Universiti Teknologi MARA

Dynamic Clustering of Parallel Coordinate Graph for Pattern Relationship in Photovoltaic Data

Liyana Fatini Binti Roslan

Final Year Project Report
Bachelor of Computer Science (Hons.)
Faculty of Computer and Mathematical Sciences

January 2019

ACKNOWLEDGEMENT

In The Name of Allah, The Most Beneficent, The Most Merciful.

The completion of this project could not have been possible without the participation and assistance of so many people. Their contributions are sincerely appreciated and gratefully acknowledged.

Special thanks to my dearest parents and family for their supports and motivations either morally, financially or physically through my whole life.

A lot of gratitude to Dr Zainura Idrus, for her endless support, kind and understanding spirit in providing direction and guidance for finishing this project. The inspiration given is the most valuable thing that shape my mind to have successfully completed this project.

Also a big thanks to my all of my friends who supported me. Thank you.

ABSTRACT

Green Energy Research Centre (GERC) is a research center in the area of photovoltaic technology. The focus of their research is in finding the best way to convert visible light into renewable electricity and to maximize the production of electricity. The data is collected every five minutes in a day thus, the amount of data become huge. With the help of data visualization, these data can be analyzed easily. Parallel coordinate graph are broadly used in numerous applications for visualizing multivariate data. This technique is helpful to reveal the relationships among the data as well as to view the data as a whole. However, the main problem when using this technique with a large dataset can cause visual clutter and over plotting where the polylines are overlapped on top of each other, makes it difficult to read the information and the visualization become inefficient to the user. Therefore, the main objective of this project is to implement an approach to overcome the problem of visual clutter in parallel coordinate graph. Bundling clustering technique has been chosen as the solution to overcome the problem, where it is applied on the parallel coordinate graph. Curve smoothness and bundling strength are the most important function in bundling as they are the global settings of bundling technique. Other than that, four visual analytics techniques are added into parallel coordinate graph to enhance the interaction between user and the system which are rendering, reordering, brushing and coloring. As a result, this technique can reduce the issue of visual clutter in parallel coordinate graph and also it can help user to have a better insight and extract knowledge directly from the graph hence, it can improve user understanding towards the data itself and helping the researchers and analyzers in a wider context.

TABLE OF CONTENTS

CONTENTS **PAGE**

CHAPTER 1

INTRODUCTION

This chapter explained the rationality and background of the study. The problems in this study are discussed and objectives are drawn based on the problems. Other than that, the details of the scope and significance of this study are also have been listed.

1.1 Background of Study

Photovoltaic technology is commonly known as a process of producing electric current through solar cells. *Electric current* is generated directly from visible light commonly known as sunlight which is converted into electrons by the photovoltaic process through various components (Park & Park, 2018). One of the main components is solar cells. The cells act as a photo detector, detecting light and other electromagnetic radiation. The cell also measures light intensity (Singh, 2013).

A research center called Green Energy Research Centre (GERC) has been established under the Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM), Malaysia. Their core research is in the area of photovoltaic technology. The focus of their research is in finding the best way to convert visible light into renewable electricity and to maximize the production of electricity. Moreover, part of their research is to predict the amount to electricity produced. The prediction is use by Tenaga Nasional Berhad in managing the production of electricity to be distributed to society. Thus, data on photovoltaic are collected and analyzed.

GERC collects the data through its various logs sensor which are installed within the photovoltaic module. Part of the prediction process is to analyze