A QUERY-DRIVEN SPATIAL DATA WAREHOUSE CONCEPTUAL SCHEMA FOR DISASTER MANAGEMENT

SAFIZA SUHANA BINTI KAMAL BAHARIN

DOCTOR OF PHILOSOPHY

2016

Faculty of Information and Communication Technology

A QUERY-DRIVEN SPATIAL DATA WAREHOUSE CONCEPTUAL SCHEMA FOR DISASTER MANAGEMENT

Safiza Suhana Binti Kamal Baharin

Doctor of Philosophy

2016

A QUERY-DRIVEN SPATIAL DATA WAREHOUSE CONCEPTUAL SCHEMA FOR DISASTER MANAGEMENT

SAFIZA SUHANA BINTI KAMAL BAHARIN

A thesis submitted
in fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2016

DECLARATION

I declare that this thesis entitle "A Query-Driven Spatial Data Warehouse Conceptual Schema for Disaster Management" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : ……

Name : Safiza Suhana Binti Kamal Baharin

Date : ……

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Doctor of Philosophy.

Signature : ……

Supervisor Name : Prof. Dr. Nanna Suryana Herman

Date : ……

DEDICATION

In the name of Allah, The Most Beneficient, The Most Merciful. May the best prayers, and the most abundant peace, be upon our exemplar The Prophet Muhammad ﷺ and upon his family and companions.

This thesis is dedicated:

To Allah ﷻ, for the guidance, remove all obstacles from and lighted my path, inspired me, and always there in every moment. Thank You Allah!

To both my parents: *Mak & Abah*, for their endless love, support and encouragement.

To my dearest husband, a partner in life: *Ahmad Kamal* who leads me through the valley of darkness with light of hope and support,

To my beloved kids whom I can't force myself to stop loving: *Muhammad AmmarFaris, Adriana Camellia and Aryssa Iris.*

And to all my family and friends.

ABSTRACT

Malaysia has experienced various types of disasters. Such events cause billions of USD and posing great challenges to a nation's government to provide better disaster management. Indeed, disaster management is an important global problem. The National Security Council's (NSC) Directive No. 20 outlines Malaysia's policy on disaster and relief management demonstrates government efforts and initiatives to efficiently respond to disasters. In this regard, decision making is a key factor for organizational success. Positive outcomes are dependent on available data that can be manipulated to provide information to the decision maker, who faces the difficult and complex task of anticipating upcoming events and analyzing multiple parameters. Disaster management involves multiple sources for data collection at various levels as well as a wide array of stakeholders. Hence, accessibility to heterogenous spatial data is challenging. It is crucial to address this problem in terms of data distribution, query operation, and the analyzation task because each resource, level, and stakeholder involved has personal preferences with regard to its format, structure, syntax, and schema.The main purpose of this research is to *support the complex decision-making process during disaster management by enriching the body of knowledge on spatial data warehousing, particularly for conceptual schema design*. A major research problem identified are the heterogeneity of a spatial resource data model, the most appropriate approach to schema design, and the level to which the schema is dependent on the given tools. These problems must be addressed as they are main roadblocks to the process of accessing and retrieving information. The existence of heterogeneous data sources and restricted accessibility to relevant information during a disaster causes several issues with spatial data warehouse design. It can be classified into three considerations namely, the need for guidelines and formalism, schema generation model and a schema design framework and finally, a generalized schema. Four strategies have been designed to address the aforementioned problems: identifying relevant requirements, creating a conceptual design framework, deriving an appropriate schema, and refining the proposed method. User queries are prioritized in the conceptual design framework. Outputs from the formalization process are used with a schema algorithm to effectively derive a generalized schema. The conceptual model framework is taken to be representative of a potential application/ system that has been developed to design a conceptual schema using the problematic heterogeneous data and a restricted approach concerning any corresponding query formalisms. In the schema derivation phase, the conceptual schema that was produced by implementing the proposed framework is presented along with the final conceptual schema. This design is then incorporated into a tool to run an experiment demonstrating that queries from a heterogeneous context are capable of performing context-appropriate conceptual schema design in generic way. Such results outshine the capabilities of a restricted design approach and could potentially answer any relevant queries in less time.

ABSTRAK

Malaysia telah mengalami pelbagai bencana. Peristiwa tersebut menuntut berbilion USD dan memberikan cabaran kepada kerajaan untuk menyediakan pengurusan bencana yang lebih baik. Pengurusan Bencana adalah masalah global yang penting. Majlis Keselamatan Negara (MKN), Arahan No. 20 menggariskan dasar mengenai Bencana dan Pengurusan Bantuan sebagai satu usaha dan inisiatif kerajaan untuk menguruskan bencana dengan cara yang berkesan. Dengan itu, membuat keputusan adalah faktor utama bagi kejayaan sesebuah organisasi. Kejayaan organisasi bergantung kepada ketersediaan data untuk menghasilkan maklumat berguna kepada pembuat keputusan yang menghadapi tugas sukar dan kompleks untuk menjangka perisitwa akan datang dan menganalisis dengan pelbagai parameter. Pengurusan bencana melibatkan pelbagai sumber data di pelbagai peringkat serta pihak berkepentingan. Dengan itu, proses capaian kepada kepelbagaian sumber data spatial menjadi suatu yang kompleks dan mencabar. Ia penting untuk ditangani dari segi pengedaran data, operasi pertanyaan dan tugas menganalisis kerana setiap sumber, peringkat dan pihak berkepentingan yang terlibat mempunyai pilihan mereka sendiri termasuk format, struktur, sintaksis dan skema.Tujuan utama kajian ini adalah untuk menyokong proses membuat keputusan yang kompleks semasa bencana melalui pengayaan ilmu pengetahuan berkaitan gudang data spatial terutamanya dalam reka bentuk skema konseptual. Masalah utama penyelidikan yang dikenalpasti adalah kepelbagaian model sumber data spatial, pendekatan reka bentuk skema yang paling sesuai, dan alat-bantuan skema. Masalah tersebut perlu ditangani segera kerana mereka adalah punca utama kesukaran dan sekatan dalam mengakses dan mendapatkan maklumat. Kewujudan sumber data yang perbagai dan akses yang terhad dalam proses membuat keputusan semasa bencana disebabkan oleh beberapa isu yang berkaitan dengan reka bentuk gudang data spatial yang diklasifikasikan kepada tiga aspek keperluan iaitu, keperluan untuk menpunyai garis panduan dan maklumat formalisasi keperluan, model penjanaan skema dan Rangka Kerja Rekabentuk Skema serta aspek terakhir adalah keperluan Skema Teritlak (umum). Empat strategi telah direka untuk masalah yang dikenalpasti iaitu keperluan, rangka kerja reka bentuk konsep, menerbitkan skema yang sesuai dan penghalusan skema. Dalam rangkakerja rekabentuk konsep, pertanyaan pengguna diberi keutamaan paling atas. Rangka kerja model konseptual di ambil untuk dijadikan sebagai wakil aplikasi/sistem bagi merekabentuk skema konseptual dengan menggunakan masalah kepelbagaian data dan pendekatan yang terhalang kedalam proses formalisasi pertanyaan. Dalam terbitan skema, skema konseptual yang dihasilkan oleh rangka kerja yang dicadangkan telah dilaksanakan. Rekabentuk ini kemudian bekerjasama dengan satu alat untuk melarikan eksperimen yang mana kemudiannya menunjukkan pertanyaan yang mempunyai kepelbagaian konteks boleh mempersembahkan rekabentuk skema konseptual berkonteks dan teritlak berbanding pendekatan terhalang, dan juga boleh menjawab pertanyaan dalam masa yang pantas.

ACKNOWLEDGEMENT

First and foremost, all praises to Allah ﷻ the Lord of the worlds for His help and bless in the whole process of the preparation, compiling and writing of this thesis. I am totally sure that this work has never came truth without His guidance, blessing, strength and help.

Special appreciation goes to my supervisor, Prof. Dr. Nanna Suryana Herman for his supervision and support. Not forgotten, I would like to express my gratitude to my co-supervisor, Dr. Abdul Samad Shibghatullah for his help.

I owe a deep dept of gratitude to the Malaysia Government on the *Skim Latihan Pensyarah IPTA (SLAI)* under Ministry of Higher Education (MOHE) and Universiti Teknikal Malaysia Melaka (UTeM) to the generous financial support and allowed me to pursue the PhD endeavor.

I am grateful to the numbers of people, En. Daeng from JPAM, En. Zul from JPBD, Melaka and to the admin and technical of FTMK staff for all the instances in which their assistance helped me along the way.

Last but not least, my deepest gratitude goes to my family and family in laws for their support they provided me through my entire life and to all my collegues, friends and people who took part in making this thesis real directly or indirectly; without whom this thesis might not be completely written.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

CCC	Contingent Control Center,
CPO	Chief Police Officer
CBQD	Context-Based Query Driven
CBQG	Context-Based Query Graph
CBQT	Context-Based Query Tree
DB	Database
DCC	District Control Centre
DDMRC	District Disaster Management and Relief Committee
DM	Decision Maker
DMRC	Disaster Management and Relief Committee
DOCC	Disaster Operation Controlling Centre
DW	Data Warehouse
ECC	Emergency Command Centre
EMD	Emergency Medical Department
ER	Entity Relationship
ETL	extraction, transformation and load
GIS	Geographical Information System
GKD	geographic knowledge discovery
ICT	Information and Communication Technology

IoT	Internet of Thing
LoD	Level of Detail
MCC	Malaysia Control Centre
NDMRC	National Disaster Management and Relief Committee
NL	Natural Language
NSC	National Security Council
OCPD	Officer in Charge of Police District,
ODS	operational data store
OLAP	On-Line Analytical Processing
OLTP	On-line Transactional Processing
OSC	On-Scene Commander
OSCP	On Scene Control Post
OWB	Oracle Warehouse Builder
R&D	Research and Development
RO	Research Objective
RP	Research Problem
RQ	Research Question
SDB	Spatial Database
SDMRC	State Disaster Management and Relief Committee
SDW	Spatial Data Warehouse
SOLAP	Spatial Online Analytical Processing
SOP	Standard Operating Procedure
SQL	Structured Query Language
UI	User Interface

LIST OF PUBLICATIONS

Safiza Suhana Kamal Baharin and Nanna Suryana Herman (2007), The Importance of Managing Context and Moving Objects Data for Emergency Management System, In:Marizan Sulaiman, Nanna Suryana Herman, et al. (Eds.); Proceedings of the International Conference on Engineering and ICT (ICEI 2007), Melaka, Malaysia

Edward Verbree and Safiza Suhana Binti Kamal Baharin, (2007), Geo-info to-go, In: Proceedings of the 4th International Symposium on LBS & TeleCartography, Hongkong.

Safiza Suhana Kamal Baharin and Sisi Zlatanova, (2008), Managing Moving Objects and Context Data for Emergency Management System, Journal Advanced Manufacturing Technology, Vol (2). No (1), Malaysia.

Zlatanova, S and Safiza Suhana Kamal Baharin (2009), Optimal Navigation of First Responders using DBMS In: Proceedings of the 4th International Symposium on LBS & TeleCartography, Hongkong.

Safiza Suhana Kamal Baharin, Abdul Samad Shibghatullah and Zahriah Othman (2009), Disaster Management in Malaysia: An Application Framework of Integrated Routing Application for Emergency Response Management System, In International Conference of Soft Computing and Pattern Recognition, Socpar09, Melaka, Malaysia

Khairul Anshar, Nanna Suryana Herman, Zahriah Othman, Safiza Suhana Kamal Baharin, (2015) Different Geospatial Data on Hybrid Map Application. Journal of Networks, North America, 10, Aug. 201 5

Suryana, N., Othman, Z., Baharin, S.S.K. The Use of Ontology Merging and Parsing Technique in Handling and Retrieving Large Spatial Data. *International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC)*. (Submitted).

Suryana, N., Othman, Z., Baharin, S.S.K. Large Spatial Data Integration Framework and Semantic Translator Engine in Resolving Semantic and Schematic Heterogeneity. *International Journal of Advanced Engineering Research and Science (IJAERS)*. (Submitted).

CHAPTER 1

INTRODUCTION

1.1 The importance of Disaster Management: An Overview

Disasters are characterized by the scope of emergency they cause. When managing an emergency exceeds a local entity's capabilities, the situation becomes a disaster. According to statistics reported in the 2012 by International Federation of Red Cross and Red Crescent Societies (IFRC) (IFRC World Disasters Report (2012), the continent most often hit by natural disasters was Asia, as shown in Figure 1.1. Disasters may occur anytime and anywhere, thus posing great challenges to a nation's government, which acts as a policy maker in order to provide better disaster management.

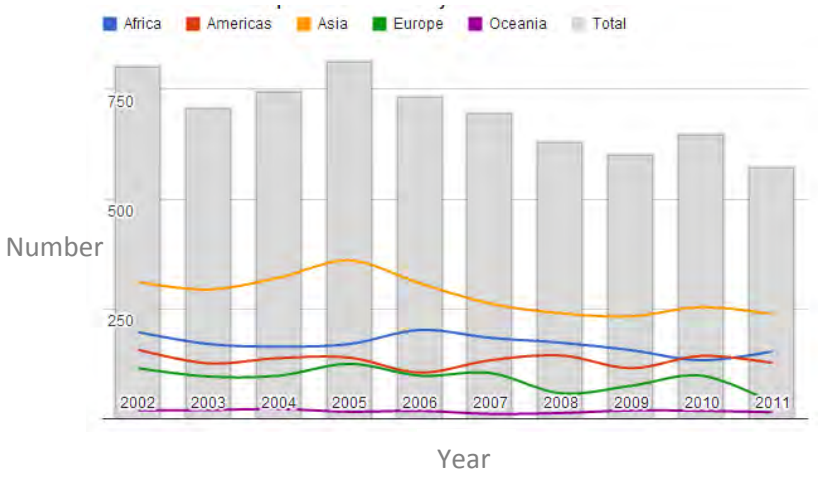

Figure 1.1 Occurrence of reported natural disasters by continent: 2002 – 2011 (*source:* IFRC World Disasters Report, 2012)