A fuzzy-based algorithm for auditors to detect elements of fraud in settled insurance claims

Jagdish Pathak
Accounting Systems and IT Auditing Area, Odette School of Business Administration, University of Windsor, Windsor, Ontario, Canada
Navneet Vidyarthi
Department of Management Sciences, University of Waterloo, Waterloo, Ontario, Canada, and
Scott L. Summers
School of Accountancy and Information Systems, Brigham Young University, Provo, Utah, USA

Abstract
Purpose – In the current global economy, the survival of an insurance company depends on its ability to respond to the customer demands. One of the demands of customers is efficient settlement of insurance claims. All insurance companies face the conflicting goals of authenticating claims and settling claims quickly. The use of human adjusters in claim settlement process leaves room for subjective judgment and the use of discretion while finalizing a claim. It has been observed that the opportunity exists for claim adjustors to settle insurance claims in favor of the claimants simply by colluding with the claimant and sacrificing the monetary interest of the insurers. The increasing cost of human experts for authentication (fraud detection) has led many companies to develop technological solutions such as expert systems to assist in the authentication of processed claims.

Design/methodology/approach – We have used fuzzy math in combination with the expert systems technology to design this model system.

Findings – We develop a fuzzy logic based expert system that can identify and evaluate whether elements of fraud are involved in insurance claims settlement.

Research limitations/implications – We could not obtain real life data from any one of the insurance companies even after various attempts in Canada. Canadian Privacy Legislation does not permit these organizations to share them with any one.

Practical implications – This expert system can help decide if settled claims are genuine or if an element of fraud might exist which needs substantive testing by an auditor. The proposed methodology has been illustrated with an example that tends to model insurance claims in general.

Originality/value – The model designed in this paper is original and carries a substantial value to internal/external auditing professional who has access to these data to train the inference engine.

Keywords Auditing, Fuzzy logic, Fraud

Paper type Research paper

1. The introduction to problem and its significance

The insurance industry is faced with two significant enemies: the ever-spiraling cost of claims and the slumping rate of return on investment (Zyl, 2003). Insurers find it difficult to raise their return on investment due to low rate of interest in the market (Greenwald, 2003). In this difficult scenario, it becomes imperative for insurers to make

Managerial Auditing Journal
Vol. 20 No. 6, 2005
pp. 632-644

0268-6902
DOI 10.1108/02686900510606119

sure that every dollar spent in claim settlement, irrespective of its size, is genuine in all respects. They must assess whether a single person or group of persons (for detailed study, please refer to Derrig and Ostaszewski, 1995) are attempting to obtain false claims through some external illegitimate means or by entering into collusion with the claim adjustors of insurance firms. Our paper is an effort in the direction to design an expert system based on fuzzy logic to help internal auditors identify claims that carry an element of fraud from amongst all settled claims. The algorithm designed in this research is tested on hypothetical data. This algorithm and methodology is also compatible with the continuous auditing paradigm. Some recent applications of fuzzy logic based techniques include pattern recognition (Bezdek, 1981), automobile body injury claims fraud (Brockett *et al.*, 1995, 1998, 2002; Tennyson and Salsas-Forn, 2002), property liability insurance claims costs (Cummins and Derrig, 1993, 1994), risk assessment of management fraud (Deshmukh and Lakshminarayana, 1998; Crocker and Tennyson, 2002). Other artificial intelligence techniques have also found applications in detecting fraud in various settings. Just to mention a few, Fanning and Cogger (1998), have reported on the development of neural network based approach for detection of management fraud using the published financial data, whereas Welch *et al.* (1998) have used genetic algorithm based classifier for modeling auditor decision behavior in fraud settings. Wheeler and Aitkin (2000) have developed algorithms for fraud detection. However, these algorithms cannot be applied in detecting fraud while processing insurance claims.

The suggested algorithm and the resultant system is expected to make use of the internal database of the insurer coupled with the external jointly supported database of major insurers, like Insurance Services Office (ISO) incorporated in the USA. The use of expert systems in the insurance industry is not new as many major insurers in developed countries of Europe and North America make use of these expert systems for various purposes (Mogel, 2003; Guidera, 2003; Hoffman, 1999).

The expert system algorithm designed in this research can be put to effective use by any outsourced fraud auditor to assess the control risk as well as detection risk to some extent of an audit engagement on any particular client. This system will evaluate and judge the work performed by call center employees as well as claim adjustors employed by insurers while assessing and deciding the claims. This system is assumed to be of tremendous importance in adjudicating the claims accepted by the insurer. Theoretically, the suggested expert system can evaluate any claim irrespective of its size, already settled by the insurers, but in this paper we have restricted it to a specific dollar amount claim threshold in the illustration, which is expected to take care of small claims arising out of insured causes.

2. The method

The fuzzy logic expert system for detecting fraud in insurance claims developed in this project is designed around the "index of ambiguity" concept. This index reflects the degree of ambiguity/incompleteness in the information furnished by the claimant and the information gathered by the claim adjustor s from various other sources. The index is used as a measure of authenticity or fraud element in the insurance claims adjustment. Because of the subjective nature of the parameters used in the evaluation of the claims, fuzzy parameters have been used in order to quantify the ambiguity associated with information. An index known as "ambiguity index" has been formulated in order to identify elements of fraud while settling such claims.

Fuzzy logic is a powerful technique for dealing with human reasoning and decision making processes which involve inexact information, approximation, uncertainty, inaccuracy, inexactness, ambiguity, vagueness, qualitative-ness, subjectivity, perception, or sources of imprecision that are non-statistical in nature. By applying fuzzy logic, we can quantify the contribution of a set of information to various parameters in terms of fuzzy membership. The foundation of fuzzy logic is fuzzy set theory, which was introduced by Zadeh (1965). During the past few years, fuzzy logic has emerged as an attractive tool for various applications ranging from finance, traffic control, cement kiln, automobile speed control, to nuclear reactor, and earthquake detections.

Unlike the classical binary values, where zero or one values are used to denote the membership, fuzzy set theory makes the use of a multi-valued membership function to denote the membership of an object. If A is the classical set of objects whose elements are denoted by x, then membership in a classical subset of X of A is often viewed as characteristic function of μ_x from A to a valuation set $\{0, 1\}$, such that $\mu_x(x) = 1$ if and only if $x \in X$, zero otherwise, where X is called a fuzzy set if the valuation set is allowed to be in the real interval of $[0, 1]$ and $\mu_x(x)$ is the grade of membership of x in X. The higher $\mu_x(x)$, the more x belongs to X. For more details and better understanding of fuzzy set theory and fuzzy logic, the readers are referred to Zimmerman (2001) and Ross (1997).

3. Architecture of the fuzzy logic based expert system

In this paper, we develop a fuzzy logic based expert system for fraud detection in insurance claims. Figure 1 shows the control mechanism of the system. The fuzzy logic based expert system consists of four components: fuzzifier, inference engine, defuzzifier, and the rule base. In the fuzzifier, crisp inputs are fuzzified into linguistic values to be associated to the input linguistic variables. After fuzzification, the inference engine refers to the knowledge/rule base containing fuzzy IF-THEN rules to

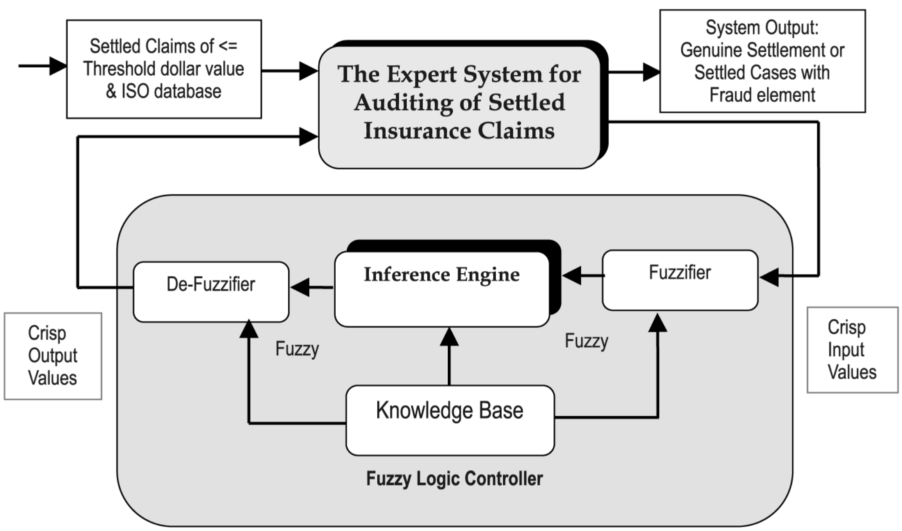

Figure 1.
The fuzzy logic based expert system for fraud detection

derive the linguistic values for the intermediate and output linguistic variables. Once the output linguistic values are available, the defuzzifier produces the final crisp values from the output linguistic values.

The fraud detection process begins by entering two sets of data; one furnished by the claimant and the other obtained by the auditor reviewing the claim. This information is obtained from a standard claim form that is used by the insurance company. The form contains different sections, each section containing a set of information such as personal information of the claimant, financial information, and nature of the claim etc. All the settled claims that are below a certain value (predetermined based on the policy of the auditors) are settled immediately, whereas claims involving high amount are considered for verification by the expert system. The expert system verifies the authenticity of the settled claim by using various measures such as: ambiguity index, degree of incomplete information of the claim, level of discretion used by the claim settlers in settling the claim. These qualitative measures are quantified and converted into linguistic variables with corresponding membership functions. For example, the ambiguity index for the information section i is given by:

$$X_1 = \frac{\left(\sum_{i=1}^{I} \sum_{j=1}^{J(i)} W_{ij} \Delta_{ij} \right)}{I},$$

where W_{ij} is the weightage or impact factor given to the jth information of the ith section, and Δ_{ij} is a 0-1 variable, where $\Delta_{ij} = 1$ if there is any deviation/difference in the information furnished by the claimant and the one obtained by the auditor, 0 otherwise. It is worthwhile noting that the information that is crucial in deciding the settlement of the claim is given higher weightage/impact factor. Also all the weights for a set of information i,

$$\sum_{j=1}^{J(i)} W_{ij}$$

add to unity. Similarly, the values of the other inputs are determined. The normalized values of these measures are used as inputs to the expert system. The degree of membership corresponding to a value of input is determined by plugging the selected input parameter into the horizontal axis and projecting vertically to the upper boundary of the membership function(s). These membership functions are designed based on information content. The input parameters are indicated by triangular fuzzy numbers because of their simplicity and good result obtained by simulation. Figure 2 shows the definition of the fuzzy sets of the input and the output functions. A rule base is then constructed based on all the applicable input parameters. For each decision several rules are fired. Table I shows a sample rule base for the system under consideration. We emphasize the fact that in real life situations, the expertise of the human auditors will be used in the construction of the rule base. These rules result in an aggregate fuzzy set that represents a particular decision regarding the processing of the claims. This fuzzy set is then converted into a crisp number, which depicts the degree of suitability of the decision regarding the processing of the claims. The rules aggregation is done using center-of-area (COA) method. Mamdani implication is used to represent the meaning of

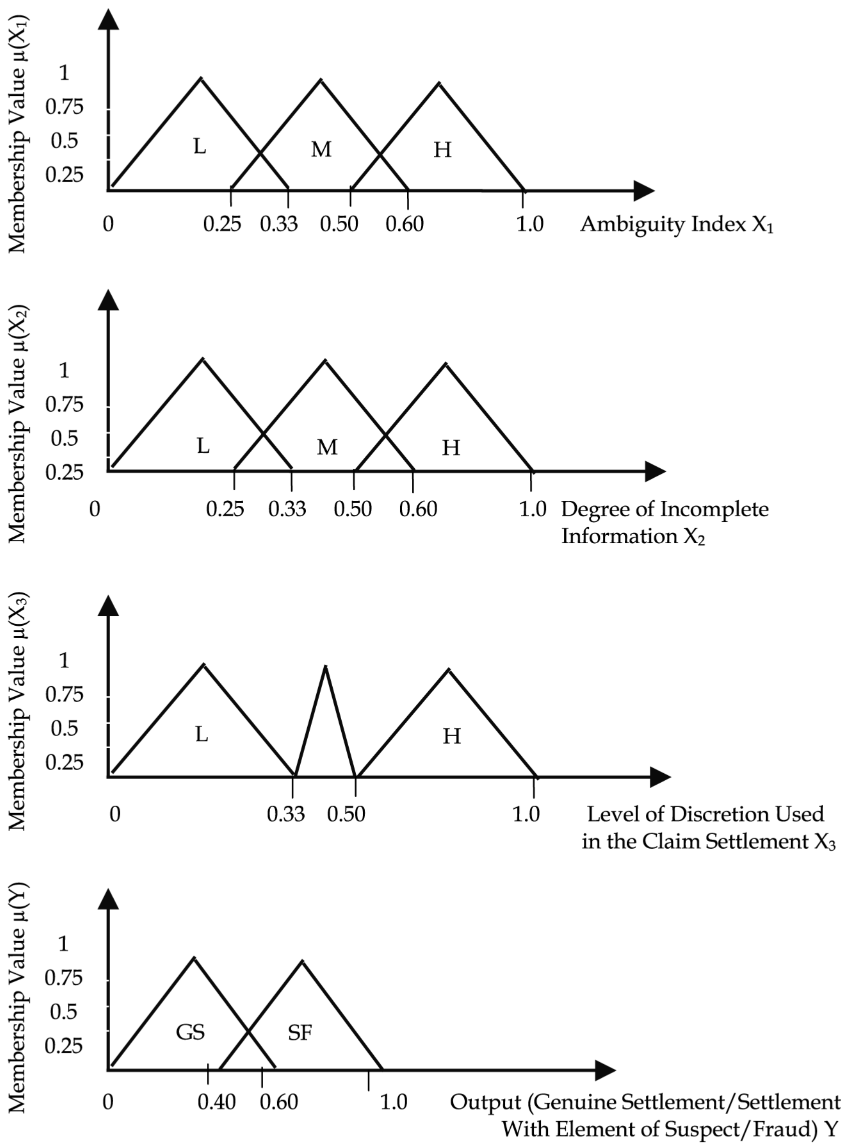

Figure 2.
Fuzzy set definition and membership functions of inputs and outputs functions

"IF-THEN" rules. In this context, the statement "if X is A then Y is B" or $A \rightarrow B$ results in a relation R such that $\mu_{\mathrm{R}}(x, y) = \min(\mu_A(x), \mu_B(y))$. This implication is precise, computationally simple, and fits various practical applications. The min operator is a natural choice for the logical AND. Bellman and Giertz (1973) have devised a set of axioms that should be satisfied by the AND operator and have proved that min operator satisfies them (Zimmerman, 2001) (Figures 3-5).

Rule No.	Inputs X_1	X_2	X_3	Output Y
1	Low	Low	Low	GS
2	Low	Low	Medium	GS
3	Low	Low	High	SF
4	Low	Medium	Low	GS
5	Low	Medium	Medium	GS
6	Low	Medium	High	SF
7	Low	High	Low	GS
8	Low	High	Medium	SF
9	Low	High	High	SF
10	Medium	Low	Low	GS
11	Medium	Low	Medium	GS
12	Medium	Low	High	SF
13	Medium	Medium	Low	GS
14	Medium	Medium	Medium	GS
15	Medium	Medium	High	SF
16	Medium	High	Low	GS
17	Medium	High	Medium	SF
18	Medium	High	High	SF
19	High	Low	Low	GS
20	High	Low	Medium	GS
21	High	Low	High	SF
22	High	Medium	Low	GS
23	High	Medium	Medium	SF
24	High	Medium	High	SF
25	High	High	Low	GS
26	High	High	Medium	SF
27	High	High	High	SF

Table I.
Sample rule base for the fuzzy logic based expert system

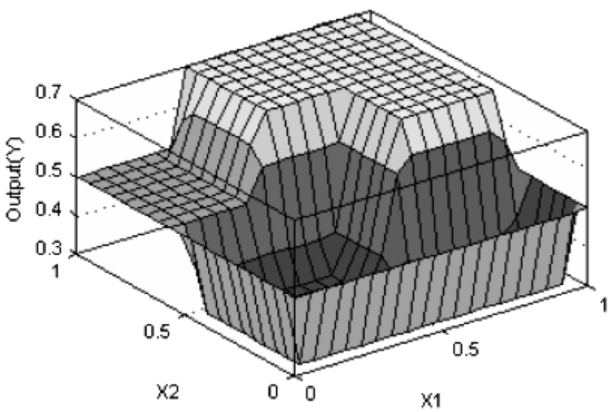

Figure 3.
Surface plot of the rule bases: X_1 and X_2 vs Y

4. Fuzzy logic based algorithm

The steps of the expert system are summarized below:

(1) *Input*. Input the crisp value of the insurance claim settlement and other information obtained.

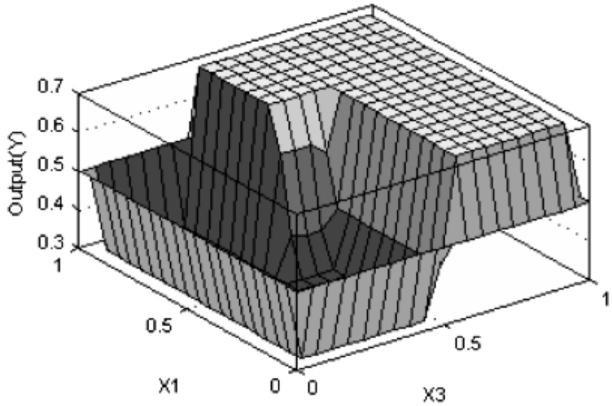

Figure 4.
Surface plot of the rule bases: X_1 and X_3 vs Y

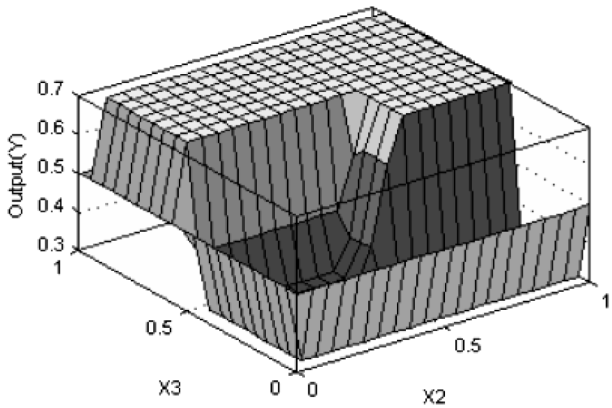

Figure 5.
Surface plot of the rule bases: X_2 and X_3 vs Y

(2) *Compare against threshold.* If the value of the claim settlement is less than certain pre-determined value as determined by the auditors (say $\alpha = \$1,000$), then the settled is considered genuine and go to Step 8.

(3) *Evaluate the authenticity of claim settlement.* Determine the ambiguity index X_1, degree of incomplete information of the claims X_2, level of discretion used by the claim settlers X_3.

(4) *Fuzzify the crisp values of inputs.* Through the use of membership functions defined for each fuzzy set for each linguistic variable (Figure 2), determine the degree of membership of a crisp value in each fuzzy set. Each of these three ambiguity indices have been divided into three fuzzy sets (LOW – L, MEDIUM – M and HIGH – H). The equations for computing memberships are:

$$\mu(X_i)_{\mathrm{L}} = \max\left\{\min\left(\frac{X_i - a_{\mathrm{L}}^{X_i}}{b_{\mathrm{L}}^{X_i} - a_{\mathrm{L}}^{X_i}}, \frac{c_{\mathrm{L}}^{X_i} - X_i}{c_{\mathrm{L}}^{X_i} - b_{\mathrm{L}}^{X_i}}\right), 0\right\} \tag{1}$$

$$\mu(X_i)_{\mathrm{M}} = \max\left\{\min\left(\frac{X_i - a_{\mathrm{M}}^{X_i}}{b_{\mathrm{M}}^{X_i} - a_{\mathrm{M}}^{X_i}}, \frac{c_{\mathrm{M}}^{X_i} - X_i}{c_{\mathrm{M}}^{X_i} - b_{\mathrm{M}}^{X_i}}\right), 0\right\} \quad (2)$$

$$\mu(X_i)_{\mathrm{H}} = \max\left\{\min\left(\frac{X_i - a_{\mathrm{H}}^{X_i}}{b_{\mathrm{H}}^{X_i} - a_{\mathrm{H}}^{X_i}}, \frac{c_{\mathrm{H}}^{X_i} - X_i}{c_{\mathrm{H}}^{X_i} - b_{\mathrm{H}}^{X_i}}\right), 0\right\} \quad (3)$$

where (a, b, c) are the vertices of the triangular membership function and L, M and H represents the fuzzy set LOW, MEDIUM, and HIGH, respectively.

(5) *Fire the rule bases that correspond to these inputs.* Every fuzzy logic based expert system uses fuzzy IF-THEN rules. A fuzzy IF-THEN rule is of the form IF $X_1 = A_1$ and $X_2 = A_2 \ldots$ and $X_n = A_n$ THEN $Y = B$, where X_i and Y are linguistic variables and A_i and B are linguistic terms. The "IF" part is known as antecedent or premise, whereas the "THEN" part is termed as a consequence or conclusion. Since all the three inputs have three fuzzy sets (LOW – L, MEDIUM – M and HIGH – H) 27 ($3 \times 3 \times 3$) fuzzy decisions are fired. There are two outputs: GENUINE SETTLEMENT – GS, CASES SETTLED WITH FRAUD ELEMENT OR ELEMENT OF SUSPECT – CF. The set "GENUINE SETTLEMENT" refers to the insurance claims which are genuinely settled without any room for ambiguity or suspect whereas "CASES SETTLED WITH FRAUD ELEMENT OR ELEMENT OF SUSPECT – CF" refers to the claims that have been settled but contain suspicious elements that need to be substantively audited.

(6) *Execute the inference engine.* Once all crisp input values have been fuzzified into their respective linguistic values, the inference engine will access the fuzzy rule base of the fuzzy expert system to derive linguistic values for the intermediate as well as the output linguistic variables. The two main steps in the inference process are aggregation and composition. Aggregation is the process of computing the values of the IF (antecedent) part of the rules while composition is the process of computing the values of the THEN (conclusion) part of the rules. During aggregation, each condition in the IF part of a rule is assigned a degree of truth based on the degree of membership of the corresponding linguistic term. From here, either the minimum (MIN) or product (PROD) of the degrees of truth of the conditions are computed to clip the degree of truth from the IF part. This is assigned as the degree of truth of the THEN part. The next step in the inference process is to determine the degrees of truth for each linguistic term of the output linguistic variable. Usually, either the maximum (MAX) or sum (SUM) of the degrees of truth of the rules with the same linguistic terms in the THEN parts is computed to determine the degrees of truth of each linguistic term of the output linguistic variable.

(7) *Defuzzification.* The last phase in the fuzzy expert system is the defuzzification of the linguistic values of the output linguistic variables into crisp values. The most common techniques for defuzzification are center-of-maximum (CoM) and CoA. CoM first determines the most typical value for each linguistic term for an output linguistic variable, and then computes the crisp value as the best compromise for the typical values and respective degrees of membership. The other common method, CoA, or sometimes called center-of-gravity (CoG), first cuts the

membership functions of each linguistic term at the degrees corresponding to the linguistic values. The superimposed areas under each cut membership function are balanced to give the compromised value. A disadvantage of this technique is the high computational demands in computing the areas under the membership functions. There are other variants of computing crisp values from linguistic values. These are mean-of-maximum (MoM), left-of-maximum (LOM) or smallest-of-maximum (SoM), right-of-maximum (ROM) or largest-of-maximum (LoM), and bisector-of-area (BoA). Details can be found in Zimmerman (2001) and Ross (1997).

(8) *Output the decisions of the expert system.* In our case, the types of the outputs are: genuine settlement of the claim and claims settled with element of suspect or fraud. The auditor further investigates into the claims settled with element of suspect or fraud.

The specific features of each controller depend on the model and performance measure. However, in principle, in all the fuzzy logic based expert system, we explore the implicit and explicit relationships within the system by mimicking human thinking and subsequently develop the optimal fuzzy control rules as well as knowledge base.

5. An illustrative example

The fuzzy logic based expert system was implemented in MATLAB 6.5. For the purpose of illustration, we consider that the insurance company uses three inputs – ambiguity index X_1, degree of incomplete information of the claims X_2, level of discretion used by the claim settlers X_3. These indices are representative of the authenticity of insurance claims settlement. These inputs represent the degree of ambiguity/incompleteness in the information furnished during various time periods or in various forms. The degree of ambiguity/incompleteness in the information and the level of discretion used by the claim settlers in deciding the settlement are used as a measure of unauthentic/fraud insurance claims.

(1) *Input.* Value of the claim = \$4,500.

(2) *Compare against threshold.* Since the value is greater than a pre-determined value (say $\alpha = \$1,000$), the authenticity of claim settled needs to be verified by the expert system.

(3) *Evaluate the authenticity of claim settlement.* The values of the inputs of the claims have to be evaluated, $X_1 = 0.27$, $X_2 = 0.55$ and $X_3 = 0.40$ (say).

(4) *Fuzzification of the crisp values of inputs.* Through the use of membership functions defined for each fuzzy set for each linguistic variable (Figure 2), the degree of membership of a crisp value in each fuzzy set is determined as follows:

$$\mu(X_1)_{\mathrm{L}} = \max\left\{\min\left(\frac{X_1 - a_{\mathrm{L}}^{X_1}}{b_{\mathrm{L}}^{X_1} - a_{\mathrm{L}}^{X_1}}, \frac{c_{\mathrm{L}}^{X_1} - X_1}{c_{\mathrm{L}}^{X_1} - b_{\mathrm{L}}^{X_i}}\right), 0\right\} = 0.86$$

$$\mu(X_1)_M = \max\left\{\min\left(\frac{X_1 - a_{\mathrm{M}}^{X_1}}{b_{\mathrm{M}}^{X_1} - a_{\mathrm{M}}^{X_1}}, \frac{c_{\mathrm{M}}^{X_1} - X_1}{c_{\mathrm{M}}^{X_1} - b_{\mathrm{M}}^{X_1}}\right), 0\right\} = 0.11$$

$$\mu(X_2)_M = \max\left\{\min\left(\frac{X_2 - a_{\mathrm{M}}^{X_2}}{b_{\mathrm{M}}^{X_2} - a_{\mathrm{M}}^{X_2}}, \frac{c_{\mathrm{M}}^{X_2} - X_2}{c_{\mathrm{M}}^{X_2} - b_{M}^{X_2}}\right), 0\right\} = 0.29$$

$$\mu(X_2)_{\mathrm{H}} = \max\left\{\min\left(\frac{X_2 - a_{\mathrm{H}}^{X_2}}{b_{\mathrm{H}}^{X_2} - a_{\mathrm{H}}^{X_2}}, \frac{c_{\mathrm{H}}^{X_2} - X_2}{c_{\mathrm{H}}^{X_2} - b_{\mathrm{H}}^{X_2}}\right), 0\right\} = 0.20$$

$$\mu(X_3)_{\mathrm{M}} = \max\left\{\min\left(\frac{X_3 - a_{\mathrm{M}}^{X_3}}{b_{\mathrm{M}}^{X_3} - a_{\mathrm{M}}^{X_3}}, \frac{c_{\mathrm{M}}^{X_3} - X_3}{c_{\mathrm{M}}^{X_3} - b_{\mathrm{M}}^{X_3}}\right), 0\right\} = 0.64$$

$$\mu(X_1)_{\mathrm{H}} = \mu(X_2)_{\mathrm{L}} = \mu(X_3)_{\mathrm{L}} = \mu(X_3)_{\mathrm{H}} = 0$$

where

$$a_{\mathrm{L}}^{X_1}, b_{\mathrm{L}}^{X_1}, c_{\mathrm{L}}^{X_1} = (0, 0.165, 0.33), \quad a_{\mathrm{M}}^{X_1}, b_{\mathrm{M}}^{X_1}, c_{\mathrm{M}}^{X_1} = (0.25, 0.425, 0.60),$$

$$a_{\mathrm{H}}^{X_1}, b_{\mathrm{H}}^{X_1}, c_{\mathrm{H}}^{X_1} = (0.50, 0.75, 1), \quad a_{\mathrm{L}}^{X_2}, b_{\mathrm{L}}^{X_2}, c_{\mathrm{L}}^{X_2} = (0, 0.165, 0.33),$$

$$a_{\mathrm{M}}^{X_2}, b_{\mathrm{M}}^{X_2}, c_{\mathrm{M}}^{X_2} = (0.25, 0.425, 0.60), \quad a_{\mathrm{H}}^{X_2}, b_{\mathrm{H}}^{X_2}, c_{\mathrm{H}}^{X_2} = (0.50, 0.75, 1),$$

$$a_{\mathrm{L}}^{X_3}, b_{\mathrm{L}}^{X_3}, c_{\mathrm{L}}^{X_3} = (0,\ 0.165,\ 0.33), \quad a_{\mathrm{M}}^{X_3}, b_{\mathrm{M}}^{X_3}, c_{\mathrm{M}}^{X_3} = (0.33, 0.44, 0.55),$$

$$a_{\mathrm{H}}^{X_3}, b_{\mathrm{H}}^{X_3}, c_{\mathrm{H}}^{X_3} = (0.50, 0.75, 1).$$

(5) *Fire the rule bases that correspond to these inputs.* Based on the value of the fuzzy membership function values for the example under consideration, the following rules apply:

Rule 5: If X_1 is LOW and X_2 is MEDIUM and X_3 is MEDIUM then Y is GENUINE SETTLEMENT (GS).

Rule 8: If X_1 is LOW and X_2 is HIGH and X_3 is MEDIUM then Y is SETTLEMENT WITH ELEMENT OF FRAUD (SF).

Rule 14: If X_1 is MEDIUM and X_2 is MEDIUM and X_3 is MEDIUM then Y GENUINE SETTLEMENT (GS).

Rule 17: If X_1 is MEDIUM and X_2 is HIGH and X_3 is MEDIUM then Y is SETTLEMENT WITH ELEMENT OF FRAUD (SF).

(6) *Execute the Inference Engine.* We use the "root sum squares" (RSS) method to combine the effects of all applicable rules, scale the functions at their respective magnitudes, and compute the "fuzzy" centroid of the composite area. This method is more complicated mathematically than other methods, but was selected for this example since it seemed to give the best weighted influence to all firing rules.

The respective output membership function strengths (range: 0-1) from the possible rules (R_1-R_{27}) are:

$$\text{“GenuineSettlement”} = \sqrt{\sum_{i \in GS} (\mu_{R_i})^2}$$

$$GS = \{1,2,4,7,8,10,11,13,16,17,19,20,22,25\}$$

$$= \sqrt{(0.29)^2 + (0.11)^2} = 0.31$$

$$\text{“Settlement with element of fraud/suspect”} = \sqrt{\sum_{i \in SF} (\mu_{R_i})^2}$$

$$SF = \{3, 5, 6, 9, 12, 14, 15, 18, 21, 23, 24, 26, 27\}$$

$$= \sqrt{(0.20)^2 + (0.11)^2} = 0.23$$

(7) *Defuzzification.* We use “fuzzy centroid algorithm” for defuzzification. The defuzzification of the data into crisp output is accomplished by combining the results of the inference process and then computing the “fuzzy centroid” of the area. The weighted strengths of each output member function are multiplied by their respective output membership function center points and summed. Finally, this area is divided by the sum of the weighted member function strengths and the result is taken as the crisp output. Figure 6 shows the

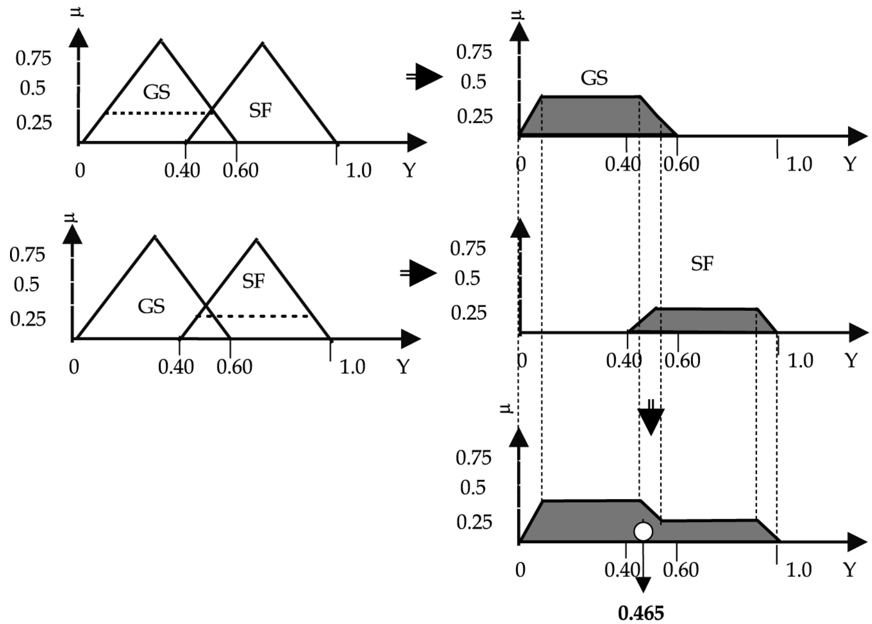

Figure 6.

shaded area and the crisp output of 0.465. The crisp output belongs to the set of GS more than the set of settlement with fraud element (SF) (as evident from its membership function). Hence, the decision in this case is GS.

(8) *Output the decisions of the expert system.* For the cases under consideration, the system indicates the case has been genuinely settled.

6. Conclusion and scope for future research

In this paper, we have reported on the development of a fuzzy based expert system for fraud detection in settled insurance claims. Our ongoing efforts will be on the improvement of the performance of the system by adjusting the membership function of the inputs. The number of points in the membership functions that might be moved and the number of inputs to the system yield a large number of possible combinations, which could be inspected. It would be interesting to tune the rule base using data from real life scenarios so that the performance of the system is optimized. We intend to use neural networks that can produce an optimum surface representing all the combination points from a few of the tested combinations. Another direction for future extension would be to take into account input factors such as: time taken by the claim adjusters to settle the claim, whether the approval of a higher authority or peers was taken into account while settling the claim, number of similar types of claims that has been filed by the claimant, etc. It is worthwhile noting that inclusion of these factors would increase the size of the rule base to the point that the tuning of the rule base using data from real life scenarios will be deemed necessary to optimize the performance of the system.

References

Bellman, R. and Giertz, M. (1973), "On the analytic formalism of the theory of fuzzy sets", *Information Sciences*, Vol. 5, pp. 149-56.

Bezdek, J.C. (1981), *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY.

Brockett, P.L., Xia, X. and Derrig, R. (1995), "Using Kohonen's self organizing feature map to uncover automobile bodily injury claims fraud", paper presented at Special Actuarial Seminar. Automobile Insurers Bureau. Boston, MA, 25 January.

Brockett, P.L., Xia, X. and Derrig, R. (1998), "Using Kohonen's self organizing feature map to uncover automobile bodily injury claims fraud", *The Journal of Risk and Insurance*, Vol. 65 No. 2, pp. 245-74.

Brockett, P.L., Derrig, R., Golden, L., Levine, A. and Alpert, M. (2002), "Fraud classification using principal component analysis of RIDITs", *The Journal of Risk and Insurance*, Vol. 69 No. 3, pp. 341-71.

Crocker, K.J. and Tennyson, S. (2002), "Insurance fraud and optimal claims settlement strategies", *Journal of Law and Economics*, Vol. 45 No. 2, pp. 469-507.

Derrig, R. and Ostaszewski, K. (1995), "Fuzzy techniques of pattern recognition in risk and claim classification", *The Journal of Risk and Insurance*, Vol. 62 No. 3, pp. 447-82.

Deshmukh, A. and Lakshminarayana, T. (1998), "A rule-based fuzzy reasoning system for assessing the risk of management fraud", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 4, pp. 231-41.

Fanning, K.M. and Cogger, K.O. (1998), "Neural network detection of management fraud using published financial data", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 7 No. 1, pp. 21-41.

Greenwald, J. (2003), "P/C insurers see strong quarter", *Business Insurance*, Vol. 37 No. 22, pp. 3-22.

Guidera, J. (2003), "COLOSSUS at the accident scene – insurers use a software program to payout claims for injuries, but lawsuits claim it's misused", *Wall Street Journal*, Eastern Edition, 2 January, C.1.

Hoffman, T. (1999), "Insurance agents tap tools to nab crooks", *Computerworld*, Vol. 33 No. 34, p. 39.

Mogel, G.S. (2003), "Claims software cited in bad faith suit", *National Underwriter*, Vol. 107 No. 21, pp. 35-6.

Ross, T.J. (1997), *Fuzzy Logic With Engineering Applications*, McGraw-Hill, Singapore.

Tennyson, S. and Pau, S.F. (2002), "Claims auditing in automobile insurance: fraud detection and deterrence objectives", *The Journal of Risk and Insurance*, Vol. 69 No. 3, pp. 289-308.

Welch, O.J., Reeves, T.E. and Welch, S.T. (1998), "Using a genetic algorithm-based classifier system for modeling auditor decision behavior in a fraud setting", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 7 No. 3, pp. 173-86.

Wheeler, R. and Aitkin, S. (2000), "Multiple algorithms for fraud detection", *Knowledge Based Systems*, Vol. 13, pp. 2-3.

Zadeh, L.A. (1965), "Fuzzy sets", *Information and Control*, Vol. 8, pp. 338-53.

Zimmerman, H.J. (2001), *Fuzzy Set Theory and its Application*, 4th ed., Kluwer Academic Publishers, London/Drodrecht/Boston, MA.

Zyl, S.V. (2003), "Canadian P/C insurers in intensive care?", *National Underwriter*, Vol. 107 No. 22, pp. 30-1.