CircularTrip and ArcTrip: Effective Grid Access Methods for Continuous Spatial Queries

by

Muhammad Aamir Cheema

A THESIS SUBMITTED IN FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTERS IN COMPUTER SCIENCE & ENGINEERING RESEARCH
(March 2007)

Supervisor: Dr. Xuemin Lin

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: **Cheema**

First name: **Muhammad** Other name/s: **Aamir**

Abbreviation for degree as given in the University calendar: **Masters by research in Computer Science and Engineering (COMPER2665)**

School: **School of Computer Science and Engineering (CSE)** Faculty: **Engineering**

Title: **CircularTrip and ArcTrip: Effective Grid Access Methods for Continuous Spatial Queries**

Abstract 350 words maximum: (PLEASE TYPE)

A k nearest neighbor query q retrieves k objects that lie closest to the query point q among a given set of objects P. With the availability of inexpensive location aware mobile devices, the continuous monitoring of such queries has gained lot of attention and many methods have been proposed for continuously monitoring the kNNs in highly dynamic environment. Multiple continuous queries require real-time results and both the objects and queries issue frequent location updates. Most popular spatial index, R-tree, is not suitable for continuous monitoring of these queries due to its inefficiency in handling frequent updates. Recently, the interest of database community has been shifting towards using grid-based index for continuous queries due to its simplicity and efficient update handling. For kNN queries, the order in which cells of the grid are accessed is very important. In this research, we present two efficient and effective grid access methods, CircularTrip and ArcTrip, that ensure that the number of cells visited for any continuous kNN query is minimum. Our extensive experimental study demonstrates that CircularTrip-based continuous kNN algorithm outperforms existing approaches in terms of both efficiency and space requirement. Moreover, we show that CircularTrip and ArcTrip can be used for many other variants of nearest neighbor queries like constrained nearest neighbor queries, farthest neighbor queries and (k+m)-NN queries. All the algorithms presented for these queries preserve the properties that they visit minimum number of cells for each query and the space requirement is low. Our proposed techniques are flexible and efficient and can be used to answer any query that is hybrid of above mentioned queries. For example, our algorithms can easily be used to efficiently monitor a (k+m) farthest neighbor query in a constrained region with the flexibility that the spatial conditions that constrain the region can be changed by the user at any time.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

THIS SHEET IS TO BE GLUED TO THE INSIDE FRONT COVER OF THE THESIS

Approval

This thesis entitled

CircularTrip and ArcTrip: Effective Grid Access Methods for Continuous Spatial Queries

by

Muhammad Aamir Cheema

has been approved by the School of Computer Science and Engineering,
The University of New South Wales, Syndey Australia.

Supervisor: Dr. Xuemin Lin

The final copy of this thesis has been examined by the signatory and I find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Signed .....

Date .....

Copyright Statement

Authenticity Statement

Originality Statement

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Muhammad Aamir Cheema

Signed .....

Date .....

Dedication

I dedicate this thesis to my parents who have always been my nearest and reverse nearest neighbors[1] [KM00] and have been so close to me that I found them with me whenever I needed. It is their unconditional love that motivates me to set higher targets. I also dedicate this thesis to my sisters (Shazia Cheema and Nadia Cheema) and brothers (Muhammad Omer Cheema, Muhammad Umair Cheema and Muhammad Ammar Cheema) who are my nearest surrounders[2] [LLL06] and have provided me with a strong love shield that always surrounds me and never lets any sadness enter inside.

[1] In a not-so-academic wording, your nearest neighbor is an object that is closest to you and your reverse nearest neighbor is any object for which you are the closest object.

[2] Your nearest surrounders are the objects that are nearest to you and surround you completely.

Acknowledgements

I am very thankful to my parents for their love, support and encouragement and for being with me on each and every step of my life. I am also very thankful to my eldest brother Muhammad Omer Cheema for being very supporting and motivating.

I wish to express my gratitude to my supervisor Dr. Xuemin Lin for his help and guidance. I feel thankful to him for his insightful advice and suggestions. I am also thankful to Dr. Wei Wang for giving me generous amount of time whenever I needed some help. I am grateful to Yidong Yuan with whom I worked during my research and learnt a lot. I feel myself very lucky that I got very friendly colleagues. In particular, I am thankful to Ying Zhang for being very helping and cooperative, Mahady Hasan for being my closest colleague (it is also true literally, he sits very close to me), Yi Luo especially for arranging group meetings, Wenjie Zhang especially for being the only other colleague who is working on moving objects and Bin Jiang especially for being the only masters student in our research group other than me.

I do not want to miss this opportunity to thank people who made my stay in Australia an enjoyable experience without which I would not have been able to concentrate on my research. First of all I am very thankful to Azeem Ahmad Cheema, my cousin who is more like brothers, for being so caring and fun loving that made my stay in Australia a memorable period of my life. I am also grateful to his family, my other brother Waseem Ahmad Cheema and his family for being very caring and affectionate. Their kids gave me so many reasons to smile. I also feel blessed that my brother and best friend Umair was also here with me and I never felt lonely in his presence. I am also thankful to Assad Mehmood for giving me company during coffee breaks.

I feel grateful to all my friends back in Pakistan for being in contact with me and encouraging me. I am thankful to my younger brother and my crime partner Ammar for keeping me informed about family events in his own special way. I am also thankful to

my best friends (they are best at finding something in me to complain about), Mohsin Hasan and Samar Cheema, for being very caring and for keeping me up-to-date of issues and events happening in Pakistan especially in our friends' circle. I am thankful to my intimate friend Irfan Ahmad for being too busy to be in contact. I am also thankful to my sweet cousin Farman Ahmad Bajwa for always remembering me and especially for always making me angry by telling me not to be angry. Finally, I am very grateful to Daisy for giving me so many valuable suggestions for this part of my thesis. The best advice she gave me was to include her name in acknowledgments.

I feel obliged to say "thanks" to so many people who have been doing so many good things for me but I am afraid the list may go forever. Same as mathematicians write all positive even numbers in very small space as $\{0, 2, 4, ..., \infty\}$, below I try to write a shorter acknowledgements. Let k be a positive integer, T be my lifetime, and $dist(p, q)$ be the distance between the hearts of any two people p and q reflecting the love the person p has for q, I feel grateful to all people who have ever been among my k nearest neighbors and/or k reverse nearest neighbors over the time period T.

Abstract

A k nearest neighbor query q retrieves k objects that lie closest to the query point q among a given set of objects P. With the availability of inexpensive location aware mobile devices, the continuous monitoring of such queries has gained lot of attention and many methods have been proposed for continuously monitoring the kNNs in highly dynamic environment. Multiple continuous queries require real-time results and both the objects and queries issue frequent location updates. Most popular spatial index, R-tree, is not suitable for continuous monitoring of these queries due to its inefficiency in handling frequent updates. Recently, the interest of database community has been shifting towards using grid-based index for continuous queries due to its simplicity and efficient update handling. For kNN queries, the order in which cells of the grid are accessed is very important. In this research, we present two efficient and effective grid access methods, CircularTrip and ArcTrip, that ensure that the number of cells visited for any continuous kNN query is minimum. Our extensive experimental study demonstrates that CircularTrip-based continuous kNN algorithm outperforms existing approaches in terms of both efficiency and space requirement. Moreover, we show that CircularTrip and ArcTrip can be used for many other variants of nearest neighbor queries like constrained nearest neighbor queries, farthest neighbor queries and $(k + m)$-NN queries. All the algorithms presented for these queries preserve the properties that they visit minimum number of cells for each query and the space requirement is low. Our proposed techniques are flexible and efficient and can be used to answer any query that is hybrid of above mentioned queries. For example, our algorithms can easily be used to efficiently monitor a $(k + m)$ farthest neighbor query in a constrained region with the flexibility that the spatial conditions that constrain the region can be changed by the user at any time.

Contents

List of Figures

List of Tables