

IDENTIFICATION AND MONITORING OF LOAD POWER QUALITY CHARACTERISTICS USING DATA MINING

This project report is presented in partial fulfillment for the award of the

Bachelor in Electrical Engineering (Hons.)

Of

UNIVERSITI TEKNOLOGI MARA (UiTM)



MUHAMMAD FIRDAUS BIN ABD RAHMAN
B. ENG (Hons.) ELECTRICAL
Faculty of Electrical Engineering
UNIVERSITI TEKNOLOGI MARA (UiTM)
40450, Shah Alam, Selangor Darul Ehsan
Malaysia

ACKNOWLEDGEMENT

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

In the name of Allah S.W.T, the most beneficent and the Merciful. Praise to Him endowment that allows me to complete this project. Syukur Allamdulillah, I have managed to complete this project and gained valuable knowledge.

I would like to take this opportunity to sincerely express my highest appreciation to my project supervisor Associate Professor Dr. Hj. Zuhaina Hj. Zakaria for her ideas, suggestion, advices and guidance until the completion of this project.

Not forget, a million of thanks also to all my friends for their moral support and guidance. May Almighty Allah bless and reward them for their generosity. Not forgotten to people who involve in this project directly or indirectly, which enable this project to be completed as required.

Finally, with my love and gratitude, I want to dedicate this thesis to my family who has supported me throughout, especially to my mother and my father.

Thank you.

ABSTRACT

The paper presents data mining application in power quality data analysis. Data mining tools are suitable for assisting in such analysis of large scale power quality monitoring data. It is to extract information that may otherwise remain hidden within the recorded data. Such information may be critical for identification and diagnoses of power quality disturbance problems, prediction of system abnormalities or failure and alarming of critical system situations.

It also shows on how clustering can be used to identify interesting patterns of harmonic measurement data and how these relate to their associated operational issues. The measured data is classified using data mining based on clustering techniques which can provide the engineers with a visually oriented method of evaluating the operational information contained within the cluster.

Such information may be critical for identification and diagnoses of power quality disturbance problems, prediction of system abnormalities and alarming of critical system situations. The monitoring program involved measurements of the harmonic currents and voltages from a substation.

From the result obtained, the capacitive load event that occur shows that by applying the clusterdata it can provide method of evaluating the operational information contained within the cluster. However, the result shows that harmonic not give much affect to the system since the energy level of harmonic that feedback to the substation is small. Only 11th harmonic of every cluster give effect to the system. The harmonics can easily be eliminated by capacitor banks since it far from fundamental voltage and current.

Keyword: data mining, clustering techniques, power quality, harmonics.

TABLE OF CONTENTS

| CONTENTS | PAGE |
|---|-------------|
| DECLARATION | i |
| DEDICATION | ii |
| ACKNOWLEDGEMENT | iii |
| ABSTRACT | iv |
| TABLE OF CONTENT | v |
| LIST OF FIGURE | ix |
| LIST OF ABBREVIATIONS | x |
| | |
| CHAPTER 1 | |
| INTRODUCTION | |
| 1.1 INTRODUCTION | 1 |
| 1.2 OBJECTIVE | 3 |
| 1.3 SCOPE OF WORK | 3 |
| 1.4 THESIS OVERVIEW | 4 |
| | |
| CHAPTER 2 | |
| DATA MINING AND CLUSTER ANALYSIS THEORY | |
| 2.1 INTRODUCTION | 5 |
| 2.2 DATA MINING | 5 |
| 2.2.1 Data Mining Process | 7 |
| 2.2.1.1 Requirement analysis | 7 |
| 2.2.1.2 Data selection and collection | 7 |
| 2.2.1.3 Cleaning and preparing data | 7 |
| 2.2.1.4 Data mining exploration and validation. | 8 |
| 2.2.1.5 Implementing, evaluating and monitoring | 8 |
| 2.2.1.6 Result visualization | 8 |

TABLE OF CONTENTS

| CONTENTS | PAGE |
|---|-------------|
| 2.3 DATA MINING TECHNIQUES | 9 |
| 2.3.1 Association rules mining | 9 |
| 2.3.2 Supervised classification | 9 |
| 2.3.3 Cluster analysis | 10 |
| 2.3.4 Web data mining | 10 |
| 2.3.5 Search engines | 10 |
| 2.3.6 Data warehousing and OLAP | 11 |
| 2.4 AIM OF CLUSTER ANALYSIS | 12 |
| 2.5 TYPES OF CLUSTER ANALYSIS METHODS | 13 |
| 2.5.1 Partitional method | 13 |
| 2.5.2 Hierarchical method | 13 |
| 2.5.3 Density-based method | 13 |
| 2.5.4 Grid-based method | 13 |
| 2.5.5 Model-based method | 14 |
| 2.6 HIERARCHICAL METHOD | 15 |
| 2.6.1 Distance type | 17 |
| 2.6.1.1 Euclidean distance | 18 |
| 2.6.1.2 Manhattan distance | 18 |
| 2.6.1.3 Chebychev distance | 18 |
| 2.6.1.4 Categorical data distance | 18 |
| 2.6.2 Linkage | 19 |
| 2.6.2.1 Single-link | 19 |
| 2.6.2.2 Complete-link | 20 |
| 2.6.2.3 Centroid | 20 |
| 2.6.2.4 Average-link | 21 |
| 2.6.3 Classification or cluster procedure | 22 |