

FUZZY CASE-BASED REASONING FOR WEATHER PREDICTION

WAN SALFARINA BINTI WAN HUSAIN

UNIVERSITI TEKNOLOGI MALAYSIA

ABSTRACT

Prediction is the process of estimation of unknown situation that refers to time-series, cross-sectional or longitudinal data. Weather prediction is the process to project how the atmosphere will evolve. Weather is known as continuous, data-intensive, multidimensional, dynamic and chaotic. The chaotic nature of atmosphere required the massive computational power in order to solve the equations that describe the atmosphere, and the incomplete understanding of weather can make the prediction become less accurate. Based on this problem, Fuzzy Case-Based Reasoning (FCBR) is introduced in solving the prediction problem. Fuzzy can have the degree of truthfulness and falsehood that can handle uncertainty of the chaotic variables of the weather. Meanwhile Case Based Reasoning (CBR) has an ability to identify the similar cases from the past using the similarity measurement technique such as Euclidean distance. CBR can reduce the knowledge acquisition task and can reason with incomplete and imprecise data or knowledge. This study is conducted to investigate how fuzzy and CBR could solve the prediction problem and how it can improve its performance. From the experiment, it shows that the Fuzzy Case Based Reasoning has improved the accuracy of the weather prediction with achievement of 87%.

ABSTRAK

Peramalan merupakan proses penganggaran situasi yang tidak diketahui berdasarkan anggaran rangkaian data masa secara bersiri, data secara keratan melintang atau secara membujur. Peramalan dalam kaji cuaca pula merupakan proses untuk menentukan tahap perkembangan atmosfera. Cuaca dikenalpasti sebagai data yang bersifat intensif, selanjar, multidimensi, dinamik dan tidak menentu. Keadaan atmosfera yang tidak menentu ini akan mengakibatkan pengiraan dalam peramalan memerlukan kos yang tinggi. Manakala pengetahuan yang kurang luas mengenai cuaca akan menyebabkan peramalan cuaca dilakukan dengan kurang tepat. Berdasarkan masalah ini, Fuzzy Case-Based Reasoning (FCBR) diperkenalkan dalam peramalan cuaca. Fuzzy dikatakan berkebolehan dalam mengatasi ketidakpastian/kesamaran di dalam cuaca kerana ia boleh diwakilkan dalam darjah kebenaran (truthfulness) dan kesalahan (falsehood). Manakala Case Based Reasoning (CBR) berkebolehan dalam mengenalpasti kes yang sama berdasarkan kes yang lepas menggunakan teknik pengiraan jarak (similarity measurement), di mana ia menyelesaikan ketidakfahaman mengenai cuaca. Hasil daripada pengkajian ini ialah mengkaji bagaimana kaedah gabungan antara fuzzy dan CBR boleh menyelesaikan masalah dalam peramalan hujan dan bagaimana ketepatan anggaran hujan dapat ditingkatkan. Dengan menggunakan teknik Fuzzy Case Based Reasoning, ketepatan dalam peramalan cuaca dianggarkan sebanyak 87%.

TABLE OF CONTENT

CHAPTER	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	xi
	LIST OF FIGURES	xii
	LIST OF ABBREVIATIONS	xiv
	LIST OF SYMBOLS	xv
	LIST OF APPENDICES	xvi
1	INTRODUCTION	
	1.1 Introduction	1
	1.2 Problem Background	2
	1.3 Problem Statement	5
	1.4 Project Aim	5
	1.5 Objectives	6
	1.6 Project Scope	6
	1.7 Significance of the project	7
	1.8 Organization of the Report	7

2 LITERATURE REVIEW

2.1	Introduction	8
2.2	Prediction problem	9
	2.2.1 Weather prediction	11
2.3	Fuzzy Expert Systems	15
	2.3.1 Fuzzy Inference and Fuzzy Expert System	20
	2.3.2 Uncertainty in expert system	26
	2.3.3 Real application using fuzzy	27
2.4	Case-based Reasoning	27
	2.4.1 Basic concepts of CBR	33
	2.4.2 CBR and information retrieve	34
	2.4.3 CBR and databases	36
	2.4.4 Real application using CBR	37
2.5	Techniques for similarities measure	39
	2.5.1 Boolean	40
	2.5.2 Probability	40
	2.5.3 Hamming distance	41
	2.5.4 Euclidean distance	42
	2.5.5 <i>k</i> -nearest neighbor	43
2.6	Hybridized technique	43
	2.6.1 CBR and Fuzzy in weather prediction	44
2.7	Chapter Summary	46

3 RESEARCH METHODOLOGY

3.1	Introduction	47
3.2	Data Collection	48
3.3	Data Modeling	49
3.4	Research framework	50
	3.4.1 Experimental on fuzzy and CBR alone	52
	3.4.2 Preprocessing process	52

3.4.3	Determining membership value	53
3.4.4	Calculate Similarity measurement	53
3.4.5	Defuzzification and prediction	54
3.4.6	New case	56
3.4.7	Determining fuzzy membership value	57
3.4.8	Applying Fuzzy and Case Retrieval	57
3.4.9	Reuse the case	58
3.3.10	Defuzzification and prediction	58
3.5	Chapter Summary	59
4	EXPERIMENTAL AND DISCUSSION	
4.1	Introduction	60
4.2	Experimental for fuzzy and CBR alone	61
4.3	Implementation	68
4.3.1	Determining the fuzzification functions	68
4.3.2	Determine threshold value	70
4.3.3	Determining the fuzzy case membership value for each case in the case-based	71
4.3.4	Applying fuzzy and retrieving the cases using the similarities	72
4.3.5	Determine the fuzzy prediction	74
4.4	Result and discussion	76
4.5	Chapter Summary	85
5	DISCUSSION AND CONCLUSION	
5.1	Introduction	86
5.2	Findings and achievements	87
5.3	Constraints and challenges	87
5.4	Contribution of the study	88
5.5	Recommendations	88
5.6	Chapter summary	89

REFERENCES	90
APPENDIX	94

CHAPTER 1

INTRODUCTION

1.1 Introduction

Prediction is a statement or claim that a particular event will occur in future. The term of prediction same as forecasting which the forecasting is the process of estimation in unknown situations. But, prediction in more generally term, and usually refers to the estimation of time series, cross-sectional or longitudinal data. Some fields of science are notorious for the difficulty of accurate prediction and forecasting, such as in natural disasters, software reliability, meteorology and so on. For example, in meteorology, prediction is used to predict the weather, atmosphere, air quality and so on that deal with its own difficulty such as dealing with various parameters in order to make the prediction.

Weather is continuous, data-intensive, multidimensional, dynamic and chaotic (Hansen and Riordan, 2001). Today, weather predictions are made by

collecting quantitative data about current state of atmosphere process to project how the atmosphere will evolve (Wikipedia, 2008).

Using the intelligent of expert, many works and researches in weather prediction using different types of artificial intelligent techniques have been done. Singh and Ganju (2005), developed a quantitative snowfall forecasting model based on the approach; past situations under similar conditions of the feature vector in feature space are used to predict the expected behavior.

Hung *et.al* (2008) used an artificial neural network model for rainfall forecasting in Bangkok. The result of forecast models show that the combination of meteorology data with rainfall data as training data has significantly improved the forecast accuracy.

In this recent years, the hybrid techniques is broadly used among researcher in many areas such as patterns recognition, making prediction, medical diagnosis and such others applications. Hybrid technique are a potentially powerful tool that may enables us to address and solve problem that are just too complex for more conventional approaches (Jackson,1999). Some of the hybrid techniques that already used by researcher are neuro-fuzzy with knowledge-based, fuzzy logic with case-based reasoning, fuzzy with artificial immune system and many others. This approach has been successfully proved by the previous researchers in their study and works.

1.2 Problem Background

Until now, several studies have been made and research papers have been published discussing the role of artificial intelligence in prediction (*Negoita et al.*, 2005). Copas and Tarling (1986) stated that a prediction study is to investigate the

events of which criterion measures (the dependencies variables) can be predicted by one or more measure factors (the predictor or independent variables). Making prediction is a complex task because it needs a lot of skill from expert. One common problem of prediction is, sometimes the prediction of the situation or application is not correct.

Avouris and Kalapanidas (1997), in the comparative study using case-based reasoning adopted algorithm, found that the failure of doing predicting is due to the fact that the algorithms cannot distinguish clearly low-level cases from the medium and higher ones. In their research, they suggested that hybridizing is the better approach in doing prediction.

The chaotic nature of atmosphere and massive computational power required to solve the equations that describe the atmosphere and incomplete understanding of atmosphere process mean that the prediction becomes less accurate.

According to Palmer and Hagedorn (2006), prediction of weather is dealing with uncertainty; observations of weather and climate are uncertain and incomplete, the model to assimilate data and predict the future are uncertain, and external effects such as volcanoes and anthropogenic greenhouse emissions are also uncertain.

Md. Zain (2001) wrote that the increasing of data of weather from time to time gave the difficulty to meteorologist to analyze and make the prediction of the weather. This problem happened because the prediction of the weather at this time is done by manually by the expert.

From the studies and researches have been done, the role of artificial intelligence could play in predicting. In (*Hung et.al*, 2008), prediction of rainfall using artificial neural network (ANN) improved the accuracy of forecasting. The studies on Neuro-Fuzzy Knowledge-Based System for air quality predictions by

(*Negoita et al.*, 2005), the behavioral of all involved networks can be improved at any moment.

Fuzzy logic is form of logic used in systems where variables can have degrees of truthfulness or falsehood (Shenker and Khoshgoftaar,1999). With fuzzy logic, the outcome of an operation can expressed imprecisely rather than as a certainty. But, fuzzy logic has a difficulty in scaling the larger problems.

Avouris and Kalapanidas (1997), CBR is broadly used in researches and studies because it based on retrieval from the database of the most similar past cases subsequently adapted to the present conditions in order to provide the new solution. (Lenz *et al.*, 1998) shared that why CBR are famous among researcher because generally, CBR cycle may be describe by four task; retrieve the most similar case or cases, reuse the information and knowledge in that case to solve problem, revise the proposed solution, and retain the parts of experience to be useful in future problem solving.

Weather prediction using case-based reasoning and fuzzy set theory by Hansen and Riordan (2001), shared that querying a large database of weather observations for past weather cases similar to a present case using a fuzzy k-nearest neighbors algorithm that is designed and tuned with the help of a weather forecasting expert which can increase the accuracy of predictions of cloud ceiling and visibility at an airport. In this study, hybridation of Fuzzy-CBR will be further investigated and tested on weather of Kluang dataset in order to analyze its performance in prediction.

1.3 Problem statement

In doing weather prediction, the expert must deal with chaotic atmosphere and must have broad knowledge of weather. This field also needs the massive computational power in solving the equations that describe atmosphere. Hansen and Riordan (2001), mentioned that the weather is continuous and data-intensive. These properties make the weather prediction is formidable. Based on this reason, Fuzzy Case Based Reasoning in weather prediction is proposed.

The hypothesis of the study can be stated as:

“How efficient is the Fuzzy Case Based Reasoning technique in increasing the accuracy of rainfall weather prediction using the observations of the past weather cases?”

1.4 Project Aim

The aim for this project is to investigate of the capability of Fuzzy-CBR technique in solving forecasting problems such as weather prediction. The strengths of each technique will be studied to produce a solution in increasing the accuracy of prediction result.

1.5 Objectives

- i. To investigate how techniques Fuzzy and Case Based Reasoning can be hybridized to improve performance of case selection in CBR utilizing Fuzzy Logic technique.
- ii. To implement hybrid Fuzzy-Case Based Reasoning with application to rainfall weather prediction.
- iii. To validate the performance of the techniques in terms of accuracy of prediction.

1.6 Project Scope

In this project, the scopes are:

- i. Techniques to be investigated are Fuzzy and CBR for predicting the rainfall that will be developed and tested using C++.
- ii. Dataset of weather prediction used is Kluang dataset (August 2000).
- iii. This project is based on the retrieval of the closest cases using Euclidean distance.
- iv. This works focus on the improvement to get better accuracy in rainfall prediction.

1.7 Significance of the Project

The importance of this study is to show that the technique using Fuzzy-CBR can improve decisions in weather prediction; By using fuzzy value for weather parameters and the selection of the similar cases to the new existing problem. The results of this experiment will be validated comparing with the results obtained by the weather prediction works using fuzzy alone.

1.8 Organization of the Report

The project consists of five chapters. The first chapter presents the introduction of the project that including the background of problem, objective, scope and also the significance of the project. Chapter two reviews the issues of prediction and weather prediction, and how techniques like fuzzy and CBR can be utilized to make good prediction. Third chapter discusses about project methodology in this project. It elaborates the detail process of the project. Chapter four presents the experimental result and analysis of the experiment conducted. Meanwhile chapter five gives the conclusion of the project and also the suggestions for the future work and improvement.

REFERENCES

- Hansen, B.K, and Riordan, D. (2001). *Weather Prediction Using Case-Based Reasoning and Fuzzy Set Theory*.
- Singh,D., Ganju,A. and Singh,A. (2005). *Weather Prediction Using nearest-neighbour model*.
- Hung,N.Q, Babel,M.S, Weesakul,S. and Tripathi, N.K . (2008). *An artificial neural network for rainfall forecasting in Bangkok, Thailand*.
- Jackson, P. (1999). *Introduction to Expert System, 3rd Edition*. Addison-Wesley.
- Negoita, M., Neagu, D., Palde, V. An Original Neuro- Fuzzy Knowledge- Based System for Air Quality Prediction: NEIKeS (Neural Explicit and Implicit Knowledge Based System. In: *Computational Intelligence: Engineering of Hybrid Systems*. New York: Springer- Verlag. 105-128; 2005
- Copas, J.B, and Tarling, R. (1986).*Criminal Careers and "Career Criminals," Volume II*.
- Palmer, T., and Hagedorn, R. (2006). *Predictability of Weather and climate, 1st Edition*. Cambridge: University Press.
- Rozilawati binti Dollah@Md. Zain (2001). *Pengelompokan data kajicuaca dalam perlombongan data: Perbandingan antara kaedah statistik dan teknik peraturan kesatuan*. Thesis Sarjana. Universiti Teknologi Malaysia, Skudai.

- Schenker, D.F, and Khoshgoftaar, T.M. (1999). *The Application of Fuzzy Case-Based Reasoning For Identifying Fault-Prone Modules*.
- Lenz ,M., Spörl, B.B, Burkhard, H.D, Wess, S. (1998). *Case-Based Reasoning Technology*.New York: Springer- Verlag Berlin Heideberg.
- Bruninghaus, S.,Ashley, K. D. (2003). *Predicting Outcomes of Case-based Legal Arguments*.
- Chen, N.Y. , Ye, C.Z, Yang, J., Geng, D.Y, Zhou,Y. Fuzzy rules to predict degree of malignancy in brain glioma. *Medical & Biological Engineering & Computing*, 2002. Vol(40): 145-152.
- Rodjito, P., Atkins, J.E., Jones, R.M. (2006). *Motion tracking and prediction using Fuzzy Logic*.
- Obaid Saad al-Sobiei, David Ardit and Gal Polat. (2005). *Predicting the risk of contractor default in Saudi Arabia utilizing artificial neural network (ANN) and genetic algorithm (GA) techniques*.
- Lorenz, E.N. (1969). *Three approaches to atmospheric predictability bulletin of American Meteorologist Society*.
- Schmidt, R. (2007). *Case-Based Reasoning in Medicine Especially an Obituary on Lothar Gierl*.
- Pintu Chandra Shill, Anwar Hossain, Md. Ali Naser Khan and S.M Masudul Ahsan (2004). *Application of Evolutionary Fuzzy System Aircraft Control*.
- Negnevitsky, M. (2005). *Artificial Intelligence: A Guide to Intelligent Systems, 2nd Edition*. Pearson Education.
- Cagnoli, B. (1998). *Fuzzy logic in volcanology*.

- Mazliham Mohd Su'ud, Loonis,P., Idris Abu Seman . Towards Automatic Recognition and Grading of Ganoderma Infection Pattern Using Fuzzy Systems. *International Journal of Biomedical Sciences*, 2004. Vol(2): 89-94
- Jones, E., and Roydhouse, A. 1995. Intelligent retrieval of archived meteorological data. *IEEE Expert*,1995. Vol(10), 50-57.
- RMKEC. (2008). *Neural Network and Fuzzy Logic Control*.
- Siller, W. and Buckley.J.J. (2005). *Fuzzy Expert Systems and Fuzzy Reasoning*. John Wiley & Sons, Inc.
- Horvitz E. and Heckerman D. (1986). *The inconsistent use of measures of certainty in artificial intelligence research*. Amsterdam: North Holland.
- Aamodt, A. and Plaza, E. Case-based reasoning: Foundation issues. Methodological variation and system approaches. *Artificial Intelligence Communication*, (1994) 39-59.
- Kolodner, J.L, Kolodner, R.M. Experience in Clinical Problem Solving: Introduction and Framework, *IEEE Transaction on Systems, Man and Cybernatic*, (1987): Vol SMC-17; N:3; 420-431.
- Shafer G. and Tversky A. (1985). *Languages and designs for probability judgment*. Cognitive Science.
- Ashley K.D and Rissland E.L (1985). *A case-based approach to modeling legal expertise*. IEEE expert.
- Bichindaritz, I. (2004). *Case Based Reasoning in Health Sciences*.
- Lynch, P. (2006). *The Emergence of Numerical Weather Prediction*. Cambridge: University Press.
- Hong, X.L, Yen, C.V. (1995). *Fuzzy Sets and Fuzzy Decision-Making*. CRC Press, Inc.

Luger, G. F (2002). *Artificial Intelligence : Structures and strategies for Complex*

Problem Solving, 4th Edition. Pearson Education.

Ackerman, S. A., Knox, J.A. (2007). *Meteorology: Understanding the*

Atmosphere, 2nd Edition. Thomson Higher Education.

Althoff, K.D, Manago, M, Boloto, M . Case-based Reasoning for Medical

Descision Support Tasks: The INRECA Approach. *Artificial Intelligence in*

Medicine Journal, 1998. Vol(12):25-41

Abdul Majid Ismail (2008). *Klimatologi*.

<http://www.kjc.gov.my>