

# **CONTENTS**

---

# CONTENTS

Chapter No.	Title	Page No.
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	Data Mining	2
1.2	Web Mining	3
1.3	Transaction Database	6
1.4	Frequent Pattern Mining	9
1.5	Association Rule Mining	10
1.6	Data Compression and Frequent Pattern Mining	13
1.7	Motivation	16
1.8	Problem Statement and Objectives	18
1.9	Layout of the dissertation	19
1.10	Chapter Summary	20
<b>2</b>	<b>REVIEW OF LITERATURE</b>	<b>21</b>
2.1	Basic Frequent Pattern Mining Algorithm	21
2.2	Applications of Frequent Pattern Mining	25
2.3	Web Usage Mining and Frequent Pattern Mining	29
2.4	Other Frequent Mining Techniques	32
2.5	Mining Compressed Frequent Patterns	42
2.6	Chapter Summary	44
<b>3</b>	<b>METHODOLOGY</b>	<b>45</b>
3.1	CT-Apriori	45
3.2	Compressed FP-Tree (CFP-Tree)	53
3.3	Frequent Pattern Mining in Web log files	63
3.4	Chapter Summary	64
<b>4</b>	<b>RESULTS AND DISCUSSION</b>	<b>65</b>
4.1	Test Datasets	65
4.2	Performance Metrics	67
4.3	Preprocessing Results	67

4.4	Results	69
4.5	Chapter Summary	77
<b>5</b>	<b>SUMMARY AND CONCLUSION</b>	<b>78</b>
	<b>BIBLIOGRAPHY</b>	
	<b>PUBLICATIONS</b>	

## LIST OF TABLES

<b>TABLE NO</b>	<b>PARTICULARS</b>	<b>PAGE NO</b>
1	EXAMPLE TRANSACTION DATABASE	14
2	TRADITIONAL FREQUENT PATTERN MINING ALGORITHMS	24
3	EXAMPLE TRANSACTION DATABASE TDB	46
4	COMPACT TRANSACTION DATABASE OF TDB	48
5	SAMPLE DATABASE	56
6	MAPPING FREQUENT ITEMS	57
7	PARAMETERS USED IN THE SYNTHETIC DATA GENERATION PROGRAM	65
8	PARAMETER SETTINGS OF SYNTHETIC DATA SETS	66
9	COMPRESSION RATIO IN TERMS OF DATABASE SIZE	69
10	COMPRESSION RATIO IN TERMS OF NUMBER OF TRANSACTIONS	71
11	AVERAGE TIME TAKEN	72

## LIST OF FIGURES

FIGURE NO	PARTICULARS	PAGE NO
1	WEB LOG FILE LOCATIONS	4
2	EXAMPLE TRANSACTION DATABASE	6
3	AN EXAMPLE OF SERVER LOG FILES	7
4	SAMPLE IIS FORMAT WEB LOG FILE	7
5	PREFIX-TREE FOR THE DATABASE IN TABLE I	14
6	FP-TREE FOR THE DATABASE IN TABLE 1.1	15
7	CT-TREE FOR THE DATABASE TDB IN TABLE 3.2	49
8	COMPACT TRANSACTION DATABASE GENERATOR	51
9	CT-APRIORI ALGORITHM	53
10	BINARY REPRESENTATION OF A TRANSACTION DATABASE	54
11	FP-TREE	56
12	CFP-TREE	56
13	EXAMPLE GLOBAL CFP-TREE	58
14	ALGORITHMS FOR CONSTRUCTING CFP- TREE	60

15	<b>MINING FREQUENT PATTERNS IN CT- PRO</b>	61
16	<b>LOCAL CFP-TREE DURING MINING PROCESS</b>	63
17	<b>FREQUENT PATTERN MINING IN WEB LOG DATA</b>	64
18	<b>WEB LOG FILE</b>	68
19	<b>REMOVAL OF IRRELEVANT TRANSACTIONS</b>	68
20	<b>CONVERTED WEB LOG FILE</b>	70
21	<b>T5I5D50K</b>	73
22	<b>T10I8D100K</b>	73
23	<b>T15I10D100K</b>	74
24	<b>T20I12D200K</b>	74
25	<b>T20I12D300K</b>	75
26	<b>WEB LOG DATA</b>	75