AdaNexus: An Improved Nexus Algorithm

A PROJECT REPORT

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
Master of Technology

IN

Faculty of Engineering

BY
Achint Chaudhary



Computer Science and Automation Indian Institute of Science Bangalore – 560 012 (INDIA)

July, 2020

Declaration of Originality

I, Achint Chaudhary, with SR No. 04-04-00-10-42-18-1-15879 hereby declare that the material presented in the thesis titled

AdaNexus: An Improved Nexus Algorithm

represents original work carried out by me in the **Department of Computer Science and Automation** at **Indian Institute of Science** during the years **2018-20**.

With my signature, I certify that:

- I have not manipulated any of the data or results.
- I have not committed any plagiarism of intellectual property. I have clearly indicated and referenced the contributions of others.
- I have explicitly acknowledged all collaborative research and discussions.
- I have understood that any false claim will result in severe disciplinary action.
- I have understood that the work may be screened for any form of academic misconduct.

Date:

Student Signature

In my capacity as supervisor of the above mentioned work. I certify that the above statements

In my capacity as supervisor of the above-mentioned work, I certify that the above statements are true to the best of my knowledge, and I have carried out due diligence to ensure the originality of the report.

Advisor Name: Advisor Signature

© Achint Chaudhary July, 2020 All rights reserved

DEDICATED TO

 $My\ Parents\ \ \ Close\ People$

For those whom I have lost, wish you can see this

Acknowledgements

I am deeply grateful to Prof. Jayant R. Haritsa fir his unmatched guidance. He was and always will be an inspiration for me for his thoughts and his perception of looking at things around us, be it academia or real life. It is my good luck, that I got a chance to work with him and obtain few chunks of knowledge from him.

I am thankful to Anshuman Dutt, Srinivas Karthik and Sanket Purandare for their guidance and discussions. I would express my gratitude towards my lab mates for suggestions during my work. My sincere thanks goes to my friends Diksha and Hemant, who have patiently listened and understood many complications, and help me bring out much clear pictures to work with. Also, CSA and CDS office staff was the team of people who make academic functioning smooth with best co-ordination possible.

Finally, I would express my gratitude to my parents, who have taught me how to stay calm and keep working on what is essential, nothing I have done or will ever do will match their support that I carry within.

Abstract

Declarative query processing in database systems often leads to sub-optimal performance due to wrong selectivity estimation from those encountered during actual execution. Plan Bouquets is a technique proposed to substitute selectivity estimation by selectivity discovery at run-time, to provide worst case performance guarantees. This is done by performing multiple partial executions for same query in an incremental fashion of cost budget from a bouquet which is compiled at very first stage.

This technique is suitable for OLAP queries as high overheads of optimizing most part of selectivity error space are amortized over multiple invocation of query in OLAP scenarios. Full space exploration overheads in past are improved upon with NEXUS, which is an algorithm, that only discovers points useful for bouquet compilation.

In this work, we proposed an adaptive version of NEXUS named AdaNEXUS, which utilize geometrical properties of contours to be discovered for bouquet. This algorithm reduces overheads of compilations empirically keeping the worst case performance complexity same. Further, we provide upper bounds for maximum cost deviation possible during bouquet compilation due to use for either NEXUS or AdaNEXUS. Evaluation of proposed system is done on TPC-DS benchmark with different scales to test system. It is demonstrated that around an order of magnitude reduction is observed in compilation overheads when compared with NEXUS, also quality of plans discovered by AdaNEXUS is better than those discovered by NEXUS.

Contents

A	cknowledgements	1				
A	bstract	i				
\mathbf{C}	ontents	iii				
$\mathbf{L}_{\mathbf{i}}$	ist of Figures	v				
$\mathbf{L}_{\mathbf{i}}$	ist of Tables	v				
1	Introduction	1				
	1.1 How to Write a Thesis: An Introduction	1				
2	Conclusion and Future Work	2				
	2.1 How to Write a Thesis: An Introduction	2				
3	Conclusion and Future Work	3				
	3.1 How to Write a Thesis: An Introduction	3				
4	Conclusion and Future Work					
	4.1 How to Write a Thesis: An Introduction	4				
5	Conclusion and Future Work					
	5.1 How to Write a Thesis: An Introduction	5				
6	Conclusion and Future Work	6				
	6.1 How to Write a Thesis: An Introduction	6				
7	Conclusion and Future Work					
	7.1 How to Write a Thesis: An Introduction	7				

CONTENTS

Bibliography 8

List of Figures

List of Tables

Introduction

A brief description of your chapter.

1.1 How to Write a Thesis: An Introduction

Conclusion and Future Work

A brief description of your chapter.

2.1 How to Write a Thesis: An Introduction

Conclusion and Future Work

A brief description of your chapter.

3.1 How to Write a Thesis: An Introduction

Conclusion and Future Work

A brief description of your chapter.

4.1 How to Write a Thesis: An Introduction

Conclusion and Future Work

A brief description of your chapter.

5.1 How to Write a Thesis: An Introduction

Conclusion and Future Work

A brief description of your chapter.

6.1 How to Write a Thesis: An Introduction

Conclusion and Future Work

A brief description of your chapter.

7.1 How to Write a Thesis: An Introduction

Bibliography

[1] Swaprava Nath. *Mechanism Design for Strategic Crowdsourcing*. PhD thesis, Indian Institute of Science, Bangalore, 2013. forthcoming. 1, 2, 3, 4, 5, 6, 7