# Hierarchical Fuzzy Case Based Reasoning with Multi-criteria Decision Making for Financial Applications

Shanu Sushmita and Santanu Chaudhury

Electrical Engineering Department, Indian Institute of Technology Delhi,
New Delhi, India - 110016
{schaudhury,shanusushmita}@gmail.com

**Abstract.** This paper presents a framework for using a case-based reasoning system for stock analysis in financial market. The unique aspect of this paper is the use of a *hierarchical* structure for case representation. The system further incorporates a *multi-criteria* decision-making algorithm which furnishes the most suitable solution with respect to the current market scenario. Two important aspects of financial market are addressed in this paper: *stock evaluation* and *investment planning*. CBR and multi-criteria when used in conjunction offer an effective tool for evaluating goodness of a particular stock based on certain factors. The system also suggests a suitable investment plan based on the current assets of a particular investor. Stock evaluation maps to a flat case structure, but investment planning offers a scenario more suited for structuring the case into successive detailed layers of information related to different facets. This naturally leads to a hierarchical case structure.

## 1 Introduction

In this paper, we propose an application framework involving a fuzzy case based reasoning system with a hierarchical case structure and a multi-criteria decision making. So far, case structures used in majority of CBR applications have had a flat case structure which, to some extent, manages to incorporate example applications that can be mapped to flat case structures. Many of the complex real world problems require a framework capable of handling non-summarized information where one component's value is dependent on several other relevant factors. Investment planning is an example application having non-summarised information, which requires classification of case features, based on different aspects, into successive layers. This segregation of information into layers leads to a hierarchical case structure. Also, combining hierarchical CBR with multi-criteria is all together a new approach towards decision making systems tailored for applications having a dynamic and complex nature.

Past did witness some of the evolving research work, which looked into the financial areas and suggested how CBR can be efficiently used as a tool to produce a suitable decision making system. Some of the work like [1] addressed the movement of the stock market and its prediction. [2] proposed the daily financial condition indicator (DFCI) monitoring financial market built on CBR, [3], in their work proposed a new learning
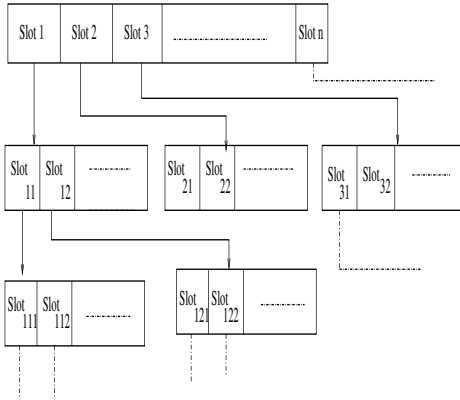
Fig(a) Case Representation
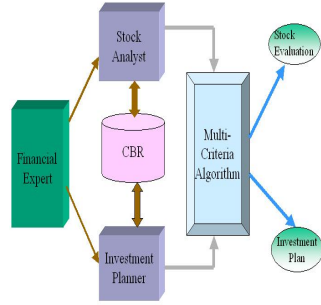
Fig(b) Overview Model

**Fig. 1.** Hierarchical Representation Of Case Structure and the System Model

technique which extracts new case vectors using Dynamic Adaptive Ensemble CBR, which again deals with the prediction of the overall stock market. Author [4] presents case-based reasoning approach for financial crises warning, which is again an approach that works on the broader level of financial issues. In other words, past research did manage to explore into the domain of financial market but, most of the research were confined to the working of the overall market situation and their prediction. Here, in this paper, we investigate use of CBR as an intelligent tool in considering and analysing an individuals stock with respect to current market scenario and investment planning of an individual. -Also, as stated before, CBR applications so far have always been mapped to flat case structures, while applications more suited to hierarchical case structures have not been explored. Although, work like [5] and [6] presented the concept of a hierarchical CBR, but the concept proposed involve reusing of multiple cases at various levels of abstraction. In [7] a CBR system with a two-level weight specification is developed. Our representation differs from the above mentioned approaches in terms of case representation style, where each case is a collection of independent sub cases.

The organisation of this paper is as follows: Section 2 gives a detailed explanation of the Case structure and working of the hierarchical case based reasoning system. Section 3 describes the implication of the proposed framework in financial market. In Section 4, we present the implementation details and the results of the two applications: Stock evaluation and Investment Planner.The last section presents the conclusion and future work.

## 2   Fuzzy Hierarchical Case Structure

**Case - Structure:** For our application, we have modified the classical flat case structure, such that each case is represented as a collection of sub-cases as shown in Fig. 1(a). Classical CBR uses symbolic and/or numeric attributes. However, many of the descriptors characterising the real world problems are associated with certain degree of

fuzziness and uncertainty. Fuzzy CBR is a methodology which uses linguistic or realistic variables for case representation. It emulates human reasoning about similarity of real-world cases, which are fuzzy. Each sub case consists of its specific attributes, which could be fuzzy or non-fuzzy or complex (consisting of a sub case).

**Working Methodology:** Working methodology of the similarity computation is based on our past work [8]. The proposed framework operates in three different stages. In the first stage, similar cases are retrieved using the indexing attributes. In the second stage, we consider each retrieved case and compute a similarity score at every subsequent level for each child slot (referred also as a *node* in the context of a tree structure) of the respective parent slot. In third stage similarity scores are combined using weighted T-Norm for fuzzy attributes, and using Weighted similarity score for crisp attributes. Figure 2 represents the similarity computation of cases having hierarchical structure. A generalised representation of this computation is given in equations 1, 2 and 3.
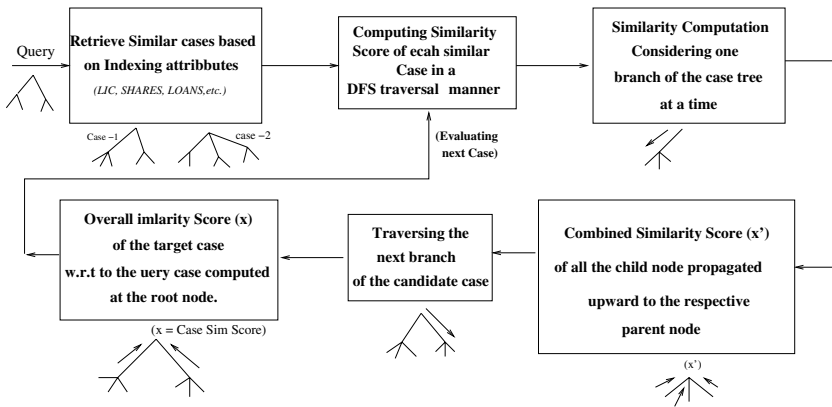


**Fig. 2.** Steps showing Similarity Score Computation of similar cases w.r.t the Query Case.(Similarity scores are combined using weighted T-Norm for the fuzzy child nodes and weighted summation for the crisp attributes.

**Similarity Score Computation:** For Crisp slots (non-fuzzy), similarity score of each node with respect to the query node is computed by exact matching, string comparison, and numerical comparison algorithms. The selection of the matching algorithm depends on the data type of the slot. At each parent slot, the aggregate similarity taking into account the non-fuzzy child slots is evaluated by a weighted summation of similarity of all the non-fuzzy child slots.

We show below the mathematical formulation for the evaluation of similarity score of a case with the hierarchical structure. Let $N_k^{\{l\}}$ denote the $k^{th}$ node at the level $\{l\}$. The node $N_k^{\{l\}}$ can have child nodes which could be fuzzy or non-fuzzy. The sets of fuzzy child nodes and crisp child nodes of $N_k^{\{l\}}$ are denoted as $fuzzyCh(N_k^{\{l\}})$ and $crispCh(N_k^{\{l\}})$ respectively. Let $s_k^{\{l\}}$ denote the aggregation of the similarity computed for the node $N_k^{\{l\}}$ and $w$ be the weight. The similarity score at level $\{l\}$, that is $s_k^{\{l\}}$ has
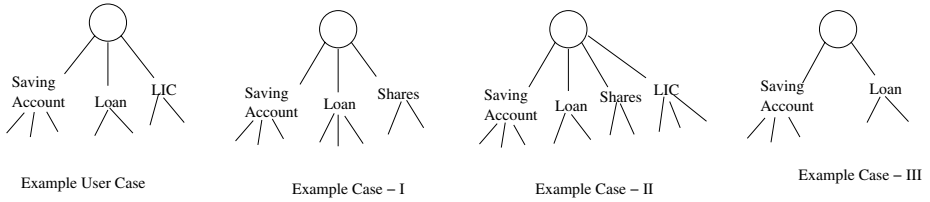
**Fig. 3.** The User case(query) represented above has three features (*Saving Account, Loan, LIC*. The database Case II has features (*Shares, Loan, LIC, Saving Account*), of which feature *Shares* is not there is the query case. This attribute(feature) and its detail expansion is not used in similarity computation. Situations where the number of similar nodes in the example cases become same, for example in case I and Case III, the computed similarity score would depend on the attribute value of the leaf nodes.

two components: fuzzy similarity score and crisp similarity score. The fuzzy component of the similarity score $s_k$ is denoted as $fuzzySim(s_k)$ and the crisp component is denoted as $crispSim(s_k)$. We have:

$$fuzzySim(s_k^{\{l\}}) \quad = \quad \frac{\sum_i s_i^{\{l+1\}} w_i}{\sum_i w_i} \tag{1}$$

$$crispSim(s_k^{\{l\}}) \quad = \quad \frac{\sum_i s_i^{\{l+1\}} w_i}{\sum_i w_i} \tag{2}$$

where the summation index $i$ varies for all the nodes $N_i^{\{l+1\}} \in crispCh(N_k^{\{l\}})$. We have:

$$s_k^{\{l\}} \quad = \quad \frac{fuzzySim(s_k^{\{l\}}) \times w_f \ + \ crispSim(s_k^{\{l\}}) \times w_c}{w_f + w_c} \tag{3}$$

where the weights $w_f$ and $w_c$ are sums of weights for fuzzy and crisp slots respectively.

Finally, we select the similar cases having a score above a threshold $\beta$. Thus we see that the similarity measure of a node is a weighted average of the similarity measure for its child nodes (both crisp and fuzzy).

**Multi-criteria Decision Making Algorithm:** Past researches have led to some useful multi-criteria decision making methods. ELECTRE [9], [10] and PRMETHEE [11] are some of the ranking function that could be used to outrank one similar case over the other based on certain parameters. In this paper, we have applied the ISD [10] method to furnish the most preferred solution, which considers the various criteria that affect the goodness of a particular stock or investment plan with respect to the current market situation. ISD($x_i$,$x_j$) measures the degree of superiority or dominance of $x_I$ over $x_j$ in respect to all multiple attributes (or criteria) and all other cases.

## 3 Application Description

The application we propose primarily focuses on two important implications of financial market: stock evaluation and investment planning. Our system is a useful tool both
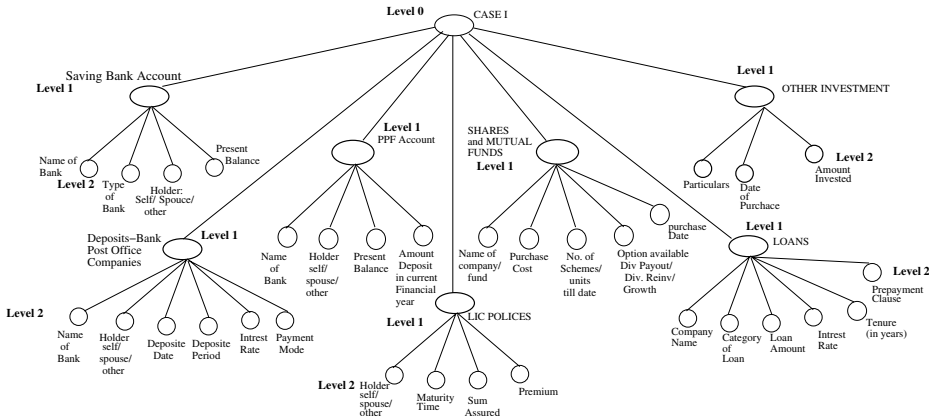
Level 0 CASE I

Saving Bank Account
Level 1

Level 1
OTHER INVESTMENT

Name of Bank
Level 2 Type of Bank
Holder: Self/ Spouse/ other
Present Balance

Level 1
PPF Account

SHARES and MUTUAL FUNDS
Level 1

Particulars
Date of Purchace
Level 2 Amount Invested

Deposits−Bank Post Office Companies
Level 1

Name of Bank
Holder self/ spouse/ other
Present Balance
Amount Deposit in current Financial year

Name of company/ fund
Purchase Cost
No. of Schemes/ units till date
Option available Div Payout/ Div. Reinv/ Growth

purchase Date
Level 1 LOANS

Level 2 Prepayment Clause

Level 2
Name of Bank
Holder self/ spouse/ other
Deposite Date
Deposite Period
Intrest Rate
Payment Mode

Level 1 LIC POLICES

Company Name
Category of Loan
Loan Amount
Intrest Rate
Tenure (in years)

Level 2 Holder self/ spouse/ other
Maturity Time
Sum Assured
Premium

**Fig. 4.** Hierarchical Representation Of Case - Investment Planner: Example cases in the case base may have structure as shown above, or they may be an instance of the case structure having fewer branches and nodes, as shown in the figure 3 (Example Case -I) and (Example Case -II). Here, example Case-I possesses attributes, *Saving Account*, *Loan* and *Shares*. However, example Case -II consists of one more attribute (branch), *LIC Policy*.

for the people who are familiar to this field (stock analyst) as well as people who are actually not finance experts, but might be interested to buy a share or have an investment plan for themselves. The overview of the proposed methodology for this application is shown in Fig. 1(b).

**Stock Analyst:** Stock analyst works by taking values for parameters, which play an important role in determining the worth of any particular stock like: - Net sale, PAT, EPS, Growth percent, RoE, RoCE, etc. These values form the problem query of the case. If the most similar case suggests the query stock values to be good in past, do these values still continues to be good with respect to today's market status? To answer this question, the selected similar cases are assessed using multi-criteria algorithm, which makes use of certain criteria to determine their fitness in the present time.

**Investment Planner:** It works by taking inputs in two stages. In the first stage it takes in the financial goals of the person. In the second stage, it considers the worth of the current financial assets of that person.Once the information for both the stages are received, an investment plan is suggested which maps various assets to different financial goals. The Investment planner uses a hierarchical case structure as shown in Fig. 4. Initial retrieval, for the similar cases, is done by taking financial goals as indexing slots. Each retrieved cases, has a similarity score of one, since initially cases are retrieved by one to one mapping of each indexing slot. Overall similarity of the Case is done at each subsequent level by using equation 1, 2 and 3.

## 4 Implementation and Results

Visual C++ is used as the front end to create user interfaces and database handling. Microsoft Access servers as the back end to store cases in case base. When the stock

analyst receives the query stock with its values for various parameters it compares the *query stock* against the stored stocks in the case base having similar values. Example stocks are selected based on qualitative or quantitative analysis, depending on the nature of the attribute. Finally the result, a suggestion regarding the query stock is given as a *"A Good Buy"* , or as a *"Not a Good buy"* or *"hold the stock"*  if already possessing it.

| Company name | Sector | Net Sales (Rs. m) | PAT (Rs. m) | EPS | EPS (Growth %) | P / E | RoE(%) | RoCE(%) | EV/ Sales | EV/ EBDIT |
|---|---|---|---|---|---|---|---|---|---|---|
| WIPRO | Information Technology | 44910 | 10964.60 | 47.8 | HIGH | 37.9 | HIGH | 39.7 | 13.9 | 47 |

**Fig. 5.** Query Stock

**Stock Analyst:** In Response to a query stock, similar stocks are retrieved. These similar stocks are compared to obtain the stock having maximum preference, based on ranking criteria. How well these stocks fit in today's market was decided by considering the various criteria:

– Time: The current the better,
– Sector of the company: Should belong to same sector: IT, Fertilizer, etc,
– Percentage holding of the company: Comparing the market share,
– Balance sheet of the company: Comparing the assets and liabilities,
– Return over capital employed(RoCE%): Stocks having similar values are preferred,
– Public holding in the company,
– Value of Rupee against Dollar: The difference should be somewhat similar.

| Company name | Sector | Net Sales (Rs. m) | PAT (Rs. m) | EPS | EPS (Growth %) | P / E | RoE(%) | RoCE(%) | EV/ Sales | EV/ EBDIT |
|---|---|---|---|---|---|---|---|---|---|---|
| WIPRO | Information Technology | 44900 | 11964.52 | 47.8 | HIGH | 37.6 | HIGH | 38.5 | 13.9 | 45 |

(a) Retrieved Case 1

| Company name | Sector | Net Sales (Rs. m) | PAT (Rs. m) | EPS | EPS (Growth %) | P / E | RoE(%) | RoCE(%) | EV/ Sales | EV/ EBDIT |
|---|---|---|---|---|---|---|---|---|---|---|
| TCS | Information Technology | 45822 | 11000.20 | 47.8 | 36.7 | 37.9 | 37.9 | 39.7 | 13.9 | 40 |

(b) Retrieved Case 2

| Company name | Sector | Net Sales (Rs. m) | PAT (Rs. m) | EPS | EPS (Growth %) | P / E | RoE(%) | RoCE(%) | EV/ Sales | EV/ EBDIT |
|---|---|---|---|---|---|---|---|---|---|---|
| POLARIS | Information Technology | 43322 | 10500.34 | 47.0 | 35.4 | 36.98 | 37.9 | 40.0 | 14.01 | 42 |

(c) Retrieved Case 3

**Fig. 6.** Shows 3 retrieved stocks for the query of Fig. 5. Stock (a) was selected since it had the maximum Dominance

| Goals | Purchase or construction of house | Major repairs or renovation | Your own Education | Education of children | Your own Marriage | Marriage of Children | Providding for medical or emergencies | Purchase of car or other vehicle | overseas Vacations | Retirments needs |
|---|---|---|---|---|---|---|---|---|---|---|
| Pririty | HIGH | LOW | LOW | LOW | HIGH | LOW | MEDIUM | HIGH | VERY HIGH | HIGH |
| Years to Go | 20 | 5 | – | -- | 2 | 25 | ANY | 2 | 1 | 30 |
| Present Value of worth | 40,00,000 | 5,00,000 | -- | -- | 2,00,000 | 10,00,000 | 50,000 | 8,00,000 | 5,00,000 | 20,00,000 |

**Fig. 7.** Financial Goals Described - Query(phase 1)

**Short-term liquidity:** Interest on deposits with banks is no more eligible for deduction from income taxes under section 80L of the IT Act. And considering considering our existing investments, we do not recommend you any fixed deposits.

**Life Insurance:** We recommend term insurance having coverage of Rs 20 Lacs coverage of Rs 20 Lacs with a term of 20 years for you.

**Housing Loan:** We recommend housing finance for 85% of the housing requirement i.e. 34 Lacs @ (Rate of Interest) 7.75% in year 2026 with a monthly EMI of Rs 40,804/–. However it needs to be reviewed periodically to ensure the rate of interest being charged is as per the market norms.

**Car Loan:** We recommend car loan for 100% of car requirement, i.e. 8 lacs @ 10% for 5 years in year 2008 with a monthly EMI of Rs 16,998/–. The same needs to be reviewed periodically to to ensure the rate of interest is charged as per market norms.

**Assumptions**

1. Growth in Savings: Your savings are assumed to grow by 8%. Your current savings have been taken at Rs. 3 Lacs per annum approx. (excluding contribution to EPF), as discussed with you.

2. Inflation and interest rates: are projected to be 5% and 7% respectively in the long term.

3. Reinvestments: Investments made will not be withdrawn except for their earmarked purpose. Likewise, earnings on investments will remain reinvested till a goal is met.

4. Existing Portfolio: We have considered your existing portfolio of Rs. 4.21 Lacs (including your Bank Balances) during the construction of the plan.

| Regular Invesment | Amount (Rs) |
|---|---|
| MIP Funds | 200000 |
| Insurance/Pension Premiums | 136170 |
| Equity/Balanced Mutual Funds | 60000 |
| Term Insurance | 7000 |
| Total (in Rs) | 403170 |

**Fig. 8.** Recommended Investment Plan based on user profile and priorities

The aforementioned parameters are considered to evaluate the dominance of one retrieved stock over another, in response to the query stock. Selected Stock which got the highest rank amongst the similar stocks is shown in Fig. 6.

**Investment Planner:** The criteria used by the investment planner to rank the retrieved similar cases in response to a query case are as follows:

- Time: current the better,
- Risk assessment: Cases having similar risk tolerance are preferred.
- Priority of the goals,
- Age factor: Cases should belong to the recommended age group,
- Number of dependants,
- Profession: Self employed or salaried.

When the query (shown in Fig. 7(user's goals input) and 9(user's current assets worth) (A) is posed to the investment planner, the case (shown in Fig. 9) (B) having second highest score was selected.After applying the multi-criteria attributes on the selected cases,the case shown in figure 9 (B)dominated the other candidate cases in terms of its similarity as well as usefulness in present time. Profile and the priority(age=34, house plans (in 5 years) , retirement plans(high priority) and child education (two children)) of the query user had the maximum similarity with the goals and profile of the selected Example case (age=40, house plan (in 3 years), retirement plan(high priority), child education(two children)).The case shown in Fig. 9 suggested the Investment plan, shown in Fig. 8, for the posed query.

Having a case base size of 250 example case for Investment Planner and over 500 for Stock Analyst, we have achieved a reasonable performance of the proposed application. The results obtained were discussed with few expert stock analysts and they found them

| Holder: Self/Spouse | Time to Maturity | Sum Assured | Premium | Table Term |
|---|---|---|---|---|
| Self | 20 | 1,00,000 | 4,422 | 107–20 |
| Self | 15 | 1,00,000 | 5,325 | 106–15 |

| Name of Bank/Post Office | Holder: Self/Spouse | Deposite Period | Deposite Amount | Interest Rate | Payment mode: monthly/yearly |
|---|---|---|---|---|---|
| ICICI | Self | 36 | 5000 | 3 | Monthly |

| Name of Bank | Type of Account | Holder: Self/Spouse | Present Balance |
|---|---|---|---|
| ICICI | Savings | Self | 30,000 |
| HDFC | Savings | Self | 2,50,000 |

| Name of Company | Loan Category House/car/personal | Loan amount | Monthly Premium | Rate of Interest | Tenure (in years) | Prepayment clause |
|---|---|---|---|---|---|---|
| ICICI | House | 1,200,000 | 13,208 | 10.75 | 15 | Partial payment |
| ICICI | House | 450,000 | 4198 | 7.5 | 20 | Partial Prepayment |

(A) Current assets Described (Query Phase –II)

| Name of Company | category of Loan House/car/personal | Loan amount | Monthly Premium | Rate of Interest | Tenure (in years) | Prepayment clause |
|---|---|---|---|---|---|---|
| ICICI | House | 1,000,000 | 13,208 | 10.75 | 17 | Partial payment |
| ICICI | House | 450,000 | 4198 | 7.5 | 20 | Partial Prepayment |
| HDFC | Car | 2,00,000 | 1,000 | 2.5 | 4 | Zero Prepayment |

| Name of Bank/Post Office | Holder: Self/Spouse | Deposite Period | Deposite Amount | Interest Rate | Payment mode: monthly/yearly |
|---|---|---|---|---|---|
| SBI | Self | 36 | 4500 | 3 | Monthly |

| Name of Bank | Type of Account | Holder: Self/Spouse | Present Balance |
|---|---|---|---|
| ICICI | Current | Self | 32,563 |
| HDFC | Savings | Spouse | 2,50,000 |

(B) Retrieved Inverstment Plan with Highest Ranking

**Fig. 9.** Query Phase(II) and the selected Investment Plan with Highest Ranking

to be satisfactory 75% of the time. For certain set of test data the prediction made by the system were also compared with the predictions made by analysts themselves.We found the comparitive results to be quite satisfying. Although, some of the results were not justified, which we conclude were due to the fact that for such kind of application dealing with fluctuating market trend a more updated and large case base would be required. The proposed suggestion fits convincingly with going market trend. Also, the Investment Planner produces logically convincing Investment plan in response to the query.

## 5 Conclusion and Future Work

Having a hierarchical case structure offers a better representation of the complex details. This provides a clearer understanding of the application while taking comprehensive details into account. As a result the proposed solution suits more appropriately with the actual requirements of the user. As a future work, we propose to look at the next few other important applications of financial market, like mutual funds.

## References

1. Kim, K.-J.: Toward Global Optimization of Case-Based Reasoning Systems for Financial Forecasting. Applied Intelligence 21, 239–249 (2004)
2. Kyong Joo Oh, T.Y.K.: Financial market monitoring by case-based reasoning. Expert Systems with Applications 32 (2007)
3. Chun, Y.J.P.S.H.: Dynamic adaptive ensemble case-based reasoning: application to stock market prediction. Expert Systems with Applications 28, 435–443 (2005)
4. Ying Wei, Y.W., Li, F., Li, F.: Case-Based Reasoning: An Intelligent Approach Applied for Financial Crises Warning. Springer, Heidelberg (2003)

5. Smyth, P.C.B., Keane, M.T.: Hierarchical case-based reasoning integrating case-based and decompositional problem-solving techniques for plant-control software design. Knowledge and Data Engineering (2001)

6. Tang, Y.: Time series extrapolation using hierarchical case-based reasoning. Signal Processing, Pattern Recognition, and Applications (2006)

7. Chi-I Hsua, P.-L.H., Chiub, C.: Predicting information systems outsourcing success using a hierarchical design of case-based reasoning (2004)

8. Singh, T., Goswami, S.C.P.S.: Distributed fuzzy case based reasoning. Elsevier:Journal of Applied Soft Computing 4, 323–343 (2004)

9. Roy, B.: The outranking approach and the foundations of ELECTRE methods ch in book: Readings in Multiple Criteria Decision Aid. In: Costa, C.A.B.e (ed.) Springer, Berlin (1990)

10. Pedro, J.S., Burstien, F.: A framework for case-based fuzzy multi-criteria decision support for tropical cyclone forecasting. In: Proceedings of the 36th Hawaii International Conference on System Sciences(HICSS), IEEE, Los Alamitos (2002)

11. Brans, J.P., Mareschal, B.: The PROMETHEE methods for MCDM; The PROMALC GAIA and BANKADVISER softwarech in book: Readings in Multiple Criteria Decision Aid. In: Costa, C.A.B.e. (ed.) Springer, Berlin (1990)