Contents

Abstract i

Declaration Sheet ii

Acknowledgements iii

Preface v

Table of Contents vi

List of Figures viii

List of Table ix

I Introduction 1
- 1.1 Background . . . 1
- 1.2 Statement of Problem . . . 3
- 1.3 Objective . . . 4
- 1.4 Scope . . . 4
- 1.5 Hypothesis . . . 4

II Literature Review 5
- 2.1 Overview . . . 5
- 2.2 Voronoi Diagram . . . 5
  - 2.2.1 First Order Voronoi Diagram . . . 6
  - 2.2.2 Higher Order Voronoi Diagram . . . 7
  - 2.2.3 Highest Order Voronoi Diagram . . . 8
- 2.3 Construction Method . . . 8
  - 2.3.1 Naive method . . . 8
  - 2.3.2 Walking method . . . 9
  - 2.3.3 Flip method . . . 9
  - 2.3.4 Incremental method . . . 9
  - 2.3.5 Divide-and-conquer method . . . 9
  - 2.3.6 Lift-up method . . . 9

2.4 FLIP and LAM Algorithm . . . 9
2.4.1 FLIP Algorithm . . . 10
2.4.2 LAM Algorithm . . . 10
2.5 Apache Spark . . . 12
2.5.1 Apache Spark Architecture . . . 12
2.5.2 Resilient Distributed Dataset (RDD) . . . 13
2.5.3 Dataframe . . . 13
2.6 Discussion . . . 14

III System Methodology and Design **15**
3.1 Overview . . . 15
3.2 Configuration . . . 15
3.3 Data Structure . . . 15
3.4 Process Overview . . . 16
3.4.1 Highest Order Voronoi Diagram Construction . . . 17
3.4.2 Highest Order Voronoi Processing on Apache Spark . . . 21
3.5 Evaluation . . . 22

IV Testing and Analysis **24**
4.1 System Testing . . . 24
4.1.1 Testing Objective . . . 24
4.1.2 Data Scenario . . . 24
4.1.3 Testing Scenario . . . 25
4.2 Testing Result and Analysis . . . 25
4.2.1 Number Of Partition Comparison on Apache Spark . . . 25
4.2.2 Execution Time Comparison . . . 26
4.2.3 Labelling Method Comparison . . . 27
4.2.4 Region Construction and Region Labelling Comparison on Apache Spark . . . 29
4.3 Summary . . . 30

V Conclusions and Recommendations **31**
5.1 Conclusions . . . 31
5.2 Future Work . . . 31

Bibliography **32**

Appendices **34**