



CEDI 2010 **VALENCIA**

7 A 10 DE SEPTIEMBRE DE 2010

III CONGRESO ESPAÑOL DE INFORMÁTICA

UNIVERSIDAD POLITÉCNICA DE VALENCIA

**Actas del III Simposio sobre Lógica Fuzzy
y Soft Computing**

| LFSC2010 | (EUSFLAT)

EDITORES

Luis Martínez, Edurne Barrenechea, Macarena Espinilla, Jesús Alcalá, Victoria López,
Manuel Mucientes, José Ángel Olivas, Rosa M^a Rodríguez

Actas del III Simposio sobre Lógica Fuzzy y Soft Computing, LFSC2010 (EUSFLAT)

Editores: Luis Martínez, Edurne Barrenechea, Macarena Espinilla Jesús Alcalá, Victoria López, Manuel Mucientes, José Ángel Olivas, Rosa M^a Rodríguez

ISBN: 978-84-92812-65-3

IBERGARCETA PUBLICACIONES, S.L., Madrid, 2010

Edición: 1^a

Impresión: 1^a

Nº de páginas: 494

Formato: 17 x 24

Materia CDU: 004 Ciencia y tecnología de los ordenadores. Informática

Reservados los derechos para todos los países de lengua española. De conformidad con lo dispuesto en el artículo 270 y siguientes del código penal vigente, podrán ser castigados con penas de multa y privación de libertad quienes reprodujeran o plagiaran, en todo o en parte, una obra literaria, artística o científica fijada en cualquier tipo de soporte sin la preceptiva autorización. Ninguna parte de esta publicación, incluido el diseño de la cubierta, puede ser reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea éste electrónico, químico, mecánico, el electro-óptico, grabación, fotocopia o cualquier otro, sin la previa autorización escrita por parte de la editorial.

Dirijase a CEDRO (Centro Español de Derechos Reprográficos), www.cedro.org, si necesita fotocopiar o escanear algún fragmento de esta obra.

COPYRIGHT © 2010 IBERGARCETA PUBLICACIONES, S.L.
info@garceta.es

Actas del III Simposio sobre Lógica Fuzzy y Soft Computing, LFSC2010 (EUSFLAT)

Derechos reservados ©2010 respecto a la primera edición en español, por LOS AUTORES

Derechos reservados ©2010 respecto a la primera edición en español, por IBERGARCETA PUBLICACIONES, S.L.

1^a Edición, 1^a Impresión

ISBN: 978-84-92812-65-3

Depósito legal: M-

Maquetación: Los Editores

Coordinación del proyecto: @LIBROTEX

Portada: Estudio Dixi

Impresión y encuadernación:

OI: 28/2010

PRINT HOUSE, S.A.

IMPRESO EN ESPAÑA -PRINTED IN SPAIN

Nota sobre enlaces a páginas web ajenas: Este libro puede incluir referencias a sitios web gestionados por terceros y ajenos a IBERGARCETA PUBLICACIONES, S.L., que se incluyen sólo con finalidad informativa. IBERGARCETA PUBLICACIONES, S.L., no asume ningún tipo de responsabilidad por los daños y perjuicios derivados del uso de los datos personales que pueda hacer un tercero encargado del mantenimiento de las páginas web ajenas a IBERGARCETA PUBLICACIONES, S.L., y del funcionamiento, accesibilidad y mantenimiento de los sitios web no gestionados por IBERGARCETA PUBLICACIONES, S.L., directamente. Las referencias se proporcionan en el estado en que se encuentran en el momento de publicación sin garantías expresas o implícitas, sobre la información que se proporciona en ellas.

Contenido

III Simposio sobre Lógica Fuzzy y Soft Computing, LFSC2010 (EUSFLAT)

SESIÓN ESPECIAL 1 Aplicaciones del Soft Computing

Organizador: Manuel Mucientes

Cálculo del Grado de la Normalidad en Sistemas Expertos de Vídeo-Vigilancia Escalables.....	3
Javier Albusac, David Vallejo, J. J. Castro-Schez, L. Jiménez-Linares	
Localización probabilística del humanoide Nao en el campo de la Robocup.....	11
Francisco J. Rdez. Lera, Juan F. García, Camino Fdez. Llamas, Vicente Matellán	
Reconstrucción volumétrica de personas utilizando la teoría de Dempster-Shafer.....	19
Luis Díaz, Rafael Muñoz, Enrique Yeguas, Rafael Medina	
Expert Fuzzy System for Prediction of Pedestrian Accidents.....	27
J. L. Castro, M. Delgado, J. Medina, M. D. Ruiz-Lozano	
Un Modelo Evolutivo para el Diseño de Materiales Compuestos Laminados Simétricos.....	35
E. Tenorio, J. I. Peláez, J. M. Doña, J. Gómez Ruiz	
Localización de Olivos en Imágenes de Alta Resolución mediante Técnicas de Razonamiento Aproximado.....	41
J. Moreno-Garcia, L. Jiménez, L. Rodriguez-Benítez, C. Solana-Cipres	

SESIÓN ESPECIAL 2 Toma de Decisiones en Ambientes de Incertidumbre

Organizadoras: Macarena Espinilla y Rocío de Andrés

A new mobile decision support system to manage dynamic decisión situations.....	51
I. J. Pérez, F. J. Cabrerizo, S. Alonso, E. Herrera-Viedma	
Modelo Lingüístico de QoS para Servicios de Red.....	59
Sergio Gramajo, Luis Martínez	
Aplicación Web para Evaluación de Desempeño 360 Grados Basada en un Modelo Lingüístico Multigranular.....	67

Macarena Espinilla, Francisco J. Martínez, Luis Martínez	
Consensus reaching with different aggregation techniques.....	75
Francisco Mata, Juan Carlos Martínez	
Valoración de proyectos públicos mediante escalas finitas: un enfoque basado en distancias.....	83
Eduarne Falcó Díaz de Cerio, José Luis García Lapresta	
Modelo para imputar valores perdidos en evaluación de Nuclear <i>Safeguards</i>	91
R.M. Rodríguez, D. Ruan, J. Liu, L. Martínez	
Aplicación de un Algoritmo de Inducción de Reglas Borrosas para la Estimación del Tiempo de Trabajo Efectivo.....	99
Francisco P. Romero, Juan Moreno-García, Arturo Peralta	
Integración de sistemas de clasificación difusos en una guía clínica ejecutable.....	107
U. Segundo, J. M. Pikatza, D. Buenestado, A. Iruetaguena, R. Barrena, J. J. García, L. Aldámiz-Echevarría, P. Sanjurjo	
Transitividad colectiva en mayorías con umbral de apoyo.....	115
Bonifacio Llamazares, Patrizia Pérez Asurmendi, José Luis García Lapresta	
Evaluación de riesgos en redes eléctricas: una aproximación difusa.....	123
Juan Gómez-Romero, Antonio Berlanga, José M. Molina, Ángel Ramos Gómez	

SESIÓN ESPECIAL 3 Lógica Difusa y Procesamiento de Imagen

<i>Organizadores: Eduarne Barrenechea, Miguel Pagola y Francisco Javier Fernández</i>	
Segmentación dinámica en escenas de tráfico sobre vídeo comprimido H.264/AVC....	133
C.J. Solana-Ciprés, L. Rguez-Benítez, L. Jiménez, J. Moreno-García	
Operadores apertura y cierre en la morfología matemática derivada de t-normas discretas. Transformaciones de Top-Hat y filtros básicos.....	141
M. González-Hidalgo, S. Massanet, J. Torrens	
Morfología Matemática sobre sistemas relacionales borrosos. Una interpretación de conceptos básicos de Lógica Borrosa mediante erosiones y dilataciones morfológicas.	149
Ramón Fuentes-González	
Generación de imágenes difusas de bordes a partir de imágenes de gradientes.....	157
C. López-Molina, H. Bustince, B. De Baets, J. Fernández, A. Jurio	

Medidas de similaridad difusas e intuicionistas de Atanassov aplicadas a reconocimiento de patrones.....	165
F. Casado, D. Paternain, M. Pagola, E. Barrenechea	
Un método orientado a píxeles para la clasificación difusa de imágenes aéreas.....	173
A. Sánchez, D. de Miguel, A. Conci, E. Nunes	

SESIÓN ESPECIAL 4 Teoría de la Lógica Fuzzy y Funciones de Agregación

<i>Organizador: Luis Martínez</i>	
Una formulación axiomática de la noción de antagonismo semántico.....	133
J. Tinguaro Rodríguez, Begoña Vitoriano, Javier Montero	
Sobre la antonimia y su extensión a los conjuntos de Atanassov.....	189
Susana Cubillo, Elena E. Castiñeira, Wilmer Montilla	
Multi-Adjoint Viewpoint on Property/Object-Oriented Concept Lattices.....	197
Jesús Medina	
Un análisis de las funciones que generalizan a las medias ponderadas y a los operadores OWA.....	205
Bonifacio Llamazares	
Sobre implicaciones borrosas basadas en uninormas que verifican la ecuación $I(x; y) = I(x; I(x; y))$	213
Sebastià Massanet, Joan Torrens	
Funciones de agregación discretas “kernel”.....	221
M. Mas, M. Monserrat, J. Torrens	

SESIÓN ESPECIAL 5 Arte y Soft Computing

<i>Organizador: Vicente Liern</i>	
Un compositor musical automático basado en una metaheurística cooperativa.....	231
G. Acampora, J. M. Cadenas, R. De Prisco, V. Loia, E. Muñoz, R. Zaccagnino	
Generación de objetos arquitectónicos usando lógica fuzzy: aplicación a diseños de Gaudí.....	239
Amadeo Monreal Pujadas, Joan Jacas Moral	
Un algoritmo de recuperación por contenidos en bases de datos de audio.....	247

Pedro Zuccarello, Teresa León, Concepción Díaz Vicedo	
Herramientas informáticas flexibles para la clase de música.....	255
José Luis Godofredo, Vicente Liern	
Una solución de consenso entre la teoría y la práctica musicales: la lógica <i>fuzzy</i>	263
Teresa León, Vicente Liern	

SESIÓN ESPECIAL 6 Aplicaciones de Soft Computing para Educación y Aprendizaje

Organizadores: Victoria López y Luis Garmendia

Control de Trenes Basado en Lógica Borrosa: una Aplicación Didáctica.....	273
Daniel Alonso Fernández, Matilde Santos, Victoria López	
Depuradora de piscina borrosa: ejemplo para el aprendizaje de razonamiento aproximado.....	281
Raúl Martín Bonilla, Luis Garmendia	
Nuevas oportunidades para la Universidad: Educación para el Desarrollo Sostenible a través de la lógica borrosa.....	287
Raquel Caro, Adela Salvador, Alfonso Garmendia	
Un sistema de recomendaciones basado en técnicas de soft computing para el filtrado de objetos de aprendizaje.....	295
Alfredo Zapata, Mateus Ferreira-Satler, Víctor H. Menéndez, Manuel E. Prieto, José Ángel Olivas, Christian L. Vidal	
E-Learning Migration: Un recomendador de plataformas.....	303
Pablo Esteve, Juan C. de Mendoza, Victoria López	
Determinación Taxonómica Inteligente de Flora Ibérica mediante Lógica Fuzzy.....	311
Diana Díaz Agrela, Luis Garmendia, Alfonso Garmendia	
Uso de Técnicas Soft Computing en el Análisis de Consumo de Marcas de Distribuidor.....	319
Jorge Caballero, Pablo Jimeno, Raúl Ortega, Victoria López	
Análisis de garantías bajo incertidumbre con EMSI.....	327
Marcos Álvarez, Jesús de Lara, Celia Padellano, Victoria López	
Técnicas de soft computing en prospectiva tecnológica.....	335
M. T. Lamata, D. A. Pelta, J.L. Verdegay	

Generando una partición fuzzy óptima para clasificar.....	343
J.M. Cadenas, M.C. Garrido, R. Martínez	

SESIÓN ESPECIAL 7 Soft Computing y Recuperación de Información en Internet

Organizadores: Francisco P. Romero, Jesús Serrano, Carlos Porcel y Antonio Gabriel López

Extracción, Análisis y Representación de Oraciones Causales y Condicionales Imperfectas a través de un proceso Semi-Automático.....	353
Cristina Puente, Alejandro Sobrino, José Ángel Olivas, Roberto Merlo	
Un sistema de recomendaciones mejorado para evitar la sobrecarga de información persistente en una biblioteca digital universitaria.....	361
La similitud borrosa en la generación de metadatos de objetos de aprendizaje.....	369
Víctor Menéndez, Manuel Prieto	
Análisis de opiniones en la red y monitorización del eWOM apartir de una red social de microblogging (Twitter).....	377
Roberto Merlo, David Contreras, Cristina Puente	
PLINIO: Observatorio de Efectos del Cambio Climático basado en la extracción inteligente de Información en Internet.....	385
Francisco P. Romero, Mateus Ferreira-Satler, José A. Olivas, Jesús Serrano-Guerrero	
Softening SQL queries to textual fields by multilingualism.....	393
Luchiana C. Brodeala Maria J. Martin-Bautista	
Técnicas de Expansión de Consultas aplicadas a repositorios Web de Historias Clínicas Electrónicas.....	401
Israel Alonso Martínez, Francisco Pascual Romero	
Fuzzy Linguistic Modelling in Information Access Systems: A Review.....	409
E. Herrera-Viedma, A.G. López-Herrera, M.J. Cobo	
Estudio del uso del factor tiempo en las folksonomías.....	417
Pedro López-Juárez, José A. Olivas	

SESIÓN ESPECIAL 8 Métodos Difusos en Minería de Datos y Extracción de Conocimiento

Organizadores: Jesús Alcalá y Daniel Sánchez

Last results on constrained microaggregation.....	
Isaac Cano, Guillermo Navarro-Arribas, Vicenç Torra	
Comparación estadística de algoritmos de aprendizaje estocásticos usando tests extendidos a datos intervalo-valorados y borrosos.....	427
José Otero, Luciano Sánchez, Inés Couso	
Influencia de un aprendizaje basado en costes lingüísticos a partir de datos de baja calidad y no balanceados respecto al preprocesamiento de balanceado de los datos.....	435 443
Ana Palacios, Luciano Sánchez, Inés Couso	
Linguistic local change comparison of time series.....	451
Rita Castillo-Ortega, Nicolás Marín, Daniel Sánchez	
Un primer estudio sobre el uso de aprendizaje sensible al coste con sistemas de clasificación basados en reglas difusas para problemas no balanceados.....	459
Victoria López, Alberto Fernández, Francisco Herrera	
Análisis del impacto del ruido en clases y atributos para Sistemas de Clasificación Basados en Reglas Difusas.....	467
José A. Sáez, Julián Luengo, Francisco Herrera	
Are fuzzy systems as interpretable (readable and understandable) as the fuzzy community usually claims?.....	
Jose M. Alonso, Luis Magdalena	
	475

Un sistema de recomendaciones mejorado para evitar la sobrecarga de información persistente en una biblioteca digital universitaria

Carlos Porcel Álvaro Tejeda Enrique Herrera-Viedma

Dept. de Informática Dept. de Ciencias de la Computación e Inteligencia Artificial

Universidad de Jaén

Universidad de Granada

23071, Jaén

18071, Granada

cporcel@ujaen.es

atejeda@decsai.ugr.es , viedma@decsai.ugr.es

Resumen

Continuamente nos saturan con una gran cantidad de información que provoca que tengamos serios problemas para acceder a información relevante, es decir, sufrimos problemas de sobrecarga de información. Los sistemas de recomendaciones han sido aplicados con éxito en numerosos ámbitos para evitar este problema, pero el número de recursos electrónicos generados diariamente continúa creciendo y el problema aparece de nuevo. Por tanto, nos enfrentamos a un problema persistente de sobrecarga de información. En este trabajo proponemos un sistema de recomendaciones mejorado para evitar el problema de la sobrecarga de información persistente que podemos encontrar en una biblioteca digital universitaria. La idea es incorporar una memoria para recordar ítems seleccionados en alguna ocasión pero por algún motivo no recomendados, y así el sistema podría incorporarlos en futuras recomendaciones para completar el panel de recomendaciones realizadas.

1. Introducción

A diario recibimos una enorme cantidad de emails, la mayoría etiquetados como *spam*. Sin embargo también recibimos algunos con información útil. El problema es que al recibir tanta información, podríamos no prestar la atención adecuada a algo que pensemos que carece de importancia y así malinterpretamos el mensaje o incluso podríamos perder dicha infor-

mación. Nos enfrentamos, pues, a un problema bien conocido por todos, la *sobrecarga de información* [20]. Se podría definir como la imposibilidad de extraer conocimiento relevante a partir de una gran cantidad de información. Esto puede significar que tengamos más información relevante de la que podamos asimilar o que nos saturemos de información que no hemos solicitado, de la cual algo podría ser relevante [4, 20].

Este problema está presente con especial importancia en el entorno de las bibliotecas digitales, donde la información se genera mucho más rápido que los usuarios puedan procesarla. Nos centramos en un contexto académico, donde las Bibliotecas Digitales Universitarias (BDU) proporcionan recursos y servicios a estudiantes, profesores y personal en un entorno donde confluyen aprendizaje, enseñanza e investigación [6]. Un servicio particularmente importante es la difusión selectiva de información [21]. Los usuarios desarrollan sus perfiles de interés y conforme se añade nuevo material a la colección, se compara con los perfiles y la BDU envía alertas a los usuarios informándoles de los ítems relevantes [18].

Debido al uso cada vez más extendido de la Web, los usuarios de una BDU necesitan herramientas que les faciliten el acceso a los recursos. Una solución muy extendida para reducir el problema de la sobrecarga de información es el uso de *sistemas de recomendaciones* [7, 24, 27]. Los sistemas de recomendaciones evalúan y filtran la gran cantidad de información disponible y asisten a los usuarios en sus

procesos de acceso a la información [3, 7, 24]. Para conseguir la personalización en el acceso a la información, se usan las preferencias de los usuarios a la hora de definir perfiles de usuarios que son aplicados como filtros sobre el flujo de documentos. Hay distintos tipos y clasificaciones de sistemas de recomendaciones, pero desde un punto de vista teórico, podríamos considerar dos [3, 7, 24]. Los *basados en contenidos*, que recomiendan información basándose en el contenido del recurso y la experiencia pasada de los usuarios cuando han recibido ítems similares. Y los *colaborativos*, que recomiendan basándose en usuarios con perfiles similares, es decir, ignorando el contenido de los recursos. En una BDU el enfoque colaborativo es muy interesante porque permite a los usuarios compartir sus experiencias [18, 25], aunque en las fases iniciales, cuando aún no hay suficientes valoraciones, habría que trabajar con el esquema basado en contenidos.

En trabajos previos ya hemos aplicado con éxito los sistemas de recomendaciones en BDU [22, 23]. Los resultados obtenidos revelan que ambas propuestas han sido satisfactorias para reducir el problema de sobrecarga de información. Sin embargo, el número de recursos electrónicos disponibles en una BDU continúa creciendo continuamente y el problema aparece de nuevo. Por tanto, encontramos un problema persistente de sobrecarga de información.

En este trabajo proponemos un sistema de recomendaciones mejorado que se basa en la memoria para evitar ese problema persistente de sobrecarga de información presente en [22, 23]. Definimos el sistema en un contexto lingüístico difuso multi-granular [5, 10, 11, 19]. El enfoque adoptado en esta propuesta es similar al presentado en [23], pero en este caso usamos una memoria para recordar recursos seleccionados pero no recomendados y de esta forma, el sistema podría incorporarlos en futuras recomendaciones para completar el panel de recursos recomendados. Como en [22, 23] el sistema recomienda tanto recursos como posibilidades de colaboración con otros usuarios, lo que permite conocer a investigadores de áreas relacionadas con los que se podría colaborar para acceder a proyectos conjuntos.

En la sección 2 trataremos los antecedentes necesarios. En la sección 3 presentamos el nuevo sistema de recomendaciones mejorado basado en memoria. Terminaremos apuntando las conclusiones obtenidas.

2. Antecedentes

2.1. Sistemas de recomendaciones

Son sistemas que ayudan a los usuarios a acceder a ítems de interés, en aquellas situaciones en las que hay que elegir entre un gran número de posibilidades [3]. Se trata de un área de investigación que ofrece herramientas para discernir entre información relevante e irrelevante, proporcionando asistencia personalizada en los procesos de acceso a la información [24].

En la literatura se pueden encontrar numerosas técnicas para generar las recomendaciones [3, 7, 24], todas ellas con sus ventajas e inconvenientes. En este trabajo proponemos el uso de un enfoque híbrido entre las técnicas basadas en contenidos y las colaborativas, para suavizar las desventajas que cada una de ellas presenta por separado y aprovecharnos de sus beneficios [16].

En este tipo de sistemas, las preferencias de información de los usuarios son usadas para definir perfiles que actúan como filtros sobre el flujo de ítems. Por tanto, la construcción de perfiles precisos, así como mantenerlos actualizados dinámicamente, es una tarea fundamental y el éxito del sistema dependerá en gran medida de ello. Para mantener este esquema, la actividad de generación de recomendaciones es seguida por una fase de realimentación en la que los usuarios valoran la relevancia de las recomendaciones suministradas, y el sistema usa estas evaluaciones para actualizar automáticamente los perfiles de los usuarios [7, 24].

Dada su utilidad, su uso es cada vez más común en diversos campos de cara a reducir la sobrecarga de información a la que nos enfrentamos hoy en día. Así por ejemplo, se aplican en portales web de comercio electrónico para mejorar las ventas [15], en sistemas de ayuda en toma de decisiones [17], en entornos

universitarios para recomendar recursos de investigación [22, 23], o incluso como herramientas de ayuda al aprendizaje [14], etc.

2.2. Enfoque lingüístico difuso 2-tupla

Hay situaciones en las que la información no puede ser valorada precisamente de forma cuantitativa, pero podría serlo de forma cualitativa. El uso de la Teoría de Conjuntos Difusos ha dado muy buenos resultados para el modelado de información cualitativa [28] y se ha demostrado su utilidad en multitud de problemas [8, 12, 13]. Se trata de una herramienta basada en el concepto de *variable lingüística* [28].

El Modelado Lingüístico Difuso (MLD) 2-tupla [9] es un modelo continuo de representación de información. Para definirlo, tenemos que establecer el modelo de representación y el modelo computacional 2-tupla para representar y agregar respectivamente la información lingüística.

Consideremos que $\mathcal{S} = \{s_0, \dots, s_g\}$ es un conjunto de términos lingüísticos con cardinalidad impar, donde la semántica asociada con cada una de las etiquetas viene dada por medio de funciones de pertenencia triangulares y consideramos todos los términos distribuidos sobre una escala sobre la que hay establecida una relación de orden total. Si mediante un método simbólico de agregación de información lingüística, obtenemos un valor $\beta \in [0, g]$, y $\beta \notin \{0, \dots, g\}$, podemos representar β como una 2-tupla (s_i, α_i) , donde s_i representa la etiqueta lingüística, y α_i es un valor numérico que expresa la traslación de β con respecto al índice de la etiqueta más cercana, i , en el conjunto de términos lingüísticos ($s_i \in \mathcal{S}$). Este modelo define un conjunto de funciones de transformación entre valores numéricos y 2-tupla: $\Delta(\beta) = (s_i, \alpha)$ y $\Delta^{-1}(s_i, \alpha) = \beta \in [0, g]$ [9].

Para establecer el modelo computacional, definimos operadores de negación, de comparación y de agregación. Usando las funciones Δ y Δ^{-1} , podemos extender cualquiera de los operadores de agregación ya definidos para trabajar con 2-tupla [9].

En función del grado de incertidumbre que un experto encargado de cualificar un fenómeno tenga sobre el mismo, el conjunto de términos lingüísticos elegido para proporcionar ese conocimiento tendrá más o menos términos. Cuando distintos expertos tienen diferentes grados de incertidumbre sobre el fenómeno o cuando un experto tiene que valorar distintos conceptos, se necesitan varios conjuntos de términos lingüísticos de diferente granularidad [10]. En tales situaciones necesitamos herramientas que nos permitan gestionar la información lingüística multi-granular. En [10] se propuso un MLD 2-tupla multi-granular basado en el concepto de jerarquía lingüística. Una *Jerarquía Lingüística*, LH , es un conjunto de niveles $l(t, n(t))$, donde cada nivel t es un conjunto de términos lingüísticos con una granularidad $n(t)$ diferente [10]. Los niveles están ordenados por granularidad, de manera que podemos definir un nivel a partir del anterior: $l(t, n(t)) \rightarrow l(t+1, 2 \cdot n(t) - 1)$. En [10] se definió una familia de funciones de transformación entre etiquetas de diferentes niveles. Para establecer el modelo computacional seleccionamos un nivel que usamos para uniformizar la información y así podemos usar los operadores definidos en el modelo 2-tupla.

2.3. Relaciones de preferencia lingüísticas difusas incompletas

Una relación de preferencia difusa P sobre un conjunto de alternativas $X = \{x_1, \dots, x_n\}$, es un conjunto difuso sobre el conjunto $X \times X$, es decir, se caracteriza por una función de pertenencia $\mu_P: X \times X \rightarrow [0, 1]$.

Cuando la cardinalidad de X es pequeña, la relación de preferencia podría ser convenientemente representada por la matriz $P = (p_{ij})$, de $n \times n$, siendo $p_{ij} = \mu_P(x_i, x_j)$ ($\forall i, j \in \{1, \dots, n\}$) interpretado como el grado de preferencia de la alternativa x_i sobre x_j , donde $p_{ij} = 1/2$ indica indiferencia, $p_{ij} = 1$ indica que x_i es totalmente preferida a x_j , y $p_{ij} > 1/2$ indica que x_i es preferida a x_j .

Sin embargo, nuestro sistema integra el MLD multi-granular basado en 2-tupla, de manera que definimos una relación de preferencia lingüística como se indica a continua-

ción. Sea $X = \{x_1, \dots, x_n\}$ un conjunto de alternativas y S un conjunto de términos lingüísticos. Una relación de preferencia lingüística $P = p_{ij} (\forall i, j \in \{1, \dots, n\})$ sobre X se define como: $\mu_P : X \times X \rightarrow S \times [0,5, 0,5)$, donde $p_{ij} = \mu_P(x_i, x_j)$ es una 2-tupla que indica el grado de preferencia de la alternativa x_i respecto x_j .

En muchos problemas de la vida cotidiana los expertos no suelen proporcionar todos los valores requeridos. Para poder modelar estas situaciones, usamos relaciones de preferencia difusas incompletas [1, 2].

Una función $f: X \rightarrow Y$ es *parcial* cuando no todos los elementos del conjunto X necesariamente se corresponden con un elemento del conjunto Y . Cuando cada elemento del conjunto X se corresponde a un elemento del conjunto Y , tenemos una función *total*.

Una relación de preferencia lingüística difusa basada en 2-tupla P sobre un conjunto de alternativas X con una función de pertenencia parcial, se trata de una *relación de preferencia lingüística difusa incompleta*.

3. Un sistema de recomendaciones mejorado basado en memoria

En esta sección presentamos un sistema de recomendaciones mejorado basado en memoria. Nuestra propuesta consiste en añadir una segunda selección sobre los recursos seleccionados en una primera ronda. De esta forma, el sistema evita la sobrecarga persistente de información presente en una BDU. Esta segunda selección se hace teniendo en cuenta las restricciones definidas por los usuarios sobre la cantidad de recursos que desean recibir y sobre la novedad de los mismos.

Normalmente, la cantidad de recursos recomendados por el sistema es mayor que la cantidad de recursos que los usuarios desean recibir. Nuestra idea es mantener todos estos recursos considerados relevantes pero que por restricciones de los usuarios, no se les han recomendado. Podrían ser útiles después, en posteriores rondas de recomendación. Por ejemplo, cuando en un momento dado la cantidad de recursos recomendados no sea suficiente

para satisfacer las restricciones de los usuarios.

Hemos optado por un enfoque de recomendación híbrido basado en un MLD multi-granular. Para conseguir una mayor flexibilidad en los procesos de comunicación del sistema, usamos diferentes conjuntos de etiquetas (S_1, S_2, \dots) para representar los distintos conceptos que tienen que ser valorados en la actividad de filtrado. Estos conjuntos de etiquetas son seleccionados de una *LH*. El número de conjuntos de etiquetas que podemos usar, está limitado por el número de niveles de la jerarquía *LH*, y por tanto, en algunos casos, dos conjuntos de etiquetas S_i y S_j pueden estar asociados al mismo nivel de *LH*, pero con diferentes interpretaciones dependiendo del concepto que se esté modelando. En nuestro sistema, vamos a considerar cinco conceptos:

- *Grado de importancia* de una disciplina con respecto al ámbito de un recurso o las preferencias de un usuario (S_1).
- *Grado de relevancia* de un recurso para un usuario (S_2).
- *Grado de compatibilidad* entre dos usuarios (S_3).
- *Grado de preferencia* de un recurso respecto de otro (S_4).
- *Número aproximado de recursos* que un usuario desea recibir (S_5).

En concreto, usaremos etiquetas seleccionadas de una *LH* de 3 niveles (de 3, 5 y 9 etiquetas), con $S_1 = S^5$, $S_2 = S^9$, $S_3 = S^9$, $S_4 = S^5$ y $S_5 = S^3$.

En la figura 1 se representa el proceso de recomendación.

3.1. Representación de los recursos

En esta fase el sistema obtiene una representación interna de los recursos según su ámbito. Usamos el modelo vectorial y una clasificación de 25 disciplinas para identificar el ámbito de un recurso, de manera que para cada recurso i vamos a usar un vector de disciplinas, VR_i ,

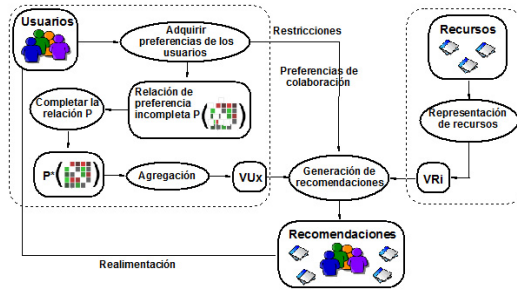


Figura 1: Esquema de recomendación.

de 25 elementos, donde en cada posición se almacena una 2-tupla $\in S_1$ que indica el grado de importancia en que el recurso cubre la disciplina correspondiente a dicha posición:

$$VR_i = (VR_{i1}, VR_{i2}, \dots, VR_{i25})$$

Estos grados de importancia son asignados por el personal de la BDU cuando se añade un nuevo recurso.

3.2. Perfiles de usuario

Los perfiles de usuario se establecen en función de restricciones sobre las recomendaciones que los usuarios quieren recibir, preferencias sobre sus temas de interés y preferencias sobre colaboraciones con otros usuarios.

Los usuarios pueden establecer dos tipos de **restricciones sobre las recomendaciones que desean recibir**:

1. *Número de recursos* que desean recibir. Pueden especificarlo de forma cuantitativa (número exacto) o de forma cualitativa (valor aproximado especificado mediante una etiqueta lingüística de S_5). Si el número de recursos considerados relevantes es mayor que el número de recursos deseados X , el sistema recomienda los X recursos más relevantes y el resto los recuerda para posteriores rondas de recomendación.
2. Restricciones acerca del *tipo de recursos* que desean recibir. Algunos usuarios sólo quieren recibir información sobre

nuevos recursos, mientras que otros podrían querer recibir información sobre recursos ya existentes considerados relevantes, pero no recomendados aún. Es más, un recurso de este tipo podría ser más interesante que uno nuevo.

Las **preferencias sobre los temas de interés** de los usuarios se adquieren siguiendo el proceso descrito en [23]. Básicamente, el sistema presenta una selección de los 4 o 5 recursos más representativos y el usuario especifica sus preferencias sobre estos recursos por medio de una relación de preferencia difusa incompleta. El usuario asigna etiquetas de S_4 , de forma que cada valor de preferencia p_{ij} representa el grado de preferencia lingüística del recurso i sobre el recurso j . Es suficiente rellenar una única fila, porque siguiendo el método propuesto en [2], completamos la relación de preferencia. Una vez completada, agregamos los vectores que representan a cada recurso ponderados según los grados de preferencia. Así obtenemos un vector VU_x de 25 posiciones que representa las preferencias sobre los temas de interés del usuario x .

Para completar el perfil, el sistema pide al usuario que exprese sus **preferencias sobre si desea colaboraciones** con otros usuarios. Sencillamente responde “Sí” o “No”.

3.3. Estrategia de recomendación basada en memoria

Fase 1. Proceso de generación de recomendaciones.

En esta fase el sistema genera las recomendaciones basándose en un proceso de cálculo de similitud entre los perfiles de los usuarios y la representación de los recursos [7]. Para calcular la similitud entre dos vectores V_1 y V_2 usamos la medida del coseno, pero definida en un contexto lingüístico de 2-tupla:

$$\sigma_l(V_1, V_2) = \Delta \left(g \times \frac{\sum_{k=1}^n (\tilde{h}_1 \times \tilde{h}_2)}{\sqrt{\sum_{k=1}^n (\tilde{h}_1)^2} \times \sqrt{\sum_{k=1}^n (\tilde{h}_2)^2}} \right) \quad (1)$$

donde g es la granularidad de S_1 , n es el número de términos usado para definir los vec-

tores, $\tilde{h}_i = \Delta^{-1}(v_{ik}, \alpha_{vik})$ y (v_{ik}, α_{vik}) es la 2-tupla de la disciplina k en el vector del recurso o del usuario (V_i). El resultado es una etiqueta de S_1 .

Cuando se inserta un nuevo recurso i se calcula su similitud, $\sigma_l(VR_i, VU_j)$, con los vectores que representan a todos los usuarios. Si $\sigma_l(VR_i, VU_j)$ es mayor que el umbral definido por el usuario, j es seleccionado para que le sea recomendado el recurso i . Para calcular la relevancia, etiqueta de S_2 , tenemos que usar la función de transformación definida en [10].

Las preferencias de colaboración se usan para clasificar a los usuarios en \mathcal{U}_C (desean colaborar) o \mathcal{U}_N (no desean colaborar). Para cada dos usuarios $x, y \in \mathcal{U}_C$ se estiman las posibilidades de colaboración calculando $\sigma_l(VU_x, VU_y)$ y transformando el resultado a una etiqueta de S_3 .

Fase 2. Segundo filtrado.

En la fase 1, el sistema ha seleccionado un número de recursos NRS_u para recomendar al usuario U . Ahora aplicamos un nuevo filtro teniendo en cuenta las restricciones impuestas por U . El número de recursos que U desea, REC_u , puede ser una cantidad exacta o una etiqueta lingüística; la idea es similar, pero en el segundo caso tenemos que usar operadores de comparación definidos en [9].

Si $NRS_u < REC_u$ significa que no se han recuperado suficientes recursos, por lo que el sistema recuerda recursos previamente seleccionados pero no recomendados. El sistema repite el proceso de la fase 1, pero incorporando estos recursos que hemos recordado.

Si la cantidad de recursos seleccionados es suficiente, $NRS_u \geq REC_u$, el sistema comprueba las restricciones de U , para ver si desea sólo nuevos recursos o también le pueden interesar recursos previos pero aún con validez y que podrían resultar más interesantes que uno nuevo. Si U desea ambos tipos de recursos, el sistema repite la fase 1, pero ahora incorporando recursos previamente seleccionados para U pero no recomendados.

Por último, el sistema envía la información correspondiente a los usuarios seleccionados junto con su grado estimado de relevancia lingüística. Para los usuarios de \mathcal{U}_C el sistema

también envía las posibilidades de colaboración junto con el grado estimado de compatibilidad.

Posteriormente, se entraría en **fase de re-almimentación**. Una vez que los usuarios hayan podido analizar los recursos y colaboraciones recomendadas por el sistema, se les pedirá que valoren la relevancia de dichas recomendaciones. El sistema usará estas valoraciones para actualizar automáticamente los perfiles de los usuarios [7, 24].

3.4. Evaluación del sistema

Actualmente tenemos implementada una versión de pruebas que trabaja con pocos usuarios, pero nos sirve para evaluar la funcionalidad del sistema y comprobar que efectivamente los recursos recomendados son de interés y se reduce el impacto de la sobrecarga de información. Cuando los usuarios reciben una recomendación la valoran, de manera que si están satisfechos asignan un valor alto.

Inicialmente hemos limitado nuestro estudio a 30 recursos de distintas áreas y 10 usuarios, que completaron el proceso de registro proporcionando sus preferencias sobre los 5 recursos más relevantes mostrados por el sistema y especificando sus restricciones sobre los recursos a recibir. Estos perfiles, junto con los recursos previamente insertados, constituyen nuestro conjunto de entrenamiento. Posteriormente añadimos 40 nuevos recursos, como conjunto de pruebas, que fueron recomendados siguiendo el nuevo enfoque propuesto. Para obtener datos para comparar, estos nuevos recursos también fueron recomendados por el personal de biblioteca. Con esta información calculamos la precisión (proporción de ítems relevantes seleccionados del total de ítems seleccionados), recall (proporción de ítems relevantes seleccionados del total de ítems relevantes) y F1 (combina con igual importancia las otras dos), ampliamente usadas en la evaluación de sistemas de recomendaciones [26]. La media para la precisión, recall y F1 es de 68.06 %, 67.13 % y 67.95 %, respectivamente, mejorando los resultados obtenidos con el sistema propuesto en [23] (67.50 %, 61.39 % y 63.51 % respectivamente).

4. Conclusiones

Las BDU proporcionan canales muy efectivos para la difusión de conocimiento y su expansión en la sociedad, pero continuamente nos saturan con una gran cantidad de información, de manera que nos enfrentamos a serios problemas de sobrecarga de información. Los sistemas de recomendaciones se usan con éxito para asistir a los usuarios en sus procesos de acceso a la información, pero el número de recursos electrónicos generados diariamente continúa creciendo y el problema de sobrecarga aparece de nuevo. Por tanto, nos enfrentamos a un problema persistente de sobrecarga de información. En este trabajo hemos propuesto un sistema de recomendaciones mejorado, basado en memoria, para evitar la sobrecarga encontrada en sistemas previos. La idea es recordar recursos previamente seleccionados pero no recomendados, para hacer una nueva selección. Esta selección se hace teniendo en cuenta las restricciones definidas por los usuarios sobre la cantidad de recursos que desean recibir y sobre la novedad de los mismos. Como principal trabajo futuro nos planteamos automatizar la representación de los recursos, de forma que el personal de la BDU no tenga que introducir manualmente los grados para cada una de las disciplinas.

Agradecimientos. Este trabajo ha sido desarrollado con la financiación de los proyectos FUZZYLING (TIN2007-61079) y PETRI (PET2007-0460).

Referencias

- [1] Alonso S., Cabrerizo F.J., Chiclana F., Herrera F., Herrera-Viedma E., *Group decision-making with incomplete fuzzy linguistic preference relations*. International Journal of Intelligent Systems, 24(2), 201-222 (2009).
- [2] Alonso S., Chiclana F., Herrera F., Herrera-Viedma E., Alcalá-Fdez J., Porcel C., *A Consistency-Based Procedure to Estimating Missing Pairwise Preference Values*. International Journal of Intelligent Systems, 23, 155-175 (2008).
- [3] Burke R., *Hybrid Web Recommender Systems*. P. Brusilovsky, A. Kobsa, and W. Nejdl (Eds.): The Adaptive Web, LNCS 4321, 377-408 (2007).
- [4] Butcher, H., *Meeting managers' information needs*. London: Aslib (1998).
- [5] Chang S.L., Wang R.C., Wang S.Y., *Applying a direct multigranularity linguistic and strategy-oriented aggregation approach on the assessment of supply performance*. European Journal of Operational Research, 177(2), 1013-1025 (2007).
- [6] Chao H., *Assessing the quality of academic libraries on the Web: The development and testing of criteria*. Library & Information Science Research, 24, 169-194 (2002).
- [7] Hanani U., Shapira B., Shoval P., *Information Filtering: Overview of Issues, Research and Systems*. User Modeling and User-Adapted Interaction, 11, 203-259 (2001).
- [8] Herrera F., Herrera-Viedma E., Verdegay J.L., *Direct approach processes in group decision making using linguistic OWA operators*. Fuzzy Sets and Systems, 79, 175-190, 1996.
- [9] Herrera F., Martínez L., *A 2-tuple fuzzy linguistic representation model for computing with words*. IEEE Transactions on Fuzzy Systems, 8(6), 746-752, 2000.
- [10] Herrera F., Martínez L., *A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision-making*. IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics, 31(2), 227-234 (2001).
- [11] Herrera-Viedma E., Cordón O., Luque M., López A.G., Muñoz A.M., *A Model of Fuzzy Linguistic IRS Based on Multi-Granular Linguistic Information*. International Journal of Approximate Reasoning, 34(3), 221-239 (2003).

- [12] Herrera-Viedma E., López-Herrera A. G., Luque M., Porcel C., *A Fuzzy Linguistic IRS Model Based on a 2-Tuple Fuzzy Linguistic Approach*. International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, 15(2), 225-250, 2007.
- [13] Herrera-Viedma E., Pasi G., López-Herrera A. G., Porcel C., *Evaluating the information quality of web sites: A methodology based on fuzzy computing with words*. Journal of the American Society for Information Science and Technology, 57(4), 538-549, 2006.
- [14] Hsu M.H., *A personalized English learning recommender system for ESL students*. Expert Systems with Applications, 34, 683-688, 2008.
- [15] Kim K.J., Ahn H., *A recommender system using GA K-means clustering in an online shopping market*. Expert systems with applications, 34, pág. 1200-1209, 2008.
- [16] Lekakos G., Caravelas P., *A hybrid approach for movie recommendation*. Multimedia tools and applications, 36, pág. 55-70, 2008.
- [17] Liang T.P., *Recommendation systems for decision support: An editorial introduction*. Decision Support Systems, 45, 385-386, 2008.
- [18] Marchionini G., *Research and Development in Digital Libraries*. http://ils.unc.edu/~march/digital_library_R_and_D.html. Last access: April, 2010.
- [19] Mata F., Martínez L., Herrera-Viedma E., *An Adaptive Consensus Support Model for Group Decision Making Problems in a Multi-Granular Fuzzy Linguistic Context*. IEEE Transactions on Fuzzy Systems, 17(2), 279-290 (2009).
- [20] Meghabghab G., Kandel A., *Search Engines, Link Analysis, and Users' Web Behavior*. Springer-Verlag Berlin Heidelberg (2008).
- [21] Morales del Castillo J.M., Pedraza-Jiménez R., Ruíz A.A., Peis E., Herrera-Viedma E., *A Semantic Model of Selective Dissemination of Information for Digital Libraries*. Information Technology and Libraries, 28(1), 22-31 (2009).
- [22] Porcel C., Moreno J.M., Herrera-Viedma E., *A multi-disciplinar recommender system to advice research resources in University Digital Libraries*. Expert Systems with Applications, 36(10), 12520-12528 (2009).
- [23] Porcel C., Herrera-Viedma E., *Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries*. Knowledge-Based Systems, 23(1), 32-39, (2010).
- [24] Reisman P., Varian H.R., *Recommender Systems*. Special issue of Communications of the ACM, 40(3), 56-59 (1997).
- [25] Ross L., Sennyey P., *The Library is Dead, Long Live the Library! The Practice of Academic Librarianship and the Digital Revolution*. The Journal of Academic Librarianship, 34(2), 145-152 (2008).
- [26] Sarwar B., Karypis G., Konstan J., Riedl J., *Analysis of recommendation algorithms for e-commerce*. Proceedings of ACM E-Commerce 2000 conference, 158-167 (2000).
- [27] Symeonidis P., Nanopoulos A., Papadopoulos A.N., Manolopoulos Y., *Collaborative recommender systems: Combining effectiveness and efficiency*. Expert Systems with Applications, 34, 2995-3013 (2008).
- [28] Zadeh L.A., *The concept of a linguistic variable and its applications to approximate reasoning*. Part I. Information Sciences, 8, 199-249, 1975. Part II, Information Sciences, 8, 301-357, 1975. Part III, Information Sciences, 9, 43-80, 1975.

- Monserrat, M., 221
Montero, Javier, 183
Montilla, Wilmer, 189
Moreno-García, Juan, 41, 99, 133
Mucientes Manuel, 1
Muñoz, E., 231
Muñoz, Rafael, 19
Navarro-Arribas, Guillermo, 427
Nunes, E., 173
Olivas, José Ángel, 295, 353, 385, 417
Ortega, Raúl, 319
Otero, José, 435
Pablo Jimeno, 319
Padellano, Celia, 327
Pagola, M., 131, 165
Palacios, Ana, 443
Pascual Romero, Francisco, 401
Paternain, D., 165
Peláez, J. I., 35
Pelta, D. A., 335
Peralta, Arturo 99
Pérez Asurmendi, Patrizia, 115
Pérez, I. J., 51
Pikatzka, J. M., 107
Porcel, Carlos, 351, 361
Prieto, Manuel E., 295, 369
De Prisco, R., 231
Puente, Cristina, 353, 377
Ramos Gómez, Ángel, 123
Rdez. Lera, Francisco J., 11
Rodríguez, R.M., 91
Rodríguez-Benítez, L., 41, 133
Romero, Francisco P., 99, 351, 385
Ruan, D., 91
Ruiz-Lozano, M.D., 27
Sáez, José A., 467
Salvador, Adela, 287
Sánchez, A., 173
Sánchez, Daniel, 425, 451
Sánchez, Luciano, 435, 443
Sanjurjo, P., 107
Santos, Matilde, 273
Segundo, U., 107
Serrano-Guerrero, Jesús, 351, 385
Sobrino, Alejandro, 353
Solana-Cipres, C., 41, 133
Tejeda, Álvaro, 361
Tenorio, E., 35
Tinguaro Rodríguez, J., 183
Torra, Vicenç, 427
Torrens, Joan, 141, 221, 213
Vallejo, David, 3
Verdegay, J. L., 335
Vidal, Christian L., 295
Vitoriano, Begoña, 183
Yeguas, Enrique, 19
Zaccagnino, R., 231
Zapata, Alfredo, 295
Zuccarello, Pedro, 247