

## Modelo de recomendación para la orientación vocacional basado en la computación con palabras

### [ Recommendation models for vocational orientation based on computing with words ]

*Milton Rafael Maridueña Arroyave<sup>1</sup>, Ailyn Febles Estrada<sup>2</sup>, and Roxana Cañizares González<sup>2</sup>*

<sup>1</sup>Universidad de Guayaquil, Guayaquil, Guayas, Ecuador

<sup>2</sup>Universidad de las Ciencias Informáticas, La Habana, Cuba

---

Copyright © 2016 ISSR Journals. This is an open access article distributed under the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abstract:** The selection of a profession suitable to expectations implies taking into account multiples factors. Despite its usefulness and high impact there are limitations in the support to vocational orientation and specifically to university major recommendation models. Among these limitations are the lack of flexible models and the dependence on historical information. In this paper a new college degree recommendation model based on the 2-tuples linguistic model and the OWA operators is presented. It includes database construction, student profiling, university degree information filtering and recommendation generation. Its implementation make possible to improve reliability. An illustrative example is shown to demonstrate the model applicability especially to information and health sciences.

**KEYWORDS:** vocational orientation, recommender systems, university major recommendation, fuzzy logic, 2-tuples linguistic model, content based recommender systems.

**RESUMEN:** La elección de una profesión que se adecúe a las expectativas de una persona implica clarificar sus aptitudes, intereses y cualidades de cada sujeto; mientras que desde la Universidad, se requiere de un bachiller idóneo respecto a las exigencias del currículo universitario. A pesar de su extendida utilidad e impacto persisten insuficiencias en el soporte informático a la orientación vocacional y en específico al tratamiento del proceso de recomendación de las carreras universitarias a cursar. Entre ellas se destacan la falta de modelos flexibles no dependientes de datos históricos y el tratamiento de la información lingüística. En el presente trabajo se propone el modelo para el proceso de recomendación de carreras universitarias basado en el modelo de representación lingüística de las 2-tuplas y operadores OWA. Incluye el proceso de obtención de la base de datos, el perfilado del estudiante, el filtrado de las carreras, la generación de recomendaciones. Su implementación posibilita mejorar la fiabilidad y la interpretabilidad de las recomendaciones de las carreras universitarias. Se desarrolla un ejemplo ilustrativo con el propósito de demostrar la aplicabilidad del modelo.

**PALABRAS-CLAVE:** orientación vocacional, sistemas de recomendación, recomendación de carreras universitarias, lógica difusa, 2-tuplas lingüísticas, sistemas de recomendación basados en contenido.

## 1 INTRODUCCIÓN

La orientación profesional, que puede coincidir con la vocacional, pretende una mayor adaptación al tipo de trabajo a realizar de acuerdo con sus aptitudes; desde la óptica de la empresa u organismo competente selecciona a la persona que mejores aptitudes posea para la ocupación a que se destina y ofrezca posibilidades de mayor rendimiento. La orientación profesional nació en Estados Unidos a comienzos del siglo XX, y se definió como una ayuda en la transición de la escuela al

trabajo. Surge como una necesidad de proporcionar a los alumnos, la información necesaria para su adecuada inserción en el mundo laboral[1]. Se entiende por orientación como el proceso de asistencia continua que se les da a las personas en todos sus aspectos, con el objetivo de potenciar la prevención y el desarrollo individual y social a lo largo de la vida[2] y tiene entre sus propósitos fundamentales la orientación de qué carrera universitaria recomendar a un determinado estudiante.

El proceso de orientación vocacional/profesional busca lograr que los estudiantes se conozcan a sí mismos, exploren su autoconocimiento, es decir: sus aficiones, aptitudes, personalidad y destrezas y se relacionen con las diversas áreas del saber, para que, llegado el momento de elegir, lo hagan con conocimiento de causa y sabiduría[2]. La literatura sobre la orientación profesional, expresa la amplitud de acepciones de este término, que han existido a lo largo de los años como se puede apreciar en[3], [4], [5],[6] ,[7], [8]. Sin embargo, existe un criterio generalizado entre ellos que avocan un planteamiento común sobre los diversos factores que inciden en la orientación profesional. Junto a los intereses existen otros aspectos que pueden influir en la elección adecuada, tales como: el nivel de aptitud intelectual, el aprovechamiento académico en su etapa de bachillerato, los rasgos de personalidad y aspectos caracterológicos, las condiciones familiares, socioeconómicas, socioculturales del estudiante y las características propias de la oferta académica de las universidades[9].

Los diferentes estudios señalan como causa importante de la deserción universitaria a la carencia de una oportuna orientación académica y personal, derivándose en las dificultades que deben afrontar los estudiantes al no cumplir los requerimientos académicos en las universidades. Galán [10] plantea que “El proceso para conocer las inclinaciones universitarias debe ofrecerse desde secundaria, con la intención de que el estudiante tenga el tiempo suficiente para explorarse, conocerse a sí mismo, informarse del abanico de carreras que existen y llevar a cabo una buena toma de decisión para su vida profesional”.

A través de la inteligencia artificial se ha tratado de imitar la capacidad del hombre para manejar conocimientos del mundo real y poder aprender de su accionar cotidiano. Una de estas aplicaciones se relaciona con la ayuda a la toma de decisiones. La toma de decisiones hace referencia al conjunto de operaciones que comprenden desde el momento en que se detecta una situación que requiere el inicio de este proceso hasta que la decisión es adoptada y ejecutada [11].

El proceso de orientación vocacional se enmarca en los modelos de ayuda a la toma de decisiones. Sin embargo las propuestas basadas en métodos de decisión multicriterio [12-14] resultan demandantes cognitivamente para el decisor manejando un número limitado de alternativas resultando más adecuado los modelos de recomendación.

Uno de los ambientes de decisión más frecuente y al cual se ajusta el proceso de recomendación de carreras, está relacionado con la toma de decisiones bajo incertidumbre [15-17]. En este contexto es necesario considerar la utilización de técnicas de softcomputing [18] por su capacidad de lidiar con la imprecisión desde el punto de vista computacional.

Los sistemas de recomendación son útiles en el proceso de toma de decisiones ya que proporcionan al usuario un conjunto de opciones que se espera satisfagan sus expectativas [19]. En el caso de la recomendación de carreras universitarias dado el enfoque actual empleado en este proceso y las condiciones de falta de información histórica se considera que este se ajusta más adecuadamente a los modelos de recomendación basado en contenido [20].

A pesar de su utilidad e impacto, persisten insuficiencias en el tratamiento del proceso de recomendación en general y de la recomendación de carreras universitarias a cursar en particular, incidiendo negativamente en la fiabilidad de los resultados obtenidos. Entre ellas se destacan:

- Los enfoques existentes basados en sistemas de recomendación se apoyan fundamentalmente en el filtrado colaborativo [21-23] o en la minería de datos, como reglas de asociación y árboles de decisión [24-26]. Sin embargo muchas veces no se cuenta con información histórica lo que imposibilita la adopción de este enfoque. Otro elemento es que por lo general se centran en la recomendación de cursos específicos y no de carreras universitarias.
- Existe falta de tratamiento adecuado de la vaguedad y la inclusión de información lingüística.

El artículo continúa de la siguiente forma: en la siguiente sección se discute lo relativo a la orientación vocacional y en la 3 Los modelos de toma de decisiones en el campo de la orientación profesional. En la sección 4 se discute el enfoque lingüístico difuso, a continuación se analizan distintos modelos de recomendación y en especial sistemas de recomendación aplicados a la orientación vocacional profesional. Se presenta el modelo propuesto y un ejemplo demostrativo en secciones 5 y 6 respectivamente. El trabajo finaliza con las conclusiones y recomendaciones de trabajo futuro.

## 2 ORIENTACIÓN VOCACIONAL

La vasta literatura pedagógica sobre la orientación, pone de manifiesto la amplitud de acepciones que existen con la orientación vocacional. A partir de un repaso histórico se destacan las aportaciones de distintos autores que se exponen a

continuación. Se considera que la orientación capacita a cada individuo para comprender sus aptitudes, intereses y rasgos personales; para desarrollarlos lo mejor posible, para relacionarlos con metas vitales y, finalmente, para alcanzar el estado de autoorientación que el ciudadano de un orden social democrático, puede desear [3]. Es un proceso de ayuda a los individuos a lograr la autocomprensión y autodirección necesarias para conseguir el máximo ajuste con el centro docente, con la familia y con la comunidad [4] .

La orientación vocacional, practicada antes de una elección profesional, tiene como finalidad el ofrecer un cuadro amplio aunque delimitado de las profesiones más acordes con los intereses y aptitudes de la persona [5]. Para los autores la orientación profesional, que puede coincidir con la vocacional, pretende una mayor adaptación al tipo de trabajo a realizar de acuerdo con sus aptitudes; desde la óptica de la empresa u organismo competente selecciona a la persona que mejores aptitudes posea para la ocupación a que se destina y ofrezca posibilidades de mayor rendimiento.

La perspectiva constructivista se ha incorporado a la orientación profesional a finales del siglo XX. El objetivo del asesoramiento constructivista es que el individuo organice y dé sentido a su experiencia a lo largo del tiempo, mediante la reflexión sobre las implicaciones del autoconocimiento anterior y actual [27].

En la actualidad y a manera de síntesis, los autores conciben la orientación como un proceso continuo de permanente asistencia al estudiante, realizado por personas técnicas y científicamente preparadas, con la participación de los miembros del centro educativo, familiares y del propio orientado; el objetivo es ayudarlo a desarrollar su condición de estudiante y las demás dimensiones personales, con la mirada puesta en el ejercicio de una actividad profesional, acorde con sus intereses y aptitudes. De esta forma se sentirá útil, satisfecho y bien integrado dentro de la sociedad.

Son diversos los factores que inciden en la orientación vocacional y profesional de los sujetos. Junto a los intereses profesionales existen otros aspectos que pueden influir en la elección adecuada. Las aptitudes pueden tener un papel relevante digno de consideración. Ya que un sujeto puede tener un gran interés en estudiar una carrera y no disponer de las aptitudes suficientes para ser capaz; también la personalidad puede influir pues hay determinadas características personales que están implícitas en ciertas profesiones.

La psicología con sus distintos instrumentos ha podido comprobar en sus investigaciones que la inteligencia es un conjunto de estrategias aprendidas o hábitos adquiridos que nos enseñan a resolver problemas. Muchos autores la han relacionado con la capacidad de adaptación y de respuesta a situaciones o problemas nuevos; otras investigaciones la relacionan con el pensamiento abstracto. También se habla de ella como la capacidad de adquirir y asimilar conocimientos.

Con respecto al tema de la personalidad, ésta se considera una variable crucial en el proceso educativo, tanto en el proceso enseñanza-aprendizaje como en la orientación académica y profesional. No puede, por tanto, descartarse su estudio, ya que la finalidad de todo proceso educativo es la maduración y realización personal [28].

Teorías en relación a la personalidad hay múltiples (Teoría del yo, Teoría del rasgo, Teoría psicodinámica, etc.) lo mismo que son muchos los instrumentos que sirven para medirla. Según [29] el protagonismo de los estudios sobre la inteligencia dejaban en un segundo plano los estudios sobre personalidad, en relación al rendimiento académico.

Los intereses también van a ser un condicionante, junto con la personalidad y las aptitudes en la orientación profesional. Los intereses tienen gran relación con las actitudes y los valores. Se relacionan con las aficiones, tendencias, gustos del sujeto y su impacto en su decisión profesional [30]. Para el estudio de los intereses tienen un papel importante los inventarios (destacando entre los primeros inventarios más conocidos los de Strong, Holland, Kuder). Los problemas de la orientación y de la adaptación profesional constituyen las fuentes de inspiración más significativas de las investigaciones sobre los intereses.

Cuando el objetivo es el descubrimiento de la vocación profesional del sujeto, es importante conocer cuáles son sus intereses pues en función de éstos se determinará en parte su futuro profesional. Es importante que el orientador sea capaz de adecuar y combinar los intereses con sus aptitudes. Puede ocurrir que el sujeto muestre preferencias por una profesión concreta y no tenga las aptitudes requeridas; también puede ocurrir que su personalidad no sea la adecuada para ese desempeño y convenga orientarle hacia otras profesiones más próximas a desempeñar con mayor éxito.

Toda esta información del estudiante, intereses, tipo de personalidad inteligencia emocional etc., constituye su perfil psicológico de estudiante a cual se puede manejar de modo subjetivo lo cual recomienda el empleo de modelos que manejen la incertidumbre y en especial los basados en lógica difusa [31].

### 3 MODELOS INFORMATICOS DE AYUDA A LA ORIENTACION VOCACIONAL

Por lo que se puede añadir a la orientación, como un proceso psicológico intencionado que implican decisiones a partir de una auto caracterización del perfil de un estudiante contra los requisitos mínimos de una carrera universitaria, proceso que incorpora como entrada las competencias que se tiene y a la vez proyecta su futuro desempeño profesional, en él pueden intervenir múltiples factores, criterios y personas, estos elementos lo determinan como un proceso de Toma de Decisión (TD) [32].

La TD describe las actividades a realizar como parte de la resolución de problemas, abarcando las etapas de definición del problema, identificación de alternativas y criterios, evaluación de dichas alternativas y la decisión final [33]. La orientación vocacional es un proceso de TD en el que intervienen en varios momentos, la intención y múltiples criterios de seres humanos, lo que hace que estos procesos ocurran bajo incertidumbre. La informática es una ciencia que trata de imitar al hombre en sus tareas mentales [34].

Según [35] en su obra "Técnicas de asesoramiento en orientación profesional" destaca una serie de sistemas informáticos que asisten a este proceso, aunque en su mayoría continúan en fase experimental: The Information System For Vocational Decisions (IVSD), Computer- Assisted Career Exploration System (CACE), The Counseling Information System (CIS), Program for Learning According to Needs (PLAN), The Experimental Education and Career Exploration System (ECESY), entre otros.

La consulta bibliográfica permiten considerar que la razón común por la que estos software no han tenido una masiva acogida en las instituciones, además de su complejidad, se debe fundamentalmente a los costosos contratos de licenciamiento, mantenimiento y soporte técnico.

#### 3.1 MODELOS DE TOMA DE DECISIONES (TD) EN EL CAMPO DE LA ORIENTACIÓN PROFESIONAL

A mediados de 1950 nació en los Estados Unidos una nueva área de investigación dentro de las ciencias computacionales llamada Inteligencia Artificial (IA). Su objetivo era conseguir que los computadores sean capaces de realizar tareas complejas independientemente de la supervisión del hombre, dotándoles para ello de un cierto grado de inteligencia para manejar conocimientos del mundo real, tomar decisiones y poder aprender de sus ejecuciones [36].

La psicología es un área de estudio que se relaciona estrechamente con el comportamiento del cerebro humano, en cuanto a su capacidad de toma de decisiones; es por esto que la implementación de un modelo que utilice técnicas de aprendizaje computacional apoyaría en gran medida a la orientación que debe desempeñar el psicólogo cuando empieza el proceso vocacional con los estudiantes, dada la similitud para representar el conocimiento y aprender de casos registrados.

Los autores consideran que un estudio integral del estudiante implicará la valoración de múltiples datos en relación a su caracterización aptitudinal y actitudinal. Y la generalización de perfiles estará dado por los patrones que se forman predictiva surjan del análisis masivo de una población objetivo.

La TD es un proceso habitual para los seres humanos en muchas actividades del mundo real [37]. Un esquema de resolución de un problema de toma de decisión generalmente consta de dos fases: (1) la fase de agregación donde se obtienen las valoraciones colectivas de cada alternativa y (2) la fase de explotación para obtener el conjunto solución de alternativas al problema [38].

La TD describe las actividades (fig 1) a realizar como parte de la resolución de problemas, abarcando las etapas de definición del problema, identificación de alternativas y criterios, evaluación de dichas alternativas y la decisión final [33]. Entonces, la elección de una futura carrera universitaria "idónea" es un proceso de TD en el que intervienen en varios momentos la demanda social y el criterio de cualificación de personas, lo que hace que estos procesos ocurran bajo incertidumbre.



Fig. 1. Actividades para la solución de un problema de toma de decisiones [39]

Uno de los ambientes de decisión más frecuente está relacionado con la toma de decisiones bajo incertidumbre [15]. En este contexto es necesario considerar la utilización de técnicas de softcomputing [18].

Los sistemas de recomendación son útiles en el proceso de toma de decisiones ya que proporcionan al usuario un conjunto de opciones que se espera satisfagan sus expectativas [19]. Partiendo de la información que recojan estos sistemas y de los algoritmos utilizados para generar las recomendaciones podemos distinguir las siguientes técnicas [40, 41]:

- Modelos de recomendación colaborativa: Agregan las valoraciones o recomendaciones de los objetos, identifican los gustos comunes de los usuarios basándose en sus valoraciones y generan una nueva recomendación teniendo en cuenta las comparaciones entre usuarios.
- Modelos de recomendación basada en contenido: Aprende un perfil de intereses de los usuarios basándose en las características presentes en los objetos que el usuario ha seleccionado.
- Modelos de recomendación basada en conocimiento: Intentan sugerir objetos haciendo inferencias sobre las necesidades de un usuario y sus preferencias.

Resulta importante para mejorar la fiabilidad de las recomendaciones conformar un perfil del estudiante basado en sus características psicológicas [42]. Este modelado permite desarrollar recomendaciones basadas en contenido teniendo en cuenta la similitud de características compartidas entre el objeto que se desea recomendar y el perfil del estudiante [43].

Los operadores de agregación son un tipo función matemática empleada para la fusión de la información. Combinan  $n$  valores en un dominio  $D$  y devuelven un valor en ese mismo dominio [44].

Denominando esas funciones  $\mathbb{C}$ [44], los operadores de agregación son funciones de forma:

$$\mathbb{C}: N^n \rightarrow N \quad (1)$$

Los operadores de agregación presentan múltiples aplicaciones en diversos dominios[45]. En la toma de decisiones su papel fundamental está en la evaluación y en la construcción de alternativas[44].

Cada una de las familias de operadores presenta características que les permiten modelar determinadas situaciones. La ponderada: la media ponderada (WA por sus siglas en inglés) posibilita asignar peso a las fuentes de información lo que permite su empleo para representar fiabilidad o importancia/preferencia. Por su parte el operador OWA (ordered weighted averaging o traducido al español media ponderada ordenada)[46] posibilita la compensación o dar peso a los datos en dependencia de sus valores. Las integrales difusas permiten modelar redundancia, complementariedad e interacciones entre criterios. Sin embargo estas últimas presentan la limitante de que sus coeficientes crecen exponencialmente con el número de criterios a ser agregados[47, 48], lo que dificulta su aplicabilidad.

El cálculo de la similitud en los sistemas de recomendación presenta falta de flexibilidad, debido a que generalmente se emplea la media ponderada de la similitud de rasgos para el cálculo de la similitud global. Sin embargo este operador no tiene en cuenta la gama (compensación, complementariedad, bipolaridad etc.) de interacciones entre rasgos que ocurren en

nuestro problema lo que limita su aplicabilidad a problemas reales[49-51] y puede ser resuelto con el uso de operadores de agregación adecuados.

A pesar de su extendida utilidad e impacto, persisten otras insuficiencias en el tratamiento del proceso de recomendación de las carreras universitarias a cursar. Entre ellas se destacan:

- Los enfoques existentes basados en sistemas de recomendación se basan fundamentalmente en el filtrado colaborativo[21-23]o en la minería de datos, como reglas de asociación y árboles de decisión [24-26] , sin embargo muchas veces no se cuenta con información histórica lo que imposibilita la adopción de este enfoque adicionalmente se centran en la recomendación de cursos específicos y no de carreras universitarias.
- Falta de tratamiento adecuado de los datos heterogéneos, distintos dominios de expresión , nivel de granularidad y origen[52, 53].

Estas debilidades descritas, hacen que en la actualidad el proceso de orientación profesional no esté sistematizado ni permita el uso de la información que se tiene del estudiante.

En el caso específico de los sistemas de recomendación las propuestas existentes se apoyan fundamentalmente en el filtrado colaborativo [21-23] o en la minería de datos, como reglas de asociación y árboles de decisión [24-26]. Sin embargo muchas veces no se cuenta con información histórica lo que imposibilita la adopción de este enfoque. Dentro de estos sistemas se destaca el sistema Degree Compass de la Austin Peay State University [54].

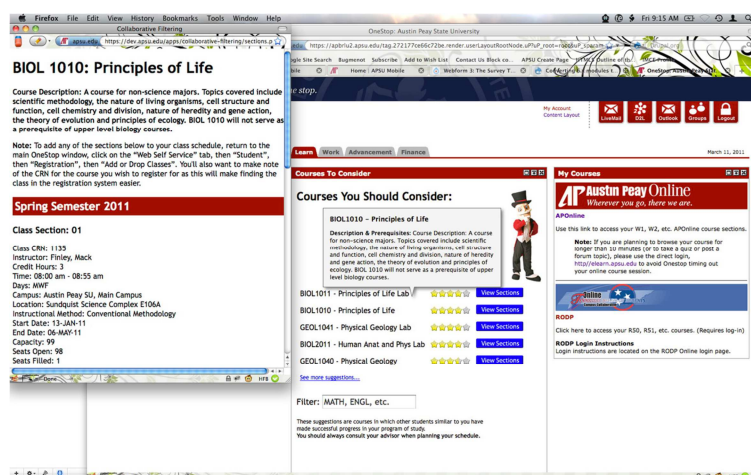


Fig. 2. Sistema de recomendación Degree Compass [54]

Sin embargo, este sistema comparte un limitante común con el resto de los sistemas estudiados y es que se centra en la recomendación de cursos específicos y no de carreras universitarias en su totalidad.

#### 4 ENFOQUE LINGÜÍSTICO DIFUSO

La teoría de los conjuntos difusos o borrosos fue introducida por Zadeh [55] en el año 1965. Esta parte de la teoría clásica de conjuntos, añadiendo una función de pertenencia o inclusión [56].

Una función de pertenencia  $\mu_a(t)$  indica el grado  $n$  en que la variable  $t$  está contenida en el concepto representado por la etiqueta  $A$  [57]. Para la definición de estas funciones de pertenencia se utilizan convenientemente ciertas familias de formas estándar, por coincidir con el significado lingüístico de las etiquetas más utilizadas. Siendo más frecuentes son triangular, trapezoidal y gaussiana (Figura 2).

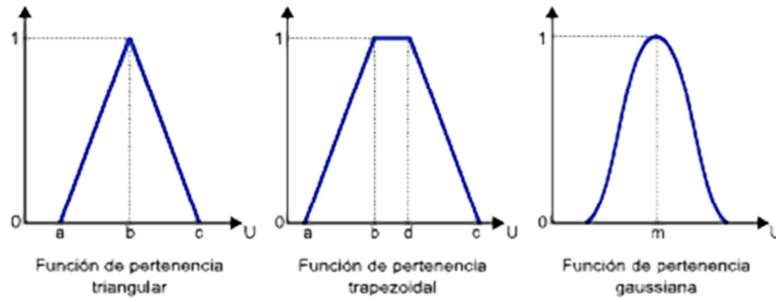


Fig. 3. Representación gráfica de las funciones de pertenencia triangular, trapezoidal y gaussiana [58]

Un elemento importante en la lógica difusa lo constituye el concepto de variable lingüísticas [59, 60]. El empleo de estas variables lingüísticas implica la necesidad de realizar procesos de computación con palabras [58].

La computación con palabras es una metodología que permite realizar un proceso de computación y razonamiento utilizando palabras pertenecientes a un lenguaje en lugar de números. Esta metodología permite crear y enriquecer modelos de decisión en los cuales la información vaga e imprecisa [61] es representada a través de variables lingüísticas.

Dentro de las propuestas existentes el modelo de representación lingüística de 2-tuplas permite realizar procesos de computación con palabras sin pérdida de información, basándose en el concepto de traslación simbólica. Este modelo de representación ha sido aplicado con éxito a los sistemas de recomendación [41, 62]. Debido a sus características y a su facilidad de uso, el autor de la presente investigación considera este modelo como el más adecuado para el manejo de información lingüística en sistemas de recomendación. .

Sea  $S = \{s_0, s_1, \dots, s_g\}$  un conjunto de términos lingüísticos y  $\beta \in [0, g]$  un valor en el intervalo de granularidad de  $S$ . La traslación simbólica de un término lingüístico,  $s_i$ , es un número valorado en el intervalo  $[-.5, .5)$  que expresa la diferencia de información entre una cantidad de información expresada por el valor  $\beta \in [0, g]$ , obtenido en una operación simbólica y el valor entero más próximo,  $i \in \{0, \dots, g\}$  que indica el índice de la etiqueta lingüística ( $s_i$ ) más cercana en  $S$  [63].

A partir de este concepto se desarrolló un modelo de representación de la información lingüística el cual hace uso de un par de valores o 2-tuplas. Este modelo de representación define un conjunto de funciones que facilitan las operaciones sobre 2-tuplas. Sea  $S = \{s_0, s_1, \dots, s_g\}$  un conjunto de términos lingüísticos y  $\beta \in [0, g]$  un valor que representa el resultado de una operación simbólica, entonces la 2-tupla lingüística que expresa la información equivalente a  $\beta$ , se obtiene usando la siguiente función:

$$\Delta: [0, g] \rightarrow S \times [-.5, .5)$$

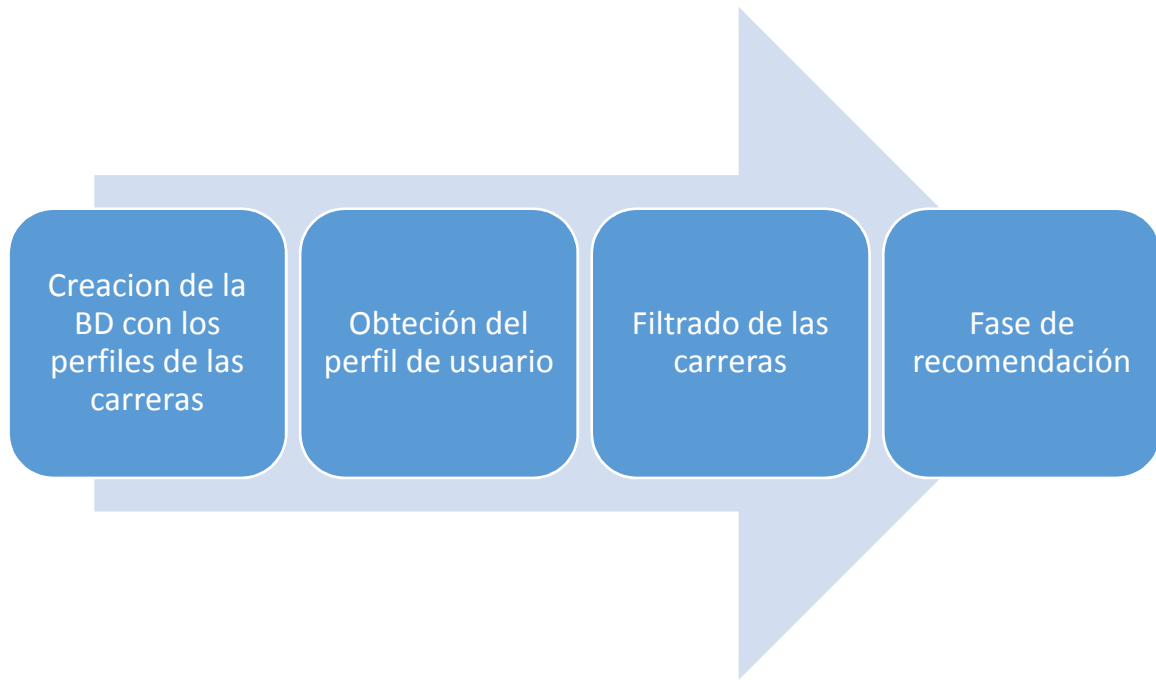
$$\Delta(\beta) = (s_i, \alpha), \text{ con } \begin{cases} s_i, & i = \text{round}(\beta) \\ \alpha = \beta - i, & \alpha \in [-.5, .5) \end{cases} \quad (1)$$

Donde round es el operador usual de redondeo,  $s_i$ , es la etiqueta con índice más cercano a  $\beta$  y  $\alpha$  es el valor de la traslación simbólica [63].

Cabe señalar que  $\Delta^{-1}: \langle S \rangle \rightarrow [0, g]$  es definida como  $\Delta^{-1}(s_i, \alpha) = i + \alpha$ . De este modo, una 2-tupla lingüística  $\langle S \rangle$  queda identificada con su valor numérico en  $[0, g]$ .

**5 MODELO PROPUESTO**

A continuación se presenta el flujo de trabajo propuesto (Figura 4). Se basa fundamentalmente en la propuesta de Cordón [41] para sistemas de recomendación basados en conocimiento adaptado a las características del dominio de aplicación y permitiendo flexibilidad en cuanto la agregación de similitud los atributos del perfil del usuario con respecto a la descripción de la carrera.



**Fig. 4. Flujo de trabajo del modelo**

La descripción detallada de cada una de sus actividades y del modelo matemático que soporta la propuesta se presenta a continuación.

### 5.1 CREACIÓN DE LA BASE DE DATOS CON LOS PERFILES DE LAS CARRERAS

Un elemento fundamental para un modelo de sistema de recomendación, es la generación de la base de datos. Cada una de las carreras  $a_i$  estará descrita por un conjunto de características que conformarán el perfil de la carrera.

$$C = \{c_1, \dots, c_k, \dots, c_l\} \quad (2)$$

En este caso se obtendrá mediante un experto o conjunto de expertos. En nuestra propuesta cada carrera universitaria será descrita por un vector de valoraciones:

$$F_{a_j} = \{v_1^j, \dots, v_k^j, \dots, v_l^j\}, j = 1, \dots, n \quad (3)$$

Las valoraciones de las características de la carrera,  $a_j$ , estarán expresadas utilizando la escala lingüística  $S$ ,  $v_k^j \in S$  donde  $S = \{s_1, \dots, s_g\}$  es el conjunto de términos lingüísticos definidos para evaluar la característica  $c_k$ , de acuerdo al grado de conocimiento de los expertos o grupos de expertos que estén realizando la evaluación.

Una vez descritas las carreras en este conjunto

$$A = \{a_1, \dots, a_j, \dots, a_n\} \quad (4)$$

serán almacenadas por el sistema en una base de datos.

### 5.2 OBTENCIÓN DEL PERFIL DEL ESTUDIANTE

Este nuevo perfil podrá corresponder con un ejemplo de carrera existente en la base de datos o la autovaloración de estudiante. En esta actividad se obtiene la información de los estudiantes y se almacena en un perfil del estudiante:

$$P_e = \{p_1^e, \dots, p_k^e, \dots, p_l^e\} \quad (5)$$

Este perfil estará integrado por un conjunto de atributos:

$$C^e = \{c_1^e, \dots, c_k^e, \dots, c_l^e\} \quad (6)$$

donde  $c_k^e \in S$



### 5.3 FILTRADO DE LAS CARRERAS

En esta actividad se filtran las carreras de acuerdo al perfil del usuario para encontrar cuáles son las más adecuadas para el estudiante.

Con este propósito se calcula la similitud entre perfil de usuario,  $P_e$  y cada carrera  $a_j$  de la base de datos. Para el cálculo de la similitud total se emplea la siguiente expresión:

$$d_j = d(P_e, a_j) = OAG_1(V_{sim}(P_e, a_j)) \tag{7}$$

Donde  $OAG_1$  es un operador(es) de agregación. En este caso se propone el empleo del operador OWA. Un operador OWA es una función  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  de dimensión  $n$  si tiene un vector asociado  $W$  de dimensión  $n$  con  $w_j \in [0, 1]$  y  $\sum_{j=1}^n w_j = 1$ , de forma tal que:

$$F(a_1, a_2, \dots, a_n) = \sum_{j=1}^n w_j b_j \tag{8}$$

donde  $b_j$  es el  $j$ -ésimo más grande de los  $a_j$ . Este operador permite una mayor flexibilidad en el proceso de agregación permitiendo expresar a compensación entre los distintos criterios

$V_{sim}(P_e, a_j) = \{sim(p_1^e, v_1^j), \dots, sim(p_k^e, v_k^j), \dots, sim(p_j^e, v_j^j)\}$  es por su parte un vector que contiene la similitud de todos los atributos del perfil del usuario con respecto a la descripción de la carrera  $a_j$ .

La función  $sim$  calcula la similitud entre los valores de los atributos del perfil de usuario y la carrera,  $a_j$ , es definida como [64]:

$$sim(p_k^e, v_k^j) = 1 - \frac{|\beta_k^{P_e} - \beta_k^{a_j}|}{g} \tag{9}$$

### 5.4 RECOMENDACIÓN

Una vez calculada la similitud entre el perfil del estudiante y el perfil de cada una de las carreras existentes en la base de datos se ordenan las carreras de acuerdo a la similitud obtenida lo cual queda representado por el siguiente vector de similitud

$$D = (d_1, \dots, d_n) \tag{10}$$

Los mejores serán aquellos que mejor satisfagan las necesidades del perfil del estudiante (con la similitud mayor).

## 6 EJEMPLO DEMOSTRATIVO

A continuación se presenta un ejemplo demostrativo, supongamos una base de datos de carreras:

$$A = \{a_1, a_2, a_3, a_4, a_5\}$$

Descrito por el conjunto de atributos

$$C = \{c_1, c_2, c_3, c_4\}$$

Los atributos se valorarán en la siguiente escala lingüística:

No	Etiquetas	Funciones de pertenencia
$s_0$	Cero (C)	(0.0,0.0,0.25)
$s_1$	Positivamente débil (PD)	(0.0,0.25,0.50)
$s_2$	Positivamente medio (PM)	(0.25,0.50,0.75)
$s_3$	Positivamente fuerte (PF)	(0.50,0.75,1)
$s_4$	Positivamente muy fuerte (PMF)	(0.75,1,1)

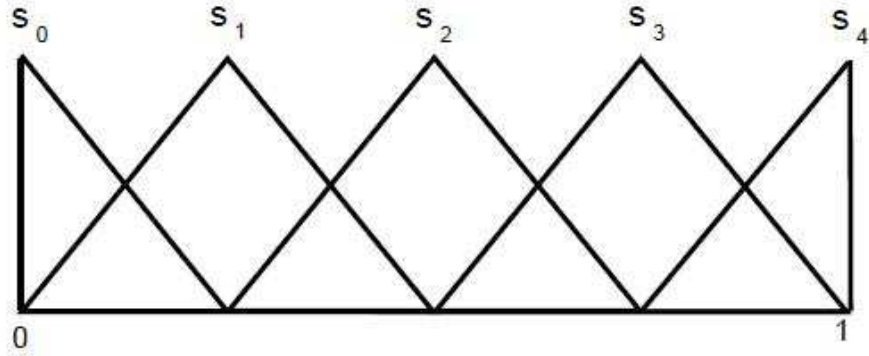


Fig. 5. Conjunto de etiquetas utilizadas

La vista de la base de datos utilizado en este ejemplo, la podemos ver en la Tabla 1

Tabla 1. Base de datos de carreras

	$c_1$	$c_2$	$c_3$	$c_4$
$a_1$	$s_2$	$s_1$	$s_4$	$s_1$
$a_2$	$s_3$	$s_3$	$s_1$	$s_2$
$a_3$	$s_1$	$s_2$	$s_3$	$s_2$
$a_4$	$s_4$	$s_1$	$s_4$	$s_1$
$a_5$	$s_3$	$s_4$	$s_3$	$s_3$

Si un usuario  $u_e$ , desea recibir la recomendaciones del sistema deberá proveer información al mismo de su valoración en las distintas características ya sea con la ayuda de un psicólogo, personalmente o mediante un ejemplo que satisfaga sus preferencias. En este caso:

$$P_e = \{s_2, s_4, s_3, s_2\}$$

El siguiente paso en nuestro ejemplo es el cálculo de la distancia entre el perfil de usuario y las carreras almacenadas en la base de datos.

Tabla 2. Similitud entre los atributos de la base de datos y el perfil de estudiante

	$c_1$	$c_2$	$c_3$	$c_4$
$a_1$	1	0,25	0,75	0,75
$a_2$	0,75	0,75	0,5	1
$a_3$	0,75	0,5	1	1
$a_4$	0,5	0,25	0,75	0,75
$a_5$	0,75	1	1	0,75

El último paso de esta fase es el cálculo de la similitud entre las carreras de la base de datos y el perfil en este caso usando el operador OWA con el siguiente vector de peso (0,4; 0,3; 0,2; 0,1)

Tabla 3. Distancia entre los productos y el perfil de usuario

$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
0,8	0,825	0,9	0,65	0,925

En la fase de recomendación se recomendará aquellas carreras que más se acerquen al perfil del estudiante. Un ordenamiento de las carreras basado en esta comparación sería el siguiente.

$$\{a_5, a_3, a_2, a_1, a_4\}$$

En caso de que el sistema recomendara las dos carreras más cercanas, estas serían las recomendaciones:

$$a_5, a_3$$

## 7 CONCLUSIONES

Las particularidades del proceso de orientación profesional resulta ser un problema de toma de decisiones complejo. A pesar del impacto a lo largo de la vida que puede tener la decisión de qué carrera cursar persisten insuficiencias en el tratamiento del proceso de recomendación de las carreras universitarias. En este trabajo se presentó un modelo de recomendación de carreras universitarias siguiendo el enfoque basado en contenido. El mismo se basa en el empleo de la computación con palabras para la construcción del perfil de usuario y la base de datos y los operadores OWA.

Trabajo futuros estarán relacionados con la inclusión de información de contexto en el modelo de recomendación, la adquisición de características psicológicas del estudiante y modelos de consenso y agregación para la creación de la base de datos a partir de múltiples expertos. Adicionalmente se trabajará en la inclusión de modelos de agregación combinando operadores. Otras áreas de trabajo futuros estarán relacionadas con el manejo de información heterogénea en los modelos y en el desarrollo de una herramienta informática.

## REFERENCIAS

- [1] Calle, A.H. and R.P. González, *Prototipo de un-sistema experto de orientación vocacional (seoriv)*. Artseduca, 2013(5): p. 92-109.
- [2] Mejía, M., *Escoja bien su carrera*. Bogotá: Intermedio. Cáp, 2000. 1.
- [3] Traxler, A.E., *La escuela y las técnicas de la conducción*. 1965, Buenos Aires:: Troquel.
- [4] Miller, T.K. and J.S. Prince, *The future of student affairs*. 1976: Jossey-Bass Incorporated Pub.
- [5] Fuente Gómez, C., *Todos los estudios y carreras*. 1994, Madrid: Planeta.
- [6] Sanchiz Ruiz, M.L., *Modelos de orientación e intervención psicopedagógica*. 2009.
- [7] Grañeras Pastrana, M. and A. Parras Laguna, *Orientación educativa: fundamentos teóricos, modelos institucionales y nuevas perspectivas*. 2012.
- [8] Cedillo, A.C., *La construcción del perfil profesional de orientador y de orientadora: estudio cualitativo basado en la opinión de sus protagonistas en Málaga*. 2010, Universidad de Málaga.
- [9] Oliveros, O. and J.R.G. Bello, *Hacia un nuevo paradigma en orientación vocacional*. Paradigma, 2012. 33(2): p. 127-141.
- [10] GALÁN, A.A. *Montero: falta de orientación vocacional causa la deserción*. 2011 [cited 2014 octubre/6]; Available from: <http://www.lajornadadeorientacion.com.mx/2011/05/19/puebla/edu307.php>.
- [11] Munier, N., *A Strategy for Using Multicriteria Analysis in Decision-making: A Guide for Simple and Complex Environmental Projects*. 2011: Springer.
- [12] Strasser, S.E., C. Ozgur, and D.L. Schroeder, *Selecting a business college major: An analysis of criteria and choice using the analytical hierarchy process*. American Journal of Business, 2002. 17(2): p. 47-56.
- [13] Jamil, R. and S.P. Jarot. *Intelligent decision support system for degree selection using AHP technique*. in *Computer and Communication Engineering (ICCCCE), 2012 International Conference on*. 2012.
- [14] Rana, S., P.K. Dey, and D. Ghosh, *Best engineering college selection through fuzzy multi-criteria decision making approach: a case study*. UNIASCIT Journal, 2012. 2(2): p. 246-256.
- [15] Merigó, J., *New extensions to the OWA operators and its application in decision making*, in *Department of Business Administration, University of Barcelona*. 2008, University of Barcelona: Barcelona.
- [16] Leyva-Vázquez, M., *Modelo de Ayuda a la Toma de Decisiones Basado en Mapas Cognitivos Difusos*. 2013, UCI. Doctor en Ciencias Técnicas: La Habana.
- [17] Pérez-Teruel, K., *Modelo de proceso de logro de consenso en mapas cognitivos difusos para la toma de decisiones en grupo*. 2014, Universidad de las Ciencias Informáticas: La Habana.
- [18] Puente Agueda, C., *Causality in Science*. Pensamiento Matemático, 2011(1): p. 12.
- [19] Leiva, J.L., et al., *Realidad aumentada y sistemas de recomendación grupales: Una nueva perspectiva en sistemas de destinos turísticos*. Estudios y perspectivas en turismo, 2014. 23(1): p. 40-59.
- [20] Lops, P., M. De Gemmis, and G. Semeraro, *Content-based recommender systems: State of the art and trends*, in *Recommender systems handbook*. 2011, Springer. p. 73-105.
- [21] Castellano, E.J. and L. Martínez. *OrieB, A Crs For Academic Orientation Using Qualitative Assessments*. in *e-Learning*. 2008.
- [22] Castellano, E.J., L. Martínez, and M. BARRANCO. *Recomendación de perfiles académicos mediante algoritmos colaborativos basados en el expediente*. in *IADIS Ibero-Americana 2007*.

- [23] Denley, T., *Degree Compass: A Course Recommendation System*. EDUCAUSE Review Online, 2013.
- [24] Bendakir, N. and E. Aïmeur. *Using association rules for course recommendation*. in *Proceedings of the AAAI Workshop on Educational Data Mining*. 2006.
- [25] Lin, C.F., et al., *Data mining for providing a personalized learning path in creativity: An application of decision trees*. *Computers & Education*, 2013. **68**: p. 199-210.
- [26] Whitten, L. and A. Sanders, *Data-Mining and Predicting Academic Success: A Recommendation System for College Students*. *Educational Technology*, 2013. **53**(4): p. 40-42.
- [27] McMahon, M., W. Patton, and M. Watson, *Developing qualitative career assessment processes*. *The Career Development Quarterly*, 2003. **51**(3): p. 194-202.
- [28] Belarmino Rimada Peña, *Inventarios de Orientación Profesional Universitaria*. 2005, México: Trillas.
- [29] Gámez, E., H. Marrero, and J.M. Díaz. *Motivación interpersonal y metas vitales en estudiantes universitarios*. 2008 [cited 2015 febrero/11]; Available from: <http://webpages.ull.es/users/egamez/capituloElenaame08.pdf>.
- [30] Moreno, M. and M. José, *Sistema experto de orientación vocacional-profesional:(un procedimiento informatizado de ayuda)*. 2002: Universidad Complutense de Madrid, Servicio de Publicaciones.
- [31] Ayyub, B. and M.M. Gupta, *Uncertainty analysis in engineering and sciences: fuzzy logic, statistics, and neural network approach*. Vol. 11. 2012: Springer Science & Business Media.
- [32] Lei, N. and S.K. Moon, *A Decision Support System for market-driven product positioning and design*. *Decision Support Systems*, 2015. **69**: p. 82-91.
- [33] Burstein, F. and C. Holsapple, *Handbook on decision support systems 2: variations*. Vol. 2. 2008: Springer Science & Business Media.
- [34] Vaquero, A., et al., *Siete Sistemas informatizados en español para el desarrollo de temas de enseñanza*. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales*, 1987.
- [35] Tolbert, E. and J.R. Sánchez, *Técnicas de asesoramiento en orientación profesional*. 1981: Oikos-tau.
- [36] Ruiz Díaz, J., *La educación imprescindible: enseñar/aprender a vivir en "Educación"*. *Educadores*, 1990.
- [37] Kim, H.D., et al. *Mining causal topics in text data: iterative topic modeling with time series feedback*. in *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*. 2013. ACM.
- [38] Dickerson, J.A., et al., *Creating Metabolic Network Models using Text Mining and Expert Knowledge*. 2003.
- [39] Herrera, F., et al., *Computing with words in decision making: foundations, trends and prospects*. 2009, Springer. p. 337-364.
- [40] Dietmar Jannach, Tutorial: Recommender Systems, in *International Joint Conference on Artificial Intelligence Beijing*, August 4, 2013.
- [41] Cerdón, L.G.P., *Modelos de recomendación con falta de información. Aplicaciones al sector turístico*. 2008, Universidad de Jaén.
- [42] Nunes, M.A.S.N. and S.A. Cerri. *Recommender systems based on human psychological reputation*. in *Recommenders' 06. com-THE PRESENT AND FUTURE OF RECOMMENDER SYSTEMS*. 2006. MyStrands.
- [43] Carrillo, G. and X. Ochoa. *Recomendación de objetos de aprendizaje basada en el perfil del usuario y la información de atención contextualizada» de*. in *LACLO 2013-Octava Conferencia Latinoamericana de Objetos y Tecnologías de Aprendizaje*. 2013.
- [44] Torra, V. and Y. Narukawa, *Modeling decisions: information fusion and aggregation operators*. 2007: Springer.
- [45] Beliakov, G., A. Pradera, and T. Calvo, *Aggregation functions: a guide for practitioners*. 2007: Springer.
- [46] Yager, R.R., *On ordered weighted averaging aggregation operators in multicriteria decisionmaking*. *Systems, Man and Cybernetics, IEEE Transactions on*, 1988. **18**(1): p. 183-190.
- [47] Yager, R.R., J. Kacprzyk, and G. Beliakov, *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*. 2011: Springer.
- [48] Grabisch, M., *The application of fuzzy integrals in multicriteria decision making*. *European Journal of Operational Research*, 1996. **89**(3): p. 445-456.
- [49] Wang, X. and D. Yeung. *Using fuzzy integral to modeling case based reasoning with feature interaction*. in *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*. 2000. IEEE.
- [50] Yager, R.R., *Soft aggregation methods in case based reasoning*. *Applied Intelligence*, 2004. **21**(3): p. 277-288.
- [51] Beliakov, G., T. Calvo, and S. James, *Aggregation of preferences in recommender systems*, in *Recommender systems handbook*. 2011, Springer. p. 705-734.
- [52] Espinilla, M., et al., *A comparative study of heterogeneous decision analysis approaches applied to sustainable energy evaluation*. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2012. **20**(supp01): p. 159-174.
- [53] Bellogín, A., I. Cantador, and P. Castells, *A comparative study of heterogeneous item recommendations in social systems*. *Information Sciences*, 2013. **221**: p. 142-169.

- [54] Denley, T., *Austin Peay State University: Degree Compass*. EDUCAUSE Review Online. Available: <http://www.educause.edu/ero/article/austin-peay-state-university-degree-compass>, 2012.
- [55] Zadeh, L.A., *Fuzzy sets*. Information and Control, 1965. **8**(3): p. 338-353.
- [56] Brio, B.M.d. and A.S. Molina, *Redes Neuronales y Sistemas Borosos*. Segunda Edición ed. 2001: Alfaomega.
- [57] Klir, G.J. and B. Yuan, *Fuzzy sets and fuzzy logic*. 1995: Prentice Hall New Jersey.
- [58] Espinilla Estévez, M., *Nuevos modelos de evaluación sensorial con información lingüística*, in *Departamento de Informática*. 2009, Universidad de Jaén: Jaen.
- [59] Berthold, M.R. and D.J. Hand, *Intelligent Data Analysis: An Introduction*. 2010: Springer.
- [60] Garcia-Cascales, M.S. and M.T. Lamata, *Nueva aproximación al método tópsis difuso con etiquetas lingüísticas*, in *ESTYLF 2010*. 2010: Huelva.
- [61] Herrera, F., et al., *Computing with words in decision making: foundations, trends and prospects*. Fuzzy Optimization and Decision Making, 2009. **8**(4): p. 337-364.
- [62] Tejeda-Lorente, Á., et al., *A quality based recommender system to disseminate information in a university digital library*. Information Sciences, 2014. **261**: p. 52-69.
- [63] Herrera, F. and L. Martínez, *A 2-tuple fuzzy linguistic representation model for computing with words*. Fuzzy Systems, IEEE Transactions on, 2000. **8**(6): p. 746-752.
- [64] Pérez-Teruel, K., M. Leyva-Vázquez, and V. Estrada-Sentí, *Mental Models Consensus Process Using Fuzzy Cognitive Maps and Computing with Words*. Ingeniería y Universidad, 2015. **19**(1): p. 7-22.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.