Case –Based Reasoning (CBR) Based on Fuzzy Set Approach For Rainfall Prediction

Nang Nwe Nwe Soe, Khin May Win
Computer University , Loikaw
dueduee@gmail.com, winn.km05@gmail.com

Abstract

Rainfall predictions and warnings are the most important services provided by the meteorological profession. Predictions are used by government and industry to protect life and to improve the efficiency of operations, and by individuals to plan a wide range of daily activities. The basic idea of this system is that CBR (case –based reasoning) solves new case by using solution to past cases. Rainfall predictions for the present case are made from the outcomes of past cases. A fuzzy set approach based methodology for knowledge acquisition is developed and used for retrieval of temporal cases in a Case –Based Reasoning (CBR) system. This system is to predict daily and monthly Rainfall Amounts (RFA) and Rainfall Type (RFT).

Keywords: Case –Based Reasoning (CBR), Fuzzy set, Prediction

1. Introduction

Prediction is a method for translating past experience into estimates of the future. Rainfall prediction becomes more and more indispensable to our lives and thus many approaches have been investigated so far to meet this high demand. Predictions are used by government and industry to protect life. Rainfall can be modeled by using various timescales including hourly, daily, and monthly timescales.

Bjarne K.Hansen and Denis Riordan.[2] showed weather prediction are very important for human lives. And they presented Fuzzy Logic knowledge acquisition is developed and used in Case –Based Reasoning (CBR) system. On that paper represented with fuzzy K-nn based prediction system, called WIND-1.

Jiawei Han and Micheline Kamber. [4] represented the classification and predictions are two forms of data analysis and that can be used to extract models describing important data classes or to predict future data trends. And they showed tat classification and prediction had many approaches: such as K-nearest neighbor (K-nn), case based reasoning (CBR), Genetic algorithms, Rough Set and Fuzzy Set techniques are introduced.

TIMOTHY J.ROSS.[7] and EARL COX. [3] analyzed different between Crisp (classical) Set and Fuzzy Set. It had represented how to define Fuzzy rules, how to use fuzzy membership function values and how to calculate using fuzzy inference technique.

In this paper, previous monthly and daily rainfall (RF) data are used to predict future Rainfall Amounts (RFA) and Rainfall Types (RFT). Case –based reasoning (CBR) is to solve new problem by reusing solution to past problems. Fuzzy technique retrieves similar case in CBR system. This system use CBR System to solve problem and fuzzy techniques retrieves similar cases by emulating domain expert who understands and interprets similar cases. Fuzzy technique is used to retrieving and reusing in CBR system.

2. Case –Based Reasoning (CBR)

Case –based reasoning is a technique for solving problems based on experience. CBR has several advantages over KBS and it is based on Human Information Processing (HIP) in some problem areas. It is recommended to developers who are challenged to reduce the knowledge acquisition tasks, avoid repeating mistakes made in the past, reason with incomplete or imprecise data, provide mean of explanation and reflect human reasoning. CBR classifiers are instanced –based. CBR has also been applied to areas such as engineering and law, where cases are either technical designs or legal rulings, respectively.

2.1. The CBR Cycle

The general CBR cycle may be described by the following four processes:

1. RETRIEVE –given a target or current problem, retrieve cases from memory that are relevant to solving it.
2. REUSE –map the solution from the previous case to the target or current problem. This may involve adapting the solution as needed to fit the new situation.

3. REVISE –having mapped the previous solution to the target or current situation, test the new solution in the real world and, if necessary, revise.
4. RETAIN –after the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory [1,6].

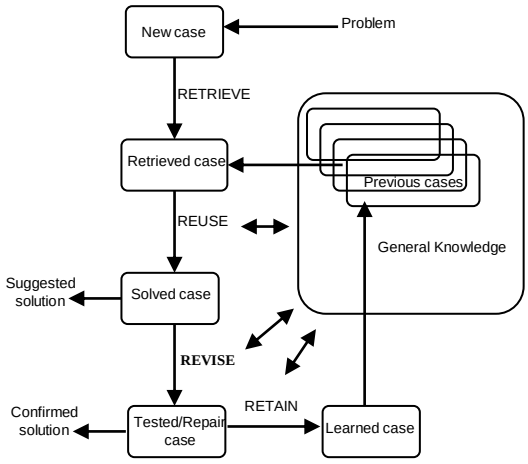

Figure1. The CBR Cycle

2.2. The CBR Process

Figure2 represents CBR process flow. New problem input enter to CBR system. Next, Case Retrieval retrieves relevant cases from historical case library and use similarity metrics to retrieve case. Case Adaptation use solution adaptation method if necessary to adapt and then test the solution that is called revise or verify.

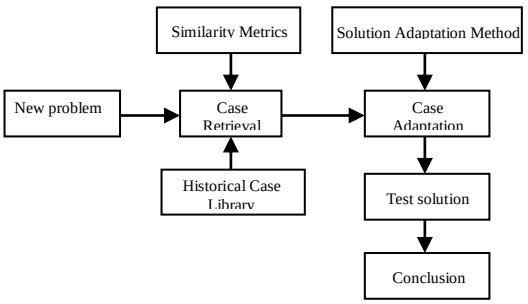

Figure2. The CBR process flow

3. Fuzzy set Approaches

The universe of discourse is the universe of all available information on a given problem.

A fuzzy set is a set containing elements that have varying degrees of membership in the set. Element in a fuzzy set, because their membership need not be complete , can also be member of other fuzzy sets on the same universe. So, fuzzy sets are vague and ambiguous.

This idea is in contrast with classical, or crisp, sets because members of a crisp set would not be members unless their membership was full, or complete, in the set (their membership is assigned a value of 1).

In this paper, Fuzzy set approach is used for retrieving previous case and estimate new prediction rainfall amounts using Mamdani (max-min) inference method.

3.1. Universe of Discourse

The complete range of values over which a variable can assume values is called the Universe of Discourse.

3.2. Linguistics variables

A fuzzy set that forms part of the term set for a variable and is used in fuzzy rules as part of a fuzzy relation is called a Linguistic variable. Rainfall types for monthly may have linguistic values such as Light Rain (LR) or Heavy Rain (HR) and Normal Rain (NR).

3.3. Fuzzy Inference technique

There are two major inference techniques in fuzzy. The first is the Mamdani (max-min) inference method and this system uses this method. Another one is called Sugeno inference method.

3.4. The Mamdani (max-min) inference method

The process of fuzzy inference system is consists of five parts. That is membership functions, Fuzzification, Rule evaluation, Aggregation of rule, and Defuzzification.

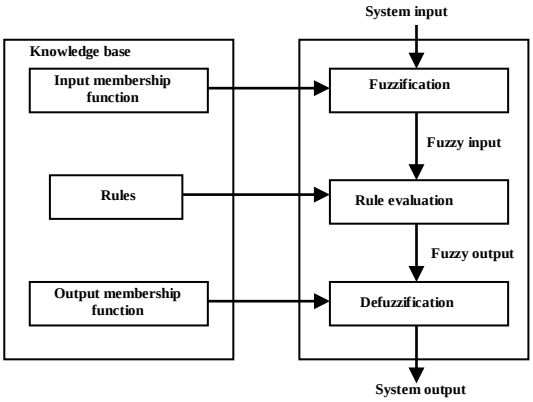

Figure3. Fuzzy inference system

3.4.1 Membership functions

Membership functions characterize the fuzziness and defuzziness in a fuzzy inference system or fuzzy set – whether the elements in the set are discrete or continuous –in a graphical form for eventual use in the mathematical formalisms of fuzzy set theory. There are four types of membership functions. They are Trapezoid, Gaussian, Singleton and Triangular membership functions. This system use Triangular membership functions.

3.4.2 Fuzzification

Fuzzification is the process of making a crisp quantity fuzzy .The process of finding the membership value of a number in the fuzzy set is called Fuzzification.

3.43 Rule Evaluation

If a given fuzzy rule has multiple antecedents, the fuzzy operation (AND or OR) is used to obtain a single number that represents the result of the antecedent evaluation. This number is applied to the consequent membership function.

Method: $\mu_{A \cap B \cap C}(x) = \text{Min}\,[\,\mu_A(x), \mu_B(x), \mu_C(x)]$

$\mu_{A \cup B \cup C}(x) = \text{Max}\,[\,\mu_A(x), \mu_B(x), \mu_C(x)]$

$\mu_A(x), \mu_B(x), \mu_C(x)$ = three fuzzy input values
$\mu_{A \cap B \cap C}(x)$ =minimum of three fuzzy inputs values
$\mu_{A \cup B \cup C}(x)$ =maximum of three fuzzy inputs values

3.4.4 Aggregation of rule outputs

Aggregation is the process of unification of the outputs of all rules. We take the membership functions of all rules. Consequents and combine them into a single fuzzy set. The input of aggregation process is the list of consequent membership functions, and the output is one fuzzy set for each output variable.

3.4.5 Defuzzification

The last step in the fuzzy inference process is Defuzzification. The process of converting the generation outcome fuzzy set to a single representative value is called Defuzzification. This system uses Center of Gravity (centroid) Defuzzification method.

3.5. Center of Gravity (COG) Method

This system uses the most popular method that is the COG or centroid method. Mathematically this COG can be expressed as:

$$COG = \frac{\Sigma^{n}_{i=1}\, \mu(x_i)\, x_i}{\Sigma \mu(x_i)}$$

i = number of values during Defuzzification

$\mu(x_i)$ = values of fuzzy number in fuzzy output membership function.

x_i = values of rainfall amounts in universe of discourse.

4. The proposed system

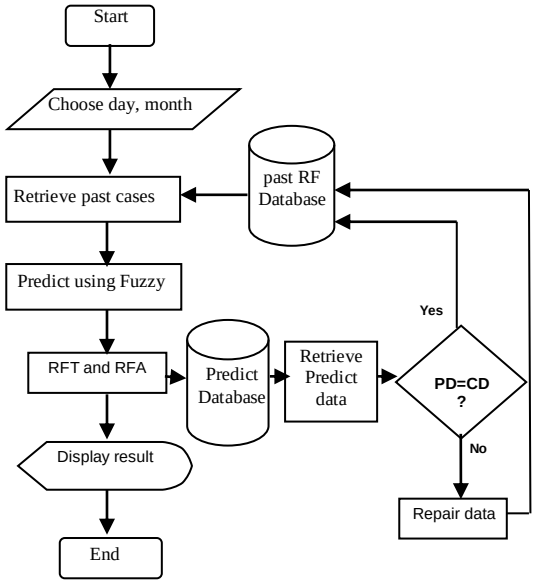

Figure 4. System flow diagram

Step 1. People who want to predict RF choose Daily or Monthly.
Step2. Then system retrieves previous three relevant cases with RFA from past RF dataset .
Step3. Fuzzified these three retrieving RFA to get fuzzy input by using triangular membership functions and calculate RFT that is used to get RFT result.
Step4. Execute these three fuzzy inputs using Mamdani (max-min) inference method and compute RFT using Fuzzy Association Rule Matrix .We get fuzzy output by aggregation of many rules to one.
Step5. If the fuzzy output is defuzzified with Center of Gravity (Centroid) defuzzification method, we get prediction rainfall amounts

(RFA) result. By reusing this RFA result, this system calculates prediction rainfall (RFT) types.

5. The System Implementation

The following are daily and monthly tables that is presented by fuzzy linguistic variables with various ranges.

Table1. Ranges for each month

Month name	LR	NR	HR
January	< 0.28 inches (7mm)	0.28 inches_0.43 inches	0.43 inches (11mm)<
February	< 0.04 inches (1mm)	0.04 inches_0.12 inches	0.12 inches (3mm)<
March	< 0.28 inches (7mm)	0.28 inches_0.43 inches	0.43 inches (11mm)<
April	< 1.30 inches (33mm)	1.30 inches_1.77 inches	1.77 inches (45mm)<
May	< 4.72 inches (120mm)	4.72 inches_5.75 inches	5.75 inches (146mm)<
June	< 5.91 inches (150mm)	5.91 inches_7.56 inches	7.56 inches (192mm)<
July	< 5.51 inches (140mm)	5.51 inches_6.77 inches	6.77 inches (172mm)<
August	< 7.09 inches (180mm)	7.09 inches_9.13 inches	9.13 inches (232mm)<
September	< 6.69 inches (170mm)	6.69 inches_7.95 inches	7.95 inches (202mm)<
October	< 4.21 inches (107mm)	4.21 inches_5.39 inches	5.39 inches (137mm)<
November	<1.57 inches (40mm)	1.57 inches_2.20 inches	2.20 inches (56mm)<
December	< 0.35 inches (9mm)	0.35 inches_0.59 inches	0.59 inches (15mm)<

Table2. Ranges for daily

Description	Notation or RFT	Ranges or RFA
Very Light Rain	VLR	<0.01 in/hrs (0.25mm/hrs)
Light Rain	LR	0.01in/hrs(0.25mm/hrs) _ 0.04 in/hrs(1.0mm/hrs)
Moderate Rain	MR	0.04in/hrs(1.0mm/hrs) _0.16 in/hrs(4.0mm/hrs)
Heavy Rain	HR	0.16in/hrs(4.0mm/hrs) _0.63in/hrs(16.0mm/hrs)
Very Heavy Rain	VHR	0.63in/hrs(16.0mm/hrs) _1.97in/hrs(450.0mm/hrs)
Extreme Rain	ER	1.97in/hrs(50.0mm/hrs)<

The following table 3 is fuzzy association rules for daily. Daily rainfall types are represented with association rules table.

Table3. Association Rules for daily

year 1	year 2	Year 3	Result
VLR	VLR	VLR	VLR
VLR	VHR	LR	LR,MR
VLR	VHR	MR	MR
LR	VLR	LR	LR
. .	. .	. .	. .
ER	ER	VHR	VHR,ER
ER	ER	ER	ER

Figure 5 is a fuzzy association rules for monthly. This monthly rainfall types are shown with association rules matrix.

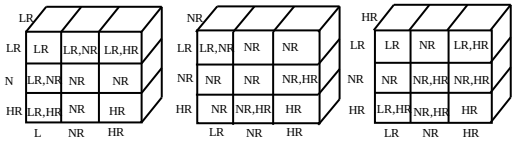

Figure5. Fuzzy association rules matrix for monthly

Table4.Linguistic variables and description for monthly

Rainfall Types (RFT)	Description
LR	Light Rain or No Rain
NR	Normal Rain
HR	Heavy Rain

Table 4 shows the linguistic variables and descriptions for monthly and table 5 show linguistic variables and descriptions for daily.

Table5. Linguistic variables and description for daily

Rainfall Types (RFT)	Description
VLR	Very Light Rain
LR	Light Rain
MR	Moderate Rain
HR	Heavy Rain
VHR	Very Heavy Rain
ER	Extreme Rain

The following figure 6 represents the calculation to get rainfall amounts and type using fuzzy Mamdani inference method in Case –Based Reasoning (CBR) system.

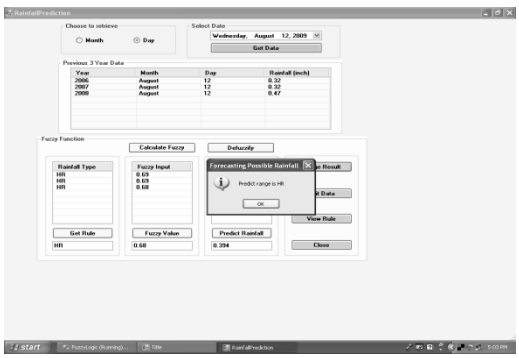

Figure6. Form of Rainfall prediction using fuzzy in CBR

The following figure 7 shows rules and ranges for daily and monthly.

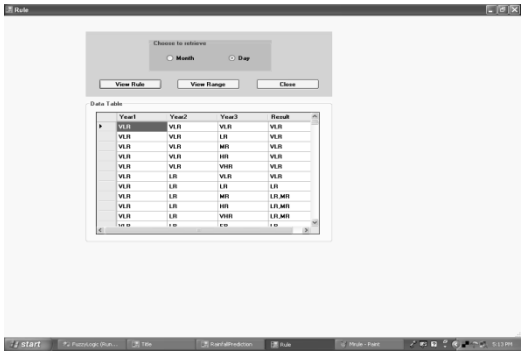

Figure7. Show rules and ranges form

6. Conclusion

In this paper, we proposed to predict new rainfall case by using previous cases. Fuzzy set approach is used in case retrieving and reusing in CBR. This paper intends to develop the monthly and daily prediction for Loikaw. The limitation of this system are the daily and monthly amounts must enter greater than and equal zero and that have at least three years rainfall data.

References

[1] A .Aamodf, E.Plaza (1994): **Case-Based Reasoning: Foundational Issues, Methodology Variations , and System Approaches AI communications**. IOS press, Vol.7:1, pp.39-59.

[2] Bjarne K.Hansen and Denis Riordan, **'Weather Prediction Using Case –Based Reasoning and Fuzzy Set Theory'**, Faculty of Computer Science, Dalhousie University, Canada.

[3] EARL COX, **'Fuzzy Modeling and Genetic Algorithm for Data Mining and EXploration'**.

[4] Jiawei Han, Micheline Kamber **: 'Data Mining Concepts and Techniques'** , Siamon Fraser University , USA.

[5] Prof. Gheorghe Tecuci, **'CS 782 Machine learning', 10 Case-Based Reasoning and Learning**, Learning Agents Laboratory, computer science department, George Mason University.

[6] Pradipta Biswas, **'Introduction to Case-Based Reasoning'**, Indian Institute of Technology Kharagpur ,India.

[7] TIMOTHY J.ROSS,**'Fuzzy Logic With Engineering Applications'**, McGraw-Hill, Inc, New York, 1995.