MODELOS DE CONJUNTOS RUGOSOS DIFUSOS TOLERANTES AL RUIDO: DEFINICIONES Y PROPIEDADES

Lynn D'eer[1], **Nele Verbiest**[1], **Chris Cornelis**[1,2], **Lluis Godo**[3]

[1] *Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281 (S9), B-9000 Gent, Bélgica,* {Lynn.Deer,Nele.Verbiest}@UGent.be

[2] *Departamento de Ciencias de la Computación y I.A., CITIC, Universidad de Granada, Calle del Periodista Daniel Saucedo Aranda s/n, 18071 Granada, España*, Chris.Cornelis@decsai.ugr.es

[3] *IIIA - CSIC, Campus de la UAB s/n, 08193 Bellaterra, España*, godo@iiia.csic.es

Resumen

Desde que a principios de los años noventa se propuso el primer modelo híbrido difuso-rugoso, muchos investigadores se han interesado en la generalización de la noción de aproximación inferior y superior de un conjunto difuso a partir de una relación difusa. Un tipo específico de estos nuevos modelos se centran en proponer aproximaciones robustas al posible ruido en los datos. En este artículo, repasamos las propuestas más prominentes, y comprobamos cuáles de las propiedades deseables del modelo estándar de los conjuntos rugosos se mantienen en estas propuestas.

Palabras Clave: conjuntos rugosos difusos, híbridización, aproximación inferior y superior, tolerancia al ruido.

1 Introducción

Los conjuntos rugosos [10] permiten caracterizar un conjunto de objetos por una aproximación inferior y superior, teniendo en cuenta una relación de equivalencia que representa la indiscernibilidad entre los objetos. Dubois y Prade propusieron en 1990 su hibridización con los conjuntos difusos [6]. Hoy en día, más de 20 años después, sigue aumentando el interés por los conjuntos rugosos difusos; esto es debido principalmente a su potencial demostrado en aplicaciones de aprendizaje automático, en particular en la selección de características [3, 8] e instancias [13].

La híbridización difusa-rugosa se ha realizado de varias maneras; el enfoque más común consiste en usar extensiones difusas de la implicación y de la conjunción booleana, junto con el uso del ínfimo y el supremo como extensiones de los cuantificadores universal y existencial. Esta idea dió lugar a la propuesta seminal en [6], y desde entonces muchos autores se han centrado en refinar dicho modelo. Se comprobó que con una elección apropriada de los operadores lógicos difusos, y con el uso de una relación de similitud (también llamada relación difusa de equivalencia) para modelar la indiscernibilidad aproximada, la mayoría de las propiedades del modelo original de los conjuntos rugosos se podían mantener [1, 9, 11]. Por otro lado, desde el punto de vista práctico, el uso de las relaciones de similitud no siempre es conveniente, por lo que se han propuesto también modelos de conjuntos rugosos difusos basados en relaciones difusas generales. En [5], se han unificado todos estos enfoques bajo el paradigma del modelo general de los conjuntos difusos rugosos basado en operadores difusos de implicación y conjunción, imponiendo restricciones mínimas a las aproximaciones.

A causa del uso del ínf y del sup, las aproximaciones basadas en implicaciones y conjunciones son sensibles a pequeños cambios en los argumentos. Esto limita su potencial de aplicación en los problemas de análisis de datos, donde típicamente hay una presencia de ruido en los datos. Por tanto, varios autores han propuesto modelos de conjuntos rugosos difusos que son robustos a la perturbación de datos, como por ejemplo [2, 4, 7, 12, 14]. Desafortunadamente, el precio a pagar por el aumento en la robustez es que los modelos correspondientes satisfacen menos propiedades teóricamente deseables. En este artículo, llevamos a cabo un análisis comparativo de cuatro modelos prominentes que son tolerantes al ruido: el modelo de β-precisión [12], el modelo de precisión variable difusa-rugosa (PVDR) [14], el modelo de conjuntos rugosos por cuantificadores difusos (CRCD) [2], y el modelo basado en los operadores OWA (Ordered Weighted Average) [4]. Evaluamos cuáles de las propiedades se mantienen en los distintos modelos, y cómo esto puede afectar en las aplicaciones que los usen.

Después de recordar los preliminares necesarios en la Sección 2, resumimos y comentamos las definiciones de las aproximaciones tolerantes al ruido en la Sección 3. En la Sección 4, las evaluamos con respecto a una lista de propiedades deseables. Finalmente, en la Sección 5, presentamos las conclusiones y unos comentarios sobre nuestro trabajo futuro.

2 Preliminares

2.1 Conectivas lógicas difusas

En este trabajo usaremos las siguientes definiciones para operadores difusos de conjunción, disyunción, negación e implicación.

Una *conjunción* es una función $\mathscr{C}\colon [0,1]^2 \to [0,1]$ que es creciente en ambos argumentos y que satisface $\mathscr{C}(0,0) = \mathscr{C}(0,1) = \mathscr{C}(1,0) = 0$ y $\mathscr{C}(1,1) = 1$. Se llama *conjunción frontera* si además cumple $\mathscr{C}(1,x) = x$ para cada x en $[0,1]$. Una conjunción frontera conmutativa y asociativa se llama *t-norma*.

Una *disyunción* es una función $\mathscr{D}\colon [0,1]^2 \to [0,1]$ que es creciente en ambos argumentos y que satisface $\mathscr{D}(1,0) = \mathscr{D}(0,1) = \mathscr{D}(1,1) = 0$ y $\mathscr{D}(0,0) = 0$. Se llama *disyunción frontera* si además cumple $\mathscr{D}(0,x) = x$ para cada x en $[0,1]$. Un disyunción frontera conmutativa y asociativa se llama *t-conorma*.

Una *negación* es una función decreciente $\mathscr{N}\colon [0,1] \to [0,1]$ que satisface $\mathscr{N}(0) = 1$ y $\mathscr{N}(1) = 0$. Se llama *involutiva* si para cada $x \in [0,1]$, $\mathscr{N}(\mathscr{N}(x)) = x$. La negación estándar $\mathscr{N}_s$ se define, para x en $[0,1]$, como $\mathscr{N}_s(x) = 1-x$.

Dadas una negación involutiva $\mathscr{N}$, una conjunción $\mathscr{C}$ y una disyunción $\mathscr{D}$, la disyunción *$\mathscr{N}$-dual de $\mathscr{C}$* es el operador definido como $\mathscr{D}_{\mathscr{C},\mathscr{N}}(x,y) = \mathscr{N}(\mathscr{C}(\mathscr{N}(x),\mathscr{N}(y)))$, y la conjunción *$\mathscr{N}$-dual de $\mathscr{D}$* es el operador definido como $\mathscr{C}_{\mathscr{D},\mathscr{N}}(x,y) = \mathscr{N}(\mathscr{D}(\mathscr{N}(x), \mathscr{N}(y)))$, para cada x,y en $[0,1]$. Se puede comprobar que el operador $\mathscr{N}$-dual de una t-norma es una t-conorma, y viceversa.

Una *implicación* $\mathscr{I}$ es una función $\mathscr{I}\colon [0,1]^2 \to [0,1]$ que satisface $\mathscr{I}(1,0) = 0$, $\mathscr{I}(1,1) = \mathscr{I}(0,1) = \mathscr{I}(0,0) = 1$ y que es decreciente en el primer argumento y creciente en el segundo argumento. Si $\mathscr{I}$ satisface $\mathscr{I}(1,x) = x$ para cada x en $[0,1]$, se llama *implicación frontera*.

Sean $\mathscr{C}$, $\mathscr{D}$ y $\mathscr{N}$ una conjunción frontera, una disyunción y una negación, respectivamente. La *S-implicación* $\mathscr{I}_{\mathscr{D},\mathscr{N}}$ basada en $\mathscr{D}$ y $\mathscr{N}$ se define, para x,y en $[0,1]$, como $\mathscr{I}_{\mathscr{D},\mathscr{N}}(x,y) = \mathscr{D}(\mathscr{N}(x),y)$. La *R-implicación* $\mathscr{I}_{\mathscr{C}}$ basada en $\mathscr{C}$ se define, para x,y en $[0,1]$, como $\mathscr{I}_{\mathscr{C}}(x,y) = \sup\{\gamma \in [0,1] \mid \mathscr{C}(x,\gamma) \le y\}$. Tanto las S-implicaciónes como las R-implicaciónes son casos particulares de implicaciones frontera.

Dadas una negación involutiva $\mathscr{N}$ y una implicación $\mathscr{I}$, la *conjunción inducida por $\mathscr{I}$ y $\mathscr{N}$* es el operador $\mathscr{C}_{\mathscr{I},\mathscr{N}}$ definido, para $x,y \in [0,1]$, como $\mathscr{C}_{\mathscr{I},\mathscr{N}}(x,y) = \mathscr{N}(\mathscr{I}(x,\mathscr{N}(y))$. No necesariamente es una t-norma.

2.2 Conjuntos Difusos y Relaciones Difusas

Un conjunto difuso A en un universo no vacío U se identifica con su función de pertenencia $A: U \to [0,1]$. La colección de todos los conjuntos difusos en U se denota por $\mathscr{F}(U)$.

Para cada α en $[0,1]$, el *conjunto (difuso) constante* $\hat{\alpha}$ se define, para x en U, como $\hat{\alpha}(x) = \alpha$. En el caso nítido, los únicos conjuntos constantes son $\emptyset$ y U.

Sean $A \in \mathscr{F}(U)$ y $\mathscr{N}$ una negación. El *$\mathscr{N}$-complemento $A^{\mathscr{N}}$ de A* se define como $A^{\mathscr{N}}(x) = \mathscr{N}(A(x))$ para todo $x \in U$.

Una relación binaria difusa R en U es un conjunto difuso en $U \times U$. R se llama *reflexiva* si $R(x,x) = 1$ para todo $x \in U$, *simétrica* si $R(x,y) = R(y,x)$ para todo $x,y \in U$, y *serial* si $\sup_{y \in U} R(x,y) = 1$ para cada $x \in U$. Cada relación difusa simétrica R cumple $R = R'$.

Dada una t-norma $\mathscr{T}$, R se llama *$\mathscr{T}$-transitiva* si para cada x, y, z en U, $\mathscr{T}(R(x,y),R(y,z)) \le R(x,z)$. Si R es reflexiva, simétrica y $\mathscr{T}$-transitiva, se llama una *relación de $\mathscr{T}$-similitud*. Cuando $\mathscr{T} =$ mín, hablamos simplemente de una *relación de similitud*. Como el mínimo es la t-norma más grande, una relación de similitud es una relación de $\mathscr{T}$-similitud para cada t-norma $\mathscr{T}$.

2.3 Aproximación Inferior y Superior en la Teoría de los Conjuntos Rugosos

2.3.1 Modelo de Pawlak

Un *espacio de aproximación clásico* es un par (U,R) donde U es un conjunto no vacío U y R es una relación de equivalencia en U. La aproximación rugosa de un conjunto nítido A en U por R es el par de conjuntos $(R{\downarrow}A, R{\uparrow}A)$ definidos, para $x \in U$, como

$$x \in R{\downarrow}A \quad \Leftrightarrow \quad (\forall y \in U)((y,x) \in R \Rightarrow y \in A), \tag{1}$$

$$x \in R{\uparrow}A \quad \Leftrightarrow \quad (\exists y \in U)((y,x) \in R \wedge y \in A). \tag{2}$$

Un par (A_1,A_2) de conjuntos en U se llama *conjunto rugoso* en (U,R) si hay un conjunto A en U tal que $A_1 = R{\downarrow}A$ y $A_2 = R{\uparrow}A$. El modelo original de Pawlak se ha generalizado de varias maneras. Abajo, nos centramos en el caso donde A y R son un conjunto difuso y una relación difusa en U respectivamente.

2.3.2 Modelo Difuso-Rugoso Basado en Implicaciones y Conjunciones

En esta sección, consideramos un *espacio de aproximación difuso*, es decir, un par (U,R) donde U es en un conjunto no vacío U y R una relación binaria difusa en U. Muchas definiciones de conjuntos rugosos difusos surgen como extensiones fieles de las ecuaciones (1) y (2) al caso $[0,1]$-valorado. En particular, Dubois y Prade trabajaron con una relación de similitud R, y remplazaron la implicación y la conjunción booleanas por la S-implicación $\mathscr{I}_{\text{máx},\mathscr{N}_s}$ (implicación de Kleene-Dienes) y el mínimo, respectivamente. En [5], se propuso el siguiente formato para las aproximaciones usando implicaciones y conjunciones generales.

Definición 1 *Sea (U,R) un espacio de aproximación difuso, A un conjunto difuso en U, $\mathscr{I}$ una implicación y $\mathscr{C}$ una conjunción. La* aproximación $(\mathscr{I},\mathscr{C})$-difusa-rugosa *de A por R es el par de conjuntos difusos $(R{\downarrow_{\mathscr{I}}}A, R{\uparrow_{\mathscr{C}}}A)$ definidos, para $x \in U$, como*

$$(R{\downarrow_{\mathscr{I}}}A)(x) = \inf_{y\in U} \mathscr{I}(R(y,x),A(y)) \tag{3}$$

$$(R{\uparrow_{\mathscr{C}}}A)(x) = \sup_{y\in U} \mathscr{C}(R(y,x),A(y)). \tag{4}$$

Un par (A_1,A_2) de conjuntos difusos en U se llama un conjunto difuso rugoso *en (U,R) si hay un conjunto difuso A en U tal que $A_1 = R{\downarrow_{\mathscr{I}}}A$ y $A_2 = R{\uparrow_{\mathscr{C}}}A$.*

La elección de la conjunción y de la implicación se hace típicamente guiada por las propiedades que se requieren a las aproximaciones correspondientes. Algunas propiedades que se imponen habitualmente se muestran en la Tabla 1: dualidad (D); inclusión de la aproximación inferior en A y de A en la aproximación superior (INC); monotonía de los conjuntos (MC) y de las relaciones (MR); interacción con la intersección y la unión (IU), idempotencia (ID), invariancia por las aproximaciones de los conjuntos vacío y universo (VU), y de los conjuntos constantes (CC). Téngase en cuenta que en un espacio de aproximación de Pawlak, todas se cumplan, salvo (CC), que se define exclusivamente para los conjuntos difusos.

Tabla 1: Propiedades de las aproximaciones inferior y superior.

Referencia	Definición
(D)	$R{\downarrow}A = (R{\uparrow}A^{\mathscr{N}})^{\mathscr{N}}$
(INC)	$R{\downarrow}A \subseteq A \subseteq R{\uparrow}A$
(MC)	$A \subseteq B \Rightarrow \begin{cases} R{\downarrow}A \subseteq R{\downarrow}B \\ R{\uparrow}A \subseteq R{\uparrow}B \end{cases}$
(MR)	$R_1 \subseteq R_2 \Rightarrow \begin{cases} R_1{\downarrow}A \supseteq R_2{\downarrow}A \\ R_1{\uparrow}A \subseteq R_2{\uparrow}A \end{cases}$
(IU)	$R{\downarrow}(A\cap B) = R{\downarrow}A \cap R{\downarrow}B$ $R{\uparrow}(A\cup B) = R{\uparrow}A \cup R{\uparrow}B$
(ID)	$R{\downarrow}(R{\downarrow}A) = R{\downarrow}A$ $R{\uparrow}(R{\uparrow}A) = R{\uparrow}A$
(VU)	$R{\downarrow}\emptyset = \emptyset = R{\uparrow}\emptyset$ $R{\downarrow}U = U = R{\uparrow}U$
(CC)	$R{\downarrow_{\mathscr{I}}}\hat{\alpha} = \hat{\alpha}$ $R{\uparrow_{\mathscr{C}}}\hat{\alpha} = \hat{\alpha}$

En lo que sigue, asumimos que (U,R), (U,R_1) y (U,R_2) son espacios de aproximación difusos, A y B son conjuntos difusos en U, $\mathscr{I}$ es una implicación, $\mathscr{C}$ una conjunción y $\mathscr{N}$ una negación *involutiva*. Se pueden comprobar los siguientes resultados (cf. [5]).

Proposición 1 *Si $\mathscr{C}$ es la conjunción inducida por $\mathscr{I}$ y $\mathscr{N}$, es decir, $\mathscr{C} = \mathscr{C}_{\mathscr{I},\mathscr{N}}$, entonces (D) se cumple.*

Esta propiedad se cumple en particular cuando el par $(\mathscr{I},\mathscr{C})$ consiste en una S-implicación $\mathscr{I}_{\mathscr{D},\mathscr{N}}$ y una conjunción $\mathscr{C}$, donde $\mathscr{D}$ es la disyunción $\mathscr{N}$-dual de $\mathscr{C}$, y también cuando $(\mathscr{I},\mathscr{C})$ consiste en una R-implicación $\mathscr{I}_{\mathscr{T}}$ y una t-norma $\mathscr{T}$ continua por la izquierda tal que $\mathscr{N} = \mathscr{N}_{\mathscr{I}_{\mathscr{T}}}$.

Proposición 2 *Si R es reflexiva, $\mathscr{I}$ es una implicación frontera y $\mathscr{C}$ es una conjunción frontera, entonces (INC) se cumple.*

Proposición 3 *(MC), (MR) y (IU) siempre se cumplen.*

Proposición 4 *Si R es reflexiva y $\mathscr{T}$-transitiva, donde $\mathscr{T}$ es una t-norma continua por la izquierda y $(\mathscr{I},\mathscr{C})$ consiste en la R-implicación $\mathscr{I}_{\mathscr{T}}$, y $\mathscr{T}$; o $(\mathscr{I},\mathscr{C})$ consiste en la S-implicación $\mathscr{I}_{\mathscr{S},\mathscr{N}}$ basada en la t-conorma $\mathscr{N}$-dual de $\mathscr{T}$ con respecto a una negación involutiva $\mathscr{N}$, y $\mathscr{T}$, (ID) se cumple.*

Proposición 5 *Si R es serial, (VU) se cumple.*

Proposición 6 *Si R es reflexiva, $\mathscr{I}$ una implicación frontera y $\mathscr{C}$ una conjunción frontera, entonces (CC) se cumple.*

3 Modelos Difusos-Rugosos Tolerantes al Ruido

Debido al uso del ínf y del sup en las ecuaciones (3) y (4), el resultado de las aproximaciones se determina por un solo elemento (en el caso de que U sea finito). Esta característica puede resultar perjudicial en un contexto de análisis de datos, considerando que algunos datos muestrales pueden ser erróneas (ruidosas), y por tanto, pueden perturbar las aproximaciones inferior y superior y así debilitar los algoritmos de aprendizaje automático que las emplean; este efecto se estudió por ejemplo en [7]. Varios autores han considerado maneras para esquivar este problema, proponiendo alternativas robustas para la aproximación inferior y superior. Abajo, discutimos cuatro enfoques bien conocidos. A lo largo de esta sección, asumimos que A es un conjunto difuso en un espacio de aproximación difuso (U,R) con U finito.

Empezamos por el modelo de β-precisión de Fernández-Salido y Murakami [12]. En esencia, este modelo remplaza el ínf y el sup por operadores de agregación menos estrictos, los cuales recordamos primero.

Definición 2 *Dados una t-norma $\mathcal{T}$, una t-conorma $\mathcal{S}$, $\beta \in [0,1]$ y $n > 1$, la quasi-t-norma $\mathcal{T}_\beta$ y la quasi-t-conorma S_β de β-precisión de orden n son funciones de $[0,1]^n$ en $[0,1]$ tal que para cada $\mathbf{x} = (x_1,\ldots,x_n)$ en $[0,1]^n$,*

$$\mathcal{T}_\beta(\mathbf{x}) = \mathcal{T}(y_1,\ldots,y_{n-m}),$$
$$\mathcal{S}_\beta(\mathbf{x}) = \mathcal{S}(z_1,\ldots,z_{n-p}),$$

donde $y_i = x_j$ si x_j es el i-ésimo elemento más grande de $\mathbf{x}$, *$z_i = x_j$ si x_j es el i-ésimo elemento más pequeño de* $\mathbf{x}$, *y*

$$m = \text{máx}\left\{ i \in \{0,\ldots,n\} \mid i \leq (1-\beta)\sum_{j=1}^{n} x_j \right\},$$
$$p = \text{máx}\left\{ i \in \{0,\ldots,n\} \mid i \leq (1-\beta)\sum_{j=1}^{n} (1-x_j) \right\}.$$

Definición 3 *Sea $\mathcal{N}$ una negación involutiva, $\mathcal{T}$ una t-norma, $\mathcal{S}$ su t-conorma $\mathcal{N}$-dual y $\beta \in [0,1]$. Dadas una implicación $\mathcal{I}$ y una conjunción $\mathcal{C}$, la aproximación β-precisión de A por R es el par de conjuntos difusos $(R{\downarrow_{\mathcal{I},\mathcal{T}_\beta}}A, R{\uparrow_{\mathcal{C},\mathcal{S}_\beta}}A)$, definidos, para x en U, como*

$$(R{\downarrow_{\mathcal{I},\mathcal{T}_\beta}}A)(x) = \mathcal{T}_\beta \underset{y \in U}{\langle} \mathcal{I}(R(y,x),A(y)) \rangle,$$
$$(R{\uparrow_{\mathcal{C},\mathcal{S}_\beta}}A)(x) = \mathcal{S}_\beta \underset{y \in U}{\langle} \mathcal{C}(R(y,x),A(y)) \rangle.$$

En otras palabras, controlados por β, los elementos más pequeños no se toman en cuenta en el cálculo de la aproximación inferior. Por analogía, los elementos más grandes no afectan a la aproximación superior. Además, el uso de t-normas y t-conormas distintas a mín y máx permiten una interacción mayor entre los argumentos a agregar.

Una solución alternativa es "ahogar" el conjunto difuso A por un nivel de incertidumbre, como pasa en el modelo de precisión variable difuso-rugoso (PVDR) de Zhao et al. [14].

Definición 4 *Dadas una implicación $\mathcal{I}$, una conjunción $\mathcal{C}$, una negación $\mathcal{N}$, y $\alpha \in [0,1)$, la aproximación de precisión variable α de A por R es el par de conjuntos difusos $(R{\downarrow_{\mathcal{I},\alpha}}A, R{\uparrow_{\mathcal{C},\alpha}}A)$, definidos, para x en U, como*

$$(R{\downarrow_{\mathcal{I},\alpha}}A)(x) = \inf_{y \in U} \mathcal{I}(R(y,x), \text{máx}\{\alpha, A(y)\}),$$
$$(R{\uparrow_{\mathcal{C},\alpha}}A)(x) = \sup_{y \in U} \mathcal{C}(R(y,x), \text{mín}\{\mathcal{N}(\alpha), A(y)\}).$$

Como tal, los elementos con grados de pertenencia muy bajos a A son suavizados por un nivel de incertidumbre α para limitar su impacto en la aproximación inferior, mientras lo contrario ocurre para la aproximación superior.

El siguiente modelo de conjuntos rugosos por cuantificadores difusos (CRCD) [2] difiere bastante de los modelos anteriores, ya que no depende de implicaciones y conjunciones, sino que emplea la noción de cuantificador difuso, extendiendo el modelo nítido de conjuntos rugosos de precisión variable (CRPV) de Ziarko [15].

Definición 5 *Un cuantificador difuso creciente regularmente es una función $Q\colon [0,1] \to [0,1]$ creciente que satisface $Q(0) = 0$ y $Q(1) = 1$.*

Definición 6 *Dado un par (Q_u, Q_l) de cuantificadores difusos crecientes regularmente, la aproximación cuantificada por (Q_u, Q_l) de A por R es el par de conjuntos difusos $(R{\downarrow_{Q_u}}A, R{\uparrow_{Q_l}}A)$ definidos, para x en U, como*

$$(R{\downarrow_{Q_u}}A)(x) = \begin{cases} Q_u\left(\frac{|Rx \cap A|}{|Rx|}\right) & \text{si } Rx \neq \emptyset \\ 1 & \text{si } Rx = \emptyset \end{cases}$$
$$(R{\uparrow_{Q_l}}A)(x) = \begin{cases} Q_l\left(\frac{|Rx \cap A|}{|Rx|}\right) & \text{si } Rx \neq \emptyset \\ 1 & \text{si } Rx = \emptyset. \end{cases}$$

donde Rx es el conjunto difuso en U definido como $Rx(y) = R(y,x)$ para y en U, y $|\cdot|$ denota cardinalidad.

En la interpretación propuesta por [2], x pertenece a $R{\downarrow_{Q_u}}A$ en el grado que "la mayoría" de los elementos relacionados con x pertenecen a A, y a $R{\uparrow_{Q_l}}A$ en el grado que "por lo menos algunos" elementos relacionados con x pertenecen a A.

El último modelo que consideramos [4], remplaza ínf y sup en las ecuaciones (3) y (4) por agregaciones de tipo OWA (promedio ponderado ordenado), cuya definición se recuerda primero.

Definición 7 *Dados una secuencia D de n valores escalares y un vector de pesos $W = \langle w_1,\ldots,w_n\rangle$ de longitud n, tal que $w_i \in [0,1]$ para cada $i \in \{1,\ldots,n\}$, y $\sum_{i=1}^{n} w_i = 1$, y dada la permutación σ de $\{1,\ldots,n\}$ tal que $d_{\sigma(i)}$ es el i-ésimo valor más grande de D, el* operador OWA *actuando en D produce el valor:*

$$\text{OWA}_W(D) = \sum_{i=1}^{n} w_i d_{\sigma(i)}.$$

Definición 8 *Dadas una implicación $\mathcal{I}$ y una conjunción $\mathcal{C}$, y vectores de pesos OWA W_1 y W_2 de longitud n, la aproximación ponderada por (W_1, W_2) de A por R es el par de conjuntos difusos $(R{\downarrow_{\mathcal{I},W_1}}A, R{\uparrow_{\mathcal{C},W_2}}A)$ definidos, para $x \in U$, como*

$$(R{\downarrow_{\mathcal{I},W_1}}A)(x) = \text{OWA}_{W_1} \underset{y \in U}{\langle} \mathcal{I}(R(y,x),A(y)) \rangle,$$
$$(R{\uparrow_{\mathcal{C},W_2}}A)(x) = \text{OWA}_{W_2} \underset{y \in U}{\langle} \mathcal{C}(R(y,x),A(y)) \rangle.$$

Variando los vectores de pesos en el operador OWA, se obtienen diferentes modelos de conjuntos rugosos difusos, como ejemplos podemos destacar aquellos propuestos en [7].

4 Propiedades

En esta sección, evaluamos los cuatro modelos robustos de la sección anterior con respecto a las propiedades de la Tabla 1. Empezamos por el model de β-precisión. Debido a la monotonicidad de las conectivas lógicas difusas implicadas, (MC) y (MR) siempre se cumplen. Para (D), tenemos la siguiente proposición.

Proposición 7 *Sea $\mathcal{T}$ una t-norma y sea $\mathcal{S}$ la t-conorma dual con respecto a la negacíon estándar $\mathcal{N}_s$. Si $\mathcal{C}$ es la conjunción inducida por $\mathcal{I}$ y $\mathcal{N}_s$, es decir, $\mathcal{C} = \mathcal{C}_{\mathcal{I},\mathcal{N}_s}$, (D) se cumple para $\downarrow_{\mathcal{I},\mathcal{T}_\beta}$ y $\uparrow_{\mathcal{C},\mathcal{S}_\beta}$.*

Esta propiedad no se cumple en general para $\mathcal{N}$ involutiva, ya que en ese caso $\mathcal{T}_\beta$ y $\mathcal{S}_\beta$ no necesariamente descartan la misma cantidad de elementos. Las propiedades restantes no son válidas. El siguiente ejemplo lo demuestra para (INC).

Ejemplo 1 *Sea $\mathcal{I}$ una implicación frontera y $\mathcal{C}$ una conjunción frontera. Sean $U = \{y_0, \ldots, y_{10}\}$, A un conjunto difuso tal que $A(y_i) = \frac{i}{10}$ para cada $i \in \{0, \ldots, 10\}$, y R una relación de similitud tal que $R(y_i, y_j) = 1$ para cada $i, j \in \{0, \ldots, 10\}$. Sean $(\mathcal{T}, \mathcal{S})$ iguales a $(\text{mín}, \text{máx})$ y $\beta = 0.8$. Tenemos que $\mathcal{I}(R(z,x), A(z)) = A(z) = \mathcal{C}(R(z,x), A(z))$ para cada $x, z \in U$. Se obtiene que $m = p = 1$, así que en la aproximación inferior (resp., superior) el valor más bajo (resp., más alto) se omite. Obtenemos para cada $x \in U$ que*

$$\underset{y\in U}{\text{mín}_{0.8}} \mathcal{I}(R(y,x), A(y)) = \underset{y\in U}{\text{mín}_{0.8}} A(y) = 0.1,$$

lo que implica que $(R{\downarrow_{\mathcal{I},\text{mín}_{0.8}}} A)(y_0) > A(y_0)$. A continuación,

$$\underset{y\in U}{\text{máx}_{0.8}} \mathcal{C}(R(y,x), A(y)) = \underset{y\in U}{\text{máx}_{0.8}} A(y) = 0.9,$$

y por lo tanto $(R{\uparrow_{\mathcal{C},\text{máx}_{0.8}}} A)(y_{10}) < A(y_{10})$.

Para el modelo PVDR, (MC), (MR) y (IU) son válidas. Las pruebas de estas propiedades, al igual que las de las proposiciones abajo, son completamente análogas a las obtenidas para el modelo basado en implicaciones y conjunciones.

Proposición 8 *Si $\mathcal{C}$ es la conjunción inducida por $\mathcal{I}$ y $\mathcal{N}$, es decir, $\mathcal{C} = \mathcal{C}_{\mathcal{I},\mathcal{N}}$, (D) se cumple para $\downarrow_{\mathcal{I},\alpha}$ y $\uparrow_{\mathcal{C},\alpha}$.*

Proposición 9 *Si R es una relación difusa reflexiva y $\mathcal{T}$-transitiva, donde $\mathcal{T}$ es una t-norma que es continua por la izquierda, y el par $(\mathcal{I}, \mathcal{C})$ consiste en la R-implicación $\mathcal{I}_\mathcal{T}$ y la t-norma $\mathcal{T}$, (ID) se cumple para $\downarrow_{\mathcal{I},\alpha}$ y $\uparrow_{\mathcal{C},\alpha}$.*

Las propiedades restantes no se cumplen en general para este modelo. Abajo consideramos un ejemplo para (VU).

Ejemplo 2 *Sean $\mathcal{I}$, $\mathcal{C}$, U y R definidos como en el Ejemplo 1, y sean $\alpha = 0.1$ y $\mathcal{N} = \mathcal{N}_s$. Obtenemos para $x \in U$ que*

$$(R{\downarrow_{\mathcal{I},\alpha}} \emptyset)(x) = \inf_{y\in U} \mathcal{I}(1, \text{máx}\{0.7, \emptyset(y)\}) = 0.1,$$
$$(R{\uparrow_{\mathcal{C},\alpha}} U)(x) = \sup_{y\in U} \mathcal{C}(1, \text{mín}\{0.3, U(y)\}) = 0.9.$$

Como el modelo nítido de CRPV de Ziarko [15] viola (D), (INC), (MR), (IU) and (ID), el modelo de CRCD, que lo extiende, sufre las mismas limitaciones. Por otro lado, como se usan cuantificadores difusos crecientes, (MC) se cumple. Finalmente, se puede comprobar que (CC) no se cumple en general, pero sí podemos garantizar la propiedad (VU).

Proposición 10 *Si R es una relación serial, (VU) se cumple para el modelo CRCD.*

Por último, el modelo basado en OWA siempre satisface las propiedades (MC) y (MR) de monotonicidad, y además cumple la dualidad (D) bajo ciertas condiciones.

Proposición 11 *Sea W_1 un vector de pesos OWA de longitud n. Si el par $(\mathcal{I}, \mathcal{C})$ consiste en una implicación $\mathcal{I}$ y la conjunción inducida por $\mathcal{I}$ y $\mathcal{N}_S$, es decir, $\mathcal{C} = \mathcal{C}_{\mathcal{I},\mathcal{N}_S}$, y W_2 es el vector de pesos OWA definido por $(W_2)_i = (W_1)_{n-i+1}$, para $i \in \{1, \ldots, n\}$, (D) se cumple para $\mathcal{N} = \mathcal{N}_S$.*

Las demás propiedades no se cumplen en general. El siguiente contraejemplo lo demuestra para (VU), (CC) y (INC).

Ejemplo 3 *Sea $U = \{x, y\}$ y R una relación de similitud tal que $R(x,y) = 0.5$. $\mathcal{I}$ y $\mathcal{C}$ se definen, para x, y en $[0,1]$, como $\mathcal{I}(x,y) = \text{mín}(1, 1 - x + y)$ y $\mathcal{C}(x,y) = \text{máx}(0, x + y - 1)$. Además, $W_1 = \langle \frac{1}{3}, \frac{2}{3} \rangle$ y $W_2 = \langle \frac{2}{3}, \frac{1}{3} \rangle$. Calculamos el grado de pertenencia de la aproximación inferior de $\emptyset = \hat{0}$ y la aproximación superior de $U = \hat{1}$ para x:*

$$(R{\downarrow_{\mathcal{I},W_1}} \hat{0})(x) = \frac{1}{6},$$
$$(R{\uparrow_{\mathcal{C},W_2}} \hat{1})(x) = \frac{5}{6}.$$

La Tabla 2 resume los resultados de nuestro análisis. Un ✓ signica que la propiedad correspondiente siempre se cumple, mientras que ✰ se usa para indicar que existen condiciones bajo las cuales la propiedad se cumple. En los casos restantes, escribimos ✗. La conclusión más importante parece ser que ninguno de los modelos tolerantes al ruido es capaz de conservar todas las propiedades del modelo original de Pawlak y de la extensión basada en implicaciones y conjunciones, incluso cuando R es una relación de similitud. Si se requiere la dualidad, solo los modelos de β-precisión, de PVDR y de OWA sirven; sin embargo, ya que en muchas aplicaciones de análisis de datos solo se usa la aproximación inferior, este problema puede ser de menor importancia. Un defecto más grave, quizás, es la violación de (MR) por el modelo CRCD: efectivamente, cuando la granularidad de los datos impuesta por la relación difusa R se hace más fina, esperamos por ejemplo que la aproximación inferior no se reduzca.

Tabla 2: Evaluación de las propiedades

propiedad	IC	β-PREC	PVDR	CRCD	OWA
(D)	✰	✰	✰	✗	✰
(INC)	✰	✗	✗	✗	✗
(MC)	✓	✓	✓	✓	✓
(MR)	✓	✓	✓	✗	✓
(IU)	✓	✗	✓	✗	✗
(ID)	✰	✗	✰	✗	✗
(VU)	✰	✗	✗	✰	✗
(CC)	✰	✗	✗	✗	✗

5 Conclusiones y Trabajo Futuro

En este artículo, hemos presentado un resumen de algunos modelos actuales de conjuntos rugosos difusos diseñados para manejar datos con ruido, y los hemos evaluado con respecto a una serie de propiedades deseables. Hemos comprobado que, en comparación con el modelo basado en implicaciones y conjunciones, menos propiedades del modelo clásico de Pawlak se mantienen, lo cual parece ser el precio lógico a pagar si queremos tener una mayor robustez. El uso práctico de cada modelo, sin embargo, se determinará por su rendimiento en aplicaciones concretas. En particular, en el contexto de análisis de datos, como trabajo futuro pensamos evaluar de forma experimental la estabilidad de las aproximaciones de las clases de decisión cuando los datos de atributos condicionales y de clase están contaminados por ruido.

Agradecimientos

Este trabajo ha sido financiado parcialmente por el Ministerio de Ciencias y Tecnología bajo el Proyecto con referencia TIN2011-28488, y por el Fondo Especial de Investigación (BOF) de la Universidad de Gante. Lluis Godo reconoce financiación parcial por parte del Proyecto MINECO con referencia TIN2012-39348-C02-01.

Referencias

[1] Boixader, D., Jacas, J., Recasens, J.: Upper and lower approximation of fuzzy sets. International Journal of General Systems **29**, 555-568 (2000)

[2] Cornelis, C., De Cock, M., Radzikowska, A.: Vaguely quantified rough sets. In: Proceedings of 11th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC2007). (2007) 87–94

[3] Cornelis, C., Hurtado Martín, G., Jensen, R., Ślęzak, D.: Attribute selection with fuzzy decision reducts. Information Sciences **180**(2) (2010) 209–224

[4] Cornelis, C., Verbiest, N., Jensen, R.: Ordered weighted average based fuzzy rough sets. In: Proceedings of the 5th International Conference on Rough Sets and Knowledge Technology (RSKT2010). (2010) 78–85

[5] D'eer, L., Verbiest, N., Cornelis, C., Godo, L.: Implication-conjunction based models of fuzzy rough sets: definitions and properties. Proceedings of RSFDGrC, Lecture Notes in Artificial Intelligence **8170** (2013) 169–179

[6] Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. International Journal of General Systems **17** (1990) 191–209

[7] Hu, Q., Zhang, L., An, S., Zhang, D., Yu, D.: On robust fuzzy rough set models. IEEE Transactions on Fuzzy Systems **20**(4) (2012) 636 – 651

[8] Jensen, R., Shen, Q.: Fuzzy-rough sets assisted attribute selection. IEEE Transactions on Fuzzy Systems **15**(1) (2007) 73–89

[9] Morsi, N., Yakout, M.: Axiomatics for fuzzy rough set. Fuzzy Sets Systems **100** (1998) 327–342

[10] Pawlak, Z.: Rough sets. International journal of computer and information sciences **11**(5) (1982) 341–356

[11] Radzikowska, A., Kerre, E.: A comparative study of fuzzy rough sets. Fuzzy Sets and Systems **126** (2002) 137–155

[12] Fernández-Salido, J.M., Murakami, S.: Rough set analysis of a general type of fuzzy data using transitive aggregations of fuzzy similarity relacións. Fuzzy Sets and Systems **139** (2003) 635–660

[13] Verbiest, N., Cornelis, C., Herrera, F.: FRPS: A fuzzy rough prototype selection method. Pattern Recognition **46**(10) (2013) 2770–2782

[14] Zhao, S., Tsang, E.C.C., Chen, D.: The model of fuzzy variable precision rough sets. IEEE Transactions on Fuzzy Systems **17**(2) (2009) 451–467

[15] Ziarko, W.: Variable Precision Rough Set Model. Journal of Computer and System Sciences **46** (1993) 39–59