S-PISNA: UNA HERRAMIENTA DE ANÁLISIS PARA GRAFOS DIFUSOS BASADA EN COMPUTACIÓN CON PALABRAS PARA LA MEJORA DE RECOMENDACIÓN A GRUPOS.

Jorge Castro Gallardo[1], **Salud Jiménez Zafra**[1], **Manuel J. Barranco García**[1], **Luis Martínez López**[1]

[1]*Universidad de Jaén, Campus Las Lagunillas, 23071, Jaén*, jcastro@ujaen.es, saludmjimenez@gmail.com, martin@ujaen.es, barranco@ujaen.es

Resumen

Los sistemas de recomendación a grupos se utilizan para realizar recomendaciones a grupos de usuarios de forma que satisfagan sus necesidades como grupo de entre un conjunto sobrecargado de alternativas. El análisis de las necesidades del grupo puede enfocarse de distintas formas atendiendo a la formación y relación del mismo. Una posibilidad es usar una representación del mismo, como si de una red social se tratase, mediante grafos difusos y así aplicar el Paradigma de Análisis Inteligente de Redes Sociales (*Paradigm of Intelligent Social Network Analysis*, PISNA) para analizar dichas necesidades. En esta contribución se presenta una herramienta software S-PISNA que implementa distintas funciones de PISNA basada en computación con palabras, que nos ayudará a mejorar procesos futuros de la recomendación a grupos.

Palabras Clave: recomendación a grupos, computación con palabras, análisis de redes sociales.

1 INTRODUCCIÓN

Continuamente nos encontramos en situaciones en las que tenemos que elegir una alternativa de entre un gran conjunto de opciones, como por ejemplo, qué película ver [1] o en qué restaurante cenar [2]. El gran número de alternativas disponibles hace que resulte interesante aplicar técnicas que ayuden a los usuarios en su búsqueda del producto más adecuado. Para ello, se utilizan los Sistemas de Recomendación (SR) [2], que personalizan los contenidos que se presentan al usuario atendiendo sus necesidades y preferencias concretas.

[1]www.netflix.com
[2]www.yelp.com

Los sistemas de recomendación descritos anteriormente calculan recomendaciones individuales, es decir, destinadas a un único usuario, pero existen actividades de cierto carácter social en los que resulta interesante obtener recomendaciones para grupos de usuarios [9] como, por ejemplo, la recomendación de películas para un grupo de amigos [11], de actividades turísticas para un grupo de viajeros [5], de noticias [7] o de música [4]. Todas estas actividades tienen en común que los miembros del grupo tienen intereses u objetivos en un momento determinado que han de tenerse en cuenta a la hora de hacer la recomendación.

Es claro que en la recomendación a grupos el éxito de las recomendaciones dependerá de una correcta identificación de las relaciones existentes entre los miembros del grupo. El análisis de las mismas puede ayudar a descubrir patrones de comportamiento en el grupo que permitan mejorar posteriormente las recomendaciones dirigidas al mismo, como por ejemplo que exista un claro líder [10] o que los miembros sean muy similares entre sí [1]. Dichas relaciones pueden representarse mediante un grafo difuso tal que, la intensidad de una relación entre individuos se corresponde con el peso de la arista que los une en el grafo. Este grafo difuso puede ser analizado de forma similar a como han sido analizadas las redes sociales en [13]. En esta contribución modelaremos el grupo de usuarios que demanda una recomendación como un grafo difuso que representa las relaciones entre ellos, de forma similar a las relaciones en una red social (por lo que utilizaremos indistintamente ambos términos). Así, dicha representación puede ser explotada posteriormente aplicando conceptos formales utilizados en el análisis de redes sociales, destacando características interesantes de la misma. Tradicionalmente, los métodos numéricos para análisis de redes sociales dan buenos resultados, pero su interpretabilidad es baja, por lo que el uso de información lingüística para los términos de análisis del grafo difuso facilita la comprensión de los conceptos, de los resultados y el ajuste de las técnicas.

Por ello, en el análisis inteligente de redes sociales [13] (en inglés *Paradigm for Intelligent Social Network Analysis, PISNA*) se utilizan términos lingüísticos. Este paradigma

permite establecer un puente entre la forma de comunicación y razonamiento de los seres humanos, mediante términos lingüísticos, y la de las máquinas, mediante símbolos formales, aplicando computación con palabras [15]. Una característica de este paradigma es que permite reinterpretar las relaciones entre los individuos del grupo aplicando términos lingüísticos, lo que dota a los sistemas que adopten este paradigma en el análisis de redes sociales un grado de flexibilidad mayor. A partir de este análisis se caracteriza el grupo para el que se pretende recomendar, lo que permite aplicar modelos más ajustados a las características concretas del mismo y mejorar las recomendaciones.

En esta contribución se presenta un prototipo de una herramienta para facilitar el análisis basado en PISNA, que permite visualizar la red social del grupo a analizar, reinterpretar dicha red para ajustar las relaciones entre los individuos mediante la computación con palabras, aplicar distintos términos formales para la caracterización de usuarios dentro de la misma e interactuar con los valores de sus parámetros para extraer conocimiento del grupo y así poder mejorar los procesos de recomendación posteriormente.

El resto de esta contribución se estructura de la siguiente manera: en primer lugar, se hace una revisión de los conceptos necesarios, presentando los conceptos utilizados de de sistemas de recomendación a grupos, de grafos difusos, de redes sociales y PISNA. Posteriormente, se presenta la herramienta desarrollada para facilitar el análisis con PISNA. Tras esto, se detalla cómo se puede aplicar PISNA a la mejora de recomendación a grupos. Para finalizar, se detallan las conclusiones y trabajo futuro de esta contribución.

2 PRELIMINARES

En esta sección se exponen los conceptos básicos de recomendación a grupos y posteriormente los conceptos utilizados en el análisis de redes sociales con PISNA.

2.1 Sistemas de recomendación a grupos

En los sistemas de recomendación tradicionales, se utilizan tres fuentes de información distintas: el conjunto de los usuarios ($U = \{u_1, \ldots, u_n\}$), el conjunto de productos ($I = \{i_1, \ldots, i_n\}$) y el conjunto de valoraciones de preferencia de los usuarios sobre los productos ($P \subseteq U \times I \rightarrow D$ con $D = \{1,2,3,4,5\}$). De especial interés es esta última (ver Tabla 1), que describe en qué medida los usuarios se sienten satisfechos con los productos que han valorado.

El problema de la recomendación consiste en encontrar el producto (o conjunto de productos) que maximice la utilidad para el usuario al que se desea recomendar:

$$recomendar(I,u) = \arg\max_{i_j \in I} utilidad(i_j,u) \qquad (1)$$

En recomendación a grupos se añade el concepto de grupo, que es un subconjunto de usuarios del sistema, llamados miembros del grupo ($G = \{m_1, \ldots, m_r\} \subseteq U$). De esta manera, en recomendación a grupos, se pueden utilizar las preferencias individuales de los miembros para generar la recomendación para el grupo.

Tabla 1: Valoraciones de preferencia de los usuarios sobre los productos.

	i_1	$\ldots$	i_j	$\ldots$	i_m
u_1	p_1^1	$\ldots$	p_j^1	$\ldots$	p_m^1
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
u_n	p_1^n	$\ldots$	p_j^n	$\ldots$	p_m^n

El problema de la recomendación a grupos consiste en encontrar el producto (o conjunto de productos) que maximizan la utilidad para el grupo:

$$recomendarGrupo(I,G) = \arg\max_{i_j \in I} [utilidad(i_j,G)] \qquad (2)$$

Dependiendo de cómo el sistema de recomendación a grupos utilice las preferencias individuales de los miembros podemos diferenciar distintas formas de calcular las recomendaciones a grupos [9]:

- Agregar las preferencias individuales para formar un perfil de grupo y recomendar como si se tratase de valoraciones individuales,
- Calcular recomendaciones individuales y agregarlas,
- Construir un modelo asociado al grupo.

2.2 Análisis de redes sociales

Dado que queremos aplicar modelos y conceptos propios de las redes sociales para analizar los grupos, en esta sección revisamos distintos conceptos básicos, su representación mediante grafos difusos y la extracción de conocimiento aplicando PISNA.

A) Redes sociales

Las redes sociales se definen como una estructura que contiene enlaces entre individuos, en la que dichos enlaces representan las relaciones interpersonales que se dan entre los individuos de la misma [6, 12]. Las redes sociales pueden ser aplicadas en problemas en los que la información se puede modelar mediante una estructura de red.

Tradicionalmente, las redes sociales han sido una herramienta para la extracción de conocimiento sobre los individuos y sus relaciones [8]. Analizando una red social se pueden obtener características interesantes, como la obtención de los líderes de la misma (usuarios influyentes) o la cohesión de un grupo de individuos de la misma.

B) Representación: Grafos difusos

Como se ha comentado anteriormente, entre los miembros de un grupo existen relaciones que pueden ser de mayor o menor intensidad, por lo que para modelarlas, una representación adecuada son los grafos difusos [13]. Un grafo difuso se compone de nodos, que corresponderán con individuos (miembros del grupo) y de aristas que unen dichos nodos, que representan las relaciones entre los mismos. Dichas aristas tienen asociado un valor en el intervalo [0,1] que representa la intensidad de la relación entre los nodos, es decir, la intensidad de la relación entre miembros. Formalmente un grafo difuso es una tupla $G =< V, E, R >$, en la que V es el conjunto de nodos del grafo, E es el conjunto de aristas del grafo y R es una relación $R : V \times V \rightarrow [0,1]$, que representa la intensidad de la relación entre dos nodos. Cuando la relación entre dos nodos es mayor que cero, se dice que los nodos están conectados.

Un concepto frecuentemente utilizado en grafos son los caminos. Un camino consiste en una sucesión de nodos que conecta el nodo inicial con el nodo final. Una propiedad interesante de estos es la longitud de un camino, que en un grafo difuso se define como:

$$\rho = v_1, v_2, \ldots, v_r \tag{3}$$

$$Longitud(\rho) = \sum_{j=1}^{r-1} \frac{1}{R(x_j, x_{j+1})}, R(x_j, x_{j+1}) > 0 \tag{4}$$

donde ρ es la sucesión de nodos y $R(x_j, x_{j+1})$ es la intensidad de la relación entre el nodo x_j y x_{j+1}.

Entre dos nodos x_i y x_j se puede calcular la distancia, como la longitud del camino más corto que los une.

Una operación interesante en grafos difusos es la composición de relaciones (ver Ecuación 5). Como la composición de relaciones es asociativa, se puede generalizar el concepto a la composición de la relación R consigo misma k veces (ver Ecuación 6), por lo que la composición de relaciones en grafos difusos resulta útil para conocer la intensidad de la relación entre dos individuos (teniendo en cuenta un número dado de saltos entre los individuos). En un grafo difuso, el valor de $R^k(x_i, x_j)$ es la intensidad máxima de los posibles caminos, de hasta k enlaces, entre x_i y x_j.

$$R_1 \diamond R_2(x,y) = \max_{y \in V} [\min(R_1(x,y), R_2(y,z)] \quad x,y,z \in V \tag{5}$$

$$R^k = \overbrace{R \diamond R \diamond \cdots \diamond R}^{k \text{ veces}} \tag{6}$$

Otro concepto utilizado en redes sociales es la detección de líderes. Se puede calcular de distintas formas, una de ellas es calculando la centralidad de los nodos. Cuanto mayor sea el valor de centralidad de un nodo, más relaciones y más fuertes serán sus relaciones con el resto de nodos, por lo que dicho nodo será un líder en la red social. Usando la representación con un grafo difuso de la red social, se puede calcular el valor de centralidad (ver Ecuación 7), que se puede generalizar teniendo en cuenta el número de enlaces (k) que se consideran en su cómputo (ver Ecuación 8)).

$$C(x_i) = \sum_{j \neq i}^{n} R(x_i, x_j) \tag{7}$$

$$C^k(x_i) = \sum_{j \neq i}^{n} R^k(x_i, x_j) \tag{8}$$

C) Paradigma de Análisis Inteligente de Redes Sociales

R. R. Yager [13] presentó PISNA para facilitar el análisis de redes sociales a través de técnicas de computación con palabras. En la Figura 1 se muestra el esquema que este paradigma sigue, sirviendo como puente entre la definición mediante la teoría de conjuntos de los conceptos de análisis de redes sociales y una representación lingüística de los mismos que facilite la comprensión de los conceptos, de los resultados y su ajuste. Mediante la computación con palabras se permite introducir cierta flexibilidad a los conceptos utilizados en el análisis, dando como resultado la obtención de conceptos formales más sencillos de entender y ajustables, a través de la aplicación de términos lingüísticos. A continuación se explican los términos lingüísticos representados según el enfoque lingüístico difuso [14].

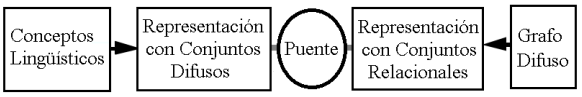

Figura 1: Paradigma de Análisis Inteligente de Redes Sociales (PISNA)

De entre los conceptos lingüísticos definidos en PISNA podemos destacar la intensidad de la relación entre nodos, que se puede definir con término *"Fuerte"*, que indica en qué medida una relación entre nodos satisface el término *"Fuerte"*. Este término define su semántica usando los parámetros α y β, que pueden ser modificados para hacer el término más o menos restrictivo (Ver Figura 2).

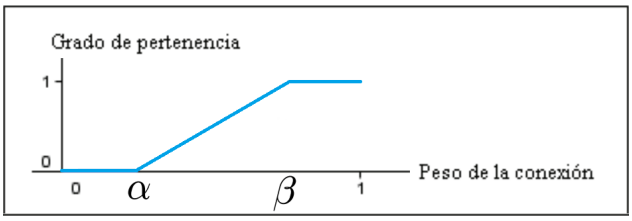

Figura 2: Representación del término *"Fuerte"*, aplicado sobre la intensidad de la relación entre dos nodos.

Así, si aplicamos este término lingüístico, el concepto de

centralidad de un nodo (ver Ecuación 8) queda como sigue:

$$C^k(x_i) = \sum_{j \neq i}^{n} S(R^k(x_i, x_j)) \quad (9)$$

Otra propiedad interesante de los grafos a definir lingüísticamente es la longitud de un camino. Por lo tanto, un término lingüístico interesante será el término *"Corto"*, cuya semántica puede definirse como se muestra en la Figura 3, para analizar la longitud de los caminos.

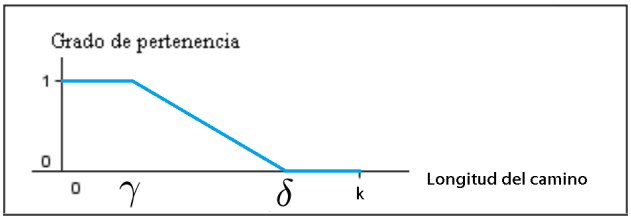

Figura 3: Representación del término *"Corto"*, aplicado sobre la longitud de un camino.

Ambos términos lingüísticos pueden ser usados conjuntamente para buscar el camino más fuerte y corto entre dos nodos. Así, este concepto se define como:

$$C_1(x_i, x_j) = \max_{k=1 \ to \ t} [SH(k) \wedge S(R^k(x_i, x_j))] \quad (10)$$

3 S-PISNA

Los conceptos revisados anteriormente se pueden aplicar en el análisis de grupos para mejorar sus recomendaciones. Por lo que, para facilitar el análisis y comprobar su funcionamiento, hemos desarrollado S-PISNA, un prototipo inicial de un software basado en PISNA que permite analizar grupos siguiendo este paradigma (ver Figura 1), visualizando los datos iniciales y los resultados intermedios.

A continuación, se detalla la arquitectura de la herramienta y las funcionalidades principales que proporciona.

3.1 ARQUITECTURA

Para la implementación del prototipo de S-PISNA se ha utilizado el lenguaje de programación Java[3], por su portabilidad entre distintas plataformas. En la Figura 4 se muestra la arquitectura del mismo. La interfaz gráfica de usuario ha sido implementada con Swing y para el manejo y, concretemente para la visualización de grafos difusos, se ha utilizado la biblioteca *Java Universal Network/Graph Framework* JUNG[4]. Se ha utilizado una base de datos MySQL para almacenar los datos de los miembros del grupo.

[3]www.oracle.com/technetwork/java

[4]jung.sourceforge.net/

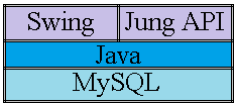

Figura 4: Arquitectura de S-PISNA

3.2 FUNCIONALIDAD

S-PISNA implementa los conceptos de PISNA para el análisis de grupos, por lo que facilita la definición de conceptos lingüísticos, el cálculo de caminos entre individuos y la obtención de los líderes de la red. A continuación se detalla cada una de estas funcionalidades.

A) DEFINICIÓN DE CONCEPTOS

Una característica que ofrece la herramienta es la posibilidad de definir los términos lingüísticos que se aplican a las propiedades relevantes del grafo.

Sobre la intensidad de la relación entre miembros se define el concepto *"Fuerte"* (ver Figura 2) con los valores de α y β que se desee experimentar. El concepto *"Fuerte"* definido modificará la intensidad de la relación entre los individuos para reinterpretarla de manera que se mantendrán aquellas relaciones con una intensidad que se considere *"Fuerte"*, según los valores de α y β. En la interfaz de S-PISNA, se puede definir el concepto *"Fuerte"*, en el panel A de la Figura 5. De manera similar se puede definir el concepto *"Corto"* sobre la longitud de los caminos entre miembros, en el panel B de la Figura 5.

Una vez hecho esto, en la parte superior derecha de la interfaz se muestra el grafo original y en la parte inferior derecha, el grafo cuyos pesos han sido modificados por el término *"Fuerte"*. Dicho grafo modificado se utiliza para calcular los líderes (ver Ecuación 9), cuyos parámetros se especifican en la parte izquierda de la interfaz. En la parte inferior izquierda, se muestra un resumen del cálculo, mostrando el valor de centralidad de los líderes calculados.

B) CAMINOS

La herramienta también facilita el cálculo de caminos entre nodos de la red. Estas funcionalidades hacen uso de los conceptos *"Fuerte"* y *"Corto"*, definidos anteriormente. El prototipo pone dispone de tres formas de calcular caminos entre individuos: cálculo de camino mínimo, camino más fuerte y camino mínimo más fuerte (ver Ecuación 10). Por medio de estas funcionalidades se puede analizar el grupo de una forma más ligada a cada individuo, estudiando si hay conexión con otros miembros y sus características.

En los paneles C, D y E de la Figura 5 se puede realizar el cálculo del camino mínimo, camino más fuerte y camino mínimo más fuerte, respectivamente. Para calcular los caminos se debe especificar el nodo inicial y final, y la herra-

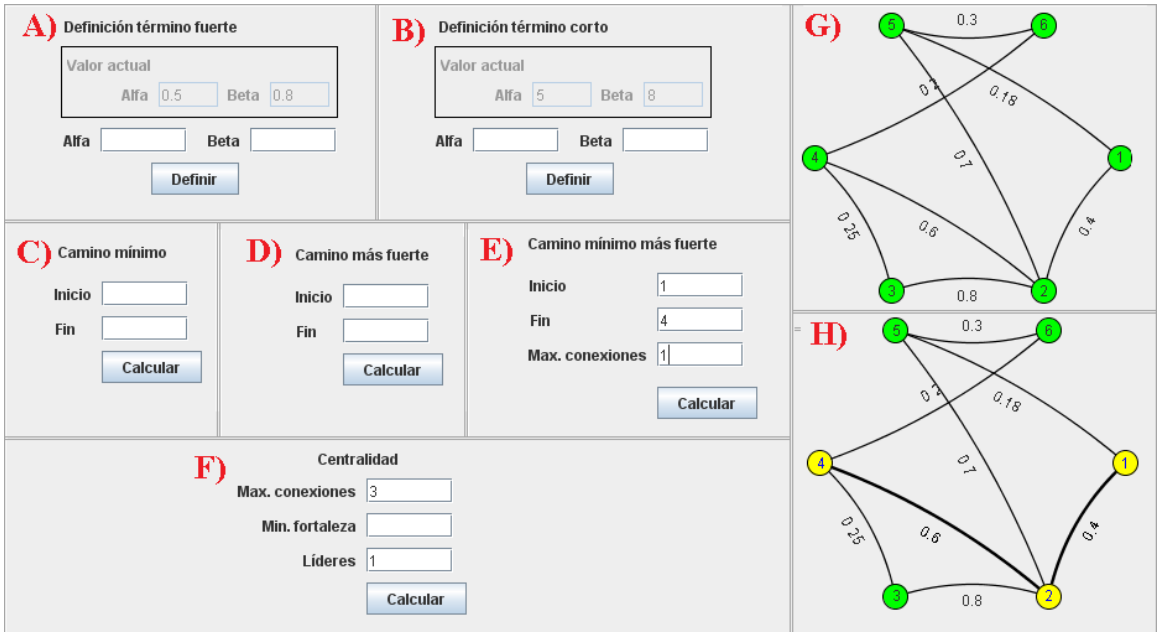

Figura 5: Interfaz de S-PISNA, que permite la visualización de redes y resultados intermedios de PISNA.

mienta calcula todos los caminos posibles. Una vez hecho esto, los ordena según su longitud, según su fortaleza o según la Ecuación 10, dependiendo del cálculo que se desee realizar. Para este último, se debe especificar también el número máximo de conexiones que el camino puede tener. El camino que más se ajusta al cálculo especificado se muestra sobre el grafo mostrado en el panel H de la Figura 5.

C) LÍDER

En el análisis de la red para recomendación a grupos se pueden aplicar métodos de detección de líderes. En la herramienta se permite analizar el grafo para mostrar el nodo con mayor valor de centralidad (ver Ecuación 9), que corresponderá con el líder del grupo.

Como se ha comentado anteriormente, se puede modificar la definición de los conceptos. Para el concepto *"Fuerte"*, se han elegido los valores $\alpha = 0.5$ y $\beta = 0.8$. La definición del término fuerte es utilizada por el cálculo de líderes, que se encuentra en el panel F de la Figura 5. Para el cálculo de líderes es necesario especificar el número máximo de conexiones a tener en cuenta (corresponde con el valor k de la Ecuación 9), la fortaleza mínima de las conexiones y el número de líderes que se desea mostrar en el grafo del panel H de la Figura 5.

4 APLICACIÓN EN RECOMENDACIÓN A GRUPOS

En esta sección se muestra la idea de como se pretende usar PISNA y S-PISNA para el análisis de grupos en sistemas de recomendación.

4.1 INTEGRACIÓN DE PISNA EN SISTEMAS DE RECOMENDACIÓN A GRUPOS

En la Figura 6 se muestra un diagrama de la arquitectura que se utilizará para la mejora de la recomendación a grupos (ver Tabla 1). A partir de la información del grupo, se calcula su grafo difuso, que es analizado para extraer características del grupo. El sistema de recomendación a grupos utiliza las preferencias individuales de los usuarios, el conjunto de usuarios que componen el grupo, su grafo difuso y las características del mismo para recomendar.

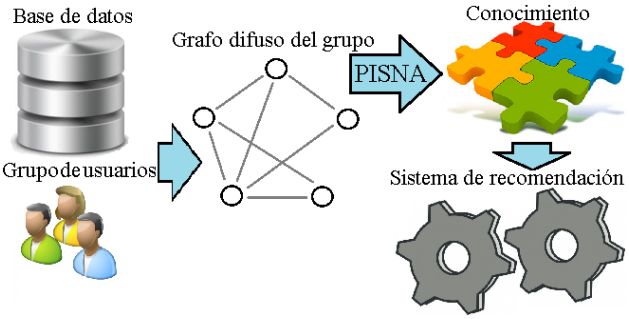

Figura 6: Sistema de recomendación a grupos con PISNA.

4.2 RECOMENDACIÓN A GRUPOS CON PISNA

El conocimiento extraído por PISNA describe las características concretas del grupo, por lo que puede ser utilizado

para la mejora de la recomendación a grupos. Una manera de utilizar esta información es aplicarla para ajustar los parámetros del sistema de recomendación a grupos que se utiliza para recomendar, de manera que puedan ajustarse las recomendaciones a cada grupo, atendiendo a sus características. Por ejemplo, se pueden utilizar distintas técnicas de recomendación a grupos atendiendo a la cohesión del mismo [1], ya que la cohesión indicará cómo de parecidas son las preferencias de los miembros del grupo.

Otro ejemplo de aplicación es diseñar un sistema que permita detectar si en el grupo existe un claro líder que tenga una estrecha relación con el resto de miembros. Aprovechando esta característica, se puede simplificar el proceso de recomendación, transformando el problema en un cálculo de recomendación a individuos [3, 10] calculando las recomendaciones para el líder y presentándola como recomendación para el grupo.

5 CONCLUSIONES

En este trabajo se ha presentado S-PISNA, un prototipo de herramienta software que implementa el paradigma de análisis inteligente de redes sociales. S-PISNA permite analizar interactivamente el grafo difuso asociado a un grupo, lo que ayudará a ajustar la recomendación a grupos por medio del análisis de sus características concretas.

También se ha presentado el modelo que se utilizará para aplicar PISNA en la mejora de la recomendación a grupos. Inicialmente, se tienen las valoraciones individuales de los miembros del grupo de usuarios que desean obtener recomendaciones. A partir de dicha información, se calcula la red social de los mismos comparando sus preferencias y ésta es analizada para obtener características que permitan definir cómo calcular las recomendaciones al grupo de manera que se ajusten más a sus características concretas.

Como trabajo futuro, se pretende seguir la arquitectura propuesta para aplicar PISNA a la recomendación a grupos, utilizando el prototipo presentado (S-PISNA) para el análisis de los grupos y mejorar el proceso de recomendación con el conocimiento extraido. También se pretende presentar un sistema de recomendación a grupos que siga esta arquitectura, evaluando la mejora aportada sobre las técnicas de recomendación a grupos actuales.

Agradecimientos

Esta contribución está financiada por los proyectos de investigación TIN2012-31263, AGR-6581 y UJA2011/12/23.

Referencias

[1] L. Ardissono, A. Goy, G. Petrone, M. Segnan, P. Torasso: Intrigue: personalized recommendation of tourist attractions for desktop and hand held devices. Applied Artificial Intell. 17, no. 8-9 : 687-714 (2003).

[2] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez: Recommender systems survey. Knowledge-Based Systems, 46: 109-132 (2013)

[3] B. Cédric, A. Brun, A. Aghasaryan, M. Bouzid, J. Picault, C. Senot: Topology of communities for the collaborative recommendations to groups. Information Systems and Economic Intelligence (2010).

[4] D. L. Chao, J. Balthrop, S. Forrest: Adaptive Radio: Achieving consensus using negative preferences. Technical Report TR-CS-2004-08, Dept. of Computer Science, The University of New Mexico (2004).

[5] I. García, L. Sebastia, E. Onaindia: On the design of individual and group recommender systems for tourism. Expert Systems with Applications 38.6: 7683-7692 (2011).

[6] C. Kadushin: Understanding Social Networks: Theories, Concepts, and Findings. Oxford University Press, USA (2011).

[7] J. K. Kim, H. K. Kim, H. Y. Oha, Y. U. Ryub: A group recommendation system for online communities. International Journal of Information Management, 30: 211-219 (2010)

[8] D. Knoke, S. Yang: Social network analysis. Vol. 154. Sage, 2008.

[9] J. Masthoff: Group recommender systems: Combining individual models. In Recommender Systems Handbook, pp. 677-702. Springer US, 2011.

[10] L. Morris, L. Hulbert, D. Abrams: An experimental investigation of group members' perceived influence over leader decisions. Group Dynamics: Theory, Research, and Practice 4, no. 2: 157 (2000).

[11] M. O'connor, D. Cosley, J. A. Konstan, and J. Riedl: PolyLens: a recommender system for groups of users. In ECSCW 2001, pp. 199-218. Springer, NL (2002).

[12] P. O'Reilly: Methodological issues in social support and social network research. Social Science and Medicine 26, 863-873 (1988).

[13] R. R. Yager: Intelligent social network analysis using granular computing. Int. Journal of Intelligent Systems, 23: 1197-1219. doi: 10.1002/int.20314 (2008).

[14] L. A. Zadeh: Fuzzy logic and approximate reasoning. Synthese 30, no. 3-4: 407-428 (1975).

[15] L. A. Zadeh: Fuzzy logic = computing with words. Fuzzy Sys., IEEE Trans. on 4, no. 2 (1996): 103-111.