK PRATAMA SOLUSINDO.



ACKNOWLEDGEMENT

I wish to thank my Almighty God, Jesus Christ. With all of His Grace and His Favour, I can finish this work.

I wish to thank my family, especially my parents who have been supporting me through all of this work.

I wish to thank Dr. Maulahikmah Galinium and Asep Sugih Suntana, Ph.D, as the advisor and co-advisor for this thesis.

I wish to thank Dr. Mohammad Soetomo, as the head of SGU MIT, who helped me a lot during the thesis work.

I wish to thank all of my colleagues and business partners who have been supporting me to finish this work.

And I wish to thank all of other parties who helped me to finish this thesis work.



SWISS GERMAN UNIVERSITY

The logo for Swiss German University (SGU) features the letters 'SGU' in a large, stylized font. The 'S' is light blue, the 'G' is a darker blue, and the 'U' is light blue. A registered trademark symbol (®) is located to the right of the 'U'. Below the logo, the words 'SWISS GERMAN UNIVERSITY' are written in a bold, blue, sans-serif font.

TABLE OF CONTENTS

STATEMENT BY THE AUTHOR.....	2
ABSTRACT.....	3
DEDICATION.....	5
ACKNOWLEDGEMENT.....	6
TABLE OF CONTENTS.....	7
LIST OF FIGURES.....	9
LIST OF TABLES.....	12
CHAPTER 1 – INTRODUCTION.....	13
1.1. Background.....	13
1.2. Research Problem.....	14
1.3. Research Objectives.....	15
1.4. Research Limitations.....	15
1.5. Research Questions.....	16
1.6. Significance of Study.....	16
1.7. Hypothesis.....	16
CHAPTER 2 – LITERATURE REVIEW.....	17
2.1. Network Marketing and Network Marketing Organizations (NMOs).....	17
2.2. Graph Theory and Graph Database.....	18
2.3. Neo4j Graph Database.....	19
2.4. Related Works.....	19
CHAPTER 3 – METHODOLOGY.....	21
3.1. Data Model Analysis.....	21
3.2. Experimental Considerations.....	23

3.3. Schema Preparation and Understanding.....	23
3.4. Modelling RDBMS Schema to GDBMS	25
3.5. Data Generation for RDBMS and GDBMS	28
3.6. Model Validity	30
3.7. Sensitivity Test.....	30
CHAPTER 4 – RESULT AND DISCUSSION	32
4.1. Performance Measurement in RDBMS and GDBMS.....	32
4.2. Calculate Time Usage of RDBMS and GDBMS	37
4.3. Calculate Storage Usage of RDBMS and GDBMS	43
4.4. Calculate CPU and Memory Usage of RDBMS and GDBMS	47
4.5. Stakeholder Interview.....	52
CHAPTER 5 – CONCLUSION AND RECOMMENDATION	53
5.1. Conclusion.....	53
5.2. Recommendation.....	54
REFERENCES	55
GLOSSARY	57
APPENDIX.....	59
CURRICULUM VITAE.....	63