Adapting TIN-layers to represent fuzzy geographic information

Jörg Verstraete, Guy De Tré, Axel Hallez
Computer Science Laboratory
Dept. of Telecommunications and Information Processing, Ghent University
Sint-Pietersnieuwstraat 41, B-9000 Gent (Belgium)
{Jorg.Verstraete,Guy.DeTre,Axel.Hallez}@telin.rug.ac.be

Abstract

One of the latest research topics in geographic information systems (GIS for short) is the modeling of fuzzy (imprecise or vague) and/or uncertain information. A GIS usually contains large amounts of data and requires very specific operations to manage (geographic) data, both of which are described in a geographic database model. In this paper, the basic concepts for a data model describing the structure of and operations on fuzzy or uncertain geographic information are presented. Specific for the presented model is the adaptation of triangulated irregular networks (TINs), commonly used in GIS (e.g. to model altitudes), for the modelling of fuzzy (geographic) data.

Keywords: triangulated irregular networks, fuzzy geographic database, fuzzy gis.

1 Introduction

An important research topic in geographic information systems is the modeling of fuzzy and/or uncertain (geographic) information [1],[2]. An interesting aspect is the modelling of vague regions which allow to describe the extent to which a given property applies for a given set of points in space (e.g. the extent to which a given location can be considered as densely populated [3]. There have been an number of publications describing techniques to model vague regions. Most of these techniques use broad boundaries, which are mostly defined by means of contour lines and traditional buffer-operations [4]. Broad boundaries defined by the use of two (e.g. the egg-yolk theory [5], [6]) or more [7], [8] contour lines do not allow for an easy definition of some of the operators: it turns out to be quite difficult to correctly define operators that result in a closed algebra (e.g. the intersection of two regions specified by contour lines is difficult —if possible— to represent as a new region defined by contour lines).

In this paper a method is presented, in which contour lines are no longer used; instead the fuzzy information is modelled by means of a TIN-based *layer*. This makes sense since a GIS usually works with layers to store different kinds of data: for instance, it could contain a layer that contains information about altitudes, another layer containing population-densities, etc. These layers can be combined (called an *overlay*), for example to calculate the result of a query. The use of Triangulated Irregular Networks allows —as presented in this paper— for a closed definition of various operators.

2 Outline of the approach

Unlike other models that use some form of contour lines [4],[5],[6],[7],[8], the presented model uses a Triangulated Irregular Network [9],[12] to represent (fuzzy) information about geographic regions. TINs are used in traditional geographic databases to model crisp geographic data, such as altitudes. In the presented model, the same TIN structure is used, but now adapted for the modelling of vague regions. To deal with fuzziness and uncertainty, fuzzy set theory will be used [10]. More specifically, membership grades (in the interval $[0, 1]$) will be associated with the points on which the TIN-based layer is built (hereafter referred to as *data-points*). The existing interpolation techniques can be maintained, but some new operators (e.g. intersection and union) need to be introduced.

The membership grades associated with the data-points of a single layer are either all interpreted as degrees of (un)certainty or all interpreted as degrees of preference [11]. If interpreted as degrees of (un)certainty, the membership grades denote the extent to which it is (un)certain that a given proposition is valid for these data-points, e.g. to model the position of an object within a given area. If interpreted as degrees of preference, the membership grades denote the extent to which a given property applies for these data-points (e.g. the extent to which a given location can be considered as densely populated). Several layers can be used to model the information about a given area; one layer representing for example the extent to which the area can be considered as densely populated and another one representing the uncertainty about the position of an object in this same area. The combination of these two layers then provides for example information about the position of an object within a (non)densely populated area.

It should be noted that the association of membership grades with data-points imposes some limitations on the model: uncertainty about the membership grades themselves cannot be modeled. For some applications (e.g. approximate temperature measurements), this feature would be beneficial. An extension of the presented approach, based on fuzzy sets of type 2 [14], is currently under research.

3 ETIN: structure and operators

3.1 Structure

The definition of a TIN is based on a triangular partition of two dimensional space. No assumption is made on the distribution an location of the vertices of the triangles [9]. In general, the structure of a TIN Tin_i is defined by three finite sets: a set P_i of points upon which the TIN is constructed, a set E_i of edges and a set T_i of triangles (the *tiles* of the TIN), i.e.

$$Tin_i = (P_i, E_i, T_i)$$

The TINs used in the presented model, are generated by means of a Delaunay triangulation [12]. A Delaunay triangulation of a set of points P_i is a triangulation of P_i with the property that no point in P_i falls in the interior of the circumcircle (circle that passes through all three vertices) of any triangle $t \in T_i$ in the triangulation. This uniquely determines the sets E_i and T_i. A variation of the Delaunay triangulation is the constrained Delaunay triangulation [13], which additionally allows the specification of a subset E of edges that need to be included in the final set E_i. The TIN structure is used because it offers some advantages when interpolating: due to the Delaunay triangulation algorithm, the tiles resemble the unilateral triangle as close as possible, thus avoiding degenerate cases (e.g. narrow, sharp triangles). The constrained Delaunay triangulation no longer has this property, and is —strictly speaking— not a Delaunay triangulation.

In the presented model, the TIN structure is extended with a mapping function f_i which characterizes a geographically dependent feature F_i, e.g. F_i might represent the feature *densely populated*. The structure of an *extended TIN* $ETin_i$ is defined by a TIN structure Tin_i and a mapping function f_i, i.e.

$$ETin_i = [(P_i, E_i, T_i), f_i]$$

where f_i is defined as

$$f_i : P_i \to [0, 1] : p(x, y) \mapsto f_i(p(x, y))$$

An ETIN-structure $ETin_i$ can be viewed as a three dimensional representation of a vague region. The X and Y axes are interpreted as the domain-axes of the two dimensional space and the Z axis represents the membership grades. These are obtained from the membership function μ_{f_i} derived (by lineair interpolation) from the mapping function f_i as follows

$$\mu_{f_i} : U \rightarrow [0,1]$$
$$p(x,y) \mapsto \begin{cases} f_i(p(x,y)) & \text{if } p(x,y) \in P_i \\ -\frac{A}{C}x - \frac{B}{C}y - \frac{D}{C} & \text{otherwise} \end{cases}$$

where U represents the two dimensional space and A, B, C and D are the parameters of the equation $Ax + By + Cz + D = 0$ of the plane going through the three points $p_1(x_1, y_1)$, $p_2(x_2, y_2)$ and $p_3(x_3, y_3)$ of the triangle in which the point $p(x, y)$ is located, i.e.

$$\begin{aligned} A &= y_1(z_2 - z_3) + y_2(z_3 - z_1) + y_3(z_1 - z_2) \\ B &= z_1(x_2 - x_3) + z_2(x_3 - x_1) + z_3(x_1 - x_2) \\ C &= x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2) \\ D &= -Ax_1 - By_1 - Cz_1 \end{aligned}$$

with $z_j = f_i(x_j, y_j)$, $j = 1, 2, 3$. The case where $C = 0$ cannot occur, as this would imply the points p_1, p_2 and p_3 to be co-lineair, in which case they would never be chosen in a Delaunay triangulation as a triangle.

3.2 Operators

As the model would be useless without appropriate operators, the definitions for the various operations on membership functions must be extended. For example, it should be possible to aggregate information contained in different ETIN-structures, as well as perform comparisons and calculate intersections and unions.

3.2.1 Approach using Delaunay triangulation

Suppose we have two ETIN-structures $ETin_1$ and $ETin_2$, which for the sake of argumentation are defined as illustrated in resp. Figure 1 and Figure 2. (As a shorthand notation, the data-points of an ETIN are denoted as $p_j(x_j, y_j, f_i(x_j, y_j))$, whereas an edge e_k connecting the points p_l and p_m is denoted as $e_k(p_l, p_m)$.) Generally, their data-points need not be on the same location, nor do they need to have the same number of data-points. As an example we shall use min and max as t-norm resp. t-conorm [1].

[1]The fact that an ETIN uses triangles has in itself nothing to do with triangular norms/conorms.

$p_1(0,0,0)$	$e_1(p_2,p_1)$
$p_2(0,100,1)$	$e_2(p_1,p_5)$
$p_3(100,0,1)$	$e_3(p_5,p_2$
$p_4(100,100,0.5)$	$e_4(p_5,p_3)$
$p_5(40,60,0.6)$	$e_5(p_3,p_4)$
	$e_6(p_4,p_5)$
	$e_7(p_1,p_3)$
	$e_8(p_4,p_2)$

Figure 1: $ETin_1$

$p_1(0,0,0)$	$e_1(p_2,p_1)$
$p_2(0,100,1)$	$e_2(p_1,p_5)$
$p_3(100,0,1)$	$e_3(p_5,p_2)$
$p_4(100,100,0.5)$	$e_4(p_5,p_3)$
$p_5(60,60,0.6)$	$e_5(p_3,p_4)$
	$e_6(p_4,p_5)$
	$e_7(p_1,p_3)$
	$e_8(p_4,p_2)$

Figure 2: $ETin_2$

The minimum of $ETin_1 = [(P_1, E_1, T_1), f_1]$ and $ETin_2 = [(P_2, E_2, T_2), f_2]$ (with associated membership functions resp. μ_{f_1} and μ_{f_2}) is by definition obtained by considering the minimum $\min(\mu_{f_1}(p(x,y)), \mu_{f_1}(p(x,y)))$ of the membership grades of each point $p(x, y)$ in the two dimensional space.

A computable definition can be derived by using the actual definitions of both ETINs. In the intersection of the triangles of both ETINs, the set P_t of relevant points is determined by calculating the equations for the edges and triangles (which are obtained through basic geometry).

P_t is obtained as the union of the set of the points that result from the intersection of the triangles in $ETin_1$ and the edges in $ETin_2$, and the set of points that result from the intersection of triangles in $ETin_2$ and the edges in $ETin_1$ (in case an intersection contains line segments, only the endpoints of these segments are considered).

The minimum will then be a new ETIN $ETin_3 = [(P_3, E_3, T_3), f_3]$, defined by the points

$$\begin{aligned} P_3 = P_t \cup & \\ & \{p(x,y) \in P_1 | \mu_{f_1}(p(x,y)) \leq \mu_{f_2}(p(x,y))\} \cup \\ & \quad \{p(x,y) \in P_2 | \mu_{f_2}(p(x,y)) \leq \mu_{f_1}(p(x,y))\} \end{aligned}$$

$p_1(0,0,0)$	$e_1(p_1,p_3)$
$p_2(0,100,1)$	$e_2(p_3,p_2)$
$p_3(100,100,0)$	$e_3(p_1,p_2)$
$p_4(100,0,1)$	$e_4(p_3,p_4)$
	$e_5(p_1,p_4)$

Figure 3: $ETin_{1'}$

$p_1(0,0,1)$	$e_1(p_2,p_1)$
$p_2(0,100,0)$	$e_2(p_1,p_4)$
$p_3(100,100,1)$	$e_3(p_4,p_2)$
$p_4(100,0,0)$	$e_4(p_4,p_3)$
	$e_5(p_3,p_2)$

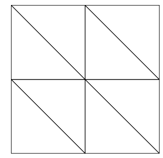
Figure 4: $ETin_{2'}$

At first sight, it seems to be sufficient to determine P_3, and use the Delaunay triangulation-algorithm to calculate the edges (and triangles). However it turns out that the result is not always what is expected: due to the Delaunay triangulation, it is possible that some generated edges will not match the minimum.

This can be illustrated with a simple example; consider two ETINs $ETin_{1'}$ and $ETin_{2'}$, as defined in Figures 3 and 4).

The minimum as obtained by calculating the set P_3 with the method described above (Figure 5), does not equal the exact minimum (Figure 6).

This is due to the fact that the Delaunay triangulation algorithm generates a set of edges (and triangles) which does not contain the expected minima. E.g. the edges $e_2(p_2,p_3)$ and $e_{12}(p_5,p_4)$ (Figure 5) are generated instead of the edges $e_2(p_1,p_6)$ and $e_{12}(p_1,p_7)$ (Figure 6). This can also be verified by considering the point $p(25,25)$ with interpolated membership grades $\mu_{f_{1'}}(p(25,25)) = \mu_{f_{2'}}(p(25,25)) = 0$. The minimum of those membership grades is clearly 0, which should be the membership grade of $p(25,25)$ in the resulting ETIN. However, in $ETin_{3'}$ the interpolated membership grade $\mu_{f_{3'}}(p(25,25))$ equals 0.5.



$p_1(50,50,0)$	$e_1(p_6,p_2)$
$p_2(50,0,0.5)$	$e_2(p_2,p_3)$
$p_3(0,50,0.5)$	$e_3(p_3,p_6)$
$p_4(100,50,0.5)$	$e_4(p_1,p_8)$
$p_5(50,100,0.5)$	$e_5(p_8,p_3)$
$p_6(0,0,0)$	$e_6(p_3,p_1)$
$p_7(100,100,0)$	$e_7(p_2,p_1)$
$p_8(0,100,0)$	$e_8(p_1,p_9)$
$p_9(100,0,0)$	$e_9(p_9,p_4)$
	$e_{10}(p_4,p_1)$
	$e_{11}(p_2,p_9)$
	$e_{12}(p_5,p_4)$
	$e_{13}(p_4,p_7)$
	$e_{14}(p_7,p_5)$
	$e_{15}(p_5,p_1)$
	$e_{16}(p_5,p_8)$

Figure 5: $ETin_{3'}$, obtained through an incorrect calculation of $\min(ETin_a, ETin_b)$

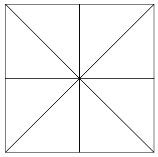
3.2.2 Approach using constrained Delaunay triangulation

A solution to this problem is to use the constrained Delaunay triangulation algorithm, hereby considering the set P_3 and a set of predefined edges E. With the understanding that E_t consists of the edges obtained through the intersection (Figure 7) of the triangles of the two ETINs $ETin_1$ and $ETin_2$ and P_t being the set of points as previously defined, E is defined as

$$E = E_t \cup \{e(p_1,p_2) | p_1 \in P_t \wedge p_2 \in P_3 \setminus P_t \wedge \\ (\exists\, e'(p'_1,p'_2) \in E_1 \cup E_2 : e(p_1,p_2) \subseteq e'(p'_1,p'_2))\}$$

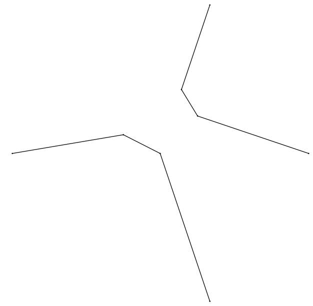
An important property of the set of edges E as defined above, is that from it a set of adjacent (non overlapping) planar polygons can be constructed. For example, with respect to Figure 8, eight adjacent planar polygons can be constructed, e.g. the one in the upper-left corner, being defined by the edges $e_1, e_2, e_{12}, e_{16}, e_{17}$ (or by the points $p_1, p_2, p_3, p_{11}, p_{13}$). The constrained Delaunay triangulation algorithm when applied to a planar polygon results in a planar triangulation thereby avoiding the problems occurring with the regular Delaunay triangulation on the set P_3.

By applying the constrained Delaunay triangulation algorithm on the sets P_3 and E, the ETIN $ETin_3 = [(P_3, E_3, T_3), f_3]$ (Figure 9) is constructed. Using the same technique to calculate



$p_1(50, 50, 0)$	$e_1(p_2, p_1)$
$p_2(50, 0, 0.5)$	$e_2(p_1, p_6)$
$p_3(0, 50, 0.5)$	$e_3(p_6, p_2)$
$p_4(100, 50, 0.5)$	$e_4(p_1, p_8)$
$p_5(50, 100, 0.5)$	$e_5(p_8, p_3)$
$p_6(0, 0, 0)$	$e_6(p_3, p_1)$
$p_7(100, 100, 0)$	$e_7(p_3, p_6)$
$p_8(0, 100, 0)$	$e_8(p_1, p_9)$
$p_9(100, 0, 0)$	$e_9(p_9, p_4)$
	$e_{10}(p_4, p_1)$
	$e_{11}(p_2, p_9)$
	$e_{12}(p_1, p_7)$
	$e_{13}(p_7, p_5)$
	$e_{14}(p_5, p_1)$
	$e_{15}(p_4, p_7)$
	$e_{16}(p_5, p_8)$

Figure 6: $ETin_{4'}$, obtained through a proper calculation of $\min(ETin_a, ETin_b)$



$p_1(37.5, 56.25, 0.5625)$	$e_1(p_1, p_2)$
$p_2(0, 50, 0.5)$	$e_2(p_1, p_3)$
$p_3(50, 50, 0.66667)$	$e_3(p_5, p_3)$
$p_4(100, 50, 0.75)$	$e_4(p_6, p_7)$
$p_5(66.66667, 0, 0.66667)$	$e_5(p_8, p_6)$
$p_6(57.1429, 71.4286, 0.571429)$	$e_6(p_4, p_7)$
$p_7(62.5, 62.5, 0.625)$	
$p_8(66.66667, 100, 0.66667)$	

Figure 7: E_t, the intersection points and edges of $(ETin_1, ETin_2)$.

the minimum $\min(ETin_{1'}, ETin_{2'})$ (Figure 3 and 4), yields the correct ETIN $ETin_{4'}$ as shown on Figure 6.

The max-conorm is defined in a completely similar way.

4 Conclusions and future work

A new approach for modelling of vague regions which allow to describe the extent to which a given property applies for a given set of points in space, is presented. Central to the approach is the use of *extended* triangulated irregular networks, which extend the traditional concept of triangulated irregular networks with a mapping function used to define a membership function over a geographic area. A technique to determine the minimum (as an example of a t-norm) of ETINs is discussed. Constrained Delaunay triangulation provides for an appropriate and correct definition of this operator.

Within the presented approach, membership grades are used to represent the extent to which a given property applies for a given set of points. In practice, it may occur that there exists uncertainty about the values of the membership grades themselves; part of the ongoing research deals with the generalization of the model by means of fuzzy sets of type 2. This extension could also be adapted to model imprecise measurements and different geographic locations (e.g. inaccurate temperature measurements). Another interesting extension would be to allow for other norms and co-norms besides min and max; this is currently also under research.

References

[1] Hans-Jürgen Zimmerman (1999). *Practical applications of fuzzy technologies.* Kluwer Academic Publishers, pages 267-300.

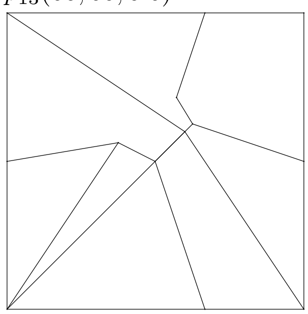
[2] Ashley Morris (2001). Why Spatial Databases Need Fuzziness. In *Proceedings of Nafips 2001*, pages 2446-2451.

[3] Markus Schneider (1999). Uncertainty management for spatial data in databases : fuzzy spatial data types. In *6th Int. Symp. on Advances in Spatial Databases (SSD)*, LNCS 1651, Springer Verlag, pages 330-351.

[4] Eliseo Clementini, Paoline Di Felice (1994). An algebraic model for spatial objects with undetermined boundaries. In *GISDATA Specialist Meeting - revised version.*

[5] Anthony G. Cohn, Nicholas Mark Gotts (1994). Spatial regions with undetermined boundaries. In *Proceedings of the Second ACM Workshop on Advances in Geographic Information Systems*, pages 52-59.

$p_1(37.5, 56.25, 0.5625)$
$p_2(0, 50, 0.5)$
$p_3(50, 50, 0.66667)$
$p_4(100, 50, 0.75)$
$p_5(66.66667, 0, 0.66667)$
$p_6(57.1429, 71.4286, 0.571429)$
$p_7(62.5, 62.5, 0.625)$
$p_8(66.66667, 100, 0.66667)$
$p_9(0, 0, 0)$
$p_{10}(100, 100, 0.5)$
$p_{11}(0, 100, 0)$
$p_{12}(100, 0, 0.5)$
$p_{13}(60, 60, 0.6)$



$e_1(p_1, p_2)$
$e_2(p_1, p_3)$
$e_3(p_5, p_3)$
$e_4(p_6, p_7)$
$e_5(p_8, p_6)$
$e_6(p_4, p_7)$
$e_7(p_9, p_2)$
$e_8(p_9, p_1)$
$e_9(p_9, p_5)$
$e_{10}(p_{10}, p_4)$
$e_{11}(p_{10}, p_8)$
$e_{12}(p_{11}, p_2)$
$e_{13}(p_{11}, p_8)$
$e_{14}(p_{12}, p_4)$
$e_{15}(p_{12}, p_5)$
$e_{16}(p_{13}, p_3)$
$e_{17}(p_{13}, p_{11})$
$e_{18}(p_{13}, p_{12})$
$e_{19}(p_{13}, p_7)$
$e_{20}(p_{13}, p_9)$

Figure 8: E, calculated minimum of $(ETin_1, ETin_2)$, prior to triangulation. This is not an actual ETIN.

[6] Nicholas Mark Gotts, Anthony G. Cohn (1995). A mereological approach to representing spatial vagueness. In *Working Papers, 9th International Workshop on Qualitative Reasoning*, pages 246-255.

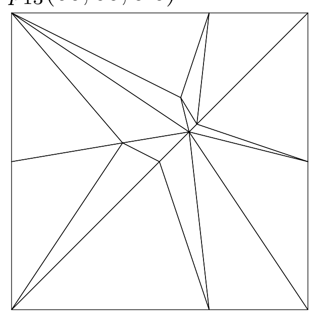
[7] J. Vertraete, B. Van Der Cruyssen, R. De Caluwe (2000). Assigning Membership Degrees to Points of Fuzzy Boundaries. In *Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Society – Nafips*, pages 444-447.

[8] A. Hallez, J. Verstraete, G. De Tré, R. De Caluwe (2002). Contourline Based Modelling of Vague Regions. Accepted for *9th International Conference IPMU 2002*.

[9] Philippe Rigaux, Michel Scholl, Agnès Voisard (2002). *Spatial Databases with Applications to GIS*. Morgan Kaufman Publishers.

[10] Didier Dubois, Henri Prade (2000). *Fundamentals of Fuzzy Sets*. Kluwer Academic Publishers.

$p_1(37.5, 56.25, 0.5625)$
$p_2(0, 50, 0.5)$
$p_3(50, 50, 0.66667)$
$p_4(100, 50, 0.75)$
$p_5(66.66667, 0, 0.66667)$
$p_6(57.1429, 71.4286, 0.571429)$
$p_7(62.5, 62.5, 0.625)$
$p_8(66.66667, 100, 0.66667)$
$p_9(0, 0, 0)$
$p_{10}(100, 100, 0.5)$
$p_{11}(0, 100, 0)$
$p_{12}(100, 0, 0.5)$
$p_{13}(60, 60, 0.6)$



$e_1(p_1, p_9)$
$e_2(p_9, p_1)$
$e_3(p_1, p_2)$
$e_4(p_3, p_9)$
$e_5(p_9, p_5)$
$e_6(p_5, p_3)$
$e_7(p_3, p_1)$
$e_8(p_1, p_{13})$
$e_9(p_{13}, p_{11})$
$e_{10}(p_{11}, p_1)$
$e_{11}(p_{11}, p_2)$
$e_{12}(p_6, p_1 1)$
$e_{13}(p_{13}, p_6)$
$e_{14}(p_6, p_8)$
$e_{15}(p_8, p_{11})$
$e_{16}(p_3, p_{13})$
$e_{17}(p_{12}, p_4)$
$e_{18}(p_4, p_{13})$
$e_{19}(p_{13}, p_{12})$
$e_{20}(p_5, p_{13})$
$e_{21}(p_7, p_8)$
$e_{22}(p_6, p_7)$
$e_{23}(p_7, p_{13})$
$e_{24}(p_4, p_7)$
$e_{25}(p_{10}, p_8)$
$e_{26}(p_7, p_{10})$
$e_{27}(p_4, p_{10})$
$e_{28}(p_5, p_{12})$

Figure 9: $ETin_3$, minimum of $(ETin_1, ETin_2)$, after triangulation.

[11] Didier Dubois, Henri Prade (2001). Possibility theory, probability theory and multiple-valued logics: A clarification. *Annals of Mathematics and Artificial Intelligence*, volume 32, pages 35-66.

[12] Jonathan Richard Shewchuk (1996). Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In *First Workshop on Applied Computational Geometry (Philadelphia, Pennsylvania)*, Association for Computing Machinery, pages 124-133.

[13] Jonathan Richard Shewchuk; *Constrained Delaunay Tetrahedralizations and Provably Good Boundary Recovery*; 2002; Submitted to the Eleventh International Meshing Roundtable.

[14] Jerry M. Mendel (2001). Uncertain Rule-Based Fuzzy Logic Systems, Introduction and New Directions. Prenctice Hall PTR.