UNIVERSITI PUTRA MALAYSIA

AN IMPROVEMENT ON EXTENDED KALMAN FILTER FOR NEURAL NETWORK TRANING

TSAN KEN YIM.

FSKTM 2005 5

AN IMPROVEMENT ON EXTENDED KALMAN FILTER FOR NEURAL NETWORK TRAINING

By

TSAN KEN YIM

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

April 2005

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirements for the degree of Master of Science

AN IMPROVEMENT ON EXTENDED KALMAN FILTER FOR NEURAL NETWORK TRAINING

By

TSAN KEN YIM

April 2005

Chairman : Associate Professor Md Nasir Sulaiman, PhD

Faculty : Computer Science and Information Technology

Information overload has resulted in difficulties of managing and processing information. Reduction of data using well-defined techniques such as rough set may provide a means to overcome this problem. Extracting useful information and knowledge from data is a major concern in information science. Artificial intelligence systems, such as neural network systems, are widely used to extract and infer knowledge from databases.

This study explored the training of a neural network inference system using the extended Kalman filter (EKF) learning algorithm. The inference accuracy, inference duration and training performance of this extended Kalman filter neural network were compared with the standard back-propagation algorithm and an improved version of the back-propagation neural network learning algorithm. It was discovered that the extended Kalman filter trained neural network required less

Dedicated to my beloved wife, my son,

my parents and the family

training cycles during the learning process; and produced similar or better inference performance when compared with other trained neural networks.

This study further improved the performance of the extended Kalman filter algorithm used in the study. The improvement was achieved with simplification on the matrix multiplications in the algorithm, through selective updates of the elements in the error covariance matrices. This improved version of the extended Kalman filter was implemented along with the three algorithms as mentioned earlier. The results obtained showed improvement in the duration for each training cycle, when compared with the original extended Kalman filter algorithm.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

SATU PEMBAIKAN DALAM PENURAS KALMAN TERLUAS UNTUK LATIHAN RANGKAIAN NEURAL

Oleh

TSAN KEN YIM

April 2005

Pengurusi : Profesor Madya Md Nasir Sulaiman, PhD

Fakulti : Sains Komputer dan Teknologi Maklumat

Informasi yang terlalu banyak telah menimbulkan pelbagai masalah dalam pengendalian dan pemprosesan maklumat. Data yang telah diringkaskan dengan teknik yang terbukti, seperti set kasar, mungkin memberikan penyelesaian kepada masalah tersebut. Pendekatan untuk memperolehi maklumat dan pengetahuan yang berguna adalah suatu isu yang utama dalam bidang sains maklumat. Sistem-sistem kecerdasan buatan, seperti sistem-sistem rangkaian neural, digunakan secara meluas untuk mendapatkan pengetahuan daripada pangkalan-pangkalan data.

Kajian ini merupakan percubaan dalam melatih sistem inferen rangkaian neural dengan menggunakan algoritma latihan penuras Kalman terluas ke atas jadual keputusan yang diringkaskan. Ketepatan inferen, jangka masa inferen dan prestasi latihan rangkaian neural dengan penuras Kalman terluas ini dibandingkan dengan algoritma-algoritma latihan perambatan balik yang biasa dan suatu versi perambatan balik yang diperbaiki. Keputusan menunjukkan bahawa rangkaian neural yang dilatih dengan penuras Kalman terluas memerlukan ulangan latihan yang kurang

berbanding dengan algoritma-algoritma lain. Pada masa yang sama, prestasi penjanaan keputusan yang diperolehi adalah tidak berbeza atau lebih baik berbanding dengan rangkaian neural terlatih yang lain.

Penyelidikan ini telah memperbaiki lagi prestasi algoritma penuras Kalman terluas. Peningkatan ini dapat dicapai dengan ringkasan pendaraban matriks dalam algoritma, melalui pengiraan nilai-nilai terpilih di dalam matriks ralat kovarians yang digunakan. Versi ringkasan penuras Kalamn terluas tersebut dilaksanakan seiring dengan tiga algoritma tersebut di atas. Keputusan yang didapati telah menunjukkan pencapaian yang lebih baik untuk jangka masa latihan yang diperlukan, apabila dibandingkan dengan algoritma penuras Kalman terluas yang asli.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor *Associate Professor Dr. Hj. Md Nasir Sulaiman*, the head of department for Computer Science, Faculty of Computer Science and Information Technology, for his patience in supervision, invaluable guidance and helpful discussion. I would also like to convey my sincere appreciations to all the members of the supervision committee, *Associate Professor Dr. Mohamed Othman* and *Associate Professor Dr. Hamidah Ibrahim* for their valuable inputs, encouragements, comments and help. This work could not be completed without them.

I would also like to express my gratitude to the Faculty of Computer Science and Information Technology, School of Graduate Studies and the Library of Universiti Putra Malaysia, for providing good assistance and environment while I was pursuing my research work.

Special thanks are due to Dr. Azuraliza Abu Bakar, who has assisted me significantly in the early stages of my research. I would also like to thank Dr. Moawia Elfaki Yahia, whose earlier research work inspired mine.

Finally, I would like to thank my wife, my parents and my family members for their motivation, patience and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BP	Back-propagation
EKF	Extended Kalman filter
PNN	Probabilistic neural network
XOR	Exclusive-OR

CHAPTER I

INTRODUCTION

Background

Making decision, from as mundane as deciding what to take for lunch, to as critical as deciding the vision of a nation, is a process that we perform for almost every task daily. Decision making process requires knowledge and information. As the amount of information is increasing exponentially over time, there is an urgent need for tools and methods to perform knowledge discovery and extraction in information-rich environments (Turban and Aronson, 2001).

Intelligent information systems, such as neural network systems, are used to facilitate the process of making decisions. Very often, these systems utilize complex and large knowledge bases. However, complex knowledge is difficult to be learned by neural network systems. Moreover, neural network may take longer learning duration if the knowledge base is large.

In addition to this, a suitable learning algorithm is critical to ensure the successful learning of knowledge for a neural network system. Learning algorithm plays a pivotal role in the efficiency and effectiveness of the neural network learning process. An appropriate learning algorithm in a neural network system provides a shorter learning duration, which is especially essential for problems that require near real-time solutions.

Literatures are widely available on the study of neural network algorithms. Many new learning algorithms as well as improvements on existing algorithms for neural networks have been introduced, some of these reference literatures can be obtained from Haykin (1999) as well as Karayiannis and Venetsanopoulos (1993).

Kalman-based algorithms (Kalman and Kwasny, 1991), have been studied since early 1990s. These algorithms are still being actively researched, improved and applied to various real-world problems, especially in real-time estimation, tracking, monitoring and control (Grewal and Andrews, 1993). The extended Kalman filter (EKF) algorithm, an extension from the Kalman filter algorithm by linearizing the system around the current parameter estimates, has served as the basis for many neural network training algorithm studies (Li *et al.*, 2002).

Studies and applications of EKF algorithms for neural networks training have shown that these algorithms are faster, more stable, simpler and generally better than other neural network training algorithms (Pérez-Ortiz *et al*, 2003; Simon, 2002; Li *et al.*, 2002; Stan and Kamen, 2000). EKF algorithms have been used to train multi-layer neural networks (Sum *et al*, 1999; Zhang and Li, 1999; Shah, Palmieri and Datum, 1992) and recurrent networks (Obradovic, 1996; Puskorius and Feldkamp, 1994).

Problem Statement

Extended Kalman filter (EKF) algorithm requires many matrix operations, thus greatly increase the computational complexity (Li *et al.*, 2002). Matrix operations, especially matrix multiplication operations, require a lot of resources especially when the matrix dimensions are big. With a large training data set, a large network configuration and multiple network inputs, training speed becomes one of the main problems for the EKF algorithm.

Objectives of the Study

This study aimed to propose an improvement on the extended Kalman filter algorithm (improved EKF) for the neural network training. This improved EKF was obtained with simplification on the matrix multiplications that involve the error covariance matrix P in the EKF algorithm. A diagonal matrix P was used in the improved EKF algorithm, and the complexity of the matrix multiplications that involved diagonal matrix P was reduced through a proposed matrix operator ***multiDiag***. The basic idea of this study was that a neural network inference system that has been trained with the improved EKF algorithm, might provide highly accurate results for faster decision making.

Scope of the Study

A set of survey data (Kaplan, 2001), which contained a total of 112 attributes for 2,569 respondents aged between 40 year-old and over 100 year-old, was used for this study. The survey data was collected for studies on demography, especially on the conditions of life in communities, for the residents in Alameda County, California, United States of America. This survey data was used in the research study as a test model to validate the idea presented earlier.

The experimentations in this study focused on comparing the inference results, using the original and simplified survey data, from four different neural network inference systems. One of the systems was the system proposed in this research, using an improved version of the EKF algorithm for the neural network training.

The simplified survey data was produced from the original data using ROSETTA, a rough set analysis tool. This simplified data, along with the original survey data, was used to construct the knowledge bases for the neural network inference systems. Comparison of results focused on the inference accuracy of the test data, inference duration taken to obtain the results, and training efficiency (in term of the number of training cycles required) among these inference systems.

Contribution of the Study

The study proposed and implemented an improved version of the EKF algorithm for training a neural network to be used in decision inference. The proposed system employed simplified knowledge produced from the original data using a rough set tool. This system was validated with test data, and the results were promising to support the objective of this study.

The contribution of this study was an improved version of the EKF algorithm. The simplification was obtained by updating only the values of the diagonal elements in the error covariance matrix. With this approach, the error covariance matrix was maintained as a diagonal matrix, and thus a simpler matrix multiplication operation was introduced for computations that involved the error covariance matrix. In other words, the proposed improved EKF algorithm provided computational savings through selective updates of only the diagonal elements in the error covariance matrix.

Comparison results between the proposed algorithm and three other neural network learning algorithms (EKF, BP and an improved BP algorithm) used for simplified knowledge were encouraging. These algorithms were used for training neural networks that employed the original as well as the simplified survey data. Experiment results showed that the proposed algorithm was on par or better than the other algorithms used in this study.

Organization of the Dissertation

The dissertation was organized in accordance with the standard structure of thesis and dissertations at Universiti Putra Malaysia. The dissertation has seven chapters, including this introductory chapter. The remaining chapters are: Literature Review, Theoretical Foundation, Improved Extended Kalman Filter, Research Methodology, Results and Discussion, as well as Conclusion and Recommendations.

Chapter II – Literature Review included literatures related to topics of the research. The related research works done previously and being carried out currently on rough set, neural networks and hybrid inference systems were reviewed and discussed.

Chapter III – Theoretical Foundation on Rough Sets Theory and Neural Network described the methods and techniques used in this research. This chapter focused on the concepts of rough sets and neural networks learning algorithms. The concepts of rough sets as a reduction technique were presented. The theoretical aspects of standard Back-Propagation (standard BP) algorithm and Improved Error Signal Back-Propagation (improved BP) algorithm were also extensively discussed.

Chapter IV – Research Methodology presented the research work done, with detail description of the methods and processes used. This chapter also described the experiments that were designed and carried out to achieve objective of the study.

Chapter V – Improved Extended Kalman Filter described the proposed algorithm for this study. This algorithm was an improved version of the EKF algorithm. The idea and illustration of the improvement were presented in this chapter. A preliminary experiment was also carried out to validate the proposed algorithm.

Chapter VI – Results and Discussion presented the inference results obtained using both the complete knowledge and simplified knowledge from the four inference systems used in the study. Detail results from the experiments were tabulated, evaluated and discussed. This chapter highlighted the main findings, and correlated these findings to the objectives of the study.

Chapter VII – Conclusion and Recommendations concluded the study with some recommendations for further research and development. The conclusion included remarks of the attainment of the objectives under study. The recommendations were presented as guidelines for future improvements.

CHAPTER II

LITERATURE REVIEW

Rough Sets Theory

Rough sets theory, introduced by Zdzislaw Pawlak in the early 1980s, is a mathematical tool to deal with knowledge, particularly imprecise and uncertain knowledge (Pawlak, 1982). This approach is based on the ability to classify data that is obtained from observation, measurement, etc. of real or abstract objects. Objects or concepts of the data are represented with a predetermined set of attributes and values in tabular formats. These tables are called Information Systems (sometimes also called knowledge representation systems, information tables, attribute-value systems or decision tables).

The main concept of rough sets theory is an indiscernibility relation, normally associated with a set of attributes. The relations among the classes of the indiscernibility relations associated with the attribute sets can be used to generate simpler decision rules. Previous study by Düntsch and Gediga in 2000 showed that attribute reduction in rough sets data analysis can be achieved by comparing the indiscernibility relations generated by sets of attributes.

One of the main advantages of rough sets theory is that it does not need any preliminary or additional information about data, unlike probability distribution in statistics, basic probability assignment in the Dempster-Shafer theory, or grade of