MODELING AND QUERYING ALTERNATIVE PATHS IN KUANTAN

MAZLINA BT MOHAMAD SALLEH

A project report submitted in partial fulfillment
of the requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computer Science and Information System
Universiti Teknologi Malaysia

NOVEMBER 2007

To my beloved mom and dad, brothers, sister and friends, thanks for the endless loves and care,
to the most beloved kids, Daneal, Lissa and Azfar , you’re light of my life
and a very special thanks
to my supervisor, Assoc. Prof. Dr.Naomie Salim for the encouragement and support

ACKNOWLEDGEMENTS

First and foremost, I would like to thank ALLAH s.w.t for all the achievements that I have gained today.Here, I would like to express my gratitude to my supervisor, Associate Professor Dr. Naomie Salim for attention, guidance, encouragement and patient through out this length of study. Not forgetting also to other examiners, Dr. Siti Zaiton Mohd Hashim and Dr. Ismail Mat Amin for the comments and guidance towards this project, En Shafry for the ideas and to all the lecturers, please have my heartfelt thanks and sincere gratitude for all your valuable kindness and may Allah reciprocate all your good deeds in the best way. I would also like to extend my deepest appreciation to the staff at Jabatan Ukur dan Pemetaan Pahang, staff at Kuantan Municipal Council (GIS Department), staff at Pahang Tourism Centre and also not forgetting En. Fikri Ismail and En.Firdaus Ali from Politeknik Sultan Haji Ahmad Shah (Polisas) for sharing of information, knowledge, ideas contribution and comments. To families and friends, the endless love will remain forever and only Allah can bless your kindness.

ABSTRACT

Route finding based on geodetic data had addressed the growing data management and analysis needs of spatial applications such as Geographic Information System (GIS). Spatial databases are prominently used in Geographic Information System (GIS) application like digital map application. This study discusses the process of modeling data from map on road network information consisting of points including the starting and ending road points and intersection between other road segments. It enables the storage of spatial dataset or the geographical information syatem (GIS) for calculation of distance between points. Route finding solution takes the distance information in the form of directed graph based on starting (source) and ending (destination) nodes for the desired paths. The graph theory algorithm used in this study is applied from the Floyd's approach and using the combination of searching techniques of breadth first search (BFS) and depth first search (DFS) strategies. The structured query language (SQL) is used for querying the database structures. A browser interface for the system makes the information dissemination easier.

ABSTRAK

Pencarian arah jalan berdasarkan data geografi telah memperlihatkan banyak keperluan analisis berkaitan aplikasi *spatial* seperti Sistem Maklumat Geografi / *Geographic Information System* (GIS). Pangkalan data *spatial* telah banyak digunakan dalam aplikasi GIS seperti aplikasi peta digital. Kajian ini membincangkan proses untuk memodelkan data berkaitan rangkaian jalan daripada maklumat peta yang mengandungi maklumat lokasi atau titik seperti titik permulaan dan pengakhiran jalan serta simpang yang terdapat pada sebuah jalan kepada jalan yang lain, seterusnya membolehkan maklumat tersebut disimpan dalam bentuk set data *spatial* untuk membolehkan aplikasi pengiraan jarak dilakukan antara titik berkaitan yang dikehendaki. Penyelesaian dalam pencarian arah jalan menggunakan maklumat jarak bagi setiap segmen jalan dari satu titik (nod) ke titik (nod) yang lain dlam bentuk graf terarah. Algoritma teori graf yang digunakan adalah berdasarkan pendekatan daripada Floyd serta kombinasi strategi carian mendatar dan mendalam. Kajian ini juga memilih bahasa pertanyaan berstruktur/ *Structured Query Language* (SQL) sebagai struktur pangkalan data. Penggunaan pelayar / *browser* sebagai antaramuka memudahkan penyampaian maklumat dan maklumbalas kepada pengguna.

TABLE OF CONTENTS

CHAPTER	TITLE		PAGE
	DECLARATION		**ii**
	DEDICATION		**iii**
	ACKNOWLEDGEMENTS		**iv**
	ABSTRACT		**v**
	ABSTRAK		**vi**
	TABLE OF CONTENTS		**vii**
	LIST OF TABLES		**x**
	LIST OF FIGURES		**xi**
	LIST OF ABBREVIATIONS		**xii**
1	**OVERVIEW**		**1**
	1.1	Introduction	1
	1.2	Problem Background	2
	1.3	Problem Statement	3
	1.4	Objectives	3
	1.5	Project Scope	3
	1.6	Significance of Study	4
	1.7	Summary	4
2	**LITERATURE REVIEW**		**5**
	2.1	Introduction	5
	2.2	Graph Theory	5

	2.2.1	Graph Characteristics and Models	6
	2.2.2	Road Network Modelling	7
	2.2.3	Graphs in Database	9
2.3	Path Algorithm		9
2.4	Searching Technique Analysis		13
	2.4.1	Breadth-First Search (BFS)	13
	2.4.2	Depth-First Search (DFS)	15
	2.4.3	Best First Search	16
2.5	Structured Query Language		16
2.6	Spatial Database		17
2.7	Digital Map		20
2.8	Kuantan Digital Map		20
2.9	Discussion		20
3	**METHODOLOGY**		**22**
3.1	Introduction		21
3.2	Research Framework		21
3.2.1	Extracting map information		23
3.2.2	Modeling selected data		25
3.2.3	Selecting Suitable Query and Algorithm		29
3.2.4	Result Analysis		32
3.3	Instrumentation		32
3.3	Conclusion		32

4	**RESULTS**		**34**
	4.1	Introduction	34
	4.2	Initial Findings	32
	4.3	Experiments	35
	4.4	Findings	40
	4.5	Conclusion	42
5	**CONCLUSION**		**42**
	REFERENCES		**45**
	APPENDICES		**47**
	Appendix A – List of nodes(places)		
	Appendix B – List of edges (source and destination)		

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Terms in Graph	8
3.1	Sample data from table location	28
3.2	Sample data from road_graph	29
3.3	List of hardware used	32
4.1	Sample of retrieval query in details	36
4.2	Sample of retrieval query in details	39
4.3	Summary of 8 query processing	41

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
2.1	Single and double direction road representation	9
2.2	Breadth-first Search (BFS)	14
2.3	Depth-first Search (DFS)	15
2.4	Example of topological data	19
2.5	Example of data storage for arc	19
2.6	Google facilities in assisting direction findings	20
3.1	Stages in database design	23
3.2	Research Framework	23
3.3	Kuantan Map	24
3.4	Map of Kuantan town	25
3.5	Kuantan town network road	26
3.6	Nodes and arc representation	27
3.7	Sample query from user through browser	31
4.1	Sample of data derived	38

LIST OF ABBREVIATIONS

GIS	Geographical Information System
SQL	Structured Query Language
RDBMS	Relational Database Management System
JUPEM	Jabatan Ukur dan Pemetaan Pahang
BFS	Breadth-First Search
DFS	Depth-First Search

CHAPTER 1

OVERVIEW

1.1 Introduction

Routes became an interesting topic when associated with travelling matters. Routes are generally associated with map, road network model and geographic data. A great number of studies have been discovered in developing the techniques for finding alternatives routes, the ‘best’ routes or the optimal routes within the road network, depending on the weights given. For example, people tend to use the same route to travel from home to office because of the road length or other conveniences but sometimes, they have to use other alternative routes due to occurrences of events such as traffic congestion, construction work or road damage. In pre-trip planning, for example, drives tend to take one path for the outbound trip and then take another path for their return due to any reason such as for different sightseeing or attending other functions at loca'tion that might intersect with the selected return path. Then, it comes the questions, *Which way should be taken?* Or, *how can I get there within that path?* Or, *which routes can be the shortest routes from A to B?*

Finding routes can be implemented by applying a set of algorithm and using a mathematical based graph theory. This study presents the model of road network based on graph structure using the MySQL database. The basic data of the road is represented using nodes as points of locations and edges or arc as path. It consists of source node as starting point and the end node as the destination. This study also embed the use of spatial data such as positioning (x and y) of certain locations. With

the development of geographic information system (GIS) technology, network transportation analysis within a GIS environment has become a common practice in many application areas (Roozbeh *et al.*, 2003). For this query, user has a definite destination in mind and desires to acquire the optimal route leading to the destination. The queries can vary and can answer few questions such as "*Which is the shortest route from Kuantan Airport to Cherating*" or "*How many routes can be taken from Teruntum Complex to Teluk Chempedak*".

Motivations for this study came from the desire of manipulating spatial GIS data that can be represented by advancement of database technology based on graph theory application. This study describes one such extension, where database technology is used to implement path queries over a graph view of relational data. Partial-path information is pre-computed and stored based on Structured Query Language (SQL). Path querying is implemented using SQL functions, thus enabling the retrieved path tables to be manipulated within SQL queries in the same way as standard relational tables.

1.2 Problem Background

Finding paths or routes has become an interest area, as part of the GIS application or Spatial Database application. Many applications can be applied based on using technique for finding direction such as in area of transportation network and travelling guideline. Most of the application on finding paths for relevant application such as map application focused on finding the best path or the shortest path acquired. The well known web services such as Google or Yahoo provide retrieval path based on the shortest path finding algorithm, comprising the best path according to the minimal distance for required length of path. In the normal situation of route finding, the path from one location to others, for example from location A to location B could produce more than one path, let say 5 paths and from the 5 retrieval paths, one path could be the shortest path in terms of minimal length

distance, another path could be the fastest route in terms of minimal time of driving, another path could be the free flow of traffic but took longer distance from the others and the others might be the path with the good sightseeing of views. Thus, it provides choices for user to choose the paths that suit their needs and requirements throughout the trip as well as to assist road users in decision making for the cost effectiveness of time and distance.

As being mentioned above, the algorithm used in most web services for path finding was based on the minimal distance that considered as ‘best’ distance whereas the consideration should be given for users who would like to know the entire possibilities path between set of location. In recent decades, road network system has become complex and congested that affect the people conveniences, thus this situation has derived this study to find the alternative solution for users in path findings through the geodetic data.

According to the JUPEM (Pahang branch), GIS application on Kuantan digital map for path finding is more directed to give one solution, that based on main road and minimal distance. For example, to find path from Teruntum Complex to Teluk Cempedak, the system compute the ‘best’ path according to the path that has high priority such as main road and has a minimal distance of length whereas, it could be has several paths that can be selected by user. This situation also has derived this study to provide the flexibility of path choices from a set of selected location.

1.3 Problem Statement

This research is to study the feasibility and effectiveness of using the query support in relational database system (RDBMS) for information retrieval on finding the routes or alternative paths between two entities with the adaptations of path

algorithms described in the literature. In this research, a framework for a data modeled, algorithm selection and queries retrieval technique is developed.

1.4 Objectives

1. To build a data model for storing Kuantan map information on road network.

2. To apply graph algorithm for finding alternative paths and their lengths between two locations using Floyd-Warshall approach towards spatial dataset.

1.5 Project Scope

1. The technique of storage and retrieval will be supported by relational database management system (RDBMS) and structured query language (SQL).

2. The data model is based on mathematical graph theory approach.

3. The graph theory algorithm is based on adaptation from Floyd approach that embedded to the SQL schema.

4. The data set and area of study is based on part of Kuantan town map, consisting 24 points of locations and 61 edges of direction.

1.6 Significance of study.

This study is necessary to support the use of mathematics graph theory and algorithm in paths finding using the set of database. This study is hoped to be another contribution in area of data storage application and information retrieval for local set data of geographic items. The use of algorithm in finding alternative paths can be an additional function to the existing application of path finding especially for Kuantan dataset.

1.7 Summary

This report consists of five chapters. This first chapter presents the overview of the project comprising the general introduction of the problem background and the scope of study. The problem background has described the current situation of path retrieval for general and Kuantan map applications, thus explain the drive of this study. Chapter 2 includes the related literature reviews towards the study, chapter 3 present the project methodology and workflow process, chapter 4 analyzes the findings and chapter 5 comprising conclusion and suggestions for future work.